IW 7697820

THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office

November 01, 2018

THIS IS TO CERTIFY THAT ANNEXED IS A TRUE COPY FROM THE RECORDS OF THIS OFFICE OF THE FILE WRAPPER AND CONTENTS OF:

APPLICATION NUMBER: ***09/736,825***
FILING DATE: ***December 14, 2000***
PATENT NUMBER: ***6,654,507***
ISSUE DATE: ***November 25, 2003***

By Authority of the
Under Secretary of Commerce for Intellectual Property
and Director of the United States Patent and Trademark Office

M. Tarver

M. TARVER
Certifying Officer
PART (2) OF (2) PART(S)

basis of a prediction error amount such that the extraction result obtained by said second object extraction section is used as an object region when the prediction error caused by said second object extraction section falls within a predetermined range, and the extraction result obtained by said first object extraction section is used as an object region when the prediction error exceeds the predetermined range.

15. The apparatus according to claim 11, wherein said second object extraction section performs inter-frame prediction in a sequence different from an input frame sequence such that a frame interval between the reference frame and the current frame as the object extraction target is set to not less than a predetermined number of frames.

16. An object extraction apparatus comprising:

a picture input section which inputs moving picture data and shape data representing an object region on a predetermined frame of a plurality of frames corresponding to the moving picture data;

a segmentation section which segments a current frame into a plurality of blocks;

a search section which searches a reference frame for a similar block, for each of the blocks, which is similar in figure represented by picture data to a current block and is larger in area than the current block, the reference frame being temporally different from the current frame;

a paste section which pastes shape data obtained by extracting and reducing shape data of each similar block from the reference frame on each block of the current frame; and

an output section which outputs the pasted shaped data as shape data of the current frame.

17. An object extraction apparatus comprising:

a picture input section which inputs picture data representing a picture including at least one object and a background surrounding the object and shape data representing the object;

a setting section which sets blocks on a contour portion of the object, and searches for a similar block, for each of the blocks, which is similar in graphic figure represented by the picture data to each block and is larger than the block, from the same picture to obtain a plurality of similar blocks;

a replacement section which replaces the shape data of each of the blocks with reduced shape data obtained by reducing the shape data of each of the similar blocks;

a repeat section which repeats the replacement by a predetermined number of times; and

an output section which outputs shape data obtained by repeating the replacement as corrected shape data.

18. The apparatus according to claim 17, further comprising a second repeat section which repeats the searching for the similar block and the predetermined number of replacements of the shape data by a plurality of times while decreasing a block size every repetition.

19. A method for extracting a moving object from an input moving picture, comprising the steps of:

determining a first background region common to a current frame containing a target object to be extracted from a moving picture signal and a first reference frame that temporally differs from the current frame on the basis of a difference between the current frame and the first reference frame;

determining a second background region common to the current frame and a second reference frame that temporally differs from the current frame on the basis of a difference between the current frame and the second reference frame; and

extracting a region, in a picture on the current frame, which belongs to neither the first background region nor the second background region as an object region, the first background region and the second background region indicating a background in the input moving picture.

20. The method according to claim 19, which comprises a step of determining pixels of the current frame as the object region when pixels of one of the first and second reference frames belongs to the object region, and determining the pixels of the current frame as the background region when the pixels of one of the first and second reference frames belongs to the background region, using a predetermined shape of the object of one of the first and second reference frames in a case of that the difference between the pixels of the current frame and the pixels the one of the first and second reference frames is small.

21. The method according to claim 20, wherein the step of determining the pixels of the current frame uses the predetermined shape of the object, when the shape of the object of one of the first and second reference frames has already been extracted, and a shape of the block of the one of the first and second reference frames which is created from the frame, from which the shape of the object has been extracted, by a block matching method, when the object region is not extracted.

22. The method according to claim 19, further comprising a step of correcting motion of a background on one of the first and second reference frames or the current frame such that the motion of the background between each of the first and second reference frames and the current frame becomes relatively zero.

23. The method according to claim 19, wherein the background region determining step includes determining the common background region using a predetermined threshold value.

24. The method according to claim 23, wherein the background region determining step includes setting the threshold value to a larger value than the predetermined threshold value when the difference of the current frame is larger than a predetermined value, and to a smaller value than it when the difference is smaller.

25. The method according to claim 23, wherein the background region determining step includes dividing the current frame into a plurality of regions, measuring a difference between each of the regions and each of corresponding regions of one of the first and second reference frames, and setting the threshold value to a larger value than a predetermined value when the difference is larger than a predetermined value and to a smaller value when it is smaller.

26. The method according to claim 19, further comprising the step of predicting a position or shape of the object on the current frame from a frame from which the object region has already been extracted, and selecting the first and second reference frames to be used by said background region determining step on the basis of the position or shape of the object on the current frame which is predicted by said predicting step.

27. The method according to claim 19, which further comprises a step of setting a figure surrounding the target object on an initial frame of the moving picture signal, and a step of setting on one of the first and second reference frames a figure surrounding a region on each input frame of

the moving picture signal which corresponds to an image inside figure of one of the first and second reference frames that temporally differs from the input frame on the basis of a correlation between the input frame and the image inside figure, and said object region extracting step extracts a region, in the image inside figure, which belongs to neither the first background region nor the second background region as an object region.

28. The method according to claim **26**, wherein said setting step sets a figure surrounding the target object on the basis of an external input.

29. A method for extracting an object from an input picture comprising the steps of:

setting a figure surrounding a target object on an initial frame of a moving picture signal;

setting on an input frame a figure surrounding a region on the input frame of the moving picture signal and corresponding to an image inside figure of a reference frame that temporally differs from the input frame on the basis of a correlation between the input frame and the image inside figure;

determining a first background region common to a current frame as an object extraction target and a first reference frame that temporally differs from the current frame on the basis of a difference between the current frame and the first reference frame, and determining a second background region common to the current frame and a second reference frame that temporally differs from the current frame on the basis of a difference between the current frame and the second reference frame, the first background region and the second background region indicating a background in the input picture

extracting a region, in the image inside figure of the current frame, which belongs to neither the first background region nor the second background region, as an object region;

extracting an object region from the image inside figure on the current frame as the object extraction target by using a method different from that used by said extracting steps; and

selectively switching the extracting steps.

30. The method according to claim **29**, further comprising the step of extracting a feature value of a picture in at least a partial region of the current frame as the object extraction target from the current frame, and wherein said switching step selectively switches said extracting steps on the basis of the extracted feature value.

31. The method according to claim **29**, wherein said extracting an object region step includes predicting a position or shape of the object on the current frame as the object, using a frame, from which an object region has already been extracted, as a reference frame, to predict a position or shape of the object on the current frame as an object extraction target from the reference frame.

32. The method according to claim **31**, wherein said extracting steps are selectively switched and used in units of blocks of each frame on the basis of a prediction error amount such that the extraction result obtained by said extracting an object region step is used as an object region when the prediction error caused by said extracting an object region step falls within a predetermined range, and the extraction result obtained by said extracting a region step is used as an object region when the prediction error exceeds the predetermined range.

33. The method according to claim **29**, wherein said extracting an object region step performs inter-frame prediction in a sequence different from an input frame sequence such that a frame interval between a reference frame and the current frame as the object extraction target is set to not less than a predetermined number of frames.

34. A method of extracting an object from an input picture comprising the steps of:

inputting moving picture data and shape data representing an object region on a predetermined frame of a plurality of frames constituting the moving picture data;

segmenting a currently processed frame into a plurality of blocks;

searching for a similar block, for each of the blocks, which is similar in figure represented by picture data to the currently processed block and is larger in area than the currently processed block, from the reference frame;

pasting shape data obtained by extracting and reducing shape data of each similar block from the reference frame on each block of the currently processed frame; and

outputting the pasted shaped data as shape data of the currently processed frame.

35. A method of extracting an object from an input picture comprising:

inputting picture data and shape data representing an object region on the picture;

setting blocks on a contour portion of the shape data;

searching for a similar block, for each of the blocks, which is similar in graphic figure represented by the picture data to each block and is larger than the block, from the same picture;

replacing the shape data of each of the blocks with shape data obtained by reducing the shape data of each of the similar blocks;

repeating the replacement by a predetermined number of times; and

outputting shape data obtained by repeating the replacement as corrected shape data.

36. The method according to claim **35**, further comprising a step of secondly repeating the searching for the similar block and the predetermined number of replacements of the shape data by a plurality of times while decreasing a block size every repetition.

* * * * *

OIPE JC73
MAR 0 5 2003
PATENT & TRADEMARK OFFICE

#6

Customer No. 01333

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of:

Jiebo Luo

AUTOMATICALLY PRODUCING AN IMAGE OF A PORTION OF A PHOTOGRAPHIC IMAGE

Serial No. 09/736,825

Filed 14 December 2000

Group Art Unit: 2625

Examiner: Aaron W. Carter

Commissioner for Patents
Washington, D.C. 20231

RECEIVED
MAR 0 6 2003
Technology Center 2600

Sir:

ASSOCIATE POWER OF ATTORNEY

I hereby appoint Patrick J. Finnan, Registration No. 39,189, and Andrew J. Aldag, Registration No. 40,483, whose address is Edell, Shaprio, Finnan & Lytle, LLC, 1901 Research Boulevard, Suite 400, Rockville, Maryland 20850-3164, as associate attorneys in the above-entitled application with full power to prosecute this application and to transact all business in the Patent Office connected therewith.

Please address all written communications to Thomas H. Close, at Eastman Kodak Company, Patent Legal Staff, Rochester, New York 14650 - 2201 and direct all telephone communications to Raymond L. Owens at 585-477-4653.

Respectfully submitted,

Attorney for Applicants
Registration No. 22,363

Raymond L. Owens/phw
Rochester, NY 14650
Telephone: (585) 477-4653
Facsimile: (585) 477-4646

A:\81595\poweraty.doc

2625

OIPE JC73
MAR 0 5 2003
PATENT & TRADEMARK OFFICE

Atty Docket: 81595RLO
0505.0018C

EDELL, SHAPIRO & FINNAN, LLC
1901 Research Boulevard
Suite 400
Rockville, Maryland 20850-3164
(301) 424-3640

In re the PATENT application of:

Jiebo LUO

Examiner: Aaron W. Carter

Serial No.: 09/736,825

Group Art Unit: 2625

Filed: December 14, 2000

For: Automatically Producing an Image of a Portion of a Photographic Image

RECEIVED
MAR 0 6 2003
Technology Center 2600

COMMISSIONER FOR PATENTS
Washington, D. C. 20231

TRANSMITTAL LETTER

Sir:

Transmitted herewith for filing in the subject application is an Associate Power of Attorney.

Respectfully submitted,

Patrick J. Finnan
Registration No. 39,189

Hand-delivered: 3/5/2003

computer storage medium having instructions stored therein causing one or more computers to perform the method of claim 1.

15. (Currently Amended) A method of producing an image of a portion of at least a portion of a photographic image onto a photographic receiver, comprising ~~the steps of~~:

a) receiving a digital image corresponding to the photographic image, the digital image ~~comprising~~ including pixels;

b) computing a belief map of the digital image~~,~~ by using the pixels of the digital image to determine a series of features~~,~~ and using such features to assign ~~the~~ a probability of ~~the~~ a location of a main subject of the digital image in the belief map;

c) determining a crop window having a shape factor and a zoom factor, the shape and the zoom ~~factor~~ factors determining a size of the crop window; and

d) locating ~~the~~ a relative optical position of a photographic image, a lens assembly, and a photographic receiver in response to the belief map and illuminating a portion of the photographic image of high subject content to produce an image of such portion ~~onto~~ on the photographic receiver.

16. (Currently Amended) The method of claim 15 wherein ~~step c) further comprises the steps of:~~ determining a crop window includes

i) computing a weighted center-of-mass of the belief map, the weighted center-of-mass weighted by the belief values of the belief map;

ii) computing weighted central moments of the belief map, relative to the center-of-mass and weighted by a weighting function of each belief value of the belief map;

iii) computing an effective rectangular bounding box according to the central moments; and

iv) determining a crop window having a shape factor and a zoom factor, the shape and the zoom factor factors determining a size of the crop window.

A3

17. (Currently Amended) The method of claim 15 wherein step d) further comprises the steps of: locating the relative optical position of a photographic image, a lens assembly, and a photographic receiver includes

i) selecting an initial position of the crop window at a location which includes the center-of-mass;

ii) using the belief values corresponding to the crop window to select the position of the crop window to include a portion of the image of high subject content in response to the belief map; and

iii) cropping the digital image according to the position of the crop window.

18. (Currently Amended) The method of claim 16 wherein ~~step d) further comprises the steps of:~~ locating the relative optical position of a photographic image, a lens assembly, and a photographic receiver includes

i) selecting a crop window of a rectangular shape and of an identical aspect ratio to the ~~(uncropped)~~ digital image; and

ii) selecting a zoom factor to determine the size of the crop window such that ~~it~~ the crop window encompasses the effective bounding box.

19. (Currently Amended) The method of claim 16 wherein the weighting function ~~in step b)~~ of computing weighted central moments of the belief map is a linear weighting function.

20. (Currently Amended) The method of claim 16 wherein the weighting function ~~in step b)~~ of computing weighted central moments of the belief map is a constant function.

21. (Currently Amended) The method of claim 17 wherein ~~step b) further comprises the steps of:~~ computing a belief map of the digital image includes

i) calculating a subject content index value for the crop window derived from the belief values;

ii) following a positioning ~~procedure~~ <u>process</u> of repeating ~~step i)~~ <u>selecting an initial position of the crop window at a location which includes the center of the mass</u> for at least two positions of the crop window; and

iii) using the subject content index values to select the crop window position.

---

22. (Original) The method of claim 15 wherein the crop window is completely within the digital image.

23. (Currently Amended) The method of claim 16 wherein ~~step b) further comprises the step of performing a~~ <u>computing a belief map of the digital image includes</u> clustering of the belief map to identify at least a cluster of highest belief values corresponding to <u>the</u> main subject, a cluster of intermediate belief values corresponding to secondary subjects, and a cluster of lowest belief values corresponding to ~~the~~ background.

24. (Currently Amended) The method of claim 23 wherein ~~said~~ clustering includes setting ~~said~~ background portions to a zero belief value.

25. (Currently Amended) The method of claim 19 further comprising positioning ~~said~~ <u>the</u> crop window such that the subject content index <u>value</u> of ~~said~~ <u>the</u> crop window is at an optimum.

26. (Currently Amended) The method of claim 17 further comprising positioning ~~said~~ the crop window such that ~~said~~ the crop window includes all of ~~said~~ main subject cluster.

27. (Currently Amended) The method of claim 26 further comprising positioning ~~said~~ the crop window to include a buffer around ~~said~~ main subject cluster.

28. (Currently Amended) A computer storage product having at least one computer storage medium having instructions stored therein causing one or more computers to perform the method of claim ~~1~~ 15.

Amendments to the Abstract:

A method of producing an image of at least a portion of a digital image, the digital image comprising pixels, comprising the steps of: that includes pixels includes

a) computing a belief map of the digital image, by using the pixels of the digital image to determine a series of features, and using such features to assign the probability of the location of a main subject of the digital image in the belief map; b) determining a crop window having a shape and a zoom factor, the shape and zoom factor determining which determine a size of the crop window; and c) cropping the digital image to include a portion of the image of high subject content in response to the belief map and the crop window.

A5

REMARKS

Favorable reconsideration of this application in view of the above amendments and the following remarks is respectfully requested. By this amendment, claims 1-7, 9-13, 15-21 and 23-28 have been amended. Applicants submit that no new matter has been added, and formal acknowledgement of such is respectfully requested. Currently, claims 1-28 are pending of which claims 1 and 15 are independent.

The Examiner is thanked for his indication of allowability of claims 2, 4-7, 9-11, 16, 18-21, and 23-25. Formal notice of such is solicited.

Claim 28 was rejected under 35 USC 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as his invention.. Claim 28 has been amended to change "claim 1" to --claim 15--, as suggested by the Examiner. Applicant submits that the amendment to claim 28 addresses the Examiner's concern, and overcomes this rejection. Accordingly, withdrawal of this rejection is respectfully requested.

Claims 1, 3, 8, 12-15, 17, 22, and 26-28 were rejected under 35 USC 102(e) as anticipated by Jia (U.S. Patent No. 6,430,320). This rejection is respectfully traversed.

Claims 1 and 15 are independent and claims 3, 8, and 12-14, and claims 17, 22, and 26-28 depend therefrom, respectively.

In general, this application relates to automatically cropping a digital image without manual techniques, but rather based on scene content as indicated by a belief map. Generally, a method of cropping a digital image having pixels to produce a cropped digital image includes developing a belief map of a photographic image by using the pixels to determine a series of features and using the features to assign a probability of a location of a main subject of the digital image in the belief map, and cropping the digital image to include main subjects indicated by the belief map to produce the cropped digital image.

In one aspect, a method of producing an image of at least a portion of a digital image includes providing a digital image having pixels, computing a belief map of the digital image, determining a crop window having a shape and a zoom factor, and cropping the digital image to include a portion of the image of high subject content in response to the belief map and the crop window. The belief map is computed by using the pixels of the digital image to determine a series of features and using the features to assign a probability of a location of a main subject of the digital image in the belief map. The shape and the zoom factors determine a size of the crop window.

In another aspect, a method of producing an image of a portion of at least a portion of a photographic image onto a photographic receiver includes receiving a digital image including pixels that corresponds to the photographic image, computing a belief map of the digital image, determining a crop window having a shape factor and a zoom factor, and locating a relative optical position of a photographic image, a lens assembly, and a photographic receiver in response to the belief map and illuminating a portion of the photographic image of high subject content to produce an image of such portion on the photographic receiver. The belief map is computed by using the pixels of the digital image to determine a series of features and using such features to assign a probability of a location of a main subject of the digital image in the belief map. The shape and the zoom factors determine a size of the crop window.

Jia relates to a system and method for automatically determining unwanted extraneous material, i.e., background or scanner background information.

Jia fails to describe of suggest the method of independent claim 1. Firstly, Jia fails to describe or suggest "computing a belief map of the digital image by using the pixels of the digital image to determine a series of features and using such features to assign a probability to a location of a main subject of the digital image in the belief map," as recited. In the instant application, the belief map is created using MSD, for example, as described in the specification at pages 6, 7, and 10.

In fact, there is no belief map in Jia. Jia addresses cropping background from a scanned document. In Jia, the background is assumed to be largely uniform, to have known characteristics, and to form a clear boundary between the background and the real image. Consequently, in Jia, a pixel of the scan line is checked against a reference background pixel, is treated as an image pixel when the pixel color is different from the color of the reference background pixel by more than a predetermined threshold value and when the color of the adjacent pixel is also different from the color of the reference background pixel by more than a predetermined threshold value, and is not cropped. (*See* Jia, col. 13, lines 38-39). Otherwise, it is cropped. Thus, there is no belief map in Jia, and no claimed "computing a belief map of the digital image" in Jia.

Further, Jia fails to describe or suggest the claimed "cropping the digital image to include a portion of the image of high subject content in response to the belief map and the crop window" of independent claim 1. Since Jia does not provide a belief map, as explained above, Jia cannot "crop[] the digital image to include a portion of the image of high subject content in response to the belief map and the crop window," as claimed (emphasis added).

Also, in the instant application, the belief map uses the pixels of the original image to determine features, uses these features to assign a probability as to the location of the main subject of the digital image in the belief map, as recited above. In Jia, however, elements

702, 710, 712, 714, 716, 718, 720 shown in Fig. 8 of Jia are the six (6) predetermined boundary pixels used by the Jia system to crop an image. The cropped image then includes only the portion of the original image located within the six (6) boundary pixels, everything else is trimmed away. (See Jia, col. 11, lines 5 - 14). Thus, Jia does not "crop the digital image to include a portion of the image of high subject content in response to the belief map and the crop window," as claimed (emphasis added). Therefore, Jia fails to describe or suggest the subject matter of independent claim 1, and independent claim 1, and claims 3, 8, and 12-14, are allowable over the cited art.

Similarly, as to independent claim 15, for at least the reasons presented above, Jia fails to describe or suggest the "computing a belief map of the digital image by using the pixels of the digital image to determine a series of features and using such features to assign a probability of a location of a main subject of the digital image in the belief map," as claimed. There is no belief map in Jia; no belief map is computed. Therefore, independent claim 15, and claims 17, 22, and 26-28 which depend therefrom, are allowable over the cited art.

Thus, Applicants submit that none of claims 1, 3, 8, 12-15, 17, 22, and 26-28 are described or suggested by Jia, and accordingly, withdrawal of this rejection is respectfully requested.

Claims 1, 8, and 14 were rejected under the judicially created doctrine of obviousness-type double-patenting as being unpatentable over claims 44-48 of co-pending Application Serial No. 09/490,915.

Firstly, Applicants do not understand this rejection. Applicants submit that this rejection is more properly a provisional double-patenting rejection as two co-pending applications are noted by the Examiner, and neither application has issued as a patent to date.

Applicants hereby submit a draft Terminal Disclaimer. Upon an indication of allowability of claims 1, 8, and 14, should such a terminal disclaimer disclaiming a portion of a patent issuing from this application, which would extend beyond the term of any patent issuing from co-pending Application Serial No. 09/490,915, become necessary, such a document will be submitted. Accordingly, withdrawal of this rejection is respectfully requested.

Applicants submit that all pending claims are in condition for allowance, and formal notice of such is solicited. If the Examiner has any questions, the Examiner is respectfully requested to contact the undersigned at the number indicated below.

Respectfully submitted,

Patrick J. Finnan
Registration No. 39,189

EDELL, SHAPIRO & FINNAN, LLC
1901 Research Boulevard, Suite 400
Rockville, Maryland 20850-3164
(301) 424-3640

Hand Delivered: 4/2/03

a

PTO/SB/25 (10-00)
Approved for use through 10/31/2002. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

TERMINAL DISCLAIMER TO OBVIATE A PROVISIONAL DOUBLE PATENTING REJECTION OVER A PENDING SECOND APPLICATION	Docket Number (Optional) 81595RLO/0505.0018C

In re Application of: Luo et al.

Application No.: 09/736,825

Filed: December 14, 2000

For: Automatically Producing An Image of a Portion of a Photographic Image

The owner*, Eastman Kodak Company, of 100 percent interest in the instant application hereby disclaims, except as provided below, the terminal part of the statutory term of any patent granted on the instant application, which would extend beyond the expiration date of the full statutory term defined in 35 U.S.C. 154 to 156 and 173 as shortened by any terminal disclaimer filed prior to the grant of any patent granted on pending second Application Number 09/490,915, filed on January 25, 2000, of any patent on the pending second application.The owner hereby agrees that any patent so granted on the instant application shall be enforceable only for and during such period that it and any patent granted on the second application are commonly owned. This agreement runs with any patent granted on the instant application and is binding upon the grantee, its successors or assigns.

In making the above disclaimer, the owner does not disclaim the terminal part of any patent granted on the instant application that would extend to the expiration date of the full statutory term as defined in 35 U.S.C. 154 to 156 and 173 of any patent granted on the second application, as shortened by any terminal disclaimer filed prior to the patent grant, in the event that any such granted patent: expires for failure to pay a maintenance fee, is held unenforceable, is found invalid by a court of competent jurisdiction, is statutorily disclaimed in whole or terminally disclaimed under 37 CFR 1.321, has all claims canceled by a reexamination certificate, is reissued, or is in any manner terminated prior to the expiration of its full statutory term as shortened by any terminal disclaimer filed prior to its grant.

Check either box 1 or 2 below, if appropriate.

1. ☐ For submissions on behalf of an organization (e.g., corporation, partnership, university, government agency, etc.), the undersigned is empowered to act on behalf of the organization.

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

2. ☑ The undersigned is an attorney or agent of record.

Signature — Date

Patrick J. Finnan, Registration No. 39,189

Typed or printed name

☑ Terminal disclaimer fee under 37 CFR 1.20(d) is included.

WARNING: Information on this form may become public. Credit card information should not be included on this form. Provide credit card information and authorization on PTO-2038.

*Statement under 37 CFR 3.73(b) is required if terminal disclaimer is signed by the assignee (owner). Form PTO/SB/96 may be used for making this statement. See MPEP § 324.

OIPE
APR 0 2 2003
PATENT & TRADEMARK OFFICE

Atty. Dckt. 0505.0018C

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re the PATENT application of

Jiebo Luo, et al.

Group Art Unit: 2625

Serial No.: 09/736,825

Examiner: Aaron W. Carter

Filed: December 14, 2000

For: AUTOMATICALLY PRODUCING AN IMAGE OF A PORTION OF A PHOTOGRAPHIC IMAGE

RECEIVED
APR 0 3 2003
Technology Center 2600

COMMISSIONER FOR PATENTS
Washington, D.C. 20231

CHANGE OF FIRM NAME

Edell, Shapiro, Finnan & Lytle, LLC attorneys of record in the subject patent application, have changed their name to:

EDELL, SHAPIRO & FINNAN, LLC
1901 Research Boulevard
Suite 400
Rockville, Maryland 20850-3164

Please note that only our firm name has been changed, our address and phone number remain unchanged. Please send all future correspondence concerning this application to the new firm name at the above address.

Respectfully submitted,

Patrick J. Finnan
Registration No. 39,189

Hand-delivered: 4/2/03

OIPE JC73
APR 0 2 2003
PATENT & TRADEMARK OFFICE

2625

Atty. Docket. No. 81595RLO
0505.0018C

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re the PATENT application of

Jiebo Luo, et al.

Group Art Unit: 2625

Serial No.: 09/736,825

Examiner: Aaron W. Carter

Filed: December 14, 2000

For: AUTOMATICALLY PRODUCING AN IMAGE OF A PORTION OF A PHOTOGRAPHIC IMAGE

RECEIVED
APR 0 3 2003
Technology Center 2600

TRANSMITTAL LETTER

COMMISSIONER FOR PATENTS
Washington, D.C. 20231

Sir:

Transmitted herewith for filing in the above-identified application is an Amendment (16 Pages);, Terminal Disclaimer (1 page); and Change of Firm Name (1 page).

It is believed that no fees are required at this time. However, Applicants hereby petition for any extension of time that may be necessary to maintain the pendency of this application. The Commissioner is hereby authorized to charge payment of any additional fees required for the above-identified application or credit any overpayment to Deposit Account No. 05-0460.

Respectfully submitted,

Patrick J. Finnan
Registration No. 39,189

EDELL, SHAPIRO & FINNAN, LLC
1901 Research Blvd., Suite 400
Rockville, Maryland 20850-3164
(301) 424-3640

Hand-delivered: 4/2/03

Official

RECEIVED 6-2-03 TC

#5

81595RLO
Customer No. 01333

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of:

Jiebo Luo

AUTOMATICALLY PRODUCING AN IMAGE OF A PORTION OF A PHOTOGRAPHIC IMAGE

Serial No. 09/736,825

Filed 14 December 2000

Group Art Unit: 2625

Examiner: Aaron W. Carter

I hereby certify that this correspondence was sent by facsimile transmission to the United States Patent and Trademark Office on the date set forth below.

Dia St. George

6/2/03
Date

Commissioner for Patents
P.O. Box 1450
Alexandria, VA. 22313-1450

Sir:

TERMINAL DISCLAIMER TO OBVIATE A DOUBLE PATENTING REJECTION OVER A PRIOR PATENT

The owner, Eastman Kodak Company of the entire interest in the instant application, hereby disclaims except as provided below, the terminal part of the statutory term of any patent granted on the instant application, which would extend beyond the expiration date of the full statutory term defined in 35 U.S.C. 154 to 156 and 173, as shortened by any terminal disclaimer, of prior Patent No. 6,545,743. The owner hereby agrees that any patent so granted on the instant application shall be enforceable only for and during such period that it and the prior patent are commonly owned. This agreement runs with any patent granted on the instant application and is binding upon the grantee, its successors or assigns.

In making the above disclaimer, the owner does not disclaim the terminal part of any patent granted on the instant application that would extend to the expiration date of the full statutory term as defined in 35 U.S.C. 154 to 156 and 173 of the prior patent, as presently shortened by any terminal disclaimer, in the event that the prior patent hereafter: expires for failure to pay a maintenance fee, is held unenforceable, is found invalid by a court of competent jurisdiction, is statutorily disclaimed in whole or terminally disclaimed under 37 CFR 1.321, has all claims canceled by a reexamination certificate, is reissued, or is in any manner

06/13/2003 MJACKSON 00000005 09736825
01 FC:1814

A:\81595\term-dis-prior patent.doc

Received from < 585 477 4646 > at 6/2/03 8:49:51 AM [Eastern Daylight Time]

#8

81595RLO
Customer No. 01333

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of:

Jiebo Luo

AUTOMATICALLY PRODUCING AN IMAGE OF A PORTION OF A PHOTOGRAPHIC IMAGE

Serial No. 09/736,825

Filed 14 December 2000

Group Art Unit: 2625

Examiner: Aaron W. Carter

I hereby certify that this correspondence was sent by facsimile transmission to the United States Patent and Trademark Office on the date set forth below.

Dia St George
Dia St. George

6/2/03
Date

Commissioner for Patents
P.O. Box 1450
Alexandria, VA. 22313-1450

Sir:

TERMINAL DISCLAIMER TO OBVIATE A DOUBLE PATENTING REJECTION OVER A PENDING SECOND APPLICATION

The owner, Eastman Kodak Company of the entire interest in the instant application, hereby disclaims except as provided below, the terminal part of the statutory term of any patent granted on the instant application, which would extend beyond the expiration date of the full statutory term defined in 35 U.S.C. 154 to 156 and 173, as shortened by any terminal disclaimer filed prior to the grant of any patent granted on pending second Application No. 09/490,915. The owner hereby agrees that any patent so granted on the instant application shall be enforceable only for and during such period that it and.any patent granted on the second application are commonly owned. This agreement runs with any patent granted on the instant application and is binding upon the grantee, its successors or assigns.

In making the above disclaimer, the owner does not disclaim the terminal part of any patent granted on the instant application that would extend to the expiration date of the full statutory term as defined in 35 U.S.C. 154 to 156 and 173 of any patent granted on the second application, as shortened by any terminal disclaimer filed prior to the patent grant on the second application, in the event that any patent granted on the second application: expires for failure to pay a maintenance fee, is held unenforceable, is found invalid by a court of competent

A:\81595\term-dis-prior appl.doc

jurisdiction, is statutorily disclaimed in whole or terminally disclaimed under 37 CFR 1.321, has all claims canceled by a reexamination certificate, is reissued, or is in any manner terminated prior to the expiration of its full statutory term as shortened by any terminal disclaimer filed prior to its grant.

[X] The undersigned is an attorney of record. (If this box is not checked do not use this form)

June 2, 2003
Date
Telephone: (585) 477-4653
Facsimile: (585) 477-4646
/das

Raymond L. Owens
Attorney of Record
Registration No. 22,363

[X] Please charge the fee to Eastman Kodak Company Deposit Account 05-0225. (A duplicate copy of this request is enclosed)

#8

FACSIMILE TRANSMITTAL

EASTMAN KODAK COMPANY
PATENT LEGAL STAFF
343 STATE STREET
ROCHESTER, NEW YORK 14650-2201

DATE: June 2, 2003

TO: Examiner Aaron W. Carter
GROUP 2625
U.S. PATENT AND
TRADEMARK OFFICE

FAX NO. 703-872-9314

FROM: RAYMOND L. OWENS
PHONE NO. 585-477-4653

FAX NO. 585-477-4646

RE: U.S. Serial No. 09/736,825 filed December 14, 2000
Jiebo Luo
Automatically Producing an Image of a Portion of a Photographic Image
ATTORNEY DOCKET NO. 81595RLO

Applicants submitted by facsimile on May 29, 2003, terminal disclaimers in view of a copending patent application and a prior patent which were requested by the Examiner in a phone conversation of that date. In reviewing these documents it has come to our attention that the copending patent application and prior patent numbers were incorrect. New terminal disclaimers with the correct numbers follow.

Total Pages Including Cover Sheet 5

Official RECEIVED 6-2-03 TC #9 N.E.

81595RLO
Customer No. 01333

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of:

Jiebo Luo

AUTOMATICALLY PRODUCING AN IMAGE OF A PORTION OF A PHOTOGRAPHIC IMAGE

Serial No. 09/736,825

Filed 14 December 2000

Group Art Unit: 2625

Examiner: Aaron W. Carter

I hereby certify that this correspondence was sent by facsimile transmission to the United States Patent and Trademark Office on the date set forth below.

Dia St. George

6/2/03
Date

Commissioner for Patents
P.O. Box 1450
Alexandria, VA. 22313-1450

Sir:

TERMINAL DISCLAIMER TO OBVIATE A DOUBLE PATENTING REJECTION OVER A PRIOR PATENT

The owner, Eastman Kodak Company of the entire interest in the instant application, hereby disclaims except as provided below, the terminal part of the statutory term of any patent granted on the instant application, which would extend beyond the expiration date of the full statutory term defined in 35 U.S.C. 154 to 156 and 173, as shortened by any terminal disclaimer, of prior Patent No. 6,545,743. The owner hereby agrees that any patent so granted on the instant application shall be enforceable only for and during such period that it and the prior patent are commonly owned. This agreement runs with any patent granted on the instant application and is binding upon the grantee, its successors or assigns.

In making the above disclaimer, the owner does not disclaim the terminal part of any patent granted on the instant application that would extend to the expiration date of the full statutory term as defined in 35 U.S.C. 154 to 156 and 173 of the prior patent, as presently shortened by any terminal disclaimer, in the event that the prior patent hereafter: expires for failure to pay a maintenance fee, is held unenforceable, is found invalid by a court of competent jurisdiction, is statutorily disclaimed in whole or terminally disclaimed under 37 CFR 1.321, has all claims canceled by a reexamination certificate, is reissued, or is in any manner

06/13/2003 MJACKSON 00000005 09736825
02 FC:1814

A:\81595\term-dis-prior patent.doc

terminated prior to the expiration of its full statutory term as shortened by any terminal disclaimer filed prior to its grant.

[X] The undersigned is an attorney of record. (If this box is not checked do not use this form)

June 2, 2003
Date
Telephone: (585) 477-4653
Facsimile: (585) 477-4646
/das

Raymond L. Owens
Attorney of Record
Registration No. 22,363

[X] Please charge the fee to Eastman Kodak Company Deposit Account 05-0225. (A duplicate copy of this request is enclosed)

Received from < 585 477 4646 > at 6/2/03 8:49:51 AM [Eastern Daylight Time]

Interview Summary	Application No.	Applicant(s)	
	09/736,825	LUO, JIEBO	
	Examiner	Art Unit	
	Aaron W Carter	2625	

X10

All participants (applicant, applicant's representative, PTO personnel):

(1) *Aaron W Carter*. (3)_____.

(2) *Raymond L. Owens, Reg. No.22,363* . (4)_____.

Date of Interview: *29 May 2003*.

Type: a)☒ Telephonic b)☐ Video Conference
c)☐ Personal [copy given to: 1)☐ applicant 2)☐ applicant's representative]

Exhibit shown or demonstration conducted: d)☐ Yes e)☒ No.
If Yes, brief description: _____.

Claim(s) discussed: _____.

Identification of prior art discussed: _____.

Agreement with respect to the claims f)☐ was reached. g)☐ was not reached. h)☒ N/A.

Substance of Interview including description of the general nature of what was agreed to if an agreement was reached, or any other comments: *Attorney agreed to submit official Terminal Disclaimer for application 09/490915 and USPN 6,545,743*.

(A fuller description, if necessary, and a copy of the amendments which the examiner agreed would render the claims allowable, if available, must be attached. Also, where no copy of the amendments that would render the claims allowable is available, a summary thereof must be attached.)

THE FORMAL WRITTEN REPLY TO THE LAST OFFICE ACTION MUST INCLUDE THE SUBSTANCE OF THE INTERVIEW. (See MPEP Section 713.04). If a reply to the last Office action has already been filed, APPLICANT IS GIVEN ONE MONTH FROM THIS INTERVIEW DATE TO FILE A STATEMENT OF THE SUBSTANCE OF THE INTERVIEW. See Summary of Record of Interview requirements on reverse side or on attached sheet.

Examiner Note: You must sign this form unless it is an Attachment to a signed Office action.

Examiner's signature, if required

Notice of Allowability	Application No. 09/736,825	Applicant(s) LUO, JIEBO	
	Examiner Aaron W Carter	Art Unit 2625	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address--
All claims being allowable, PROSECUTION ON THE MERITS IS (OR REMAINS) CLOSED in this application. If not included herewith (or previously mailed), a Notice of Allowance (PTOL-85) or other appropriate communication will be mailed in due course. **THIS NOTICE OF ALLOWABILITY IS NOT A GRANT OF PATENT RIGHTS.** This application is subject to withdrawal from issue at the initiative of the Office or upon petition by the applicant. See 37 CFR 1.313 and MPEP 1308.

1. ☒ This communication is responsive to 04/02/2003.
2. ☒ The allowed claim(s) is/are 1-28.
3. ☒ The drawings filed on 16 March 2001 are accepted by the Examiner.
4. ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
   a) ☐ All b) ☐ Some* c) ☐ None of the:
   1. ☐ Certified copies of the priority documents have been received.
   2. ☐ Certified copies of the priority documents have been received in Application No. _____ .
   3. ☐ Copies of the certified copies of the priority documents have been received in this national stage application from the International Bureau (PCT Rule 17.2(a)).

   * Certified copies not received: _____.
5. ☐ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. § 119(e) (to a provisional application).
   (a) ☐ The translation of the foreign language provisional application has been received.
6. ☐ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. §§ 120 and/or 121.

Applicant has THREE MONTHS FROM THE "MAILING DATE" of this communication to file a reply complying with the requirements noted below. Failure to timely comply will result in ABANDONMENT of this application. **THIS THREE-MONTH PERIOD IS NOT EXTENDABLE.**

7. ☐ A SUBSTITUTE OATH OR DECLARATION must be submitted. Note the attached EXAMINER'S AMENDMENT or NOTICE OF INFORMAL PATENT APPLICATION (PTO-152) which gives reason(s) why the oath or declaration is deficient.

8. ☐ CORRECTED DRAWINGS must be submitted.
   (a) ☐ including changes required by the Notice of Draftsperson's Patent Drawing Review (PTO-948) attached
   1) ☐ hereto or 2) ☐ to Paper No. _____.
   (b) ☐ including changes required by the proposed drawing correction filed _____, which has been approved by the Examiner.
   (c) ☐ including changes required by the attached Examiner's Amendment / Comment or in the Office action of Paper No. _____.

Identifying indicia such as the application number (see 37 CFR 1.84(c)) should be written on the drawings in the front (not the back) of each sheet.

9. ☐ DEPOSIT OF and/or INFORMATION about the deposit of BIOLOGICAL MATERIAL must be submitted. Note the attached Examiner's comment regarding REQUIREMENT FOR THE DEPOSIT OF BIOLOGICAL MATERIAL.

Attachment(s)

1☐ Notice of References Cited (PTO-892)
3☐ Notice of Draftperson's Patent Drawing Review (PTO-948)
5☐ Information Disclosure Statements (PTO-1449), Paper No. _____.
7☐ Examiner's Comment Regarding Requirement for Deposit of Biological Material

2☐ Notice of Informal Patent Application (PTO-152)
4☒ Interview Summary (PTO-413), Paper No. 10 .
6☐ Examiner's Amendment/Comment
8☒ Examiner's Statement of Reasons for Allowance
9☐ Other .

DETAILED ACTION

1. This action is responsive to papers filed on April 2, 2003.

Response to Amendment

2. In response to the applicant's amendment received on April 2, 2003, all requested changes to the claims have been entered.

Response to Arguments

3. Applicant's arguments, see Amendment A, filed April 2, 2003, with respect to claims have been fully considered and are persuasive. The 35 USC 102(e), 112 and Double Patenting rejections of claims has been withdrawn.

Allowable Subject Matter

4. Claims 1-28 are allowed.

5. The following is an examiner's statement of reasons for allowance: With respect to claims 1- 28, none of the prior art, teach or fairly suggest computing a belief map of the digital image, by using the pixels of the digital image to determine a series of features, and using such features to assign a probability of a location of a main subject of the digital image in the belief map. Although Jia (US 6,430,320) does disclose a method of automatically cropping an image according to a belief of where in the image a main subject is located (column 13, lines 25-35 and

Fig. 8) he does not teach or fairly suggest the use of a belief map in which a probability of the location of the main subject is assign.

Any comments considered necessary by applicant must be submitted no later than the payment of the issue fee and, to avoid processing delays, should preferably accompany the issue fee. Such submissions should be clearly labeled "Comments on Statement of Reasons for Allowance."

Contact Information

6. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Aaron W. Carter whose telephone number is 703.306.4060. The examiner can normally be reached by telephone between 8am - 4:30pm (Mon. – Fri.).

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Bhavesh Mehta can be reached on 703.308.5246. The fax phone number for the organization where the application or proceeding is assigned is 703.872.9314 for regular communications.

Any inquiry of a general nature or relating to the status of this application or proceeding should be directed to the receptionist whose telephone number is 703.306.0377.

Aaron W. Carter
Examiner
Art Unit 2625

AWC
awc
June 16, 2003

BHAVESH M. MEHTA
SUPERVISORY PATENT EXAMINER
TECHNOLOGY CENTER 2600

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

NOTICE OF ALLOWANCE AND FEE(S) DUE

7590 06/18/2003

Patent Legal Staff
Eastman Kodak Company
343 State Street
Rochester, NY 14650-2201

EXAMINER	
CARTER, AARON W	
ART UNIT	CLASS-SUBCLASS
2625	382-282000

DATE MAILED: 06/18/2003

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/736,825	12/14/2000	Jiebo Luo	81595WFN	8670

TITLE OF INVENTION: AUTOMATICALLY PRODUCING AN IMAGE OF A PORTION OF A PHOTOGRAPHIC IMAGE

APPLN. TYPE	SMALL ENTITY	ISSUE FEE	PUBLICATION FEE	TOTAL FEE(S) DUE	DATE DUE
nonprovisional	NO	$1300	$300	$1600	09/18/2003

THE APPLICATION IDENTIFIED ABOVE HAS BEEN EXAMINED AND IS ALLOWED FOR ISSUANCE AS A PATENT. PROSECUTION ON THE MERITS IS CLOSED. THIS NOTICE OF ALLOWANCE IS NOT A GRANT OF PATENT RIGHTS. THIS APPLICATION IS SUBJECT TO WITHDRAWAL FROM ISSUE AT THE INITIATIVE OF THE OFFICE OR UPON PETITION BY THE APPLICANT. SEE 37 CFR 1.313 AND MPEP 1308.

THE ISSUE FEE AND PUBLICATION FEE (IF REQUIRED) MUST BE PAID WITHIN THREE MONTHS FROM THE MAILING DATE OF THIS NOTICE OR THIS APPLICATION SHALL BE REGARDED AS ABANDONED. THIS STATUTORY PERIOD CANNOT BE EXTENDED. SEE 35 U.S.C. 151. THE ISSUE FEE DUE INDICATED ABOVE REFLECTS A CREDIT FOR ANY PREVIOUSLY PAID ISSUE FEE APPLIED IN THIS APPLICATION. THE PTOL-85B (OR AN EQUIVALENT) MUST BE RETURNED WITHIN THIS PERIOD EVEN IF NO FEE IS DUE OR THE APPLICATION WILL BE REGARDED AS ABANDONED.

HOW TO REPLY TO THIS NOTICE:

I. Review the SMALL ENTITY status shown above.

If the SMALL ENTITY is shown as YES, verify your current SMALL ENTITY status:

A. If the status is the same, pay the TOTAL FEE(S) DUE shown above.

B. If the status is changed, pay the PUBLICATION FEE (if required) and twice the amount of the ISSUE FEE shown above and notify the United States Patent and Trademark Office of the change in status, or

If the SMALL ENTITY is shown as NO:

A. Pay TOTAL FEE(S) DUE shown above, or

B. If applicant claimed SMALL ENTITY status before, or is now claiming SMALL ENTITY status, check the box below and enclose the PUBLICATION FEE and 1/2 the ISSUE FEE shown above.

☐ Applicant claims SMALL ENTITY status.
See 37 CFR 1.27.

II. PART B - FEE(S) TRANSMITTAL should be completed and returned to the United States Patent and Trademark Office (USPTO) with your ISSUE FEE and PUBLICATION FEE (if required). Even if the fee(s) have already been paid, Part B - Fee(s) Transmittal should be completed and returned. If you are charging the fee(s) to your deposit account, section "4b" of Part B - Fee(s) Transmittal should be completed and an extra copy of the form should be submitted.

III. All communications regarding this application must give the application number. Please direct all communications prior to issuance to Box ISSUE FEE unless advised to the contrary.

IMPORTANT REMINDER: Utility patents issuing on applications filed on or after Dec. 12, 1980 may require payment of maintenance fees. It is patentee's responsibility to ensure timely payment of maintenance fees when due.

PTOL-85 (REV. 05-03) Approved for use through 04/30/2004.

PART B - FEE(S) TRANSMITTAL

Complete and send this form, together with applicable fee(s), to: **Mail** **Mail Stop ISSUE FEE**
Commissioner for Patents
Alexandria, Virginia 22313-1450
Fax **(703)746-4000**

INSTRUCTIONS: This form should be used for transmitting the ISSUE FEE and PUBLICATION FEE (if required). Blocks 1 through 4 should be completed where appropriate. All further correspondence including the Patent, advance orders and notification of maintenance fees will be mailed to the current correspondence address as indicated unless corrected below or directed otherwise in Block 1, by (a) specifying a new correspondence address; and/or (b) indicating a separate "FEE ADDRESS" for maintenance fee notifications.

CURRENT CORRESPONDENCE ADDRESS (Note: Legibly mark-up with any corrections or use Block 1)

7590 06/18/2003

Patent Legal Staff
Eastman Kodak Company
343 State Street
Rochester, NY 14650-2201

Note: A certificate of mailing can only be used for domestic mailings of the Fee(s) Transmittal. This certificate cannot be used for any other accompanying papers. Each additional paper, such as an assignment or formal drawing, must have its own certificate of mailing or transmission.

Certificate of Mailing or Transmission

I hereby certify that this Fee(s) Transmittal is being deposited with the United States Postal Service with sufficient postage for first class mail in an envelope addressed to the Box Issue Fee address above, or being facsimile transmitted to the USPTO, on the date indicated below.

(Depositor's name)

(Signature)

(Date)

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/736,825	12/14/2000	Jiebo Luo	81595WFN	8670

TITLE OF INVENTION: AUTOMATICALLY PRODUCING AN IMAGE OF A PORTION OF A PHOTOGRAPHIC IMAGE

APPLN. TYPE	SMALL ENTITY	ISSUE FEE	PUBLICATION FEE	TOTAL FEE(S) DUE	DATE DUE
nonprovisional	NO	$1300	$300	$1600	09/18/2003

EXAMINER	ART UNIT	CLASS-SUBCLASS
CARTER, AARON W	2625	382-282000

1. Change of correspondence address or indication of "Fee Address" (37 CFR 1.363).

- ❑ Change of correspondence address (or Change of Correspondence Address form PTO/SB/122) attached.
- ❑ "Fee Address" indication (or "Fee Address" Indication form PTO/SB/47; Rev 03-02 or more recent) attached. **Use of a Customer Number is required.**

2. For printing on the patent front page, list (1) the names of up to 3 registered patent attorneys or agents OR, alternatively, (2) the name of a single firm (having as a member a registered attorney or agent) and the names of up to 2 registered patent attorneys or agents. If no name is listed, no name will be printed.

1 ____________
2 ____________
3 ____________

3. ASSIGNEE NAME AND RESIDENCE DATA TO BE PRINTED ON THE PATENT (print or type)

PLEASE NOTE: Unless an assignee is identified below, no assignee data will appear on the patent. Inclusion of assignee data is only appropriate when an assignment has been previously submitted to the USPTO or is being submitted under separate cover. Completion of this form is NOT a substitute for filing an assignment.

(A) NAME OF ASSIGNEE (B) RESIDENCE: (CITY and STATE OR COUNTRY)

Please check the appropriate assignee category or categories (will not be printed on the patent) ❑ individual ❑ corporation or other private group entity ❑ government

4a. The following fee(s) are enclosed:

- ❑ Issue Fee
- ❑ Publication Fee
- ❑ Advance Order - # of Copies ____________

4b. Payment of Fee(s):

- ❑ A check in the amount of the fee(s) is enclosed.
- ❑ Payment by credit card. Form PTO-2038 is attached.
- ❑ The Commissioner is hereby authorized by charge the required fee(s), or credit any overpayment, to Deposit Account Number ____________(enclose an extra copy of this form).

Commissioner for Patents is requested to apply the Issue Fee and Publication Fee (if any) or to re-apply any previously paid issue fee to the application identified above.

(Authorized Signature) (Date)

NOTE: The Issue Fee and Publication Fee (if required) will not be accepted from anyone other than the applicant; a registered attorney or agent; or the assignee or other party in interest as shown by the records of the United States Patent and Trademark Office.

TRANSMIT THIS FORM WITH FEE(S)

PTOL-85 (REV. 05-03) Approved for use through 04/30/2004. OMB 0651-0033 U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/736,825	12/14/2000	Jiebo Luo	81595WFN	8670

7590 06/18/2003

Patent Legal Staff
Eastman Kodak Company
343 State Street
Rochester, NY 14650-2201

EXAMINER	
CARTER, AARON W	
ART UNIT	**PAPER NUMBER**
2625	

DATE MAILED: 06/18/2003

Determination of Patent Term Adjustment under 35 U.S.C. 154 (b)

(application filed on or after May 29, 2000)

The patent term adjustment to date is 334 days. If the issue fee is paid on the date that is three months after the mailing date of this notice and the patent issues on the Tuesday before the date that is 28 weeks (six and a half months) after the mailing date of this notice, the term adjustment will be 334 days.

If a continued prosecution application (CPA) was filed in the above-identified application, the filing date that determines patent term adjustment is the filing date of the most recent CPA.

Applicant will be able to obtain more detailed information by accessing the Patent Application Information Retrieval (PAIR) system. (http://pair.uspto.gov)

Any questions regarding the patent term extension or adjustment determination should be directed to the Office of Patent Legal Administration at (703)305-1383.

PTOL-85 (REV. 05-03) Approved for use through 04/30/2004.

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/736,825	12/14/2000	Jiebo Luo	81595WFN	8670

7590 06/18/2003

Patent Legal Staff
Eastman Kodak Company
343 State Street
Rochester, NY 14650-2201
UNITED STATES

EXAMINER	
CARTER, AARON W	
ART UNIT	PAPER NUMBER
2625	11

DATE MAILED: 06/18/2003

Notice of Fee Increase on January 1, 2003

If a reply to a "Notice of Allowance and Fee(s) Due" is filed in the Office on or after January 1, 2003, then the amount due will be higher than that set forth in the "Notice of Allowance and Fee(s) Due" since there will be an increase in fees effective on January 1, 2003. See Revision of Patent and Trademark Fees for Fiscal Year 2003; Final Rule, 67 Fed. Reg. 70847, 70849 (November 27, 2002).

The current fee schedule is accessible from: http://www.uspto.gov/main/howtofees.htm.

If the issue fee paid is the amount shown on the "Notice of Allowance and Fee(s) Due," but not the correct amount in view of the fee increase, a "Notice to Pay Balance of Issue Fee" will be mailed to applicant. In order to avoid processing delays associated with mailing of a "Notice to Pay Balance of Issue Fee," if the response to the Notice of Allowance and Fee(s) due form is to be filed on or after January 1, 2003 (or mailed with a certificate of mailing on or after January 1, 2003), the issue fee paid should be the fee that is required at the time the fee is paid. If the issue fee was previously paid, and the response to the "Notice of Allowance and Fee(s) Due" includes a request to apply a previously-paid issue fee to the issue fee now due, then the difference between the issue fee amount at the time the response is filed and the previously paid issue fee should be paid. See Manual of Patent Examining Procedure, Section 1308.01 (Eighth Edition, August 2001).

Questions relating to issue and publication fee payments should be directed to the Customer Service Center of the Office of Patent Publication at (703) 305-8283.

PTOL-85 (REV. 05-03) Approved for use through 04/30/2004.

SEP 1 0 2003

PART B - FEE(S) TRANSMITTAL

Complete and send this form, together with applicable fee(s), to: **Mail** **Mail Stop ISSUE FEE**
Commissioner for Patents
Alexandria, Virginia 22313-1450
Fax (703)746-4000

INSTRUCTIONS: This form should be used for transmitting the ISSUE FEE and PUBLICATION FEE (if required). Blocks 1 through 4 should be completed where appropriate. All further correspondence including the Patent, advance orders and notification of maintenance fees will be mailed to the current correspondence address as indicated unless corrected below or directed otherwise in Block 1, by (a) specifying a new correspondence address; and/or (b) indicating a separate "FEE ADDRESS" for maintenance fee notifications.

CURRENT CORRESPONDENCE ADDRESS (Note: Legibly mark-up with any corrections or use Block 1)

7590 06/18/2003

Patent Legal Staff
Eastman Kodak Company
343 State Street
Rochester, NY 14650-2201

Note: A certificate of mailing can only be used for domestic mailings of the Fee(s) Transmittal. This certificate cannot be used for any other accompanying papers. Each additional paper, such as an assignment or formal drawing, must have its own certificate of mailing or transmission.

Certificate of Mailing or Transmission

I hereby certify that this Fee(s) Transmittal is being deposited with the United States Postal Service with sufficient postage for first class mail in an envelope addressed to the Box Issue Fee address above, or being facsimile transmitted to the USPTO, on the date indicated below.

June P. Carfagna (Depositor's name)
June P. Carfagna (Signature)
September 8, 2003 (Date)

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/736,825	12/14/2000	Jiebo Luo	81595WFN	8670

TITLE OF INVENTION: AUTOMATICALLY PRODUCING AN IMAGE OF A PORTION OF A PHOTOGRAPHIC IMAGE

APPLN. TYPE	SMALL ENTITY	ISSUE FEE	PUBLICATION FEE	TOTAL FEE(S) DUE	DATE DUE
nonprovisional	NO	$1300	$300	$1600	09/18/2003

EXAMINER	ART UNIT	CLASS-SUBCLASS
CARTER, AARON W	2625	382-282000

1. Change of correspondence address or indication of "Fee Address" (37 CFR 1.363).

☐ Change of correspondence address (or Change of Correspondence Address form PTO/SB/122) attached.

☐ "Fee Address" indication (or "Fee Address" Indication form PTO/SB/47; Rev 03-02 or more recent) attached. **Use of a Customer Number is required.**

2. For printing on the patent front page, list (1) the names of up to 3 registered patent attorneys or agents OR, alternatively, (2) the name of a single firm (having as a member a registered attorney or agent) and the names of up to 2 registered patent attorneys or agents. If no name is listed, no name will be printed.

1 Raymond L. Owens
2 ______
3 ______

3. ASSIGNEE NAME AND RESIDENCE DATA TO BE PRINTED ON THE PATENT (print or type)

PLEASE NOTE: Unless an assignee is identified below, no assignee data will appear on the patent. Inclusion of assignee data is only appropriate when an assignment has been previously submitted to the USPTO or is being submitted under separate cover. Completion of this form is NOT a substitute for filing an assignment.

(A) NAME OF ASSIGNEE (B) RESIDENCE: (CITY and STATE OR COUNTRY)

EASTMAN KODAK COMPANY
343 STATE STREET, ROCHESTER, NY 14650-2201

Please check the appropriate assignee category or categories (will not be printed on the patent) ☐ individual ☒ corporation or other private group entity ☐ government

4a. The following fee(s) are enclosed:

☒ Issue Fee
☒ Publication Fee
☐ Advance Order - # of Copies ______

4b. Payment of Fee(s):

☐ A check in the amount of the fee(s) is enclosed.
☐ Payment by credit card. Form PTO-2038 is attached.
X☐ The Commissioner is hereby authorized by charge the required fee(s), or credit any overpayment, to Deposit Account Number 05-0225 (enclose an extra copy of this form).

Commissioner for Patents is requested to apply the Issue Fee and Publication Fee (if any) or to re-apply any previously paid issue fee to the application identified above.

(Authorized Signature) Thomas H. Close/jpc (Date) 9/8/03

NOTE: The Issue Fee and Publication Fee (if required) will not be accepted from anyone other than the applicant; a registered attorney or agent; or the assignee or other party in interest as shown by the records of the United States Patent and Trademark Office.

This collection of information is required by 37 CFR 1.311. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 12 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, Alexandria, Virginia 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Commissioner for Patents, Alexandria, Virginia 22313-1450.

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

09/11/2003 KBETEHA2 00000025 050225 09736825
01 FC:1501 1300.00 DA
02 FC:1504 300.00 DA

TRANSMIT THIS FORM WITH FEE(S)

PTOL-85 (REV. 05-03) Approved for use through 04/30/2004. OMB 0651-0033 U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Attorney Docket No. 0505.0018C

PATENT

2625

OIPE JC139 SEP 2 9 2003 PATENT & TRADEMARK OFFICE

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

#12 Prior Art mas 4/5/04

In re application of

Jiebo Luo, et al.

Examiner: Carter, Aaron W.

Serial No.: 09/736,825

Art Unit: 262

RECEIVED OCT 0 7 2003 Technology Center 2600

Confirmation No.: 8670

Filed: December 14, 2000

For: Automatically Producing an Image of a Portion of a Photographic Image

---

Commissioner for Patents
P.O. Box 1450
Alexandria, Virginia 22313-1450

INFORMATION DISCLOSURE STATEMENT UNDER 37 C.F.R. §1.97(d)

Pursuant to the duty imposed by 37 C.F.R. §1.56 to disclose information which may be material to the patentability of the above-identified patent application, the Applicant(s) would like to direct the Examiner's attention to the documents listed on the enclosed Information Disclosure Citation Form (PTO/SB/08). In accordance with this duty of disclosure, Applicant(s) hereby submits the following information in conformance with 37 C.F.R. §§1.97 and 1.98.

☒ Pursuant to 37 C.F.R. §1.98, a copy of each document cited in the attached Form PTO/SB/08A is enclosed.

☐ The subject Information Disclosure Statement is being filed without copies of the listed U.S. Patents and/or U.S. patent application publications pursuant to the waiver of the requirement under 37 C.F.R. §1.98(a)(2)(i) for applications filed after June 30, 2003. Copies of the listed foreign patent documents and/or non-patent literature are enclosed herewith.

☐ No copies of the publications listed on the attached Form PTO/SB/08A are being provided pursuant to 37 C.F.R. §1.98(d) because the publications were previously cited by or submitted to the Office in prior Application

10/02/2003 SFELEKE1 00000037 09736825
01 FC:1806 180.00 OP

Serial No. _____ to which the above-identified application claims priority under 35 U.S.C. §120.

☒ Each document cited on the attached Form PTO/SB/08A was cited in a foreign search or examination report in a corresponding international patent application.

☐ Enclosed is a copy of a non-English publication(s) _____. Pursuant to §609 of the M.P.E.P., Applicant submits the attached foreign search or examination report, which cites such non-English language publication(s).

☐ Enclosed is a copy of a non-English publication(s) _____. English language publication _____ (copy enclosed) claims priority from this non-English publication.

☐ Enclosed is an explanation of non-English publication(s) _____ for which an English translation is not available.

☐ Enclosed is an English translation of non-English publication(s) _____ cited in the attached Form PTO/SB/08A.

☐ Enclosed is a copy of pending patent Application Serial No. _____.

This Information Disclosure Statement is filed after the period specified in 37 C.F.R. § 1.97(c), but on or before the payment of the issue fee.

In accordance with 37 C.F.R. §1.97(d) also enclosed is:

☒ Fee under 37 C.F.R. §1.17(p) in the amount of $180.00; and

☐ Statement as specified in 37 C.F.R. §1.97(e):

☐ Each item of information contained in the Information Disclosure Statement cited herein was first cited in a communication from a foreign patent office in a counterpart foreign application not more than three months prior to the filing date of the Information Disclosure Statement; or

☐ No item of information contained in the Information Disclosure Statement submitted herewith was cited in a communication from a foreign patent office in a counterpart foreign application, and, to the knowledge of the undersigned, having made a reasonable inquiry, no item of information contained in the Information Disclosure Statement was known to any individual designated in 37

C.F.R. §1.56(c) more than three months prior to the filing date of the Information Disclosure Statement.

It is respectfully requested that the Examiner consider the above-noted information and return an initialed copy of the attached Form PTO/SB/08A to the undersigned.

Respectfully submitted by
EDELL, SHAPIRO & FINNAN, LLC

Dated: September 29, 2003

By: [signature]
Heather Morin
Reg. No. 37,336

EDELL, SHAPIRO & FINNAN, LLC
CUSTOMER NO. 27896
1901 Research Boulevard, Suite 400
Rockville, MD 20850
(301) 424-3640

Please type a plus sign (+) inside this box → +

PTO/SB/08A (08-00)

Substitute for form 1449A/PTO				*Complete if Known*	
INFORMATION DISCLOSURE STATEMENT BY APPLICANT				**Application Number**	09/736,825
(*use as many sheets as necessary*)				**Filing Date**	December 14, 2000
				First Named Inventor	Jiebo, et al.
				Group Art Unit	2625
				Examiner Name	Aarron W, Carter
Sheet	1	of	2	Attorney Docket Number	0505.0018C

U.S. PATENT DOCUMENTS

Examiner Initials*	Cite No.[1]	U.S. Patent Document Number	Kind Code[2] (*if known*)	Name of Patentee or Applicant of Cited Document	Date of Publication of Cited Document MM-DD-YYYY
		5,732,161		Kuroda, et al.	03/24/1998

FOREIGN PATENT DOCUMENTS

Examiner Initials*	Cite No.[1]	Foreign Patent Document Office[3]	Number[4]	Kind Code[5] (if known)	Name of Patentee or Applicant of Cited Document	Date of Publication of Cited Document MM-DD-YYYY	T[6]
		EP	1 158 464	A1	Eastman Kodak Company	11/28/2001	
		EP	1 017 019	A2	Eastman Kodak Company	07/05/2000	
		EP	0 975 146	A1	Xerox Corporation	01/26/2000	

Examiner Signature		Date Considered	

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

[1] Unique citation designation number.
[2] See attached Kinds of U.S. Patent Documents.
[3] Enter Office that issued the document, by the two-letter code (WIPO Standard ST.3).
[4] For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document.
[5] Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST. 16 if possible.
[6] Applicant is to place a check mark here if English language Translation is attached.

Substitute for form 1449B/PTO

INFORMATION DISCLOSURE STATEMENT BY APPLICANT

(use as many sheets as necessary)

Sheet	2	of	2

Complete if Known	
Application Number	09/736,825
Filing Date	December 14, 2000
First Named Inventor	Jiebo, et al.
Group Art Unit	2625
Examiner Name	Aarron W, Carter
Attorney Docket Number	0505.0018C

OTHER DOCUMENTS – NON PATENT LITERATURE DOCUMENTS

Examiner Initials*	Cite No.[1]	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc.), date, page(s), volume-issue number(s), publisher, city and/or country where published.	T[2]
		Communication pursuant to Article 96(2) EPC; European Patent Office; June 4, 2003.	

Examiner Signature		Date Considered	

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

[1] Unique citation designation number.
[2] Applicant is to place a check mark here if English language Translation attached.

EPA/EPO/OEB
D-80298 München
+49 89 2399-0
TX 523 656 epmu d
FAX +49 89 2399-4465

Europäisches Patentamt | **European Patent Office** | **Office européen des brevets**

Generaldirektion 2 | Directorate General 2 | Direction Générale 2

Telephone Numbers:

Primary Examiner (substantive examination) +49 89 2399-8631

Formalities Officer / Assistant (Formalities and other matters) +49 89 2399-2432

Parent, Yves
KODAK INDUSTRIE,
Département Brevets,
CRT - Zone Industrielle
71102 Chalon-sur-Saône Cedex
FRANCE

BREVETS
10 JUIN 2003
EW

Application No.	Ref.	Date
01 204 654.6-2202	81595	04.06.2003

Applicant
EASTMAN KODAK COMPANY

ENTERED PC MASTER

Communication pursuant to Article 96(2) EPC

The examination of the above-identified application has revealed that it does not meet the requirements of the European Patent Convention for the reasons enclosed herewith. If the deficiencies indicated are not rectified the application may be refused pursuant to Article 97(1) EPC.

You are invited to file your observations and insofar as the deficiencies are such as to be rectifiable, to correct the indicated deficiencies within a period

of 4 months

from the notification of this communication, this period being computed in accordance with Rules 78(2) and 83(2) and (4) EPC.

One set of amendments to the description, claims and drawings is to be filed within the said period on separate sheets (Rule 36(1) EPC).

Failure to comply with this invitation in due time will result in the application being deemed to be withdrawn (Article 96(3) EPC).

EUROPÄISCHES PATENTAMT - EUROPEAN PATENT OFFICE - OFFICE EUROPÉEN DES BREVETS

MOORHOUSE D C
Primary Examiner
for the Examining Division

Enclosure(s): 7 page/s reasons (Form 2906)

The examination is being carried out on the following application documents:

Description, pages:

1-16 as originally filed

Claims, No.:

1-28 as originally filed

Drawings, No.:

1-18 as originally filed

1. The following documents are cited :

   D1 : EP-A-1 158 464

   D2 : EP-A-1 017 019

   D3 : EP-A-0 975 146

   D4 : US-A-5 732 161

2. Claims 1 to 7, 9, 14 to 21 and 23 lack clarity and / or support in the description, and are therefore not allowable according to Article 84 EPC. The objections are set out in detail below.

   Claims 1 and 15

   It is not clear from step (b) whether the "probability" is assigned separately to each pixel, or is applied to some other sized region, such as on an image block basis.

It is not clear what is meant by the following wording used in step (c): "a crop window having ... a zoom factor". A zoom factor would appear to define an operation to be carried out on a window, rather than defining the window itself. Moreover, it is not clear **how** the said factors determine the size of the crop window, since the said size us surely dependent on the size of the main subject within the image.

In step (d), it is not clear whether "high subject content" refers to the previously defined main subject, or to some other unspecified subject (as well).

Claims 2 and 16

In steps (I) and (ii), it is not at all clear, of what the centre of mass or weighted central moments is/are calculated, since there is no mention previously of the clustering operation. Thus, according to the present wording, the said steps could be carried out on the entire image.

In step (iv), it is not clear, **how** the crop window is determined. Moreover, the wording of this feature is objected to for reasons set out above with respect to claims 1 and 15, step (c).

Claims 3 and 17

It is nor clear, how a location can "include" the centre of mass.

In step (ii), since step (I) specifies selecting an initial position, then it would appear that step (ii) should specify "using ... to **move** the position ...".

Claims 4 and 18

The wording "uncropped" should be removed from parentheses (cf. Rule 29 (7) EPC).

Further, it is not clear where in the description and drawings step (ii) finds support.

Claims 5 and 19

These claims should refer back to step **(ii)** of claims 2 and 16 respectively.

Claims 6 and 20
These claims should refer back to step (ii) of claims 2 and 16 respectively.

Moreover, it is not clear where the feature in these claims is supported in the description and drawings.

Claims 7 and 21
These claims should refer back to step (ii) of claims 3 and 17 respectively.

Moreover, step (iii) covers more than selecting the crop window to optimize the subject content. However, no other possibilities have been disclosed in the originally filed application documents. Thus, claim 7, step (iii) lacks support in the description.

Claims 9 and 23
These claims should refer back to step (ii) of claims 2 and 16 respectively.

Claims 14 and 28
It is not at all clear, where in the description there is support for the use of more than one computer to perform the method of claim 1.

3. The above lacks of clarity and support notwithstanding, as full an examination as is possible in the circumstances has been carried out with respect to the requirements set out in Articles 52 to 56 EPC.

4. The subject-matter of claims 1, 3, 7 to 15, 17 and 21 to 28 lacks novelty with respect to the disclosure of document D1, inasfar as the same contracting states (DE, FR, GB) are designated. These claims are therefore not allowable according to Articles 52 (1), 54 (3) and 54 (4) EPC.

In detail, document D1 discloses:

- A method of producing an image of at least a portion of a digital image, comprising the steps of:
- a) providing a digital image having pixels (32);

- b) computing a belief map of the digital image (Paragraph [0026], method step 200 / Figure 3, claim 2);
- c) determining a crop window having a shape and a zoom factor (Paragraph [0026], step 201 / Figure 3, claim 4); and
- d) cropping the digital image to include a portion of the image of high subject content in response to the belief map and the crop window (Figure 4, paragraphs [0038] and [0039], claim 10).

Thus, the subject-matter of claim 1 is known from document D1.

The features of claim 3 are also known from document D1, in particular claims 5 and 6 and steps 204 to 207 in Figure 3.

Likewise, the features of claim 7 are known from document D1, in particular claims 5 and 6 and steps 204 to 207 in Figure 3.

The subject-matter of claim 8 is known from document D1, Figure 3, step 205; paragraph [0035] and claim 7.

The subject-matter of claim 9 is known from document D1, Figure 3, step 202; paragraphs [0027] and [0041] to [0045] and claim 8.

The subject-matter of claim 10 is known from document D1, paragraph [0027] and claim 9.

The subject-matter of claim 11 is known from document D1, Figure 3, step 210; Figure 4, step 306; Figure 5, step 406; and paragraphs [0032], [0033] and [0054].

The subject-matter of claim 12 is known from document D1, paragraph [0055].

The subject-matter of claim 13 is known from document D1, paragraph [0056].

The subject-matter of claim 14 is known from document D1 paragraphs [0050] and [0057].

The subject-matter of independent claim 15 is known from document D1. In addition to the passages cited in the context of claim 1, attention is also drawn to paragraphs [0010], [0023] and [0034] and claims 1 to 3.

The features specified in claims 17 and 21 to 28 are the same as those specified in claims 3 and 7 to 14 respectively, and thus the subject-matter of claims 17 and 21 to 28 lacks novelty, for reasons set out above.

5. The subject-matter of claim 1 lacks an inventive step, and is thus not allowable according to Articles 52 (1) and 56 EPC.

Document D2 discloses steps (a) and (b) of claim 1. It does not disclose steps (c) and (d). However, these steps merely represent an obvious use of the technique disclosed in document D2, especially as they are traditionally done after performing a procedure similar to steps (a) and (b) manually.

Thus, the subject-matter of claim 1 is rendered obvious by the disclosure of document D2 taken alone.

Moreover, it is well known that the zoom of the cropped area and the shape thereof influence the size of the cropped area and vice versa - see, for example, document D4, column 10, line 65 to column 11, line 13.

Thus, the subject-matter of claim 1 is rendered obvious by the disclosures of documents D2 and D4 in combination.

The additional features of claims 2, 4, 6, 16, 18 and 20 are rendered obvious by the disclosure of document D3, Figures 6 and 7 and paragraphs [0048] to [0054].

The use of the MSD technique of document D2 in positioning a film sample is, by analogy with the arguments presented above with respect to cropping, is an obvious suggestion for one skilled in the photographic arts to make. Thus, the subject-matter of claim 15 is rendered obvious by the disclosure of document D2.

The additional features of claims 16, 18 and 20 are rendered obvious by the disclosure of document D3, Figures 6 and 7 and paragraphs [0048] to [0054].

6. If, despite the above objections, the Applicant nevertheless wishes to proceed further with the present application, then when filing amendments to overcome the said objections, the following points should also be noted :-

The ultimately adopted independent claims should be divided in the proper two-part form based on D2. (Rule 29 (1) (a) and (b) EPC). This requires inclusion in the preamble, of at least those features known in combination from D2 as set out above.

The opening pages of the description should be augmented to provide an acknowledgement of the background art (Rule 27 (1) (b) EPC) as disclosed in D1, D3 and D4.

The consistory clause forming the disclosure of the invention(s) as claimed, should be brought into the agreement with the claim(s) of broadest scope (Rule 27 (1) (c) EPC).

The "inclusion by reference" on page 6 must be deleted (Guidelines, C-II, 4.18).

Bracketed reference numerals should be included throughout the claims as this would improve their intelligibility (Rule 29 (7) EPC).

The references to application documents should be replaced by published equivalents where they occur on pages 3 and 6.

Care should be taken to avoid giving rise to further objections by the inadvertent addition of subject-matter (Article 123 (2) EPC), and to ensure that the claims relate to a single invention (Article 82 EPC).

The Applicant is requested to file amendments by way of replacement pages. The Applicant should also take account of the requirements of Rule 36 (1) EPC. In particular, fair copies of the amendments should be filed in triplicate. Handwritten amendments should be avoided according to Rules 36 (1) and 35 (10) EPC, in particular with a view to avoiding mistakes in printing any patent which might result from the present Application.

Attorney Docket No. 0505.0018C

PATENT

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re application of

Jiebo Luo, et al.

Examiner: Carter, Aaron W.

Serial No.: 09/736,825

Art Unit: 2625

Confirmation No.: 8670

Filed: December 14, 2000

For: Automatically Producing an Image of a Portion of a Photographic Image

---

Commissioner for Patents
P.O. Box 1450
Alexandria, Virginia 22313-1450

STATUS INQUIRY

Applicants respectfully request that the Patent and Trademark Office notify Applicants' representative of the current status of the above-identified application. The last communication in the application was an Amendment filed by Applicants on April 2, 2003.

Dated: September 29, 2003

Respectfully submitted by
EDELL, SHAPIRO & FINNAN, LLC

By: [signature]

Heather Morin
Reg. No. 37,336

EDELL, SHAPIRO & FINNAN, LLC
CUSTOMER NO. 27896
1901 Research Boulevard, Suite 400
Rockville, MD 20850
(301) 424-3640

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/736,825	12/14/2000	Jiebo Luo	81595WFN	8670

7590 04/15/2004

Patent Legal Staff
Eastman Kodak Company
343 State Street
Rochester, NY 14650-2201

EXAMINER	
CARTER, AARON W	
ART UNIT	PAPER NUMBER
2625	13

DATE MAILED: 04/15/2004

Please find below and/or attached an Office communication concerning this application or proceeding.

PTO-90C (Rev. 10/03)

DETAILED ACTION

Information Disclosure Statement

Applicant's information disclosure statement of September 29, 2003 was filed after the issue fee was paid. Information disclosure statements filed after payment of the issue fee will not be considered, but will be placed in the file. However, the application may be withdrawn from issue in order to file a request for continued examination (RCE) under 37 CFR 1.114 upon the grant of a petition under 37 CFR 1.313(c)(2), or a continuing application under 37 CFR 1.53(b) (or a continued prosecution application (CPA) under 37 CFR 1.53(d) if the CPA is for a design patent and the prior application of the CPA is a design application) upon the grant of a petition filed under the provisions of 37 CFR 1.313(c)(3). Alternatively, the other provisions of 37 CFR 1.313 may apply, e.g., a petition to withdraw the application from issue under the provisions of 37 CFR 1.313(c)(1) may be filed together with an unequivocal statement by the applicant that one or more claims are unpatentable over the information contained in the statement. The information disclosure statement would then be considered upon withdrawal of the application from issue under 37 CFR 1.313(c)(1).

Conclusion

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Aaron W Carter whose telephone number is (703) 306-4060. The examiner can normally be reached on 7am - 3:30 am (Mon. - Fri.).

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Bhavesh Mehta can be reached on (703) 308-5246. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

awc

BHAVESH M. MEHTA
SUPERVISORY PATENT EXAMINER
TECHNOLOGY CENTER 2600

PATENT APPLICATION BASED ON: Docket Number 79617DMW

Inventor(s): Jiebo Luo
Robert T. Gray

Attorney: David M. Woods

Document ID: \DOCKETS\79617\APPL.DOC

METHOD FOR AUTOMATICALLY CREATING CROPPED AND ZOOMED VERSIONS OF PHOTOGRAPHIC IMAGES

EASTMAN KODAK COMPANY

RESTRICTED INFORMATION

358/453

345/ 118 1551

"Express Mail" mailing label number EL267111692US

Date of Deposit January 25, 2000

I hereby certify that this paper or fee is being deposited with the United States Postal Service "Express Mail Post Office to "Addressee" service under 37 CFR 1.10 on the date indicated above and is addressed to the Commissioner of Patents and Trademarks, Washington, D.C. 20231

Robin G. Reeves
(Typed or printed name of person mailing paper of fee)

Robin G. Reeves
(Signature of person mailing paper or fee)

METHOD FOR AUTOMATICALLY CREATING CROPPED AND ZOOMED VERSIONS OF PHOTOGRAPHIC IMAGES

BACKGROUND OF THE INVENTION

Field of the Invention

The invention relates generally to the field of digital image processing and digital image understanding and more particularly to a process and system for automatically creating cropped and zoomed versions of photographic images.

Description of the Related Art

For many decades, traditional commercial photo-finishing systems have placed limits on the features offered to consumers to promote mass production. Among those features that are unavailable conventionally, zooming and cropping have been identified by both consumers and photofinishers as extremely useful additional features that could potentially improve the quality of the finished photos and the subsequent picture sharing experiences. With the advent of, and rapid advances in digital imaging, many of the technical barriers that existed in traditional photography no longer stand insurmountable.

Hybrid and digital photography provide the ability to crop undesirable content from a picture, and magnify or zoom the desired content to fill the entire photographic print. In spite of the fact that some traditional cameras with zoom capability provide consumers greater control over composing the desired scene content, studies have found that photographers may still wish to perform a certain amount of cropping and zooming when viewing the finished photograph at a later time. Imprecise viewfinders of many point-and-shoot cameras, as well as simply second-guessing their initial compositions, are factors in the desirability of zoom and crop. In addition, it maybe desirable to use some other regular border templates such as ovals, heart shapes, squares, etc. In another scenario, some people commonly referred to as "scrapbookers" tend to margins around the borders of an image. It does not examine the overall content of the image. In practice, the XV program is effective in cropping out the dark border generated due to imprecise alignment during the scanning process. However, disastrous results can often be produced due to the apparent lack of scene understanding. In some extreme cases, the entire image can be cropped.

Another conventional system, described by Bollman et al. in U.S. patent No. 5,978,519 (incorporated herein by reference), provides a method for cropping images based upon the different intensity levels within the image. With this system, an image to be cropped is scaled down to a grid and divided into non-overlapping blocks. The mean and variance of intensity levels are calculated for each block. Based on the distribution of variances in the blocks, a threshold is selected for the variance. All blocks with a variance higher than the threshold variance are selected as regions of interest. The regions of interest are then cropped to a bounding rectangle. However, such a system is only effective when uncropped images contain regions where intensity levels are uniform and other regions where intensity levels vary considerably. The effectiveness of such a system is expected to be comparable to that of the XV program. The difference is that the XV program examines the image in a line by line fashion to identify uniform areas, while U.S. patent No. 5,978,519 examines the image in a block by block fashion to identify uniform areas. In summary, both techniques cannot deal with images with non-uniform background.

Problems to be Solved by the Invention

The major drawback of conventional techniques is that they do not provide a system for having photographs automatically cropped or zoomed based upon the main subject in the image except, using expensive manual techniques.

SUMMARY OF THE INVENTION

The invention comprises a method and computer program for cropping a digital image that includes inputting a belief map of an image, selecting a zoom factor and a crop window, clustering regions of the belief map to produces a higher sum of beliefs or repeat the positioning process and the moving processes with a rotated image and determine if the rotated image produces a higher sum of beliefs. Similarly, the selector, window mover and cropper can repeat the processes with a second crop window and determine if the second crop window produces a higher sum of beliefs.

ADVANTAGES OF THE INVENTION

One advantage of the invention lies in the ability to automatically crop and zoom photographic images based upon the scene contents. With the invention of the main subject of the image is identified and the cropping and zooming is performed around this main subject. Therefore, the invention produces high-quality zoomed or crops images automatically, regardless whether the background is uniform or not.

BRIEF DESCRIPTION OF THE DRAWINGS

The foregoing and other objects, aspects and advantages will be better understood from the following detailed description of a preferred embodiment of the invention with reference to the drawings, in which:

Figure 1 is a schematic architectural diagram of an embodiment of the invention;

Figure 2 is a schematic architectural diagram of an embodiment of the invention;

Figure 3 is a schematic architectural diagram of an embodiment of the invention;

Figure 4 is a schematic architectural diagram of an embodiment of the invention;

Figures 5 illustrates the application of the invention to a simulated photograph;

Figure 6 illustrates the application of the invention to a simulated photograph;

on 31 December 1998 in the names of Jiebo Luo, Stephen Etz and Amit Singhal, and entitled "Method for Automatic Determination of Main Subjects in Photographic Images" incorporated herein by reference. Main subject detection provides a measure of saliency or relative importance for different regions that are associated with different subjects in an image. Main subject detection enables a discriminative treatment of the scene content for a number of applications related to consumer photographic images, including automatic crop and zoom.

Conventional wisdom in the field of computer vision, which reflects how a human observer would perform such tasks as main subject detection and cropping, calls for a problem-solving path via object recognition and scene content determination according to the semantic meaning of recognized objects. However, generic object recognition remains a largely unsolved problem despite decades of effort from academia and industry.

The MSD system is built upon mostly low-level vision features with semantic information integrated whenever available. This MSD system has a number of sub-tasks, including region segmentation, perceptual grouping, feature extraction, and probabilistic and semantic reasoning. In particular, a large number of features are extracted for each segmented region in the image to represent a wide variety of visual saliency properties, which are then input into a tunable, extensible probability network to generate a belief map containing a continuum of values.

Using MSD, regions that belong to the main subject are generally differentiated from the background clutter in the image. Thus, automatic zoom and crop becomes possible. Automatic zoom and crop is a nontrivial operation that was considered impossible for unconstrained images, which do not necessarily contain uniform background, without a certain amount of scene understanding. In the absence of content-driven cropping, conventional systems have concentrated on simply using a centered crop at a fixed zoom (magnification) factor, or removing the uniform background touching the image borders. The centered crop has been found unappealing to customers.

The output of MSD used by the invention is a list of segmented regions ranked in descending order of their likelihood (or belief) as potential main

To reduce the degrees of freedom in determining the amount of crop, in particular for making photographic prints, in one embodiment, the invention restricts the set of allowable zoom factors {e.g., to1.5x, 2x, 3x, 4x, etc.}. This is based on the findings in the customer focus studies. Although one ordinarily skilled in the art would recognize that the invention can be used with any zoom factor. In one example, the default zoom factor is set at 1.5X.

The subject matter of the present invention relates to digital image understanding technology, which is understood to mean technology that digitally processes a digital image to recognize and thereby to assign useful meaning to human understandable objects, attributes or conditions and then to utilize the results obtained in the further processing of the digital image.

A block diagram of the overall sky detection system (e.g., the digital image understanding technology) is shown in Figure 1. First, a digital image 10 is digitally processed 20. The results 30 obtained from processing step 20 are used along with the original digital image 10 in an image modification step 40 to produce a modified image 50.

A more specific block diagram of the inventive sky detection process is shown in Figure 2, which is discussed in relation to Figures 5-11. Figures 5-11 illustrate the inventive process being applied to an original image shown in Figure 5.

In item 200, the image is input and a belief map is created using MSD. The invention selects a zoom factor (e.g. 1.5X) and a cropped window, as shown in item 201 (e.g. item 60 in Figure 6). This zoom factor can be selected by an operator, or by an automatic method based directly on the main subject belief map (e.g., an estimate of the size of the main subject). The crop window is typically a rectangular window with a certain aspect ratio.

In item 202, regions of the belief map are clustered and the lowest belief cluster (e.g., the background belief) is set to zero using a predefined threshold. As discussed in greater detail below, sections of the image having a belief value below a certain threshold are considered background sections. In item 202 such sections are given a belief of zero for purposes of this embodiment of the invention.

Provided that the primary subjects are included, moving the cropping window so that more of the secondary subjects are included would increase the sum of belief values within the cropping window. Recall that the primary subjects are indicated by the highest belief values and the secondary subjects are indicated by belief values lower than those of the primary subjects but higher than those of the background subjects. The goal is to find the cropping window that has the highest sum of belief values while ensuring that the primary subjects are completely included in the cropping window, i.e.,

$$\tilde{w} = \max_{w \in W} sum(w),$$

where W denotes the set of all possible cropping windows that satisfy all the aforementioned constraints (e.g., those that are completely within the uncropped image and those that encompass the entire primary subjects). Then, in item 212, the image is actually cropped according to the cropped window calculated above, producing the cropped image 214 shown in Figure 10.

If decision box 208 does not produce an acceptable solution, the image is rotated (e.g., by 90 degrees), as shown in item 206, and the processing returns to item 205. At the second and greater pass through decision box 209, if the cropped window has already been rotated and there is still not an acceptable solution (from decision box 208), then the sum of the beliefs for the centroid-based cropped window are computed for both orientations (e.g., original and 90 degrees), as shown in item 211. Then, the image is cropped according to the final position of the higher belief of the two different orientations, as shown in item 213, producing the cropped image, as shown in item 215.

As would be known by one ordinarily skilled in the art given this disclosure, the image can be rotated at any number of potential angles (e.g. 45°,180°, etc.) in item 206. Further, the decision box 209 can allow multiple passes of the process described in items 205, 207 at different rotation angles in an attempt to find the orientation having the highest sum of beliefs.

The simulated image example shown in Figures 5-11 illustrates the progress the invention makes as it moves through the process shown in Figure 2. One could formulate the problem as a global exhaustive search for the best (medium) to capture secondary subjects, or uncertainty, or salient regions of background. Therefore, the invention can perform a k-means clustering with k = 3 on the MSD belief map to "quantize" the beliefs. Consequently, the belief for each region is replaced by the mean belief of the cluster in that region. Note that a k-means clustering with k = 2 essentially produces a binary map with two clusters, "high" and "low," which is undesirable for cropping based on earlier discussion.

There are two major advantages in performing such clustering or quantization. First, clustering helps background separation by grouping low-belief background regions together to form a uniformly low-belief (e.g., zero belief) background region. Second, clustering helps remove noise in belief ordering by grouping similar belief levels together. The centroiding operation does not need such quantization (nor should it be affected by the quantization). The main purpose of the quantization used here is to provide a threshold for the background.

The k-means clustering effectively performs a multi-level thresholding operation to the belief map. After clustering, two thresholds can be determined as follows:

$$threshold_{low} = (C_{low} + C_{med})/2, \quad threshold_{high} = (C_{med} + C_{high})/2$$

where $\{C_{low}, C_{med}, C_{high}\}$ is the set of centroids (average belief values) for the three clusters, and $threshold_{low}$ and $threshold_{high}$ are the low and high thresholds, respectively.

Regions with belief values below the lower threshold are considered "background" and their belief values are set to zero in items 202, 302 and 402 discussed above. Regions with belief values above the higher threshold are considered part of the main subject and need to be included in their entirety, whenever possible. Regions with intermediate belief values (e.g., less than or equal to the higher threshold and greater than or equal to the lower threshold) are considered part of the "secondary subject" and will be included as a whole or partially, if possible, to maximize the sum of main subject belief retained by the cropping window. Note that the variance statistics on the three clusters can be used to set the thresholds more accurately to reflect cluster dispersions.

the algorithm leads to more interesting cropped pictures than simply placing the main subjects (players) in the center of the cropped image (Figure 18). The invention was able to do so because the trees are indicated to be of secondary importance based on the belief map. It is obvious that the art taught by Bradley and Bollman in U.S. Patent 5,978,519 would not be able to produce such a nicely cropped image. In fact, both Bradley and Bollman (U.S. Patent 5,978,519) would at best remove the lower lawn portion of the picture and keep the tree branches in the upper-left of the uncropped image (Figure 18).

While the overall methodology of the invention is described above, the invention can be embodied in any number of different types of systems and executed in any number of different ways, as would be known by one ordinarily skilled in the art. For example, as illustrated in Figure 12, a typical hardware configuration of an information handling/computer system in accordance with the invention preferably has at least one processor or central processing unit (CPU) 1200. For example, the central processing unit 1200 could include various image/texture processing units, mapping units, weighting units, adders, subtractors, comparators, selectors, window movers, croppers, etc. Alternatively, as would be known by one ordinarily skilled in the art given this disclosure, multiple specialized CPU's (or other similar individual functional units) could perform the same processing, mapping, weighting, adding, subtracting, comparing, etc.

The CPU 1200 is interconnected via a system bus 1201 to a random access memory (RAM) 1202, read-only memory (ROM) 1203, input/output (I/O) adapter 1204 (for connecting peripheral devices such as disk units 1205 and tape drives 1206 to the bus 1201), communication adapter 1207 (for connecting an information handling system to a data processing network) user interface adapter 1208 (for connecting a peripherals 1209, 1210 such as a keyboard, mouse, microphone speaker and/or other user interface device to the bus 1201), a printer 1212, and display adapter 1213 (for connecting the bus 1201 to a display device 1214). The invention could be implemented using the structure shown in Figure 12 by including the inventive method within a computer program stored on the storage device 1205. Such a computer program would act

WHAT IS CLAIMED IS:

1. A method of cropping a digital image comprising:

inputting a belief map of a photographic image, said belief map comprising a plurality of belief values, each belief value at each location in said belief map indicating an importance of a photographic subject at said location, wherein a photographic subject having a highest belief value comprises a main subject;

selecting a crop window having a shape and a zoom factor, said shape and zoom factor determining a size of said crop window;

positioning said crop window such that said crop window is centered at a center-of-mass of said main subject;

moving said crop window such that said crop window is completely within said image;

moving said crop window such that a sum of belief of said crop window is at a maximum; and

cropping said image according to said crop window.

2. The method in claim 1, further comprising moving said crop window such that said crop window includes all of at least one main subject.

3. The method in claim 2, further comprising moving said crop window to include a buffer around said main subject.

4. The method in claim 1, further comprising clustering regions of said belief map to identify background subjects and secondary subjects.

5. The method in claim 4, wherein said clustering includes setting said background portions to a zero belief value.

6. The method in claim 1, further comprising repeating said moving processes with a rotated image and determining if said rotated image

setting portions of said belief map having a lowest belief value to a zero belief value.

14. The method in claim 9, further comprising moving said crop window such that said crop window is completely within said image.

15. The method in claim 9, further comprising moving said crop window such that a sum of belief values of said crop window is at a maximum.

16. The method in claim 9, further comprising moving said crop window such that said crop window includes all of said main subject.

17. The method in claim 9, further comprising moving said crop window to include a buffer around said main subject.

18. The method in claim 9, further comprising repeating said moving processes with a rotated image and determining if said rotated image produces a higher sum of beliefs.

19. The method in claim 9, further comprising repeating said positioning process and said moving processes with a rotated image and determining if said rotated image produces a higher sum of beliefs.

20. The method in claim 9, further comprising repeating said processes with a second crop window and determining if said second crop window produces a higher sum of beliefs.

21. A method of using a computer program operating on a computer to crop an image comprising:

using said computer program to input a belief map of a photographic image, said belief map comprising a plurality of belief values, each computer program to move said crop window to include a buffer around said main subject.

29. The method in claim 21, further comprising using said computer program to repeat said move processes with a rotated image and using said computer program to determine if said rotated image produces a higher sum of beliefs.

30. The method in claim 21, further comprising using said computer program to repeat said position process and said move processes with a rotated image and using said computer program to determine if said rotated image produces a higher sum of beliefs.

31. The method in claim 21, further comprising using said computer program to repeat said processes with a second crop window and using said computer program to determine if said second crop window produces a higher sum of beliefs.

32. A system for cropping images comprising:

an input receiving a belief map of a photographic image, said belief map comprising a plurality of belief values, each belief value at each location in said belief map indicating an importance of a photographic subject at said location, wherein a photographic subject having a highest belief value comprises a main subject;

a selector choosing a crop window;

a window mover positioning said crop window such that said crop window is centered around said main subject; and

a cropper modifying said image according to said crop window.

33. The system in claim 32, further comprising a second selector selecting a zoom factor.

6-22-00

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of:

Jiebo Luo, et al

METHOD FOR AUTOMATICALLY CREATING CROPPED AND ZOOMED VERSIONS OF PHOTOGRAPHIC IMAGES

Serial No. US 09/490,915

Filed 25 January 2000

Group Art Unit 2851
Examiner not yet assigned

I hereby certify that this correspondence is being deposited today with the United States Postal Service as first class mail in an envelope addressed to Commissioner for Patents, Washington, D.C. 20231.

Joyce VanDam

Date: **April 20, 2000**

RECEIVED MAY -3 2000 TC 2700 MAIL ROOM

Commissioner for Patents
Washington, D.C. 20231

Sir:

PRELIMINARY AMENDMENT

Prior to examination of the above-identified application, please amend the claims as follows:

In the Claims:

Please add new claims 43-48:

a.

-- 43. A method of cropping a digital image having pixels to produce a cropped digital image, comprising the steps of:

developing a belief map of a photographic image, by using such pixels to determine a series of features, and using such features to assign the probability of the location of a main subject of the digital image in the belief map; and

cropping the digital image to include main subjects indicated by the belief map to produce the cropped digital image.

/27/2000 SLUANG 00000155 050225 09490915
FC:102 312.00 CH
FC:103 108.00 CH

iii) cropping the digital image at the high probability subject content position to produce the cropped digital image.

46. The method of claim 45 wherein the crop window is completely within the digital image.

47. A method of cropping a digital image having pixels to produce a cropped digital image, comprising the steps of:

a) developing a belief map of a photographic image, by using such pixels to determine a series of features, and using such features to assign the probability of the location of a main subject of the digital image in the belief map;

b) performing a clustering of the belief map to identify at least a cluster of highest belief values corresponding to man subject, a cluster of intermediate belief values corresponding to secondary subjects, and a cluster of lowest belief values corresponding to the background;

c) using the belief map to determine the center of the mass of the belief map;

d) positioning a cropped window at a location which includes the center of mass;

e) cropping the digital image to include at least the main subject found in step b) to produce the cropped digital image by:

i) selecting the crop window to have a shape and a zoom factor, said shape and zoom factor determining a size of said crop window;

ii) moving the cropped window to a plurality of positions and using the belief map values to select the position which has a high probability of subject content; and

iii) cropping the digital image at the high probability subject content position to produce the cropped digital image.

48. The method of claim 47 wherein the crop window is completely within the digital image. --

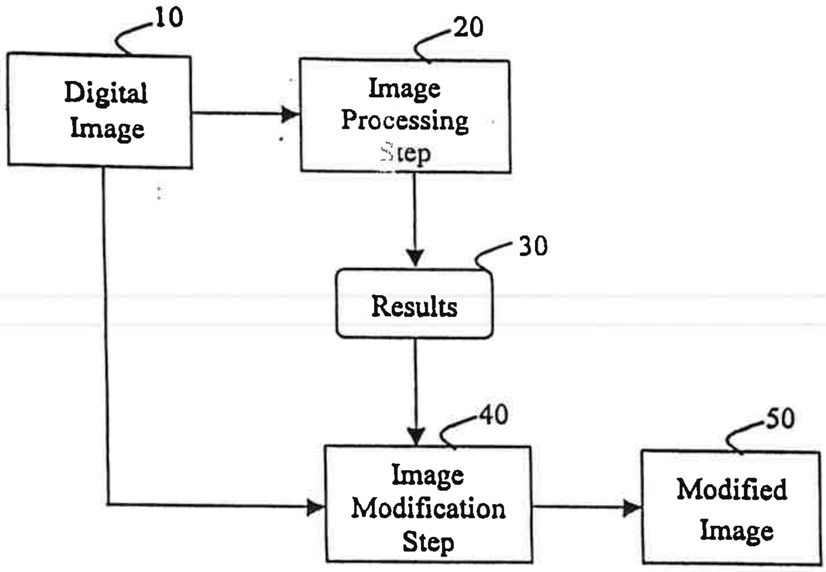

10
Digital Image
20
Image Processing Step
30
Results
40
Image Modification Step
50
Modified Image

Fig.1

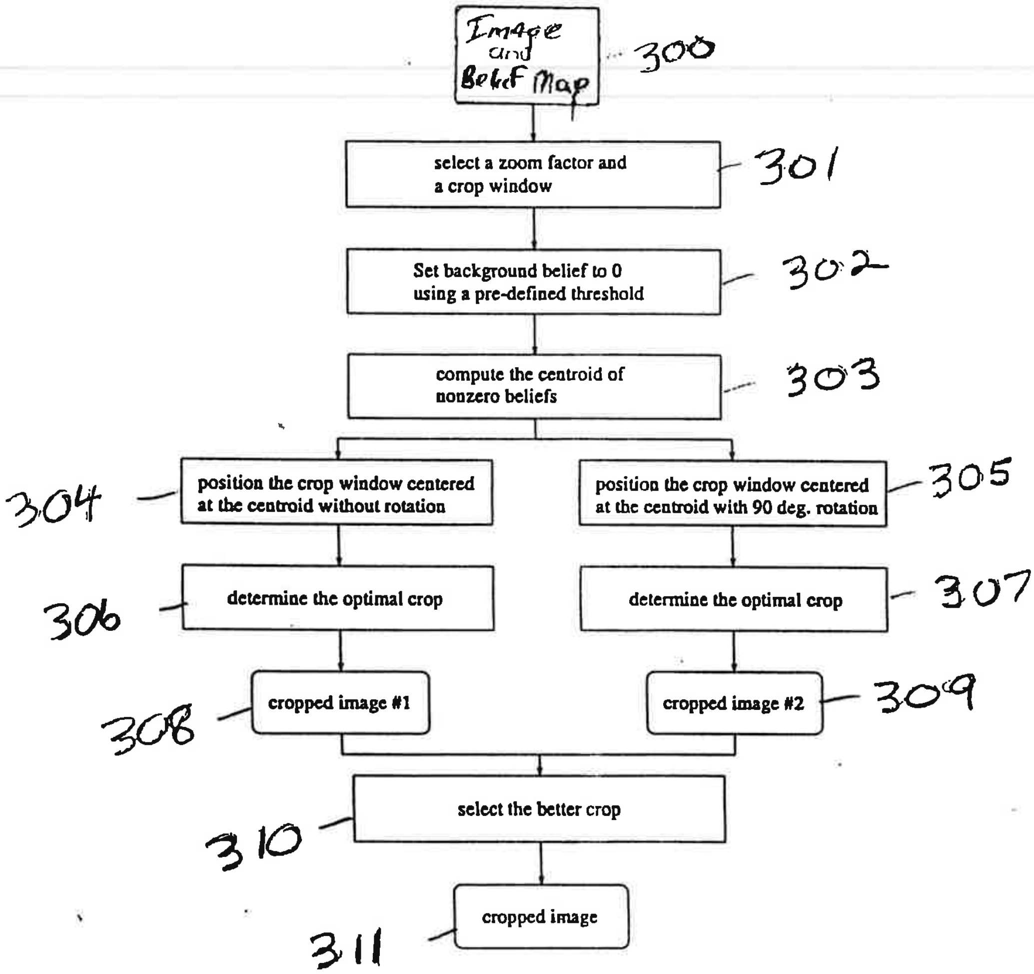


Image and Belief Map
300
select a zoom factor and a crop window
301
Set background belief to 0 using a pre-defined threshold
302
compute the centroid of nonzero beliefs
303
position the crop window centered at the centroid without rotation
304
position the crop window centered at the centroid with 90 deg. rotation
305
determine the optimal crop
306
determine the optimal crop
307
cropped image #1
308
cropped image #2
309
select the better crop
310
cropped image
311

Figure 3

Figure 5

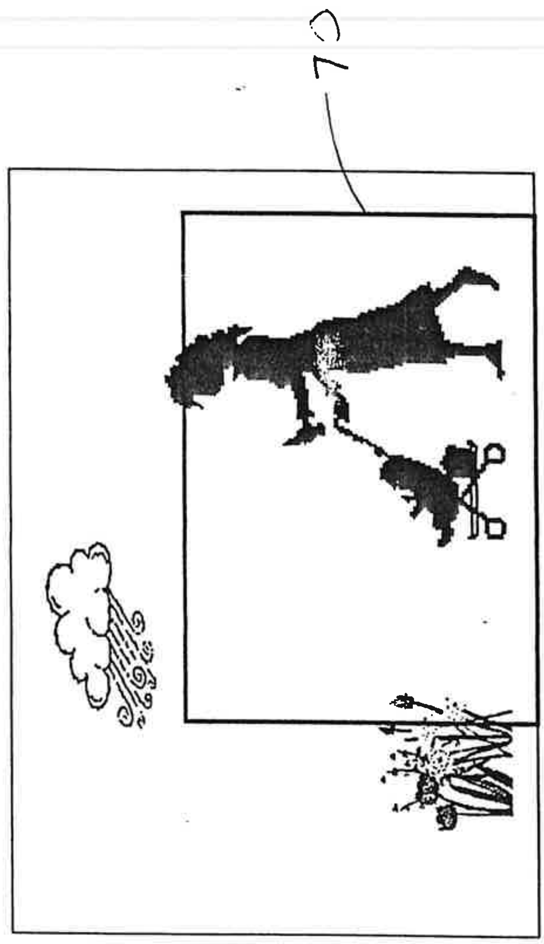


Figure 7

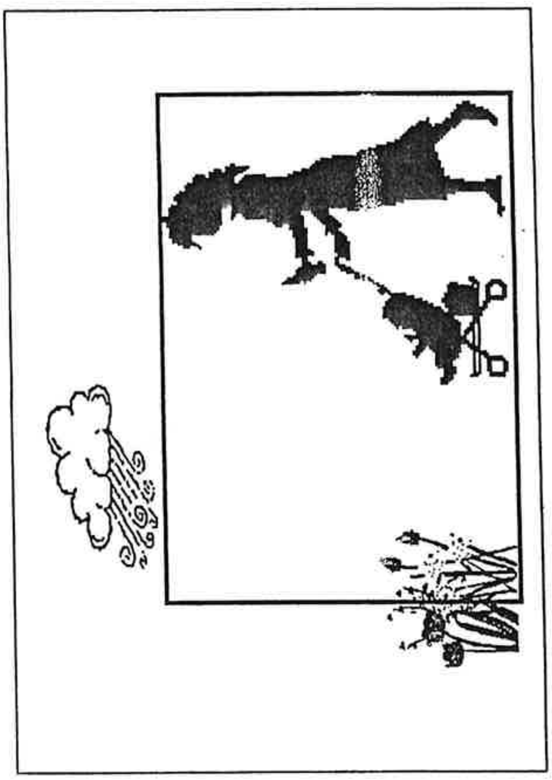


Figure 9

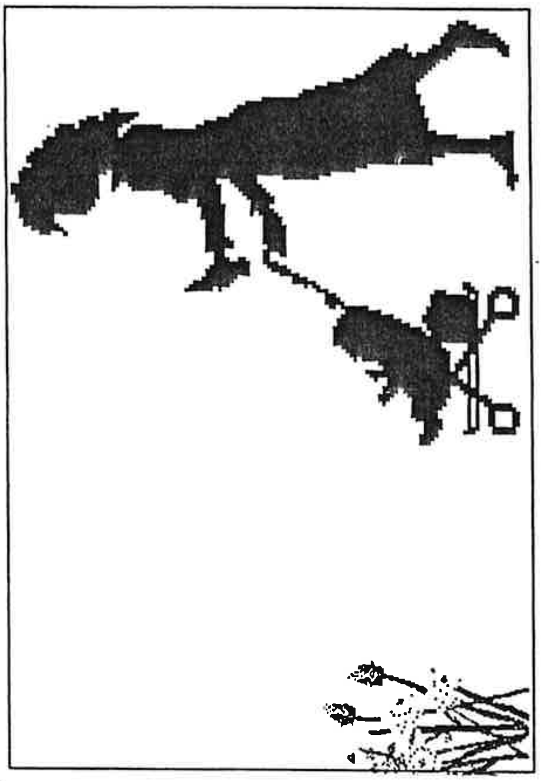


Figure 11

09849015 .012502

Figure 13

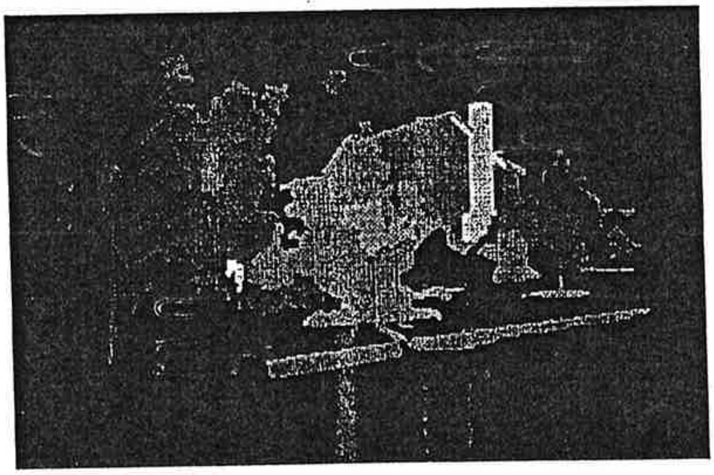


Figure 14

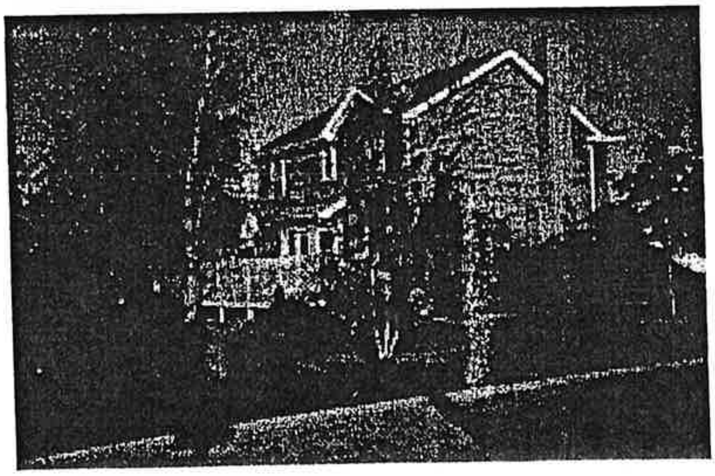


Figure 15

PATENT APPLICATION BASED ON:

Docket No:	81595/WFN
Inventors:	Jiebo Luo
Attorney:	William F. Noval

AUTOMATICALLY PRODUCING AN IMAGE OF A PORTION OF A PHOTOGRAPHIC IMAGE

Commissioner for Patents
Attn: Box Patent Application
Washington, DC 20231

Express Mail Label No: EL267106180US
Date: December 14, 2000

cameras, as well as simply second-guessing their initial compositions, are factors in the desirability of zoom and crop. In addition, it may be desirable to use some other regular border templates such as ovals, heart shapes, squares, etc. In another scenario, some people commonly referred to as "scrapbookers" tend to perform more aggressive crop in making a scrapbook, e.g., cutting along the boundary of objects.

There are significant differences in objectives and behaviors between these two types of cropping, namely album-making and scrapbook making, with the latter more difficult to understand and summarize. The invention described below performs automatic zooming and cropping for making photographic prints. One customer focus group study indicated that it would be beneficial to provide customers a double set of prints -- one regular and one zoom. Moreover, it is preferred that the cropping and zooming be done automatically. Most customers do not want to think about how the zooming and cropping is being done as long as the content and quality (e.g., sharpness) of the cropped and zoomed pictures is acceptable.

There has been little research on automatic zoom and crop due to the apparent difficulty involved in performing such a task. None of the known conventional image manipulation software uses scene content in determining the automatic crop amount. For example, a program entitled "XV", a freeware package developed by John Bradley at University of Pennsylvania, USA (Department of Computer and Information Science), provides an "autocrop" function for manipulating images and operates in the following way:

the program examines a border line of an image, in all of the four directions, namely from the top, bottom, left and right sides;

the program checks the variation within the line. In grayscale images, a line has to be uniform to be cropped. In color images, both the spatial correlation and spectral correlation have to be low, except for a small percentage of pixels, for the line to be qualified for cropping. In other words, a line will not be cropped if it contains a significant amount of variation;

if a line along one dimension passes the criterion, the next line (row or column) inward is then examined; and magnification of prints made from film originals utilizing a fixed optical lens instead of zoom lens. In U.S. Patent 5,872,619, Stephenson et al. describe a method of printing photographs from a processed photographic filmstrip having images of different widths measured longitudinally of the filmstrip and having heights measured transversely of the filmstrip. This method uses a photographic printer having a zoom lens and a printing mask to provide printed images having a selected print width and a selected print height. In U.S. Patent 4,809,064, Amos et al. describe an apparatus for printing a selected region of a photographic negative onto a photosensitive paper to form an enlarged and cropped photographic print. This apparatus includes means for projecting the photographic negative onto first and second zoom lenses, each of the zoom lenses having an adjustable magnification. In U.S. Patent 5,872,643, Maeda et al. describe a film reproducing apparatus that can effectively perform zoom and crop. This apparatus includes an image pick-up device which picks up a film frame image recorded on a film to generate image data, an information reader which reads information about photographing conditions of the film frame image, and a reproducing area designator which designates a reproducing area of the film frame image. However, the reproducing area of the film frame image is determined based on pre-recorded information about the position of the main object, as indicated by which zone of the photograph the automatic focusing (AF) operation in the camera was on – part of the recorded information about photographing conditions. In all the above-mentioned optical printing systems, the position of the photographic film sample and magnification factor of the relay lens are pre-selected.

SUMMARY OF THE INVENTION

According to the present invention, there is provided a solution to the problems of the prior art. It is an object of the present invention to provide a method for producing a portion of a photographic image by identifying the main subject of the photographic image.

According to a feature of the present invention, there is provided a method of producing an image of at least a portion of a digital image,

Fig. 6 illustrates the application of the invention to a simulated photograph;

Fig. 7 illustrates the application of the invention to a simulated photograph;

Fig. 8 illustrates the application of the invention to a simulated photograph;

Fig. 9 illustrates the application of the invention to a simulated photograph;

Fig. 10 illustrates the application of the invention to a simulated photograph;

Fig. 11 illustrates the application of the invention to a simulated photograph;

Fig. 12 illustrates the application of the invention to a simulated photograph;

Fig. 13 is an exemplary uncropped photograph;

Fig. 14 is a belief map of the image shown in FIG. 13;

Fig. 15 is a cropped version of the image shown in FIG. 13;

Fig. 17 is a belief map of the image shown in FIG. 16; and

Fig. 18 is a cropped version of the image shown in FIG. 16.

DETAILED DESCRIPTION OF THE INVENTION

The invention automatically zooms and crops digital images according to an analysis of the main subject in the scene. Previously, a system for detecting main subjects (e.g., main subject detection or "MSD") in a consumer-type photographic image from the perspective of a third-party observer has been developed and is described in U.S. Patent Application Serial No. 09/223,860, filed December 31, 1998, the disclosure of which is incorporated herein by reference. Main subject detection provides a measure of saliency or relative importance for different regions that are associated with different subjects in an image. Main subject detection enables a discriminative treatment of the scene content for a number of applications related to consumer photographic images, including automatic crop and zoom.

To some extent, this belief map reflects the inherent uncertainty for humans to perform such a task as MSD because different observers may disagree on certain subject matter while agreeing on other subject matter in terms of main subjects. However, a binary decision, when desired, can be readily obtained by using an appropriate threshold on the belief map. Moreover, the belief information may be very useful for downstream applications. For example, different weighting factors can be assigned to different regions (subject matters) in determining the amount of crop.

For determination of crop, the invention uses the main subject belief map instead of a binarized version of the map to avoid making a bad cropping decision that is irreversible. Furthermore, using the continuous values of the main subject beliefs helps trade-off different regions under the constraints encountered in cropping. A binary decision on what to include and what not to include, once made, leaves little room for trade-off. For example, if the main subject region is smaller than the crop window, the only reasonable choice, given a binary main subject map, is to leave equal amounts of margin around the main subject region. On the other hand, secondary main subjects are indicated by lower belief values in the main subject belief map, and can be included according to a descending order of belief values once the main subject of highest belief values are included. Moreover, if an undesirable binary decision on what to include/exclude is made, there is no recourse to correct the mistake. Consequently, the cropping result becomes sensitive to the threshold used to obtain the binary decision. With a continuous-valued main subject belief map, every region or object is associated with a likelihood of being included or a belief value in its being included.

To reduce the degrees of freedom in determining the amount of crop, and to limit the amount of resolution loss incurred in the zoom process, in particular for making photographic prints, in one embodiment, the invention restricts the set of allowable zoom factors to the range of [1.2, 4]. This is based on the findings in the customer focus studies. Those skilled in the art would recognize that the present invention could be used with any the zoom factor.

in the form of a film sample position 9. The photographic film sample is positioned in a gate device 36 which holds the film negative in place during the exposure. The gate device 36 receives the film sample position 9 to position the photographic film sample to adjust which portion of the imaging area of the photograph will be printed.

Referring to Fig. 1a, a lamp house 34 provides the illumination source which is transmitted through the photographic film sample 31 and focused by a lens 12 onto photographic paper 38. The time integration device 13 opens and closes a shutter for a variable length of time allowing the focused light from the lamp house 34 to expose the photographic paper 38. The exposure control device 16 receives a brightness balance value from the digital image processor 20. The exposure control device 16 uses the brightness balance value to regulate the length of time the shutter of the time integration device stays open.

A block diagram of the inventive cropping process (e.g., the digital image understanding technology) is shown in Fig. 3, which is discussed in relation to Figs. 5-12. Figs. 5-12 illustrate the inventive process being applied to an original image shown in Fig. 5.

In item 200, the belief map is created using MSD. The present invention automatically determines a zoom factor (e.g. 1.5X) and a crop window 80 (as shown in Fig. 7), as referred to in item 201 of Fig. 3. This zoom factor is selected by an automatic method based directly on the main subject belief map (e.g., an estimate of the size of the main subject). The crop window is typically a rectangular window with a certain aspect ratio. After the zoom factor is determined by the digital image processor 20, the value of the zoom factor is used subsequently by the digital image processor 20 shown in Fig. 1. In Fig. 1a, the zoom factor is used to communicate with the lens 12 to adjust the lens magnification setting. This adjustment allows the lens 12 to image the appropriate size of the photographic film sample 31 onto the photographic paper 38.

In item 201, regions of the belief map are clustered and the lowest belief cluster (e.g., the background belief) is set to zero using a predefined threshold. As discussed in greater detail below, sections of the image having a belief value below a certain threshold are considered background sections. In

Fig. 6 illustrates that the effective bounding rectangle 70 is centered at approximately the top of the boy's head and approximately encompasses the region of high subject content. In general, the aspect ratio of the original image is maintained. Therefore, a crop window 80 is determined in item 303 such that it is the smallest rectangle of the original aspect ratio that encompasses the effective MBR 70.

In item 204, the initial position of the crop window p 80 is centered at the centroid, as shown in Fig. 7.

The crop window is 80 then moved so that the entire crop window is within the original image (e.g. item 205) as shown in Fig. 8. In item 206, the crop window 80 is moved again so that all the regions of the highest belief values ("main subject") are included within the crop window and to create a margin 81, as shown in FIG. 9. This process (e.g., 206) captures the entire subject of interest. Therefore, as shown in Fig. 9, the top of the woman's head is included in the crop window. Compare this to Fig. 8 where the top of the woman's head was outside the crop window.

Decision box 207 determines whether an acceptable solution has been found, i.e., whether it is possible to include at least the regions of the highest belief values in the crop window.

If an acceptable solution exists, the window is again moved, as shown in item 208, to optimize a subject content index for the crop window. The preferred embodiment of the present invention defines the subject content index as the sum of belief values within the crop window. It should be noted that the present invention specifies higher numerical belief values corresponding to higher main subject probability. Therefore, finding a numerical maximum of the sum of the belief values is equivalent to finding an optimum of the subject content index. This is shown in Fig. 10 where the secondary objects (e.g. flowers) are included within the crop window 80 to increase the sum of beliefs. The sum of beliefs for a crop window is computed as follows.

$$sum(w) = \sum_{(x,y)\in w} bel(x,y),$$

One could formulate the problem as a global exhaustive search for the best solution. The procedure used in the invention is considered a "greedy" searching approach and is certainly more efficient than conventional processes.

The invention utilizes a built-in "k-means" clustering process to determine proper thresholds of MSD beliefs for each application. The invention also uses clustering, as discussed below to enhance the cropping process. In one preferred embodiment, it is sufficient to use three levels to quantize MSD beliefs, namely "high", "medium", and "low." As would be known by one ordinarily skilled in the art, the invention is not limited to simply three levels of classification, but instead can utilize a reasonable number of classification levels to reduce the (unnecessary) variation in the belief map. These three levels allow for the main subject (high), the background (low), and an intermediate level (medium) to capture secondary subjects, or uncertainty, or salient regions of background. Therefore, the invention can perform a k-means clustering with $k = 3$ on the MSD belief map to "quantize" the beliefs. Consequently, the belief for each region is replaced by the mean belief of the cluster in that region. Note that a k-means clustering with $k = 2$ essentially produces a binary map with two clusters, "high" and "low," which is undesirable for cropping based on earlier discussion.

There are two major advantages in performing such clustering or quantization. First, clustering helps background separation by grouping low-belief background regions together to form a uniformly low-belief (e.g., zero belief) background region. Second, clustering helps remove noise in belief ordering by grouping similar belief levels together. The centroiding operation does not need such quantization (nor should it be affected by the quantization). The main purpose of the quantization used here is to provide a threshold for the background.

The k-means clustering effectively performs a multi-level thresholding operation to the belief map. After clustering, two thresholds can be determined as follows:

A computer program product may include one or more storage medium, for example; magnetic storage media such as magnetic disk (such as a floppy disk) or magnetic tape; optical storage media such as optical disk, optical tape, or machine readable bar code; solid-state electronic storage devices such as random access memory (RAM), or read-only memory (ROM); or any other physical device or media employed to store a computer program having instructions for practicing a method according to the present invention.

While the invention has been described in terms of preferred embodiments, those skilled in the art will recognize that the invention can be practiced with modification within the spirit and scope of the appended claims.

The subject matter of the present invention relates to digital image understanding technology, which is understood to mean technology that digitally processes a digital image to recognize and thereby assign useful meaning to human understandable objects, attributes or conditions and then to utilize the results obtained in the further processing of the digital image.

The invention has been described in detail with particular reference to certain preferred embodiments thereof, but it will be understood that variations and modifications can be effected within the spirit and scope of the invention.

iii) cropping the digital image according to the position of the crop window.

4. The method of claim 2 wherein step d) further comprises the steps of:

i) selecting a crop window of a rectangular shape and of an identical aspect ratio to the (uncropped) digital image; and

ii) selecting a zoom factor to determine the size of the crop window such that it encompasses the effective bounding box.

5. The method of claim 2 wherein the weighting function in step b) is a linear weighting function.

6. The method of claim 2 wherein the weighting function in step b) is a constant function.

7. The method of claim 3 wherein step b) further comprises the steps of:

i) calculating a subject content index for the crop window derived from the belief values;

ii) following a positioning procedure of repeating step i) for at least two positions of the crop window; and

iii) using the subject content index values to select the crop window position.

8. The method of claim 1 wherein the crop window is completely within the digital image.

9. The method of claim 2 wherein step b) further comprises the step of performing a clustering of the belief map to identify at least a cluster of highest belief values corresponding to main subject, a cluster of intermediate

image, a lens assembly, and a photographic receiver in response to the belief map and illuminating a portion of the photographic image of high subject content to produce an image of such portion onto the photographic receiver.

16. The method of claim 15 wherein step c) further comprises the steps of:

i) computing a weighted center-of-mass of the belief map, weighted by the belief values of the belief map;

ii) computing weighted central moments of the belief map, relative to the center-of-mass and weighted by a weighting function of each belief value of the belief map;

iii) computing an effective rectangular bounding box according to the central moments; and

iv) determining a crop window having a shape and a zoom factor, the shape and zoom factor determining a size of the crop window.

17. The method of claim 15 wherein step d) further comprises the steps of:

i) selecting an initial position of the crop window at a location which includes the center of mass;

ii) using the belief values corresponding to the crop window to select the position of the crop window to include a portion of the image of high subject content in response to the belief map; and

iii) cropping the digital image according to the position of the crop window.

18. The method of claim 16 wherein step d) further comprises the steps of:

i) selecting a crop window of a rectangular shape and of an identical aspect ratio to the (uncropped) digital image; and

ii) selecting a zoom factor to determine the size of the crop

26. The method of claim 17 further comprising positioning said crop window such that said crop window includes all of said main subject cluster.

27. The method of claim 26 further comprising positioning said crop window to include a buffer around said main subject cluster.

28. A computer storage product having at least one computer storage medium having instructions stored therein causing one or more computers to perform the method of claim 1.

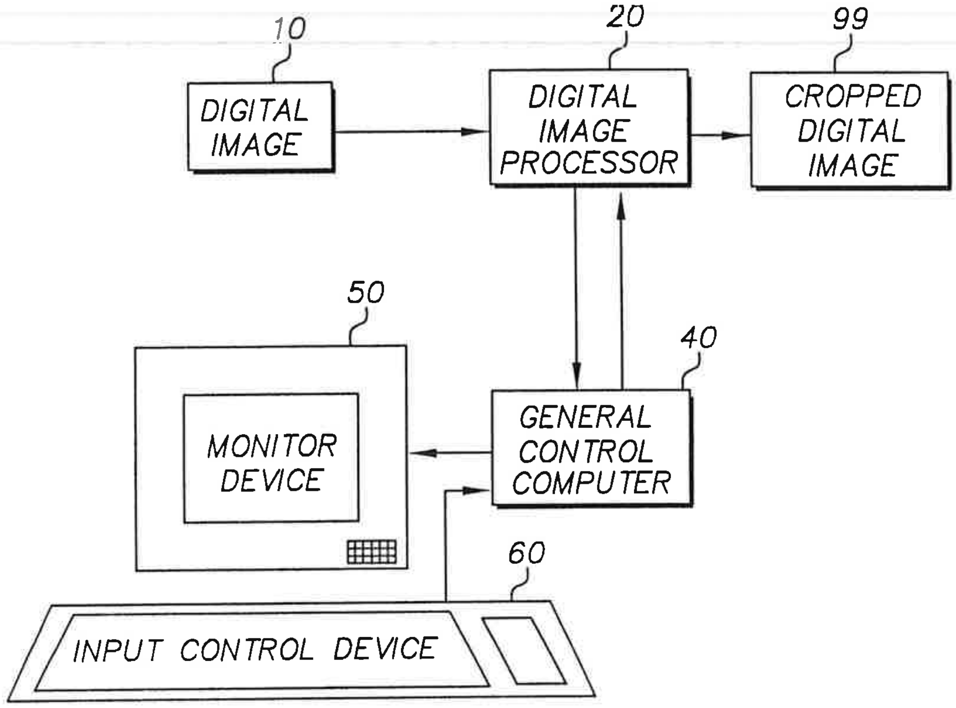

10
DIGITAL IMAGE
20
DIGITAL IMAGE PROCESSOR
99
CROPPED DIGITAL IMAGE
50
MONITOR DEVICE
40
GENERAL CONTROL COMPUTER
60
INPUT CONTROL DEVICE

FIG. 1

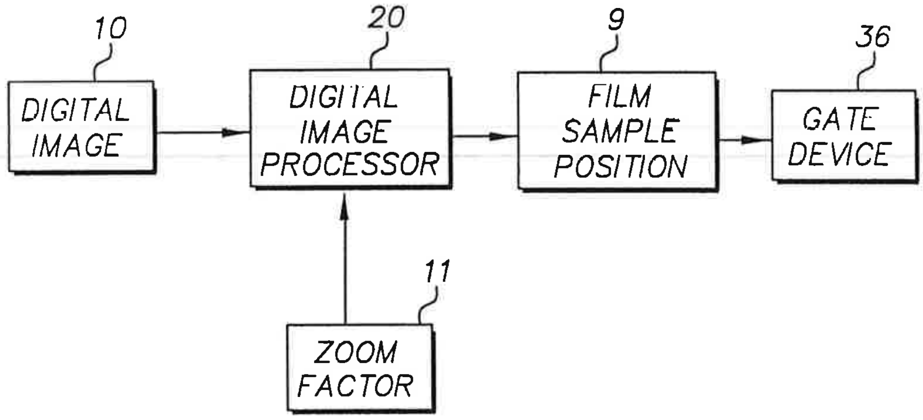

10
DIGITAL IMAGE
20
DIGITAL IMAGE PROCESSOR
9
FILM SAMPLE POSITION
36
GATE DEVICE
11
ZOOM FACTOR

FIG. 2

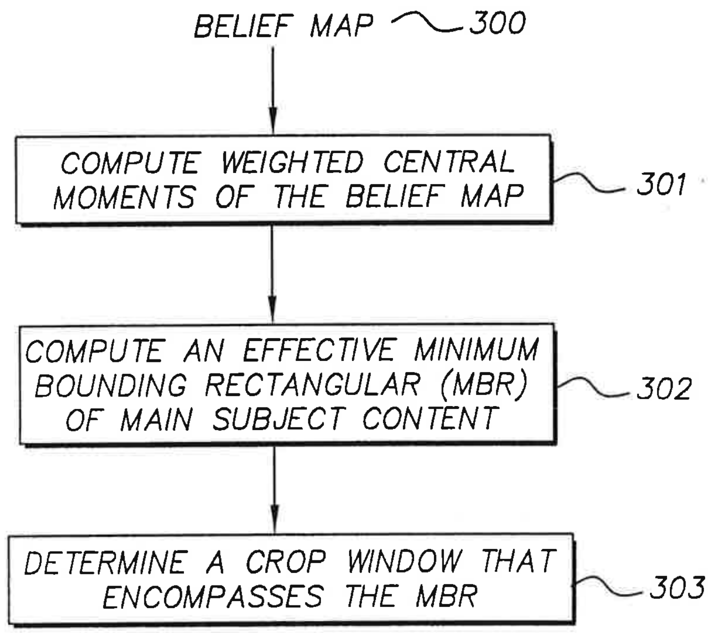

BELIEF MAP
300
COMPUTE WEIGHTED CENTRAL MOMENTS OF THE BELIEF MAP
301
COMPUTE AN EFFECTIVE MINIMUM BOUNDING RECTANGULAR (MBR) OF MAIN SUBJECT CONTENT
302
DETERMINE A CROP WINDOW THAT ENCOMPASSES THE MBR
303

FIG. 4

Fig. 5

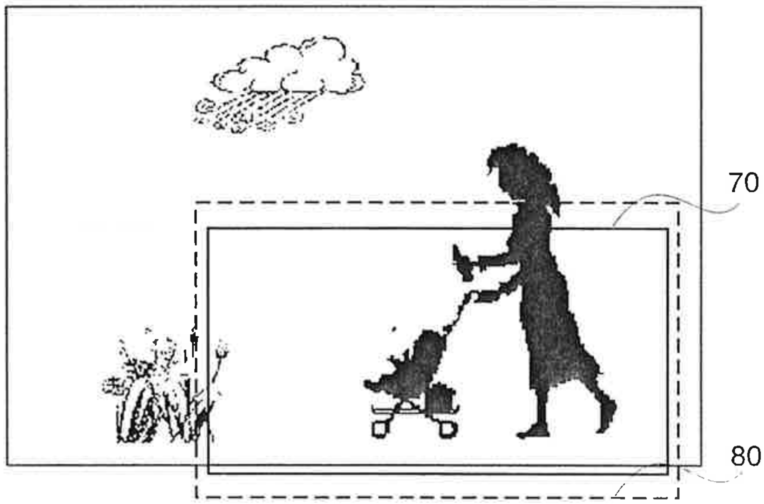


Fig. 6

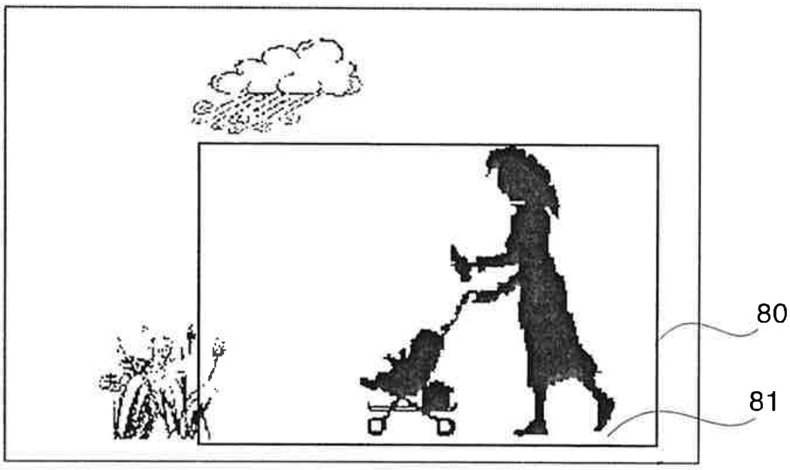


Fig. 9

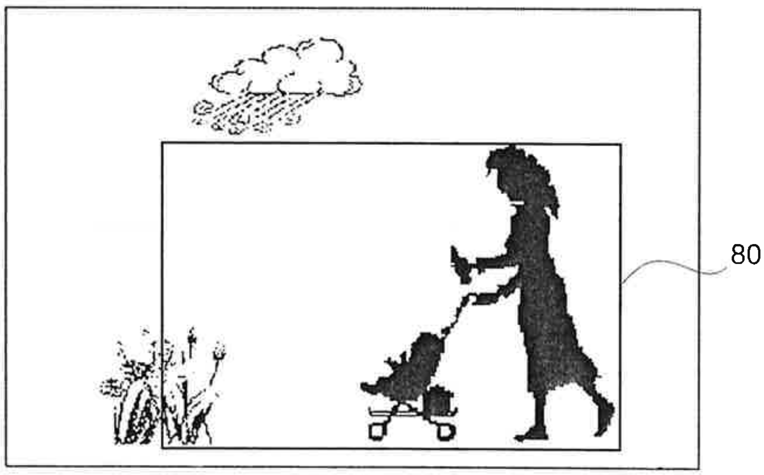


Fig. 10

Fig. 13

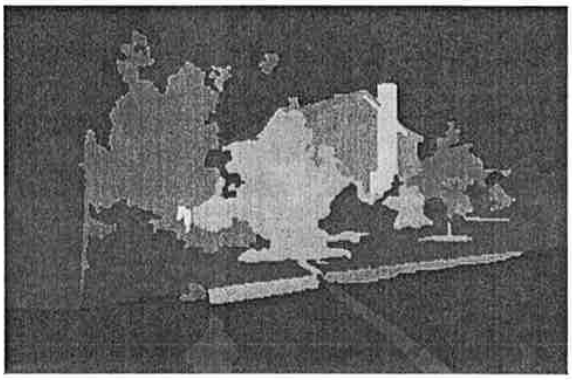


Fig. 14

Fig. 15

Summary of Record of Interview Requirements

Manual of Patent Examining Procedure (MPEP), Section 713.04, Substance of Interview Must be Made of Record
A complete written statement as to the substance of any face-to-face, video conference, or telephone interview with regard to an application must be made of record in the application whether or not an agreement with the examiner was reached at the interview.

Title 37 Code of Federal Regulations (CFR) § 1.133 Interviews
Paragraph (b)
In every instance where reconsideration is requested in view of an interview with an examiner, a complete written statement of the reasons presented at the interview as warranting favorable action must be filed by the applicant. An interview does not remove the necessity for reply to Office action as specified in §§ 1.111, 1.135. (35 U.S.C. 132)

37 CFR §1.2 Business to be transacted in writing.
All business with the Patent or Trademark Office should be transacted in writing. The personal attendance of applicants or their attorneys or agents at the Patent and Trademark Office is unnecessary. The action of the Patent and Trademark Office will be based exclusively on the written record in the Office. No attention will be paid to any alleged oral promise, stipulation, or understanding in relation to which there is disagreement or doubt.

The action of the Patent and Trademark Office cannot be based exclusively on the written record in the Office if that record is itself incomplete through the failure to record the substance of interviews.

It is the responsibility of the applicant or the attorney or agent to make the substance of an interview of record in the application file, unless the examiner indicates he or she will do so. It is the examiner's responsibility to see that such a record is made and to correct material inaccuracies which bear directly on the question of patentability.

Examiners must complete an Interview Summary Form for each interview held where a matter of substance has been discussed during the interview by checking the appropriate boxes and filling in the blanks. Discussions regarding only procedural matters, directed solely to restriction requirements for which interview recordation is otherwise provided for in Section 812.01 of the Manual of Patent Examining Procedure, or pointing out typographical errors or unreadable script in Office actions or the like, are excluded from the interview recordation procedures below. Where the substance of an interview is completely recorded in an Examiners Amendment, no separate Interview Summary Record is required.

The Interview Summary Form shall be given an appropriate Paper No., placed in the right hand portion of the file, and listed on the "Contents" section of the file wrapper. In a personal interview, a duplicate of the Form is given to the applicant (or attorney or agent) at the conclusion of the interview. In the case of a telephone or video-conference interview, the copy is mailed to the applicant's correspondence address either with or prior to the next official communication. If additional correspondence from the examiner is not likely before an allowance or if other circumstances dictate, the Form should be mailed promptly after the interview rather than with the next official communication.

The Form provides for recordation of the following information:
- Application Number (Series Code and Serial Number)
- Name of applicant
- Name of examiner
- Date of interview
- Type of interview (telephonic, video-conference, or personal)
- Name of participant(s) (applicant, attorney or agent, examiner, other PTO personnel, etc.)
- An indication whether or not an exhibit was shown or a demonstration conducted
- An identification of the specific prior art discussed
- An indication whether an agreement was reached and if so, a description of the general nature of the agreement (may be by attachment of a copy of amendments or claims agreed as being allowable). Note: Agreement as to allowability is tentative and does not restrict further action by the examiner to the contrary.
- The signature of the examiner who conducted the interview (if Form is not an attachment to a signed Office action)

It is desirable that the examiner orally remind the applicant of his or her obligation to record the substance of the interview of each case. It should be noted, however, that the Interview Summary Form will not normally be considered a complete and proper recordation of the interview unless it includes, or is supplemented by the applicant or the examiner to include, all of the applicable items required below concerning the substance of the interview.

A complete and proper recordation of the substance of any interview should include at least the following applicable items:
1) A brief description of the nature of any exhibit shown or any demonstration conducted,
2) an identification of the claims discussed,
3) an identification of the specific prior art discussed,
4) an identification of the principal proposed amendments of a substantive nature discussed, unless these are already described on the Interview Summary Form completed by the Examiner,
5) a brief identification of the general thrust of the principal arguments presented to the examiner,
   (The identification of arguments need not be lengthy or elaborate. A verbatim or highly detailed description of the arguments is not required. The identification of the arguments is sufficient if the general nature or thrust of the principal arguments made to the examiner can be understood in the context of the application file. Of course, the applicant may desire to emphasize and fully describe those arguments which he or she feels were or might be persuasive to the examiner.)
6) a general indication of any other pertinent matters discussed, and
7) if appropriate, the general results or outcome of the interview unless already described in the Interview Summary Form completed by the examiner.

Examiners are expected to carefully review the applicant's record of the substance of an interview. If the record is not complete and accurate, the examiner will give the applicant an extendable one month time period to correct the record.

Examiner to Check for Accuracy

If the claims are allowable for other reasons of record, the examiner should send a letter setting forth the examiner's version of the statement attributed to him or her. If the record is complete and accurate, the examiner should place the indication, "Interview Record OK" on the paper recording the substance of the interview along with the date and the examiner's initials.

#3

OIPE JC183
MAR 1 6 2001
PATENT & TRADEMARK OFFICE

81595WFN
Customer No. 01333

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of: Jiebo Luo AUTOMATICALLY PRODUCING AN IMAGE OF A PORTION OF A PHOTOGRAPHIC IMAGE Serial No. 09/736,825 Filed December 14, 2000	Group Art Unit: 2621 Batch: Allowed: Examiner: I hereby certify that this correspondence is being deposited today with the United States Postal Service as first class mail in an envelope addressed to Commissioner for Patents, Washington, D.C. 20231. Laurie A. Wurtz Laurie A. Wurtz March 14, 2001 Date

Commissioner for Patents
Washington, D.C. 20231

Sir:

LETTER TO THE OFFICIAL DRAFTSPERSON

Enclosed are 10 sheets of formal drawings depicting Figure(s) 1-18. Please substitute these drawings for those currently on file in the subject application. These drawings correct the informalities noted in the Notice to File Corrected Application Papers mailed February 7, 2001.

The Commissioner is hereby authorized to charge any fees in connection with this communication to Eastman Kodak Company Deposit Account No. 05-0225. **A duplicate copy of this letter is enclosed.**

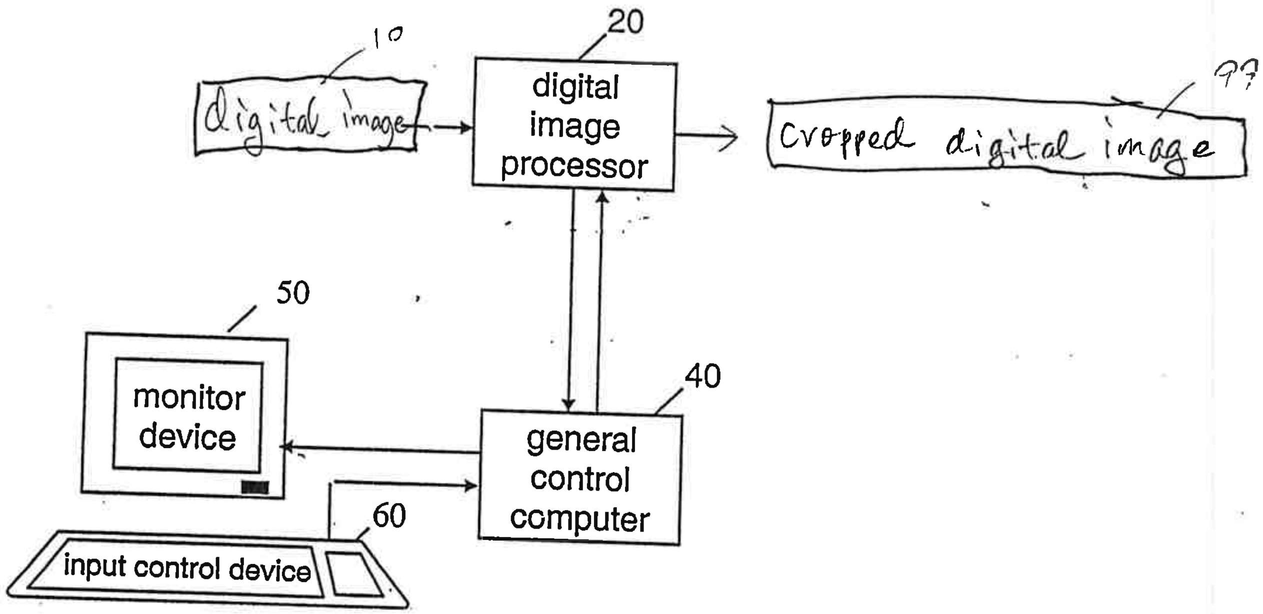
Respectfully submitted,

William F. Noval
Attorney for Applicant
Registration No. 22,049

William F. Noval/law
Telephone: (716) 477-5272
Facsimile: (716) 477-4646

E:\grp\THC\WFN\dockets\81595\drftprsn.doc

Figure 1



10
digital image
20
digital image processor
Cropped digital image
40
general control computer
50
monitor device
60
input control device

Figure 1a

(printer exposure configuration)

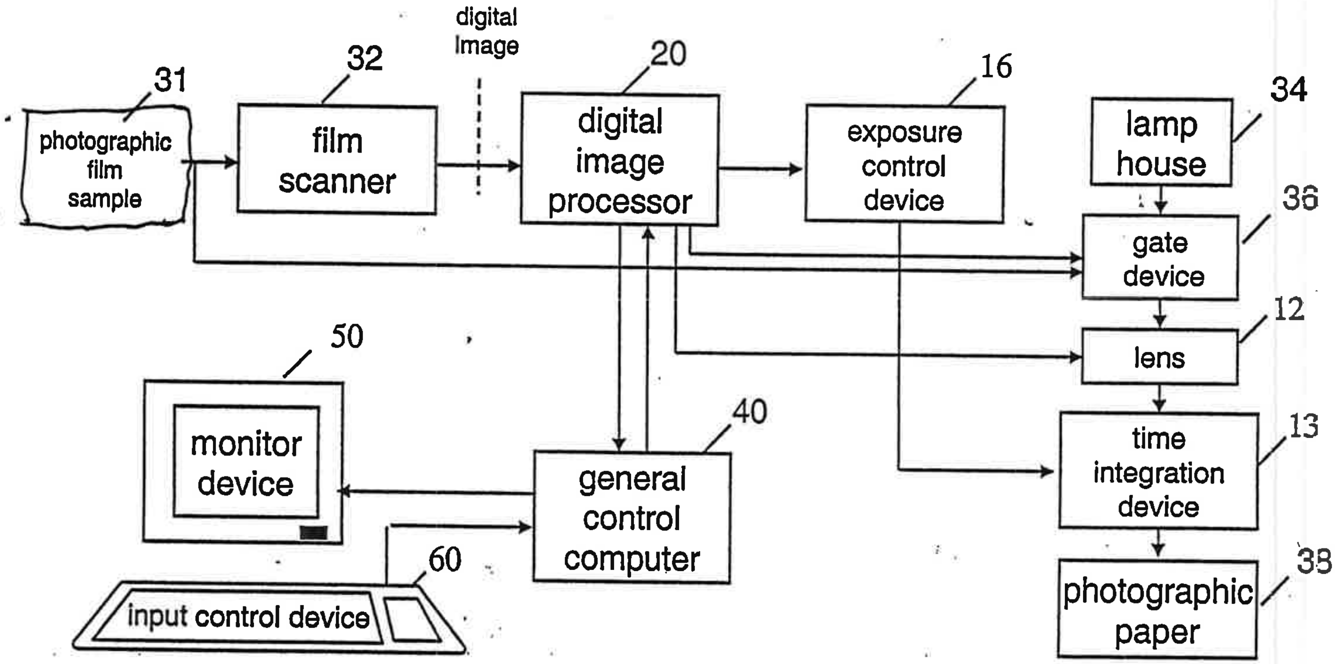
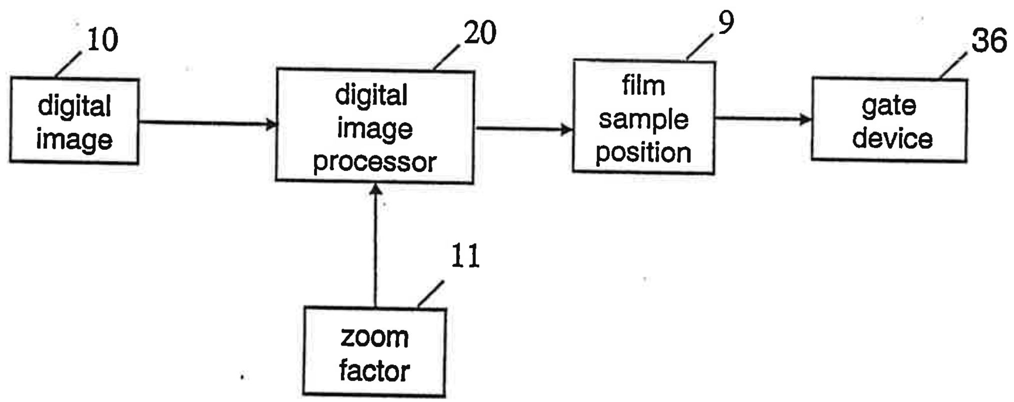


Figure 2

(image processor)



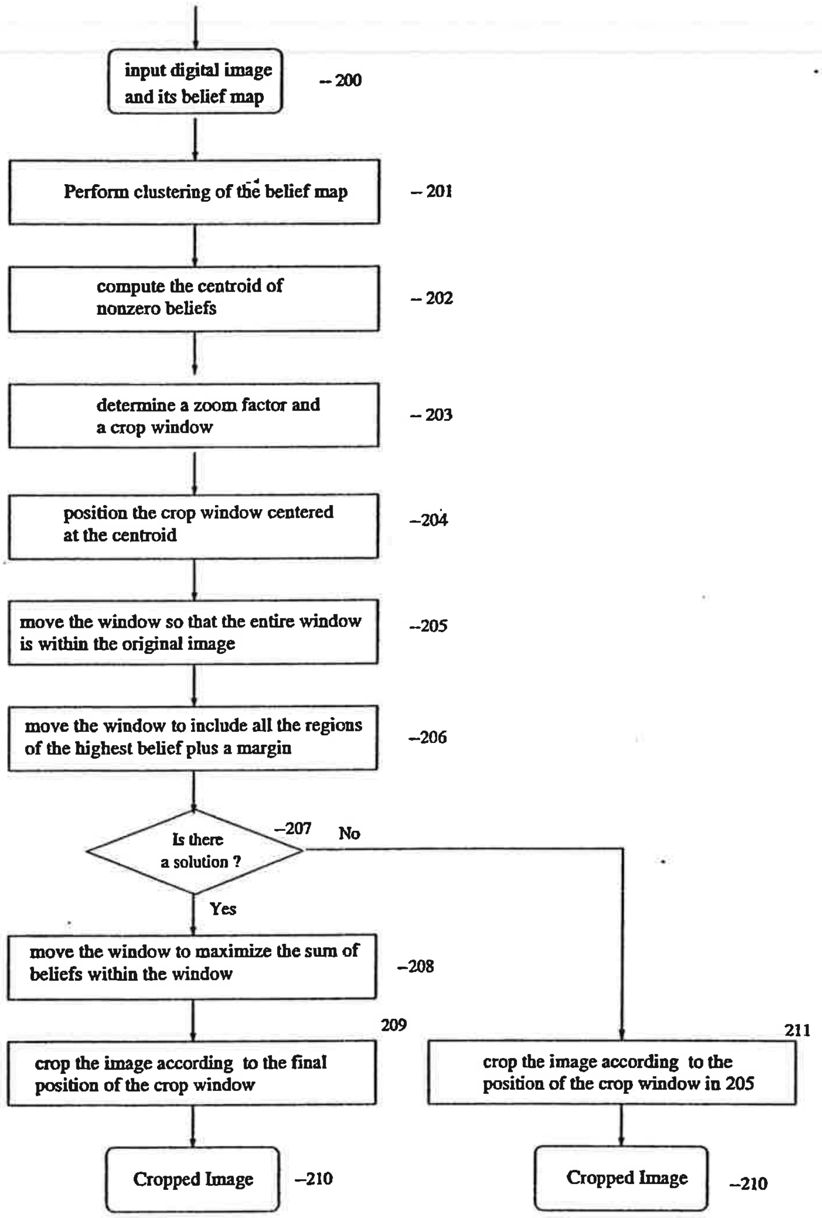

input digital image and its belief map
200
Perform clustering of the belief map
201
compute the centroid of nonzero beliefs
202
determine a zoom factor and a crop window
203
position the crop window centered at the centroid
204
move the window so that the entire window is within the original image
205
move the window to include all the regions of the highest belief plus a margin
206
Is there a solution ?
207
No
Yes
move the window to maximize the sum of beliefs within the window
208
209
crop the image according to the final position of the crop window
211
crop the image according to the position of the crop window in 205
Cropped Image
210
Cropped Image
210

FIG. 3

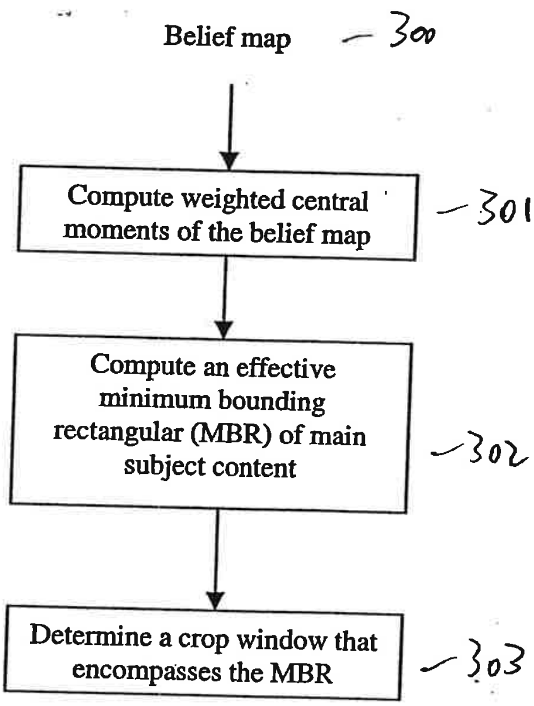

Belief map
300
Compute weighted central moments of the belief map
301
Compute an effective minimum bounding rectangular (MBR) of main subject content
302
Determine a crop window that encompasses the MBR
303

FIG. 4

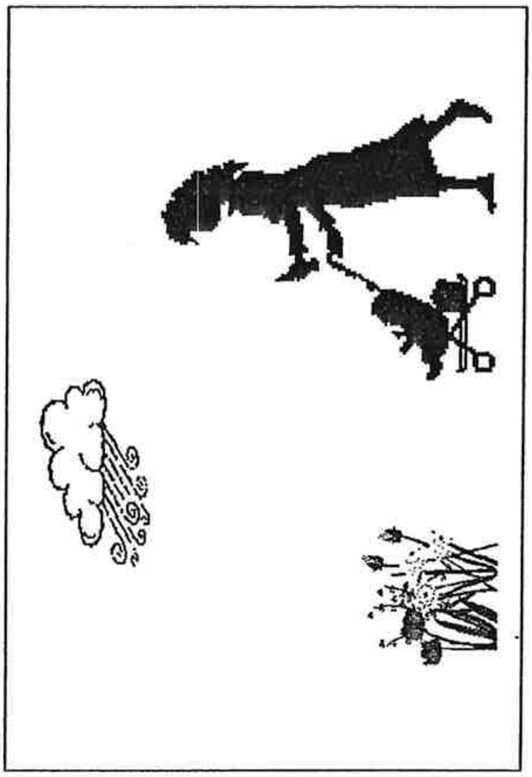


FIG. 5

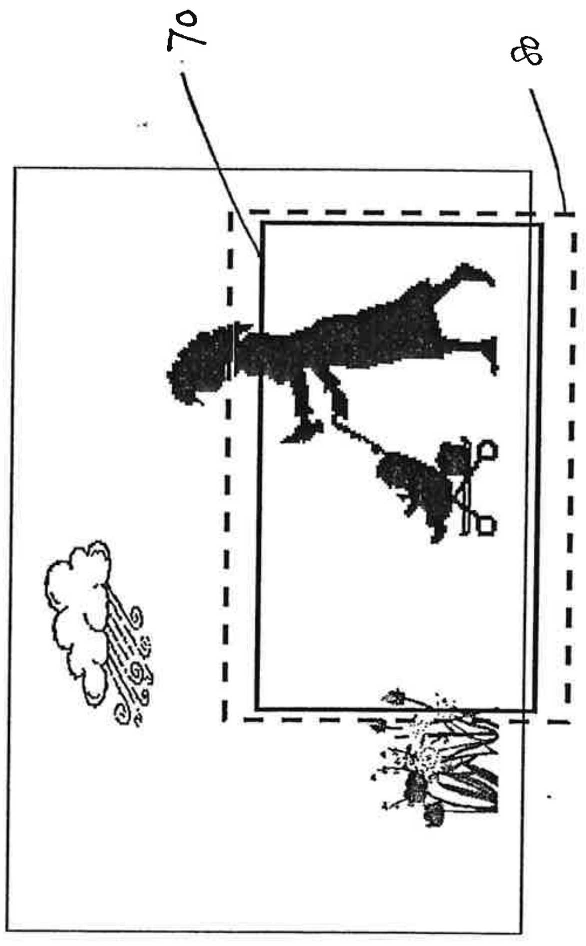


FIG. 6

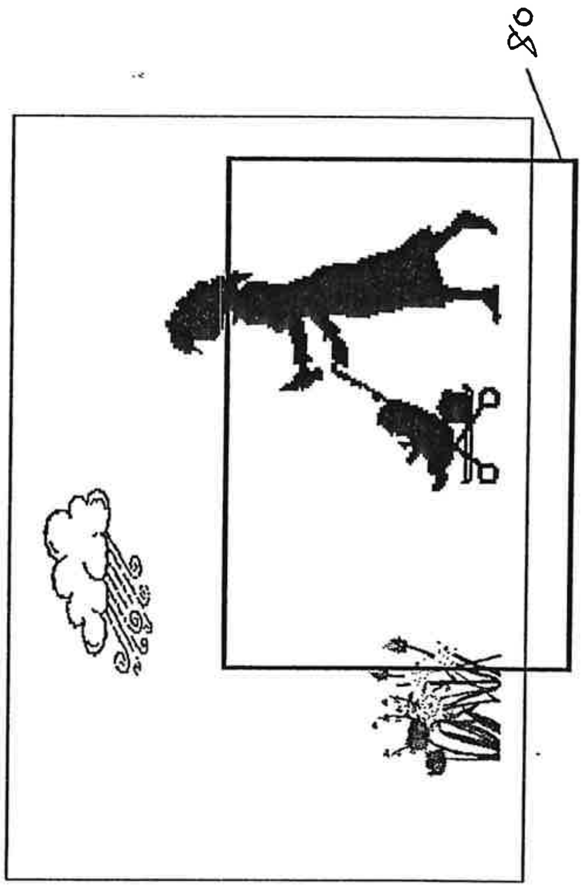


FIG. 7

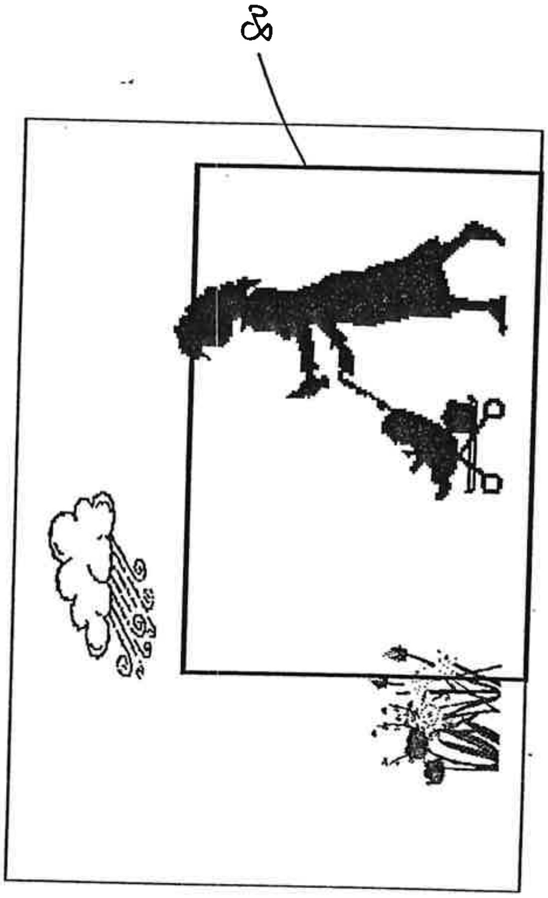


FIG. 8

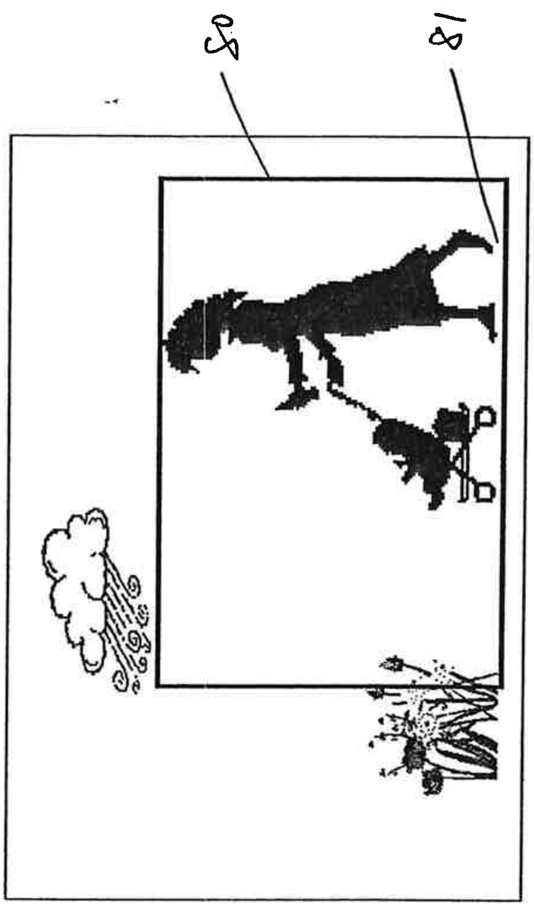


FIG. 9

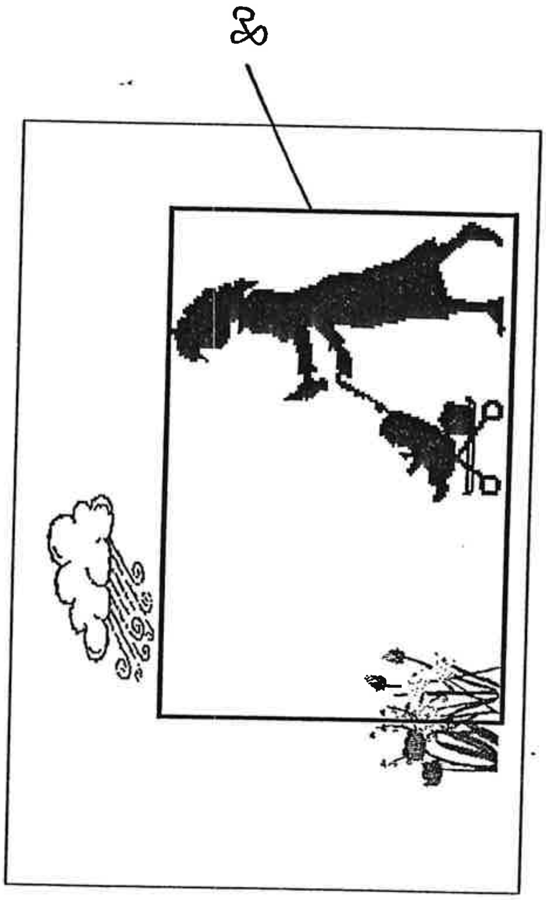


Fig. 10

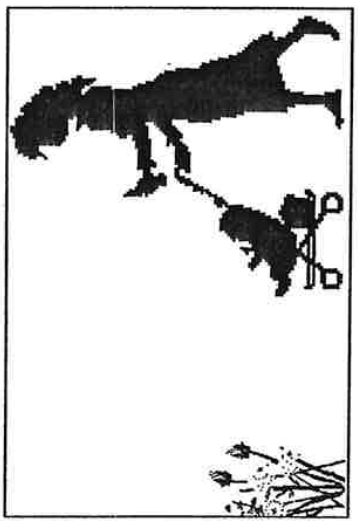


FIG. 11

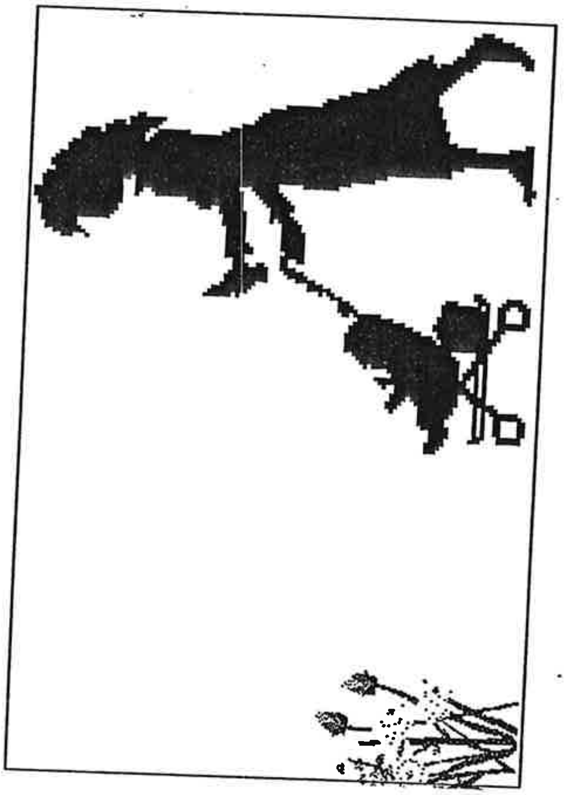


FIG. 12

Figure 13

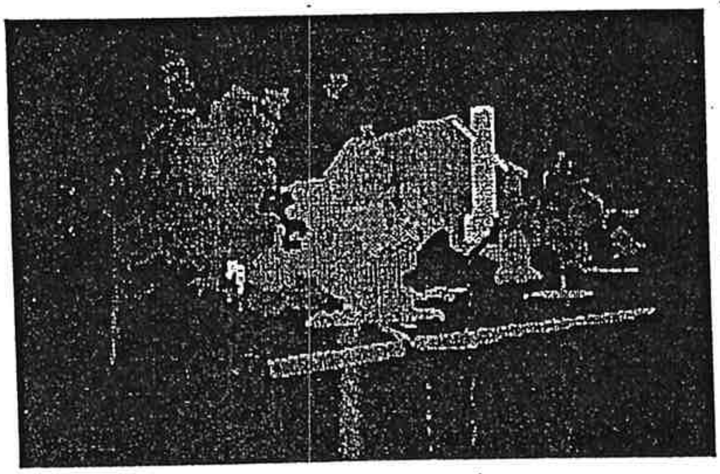


Figure 14

Figure 15

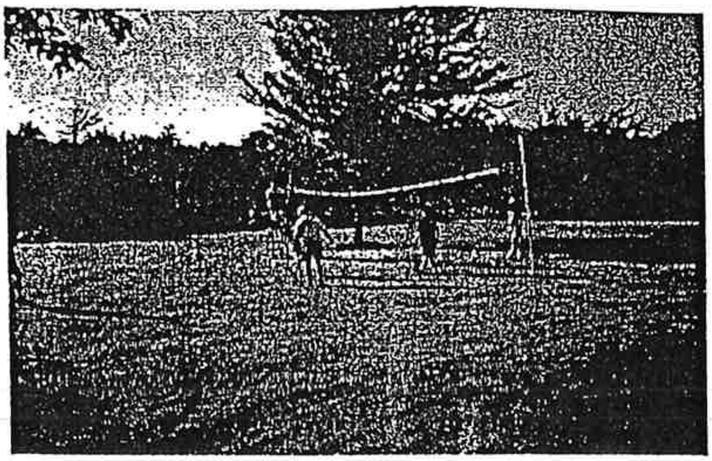


Figure 16

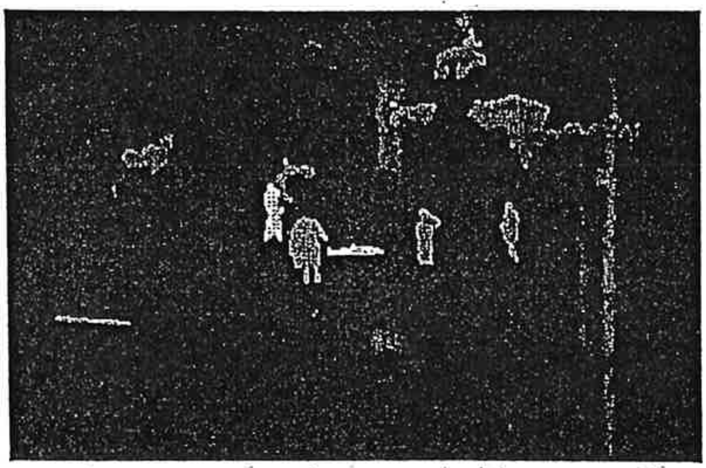


Figure 17

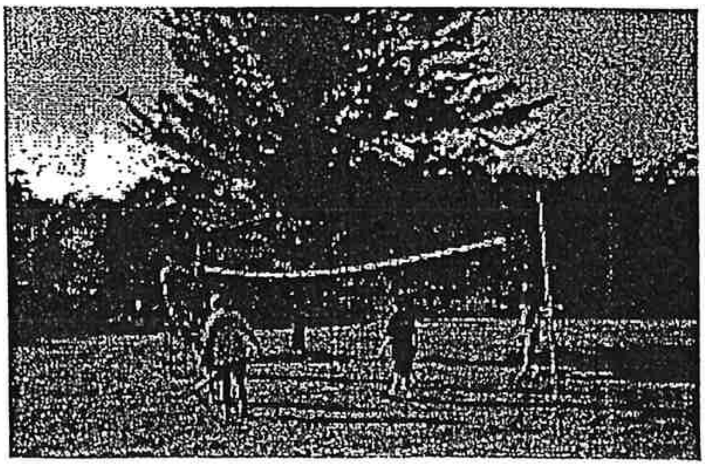


Figure 18

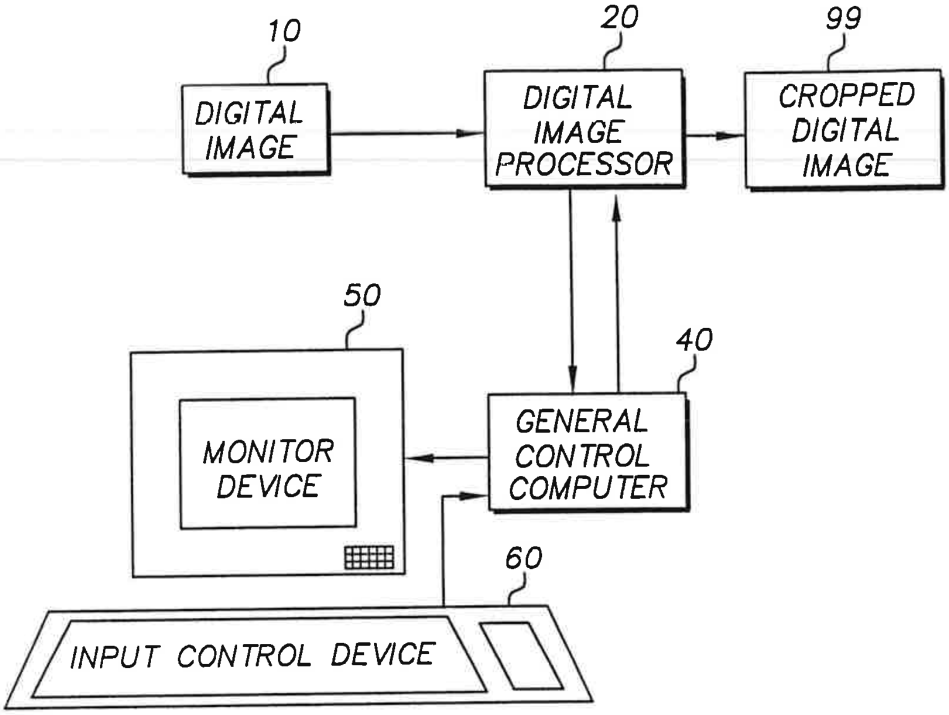

10
DIGITAL IMAGE
20
DIGITAL IMAGE PROCESSOR
99
CROPPED DIGITAL IMAGE
50
MONITOR DEVICE
40
GENERAL CONTROL COMPUTER
60
INPUT CONTROL DEVICE

FIG. 1

PRINT OF DRAWINGS
AS ORIGINALLY FILED

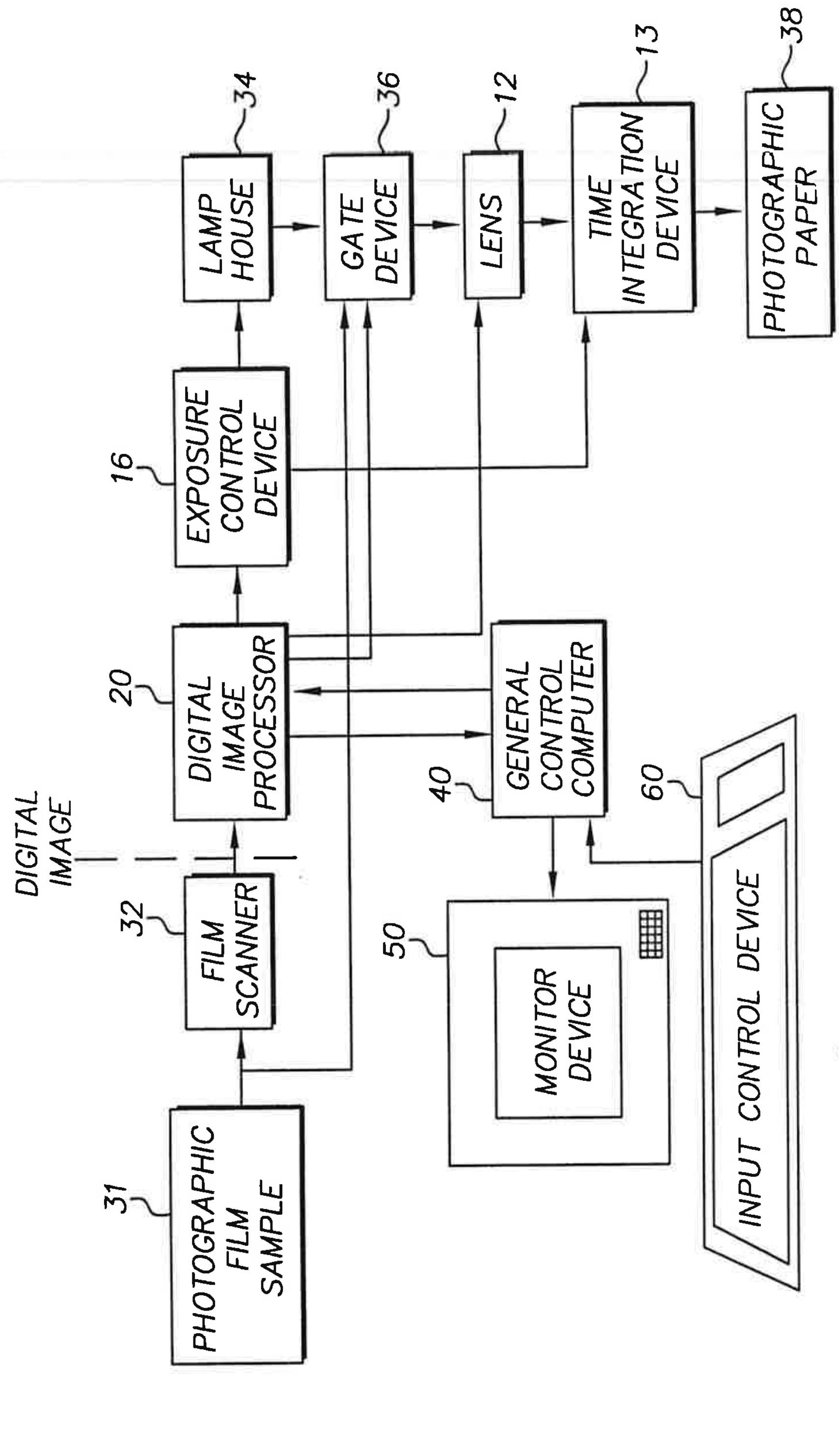


FIG. 1A

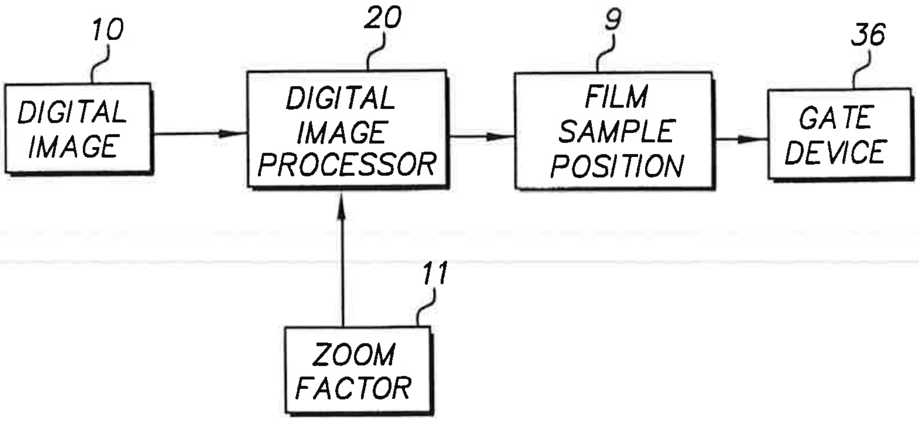

10
DIGITAL IMAGE
20
DIGITAL IMAGE PROCESSOR
9
FILM SAMPLE POSITION
36
GATE DEVICE
11
ZOOM FACTOR

FIG. 2

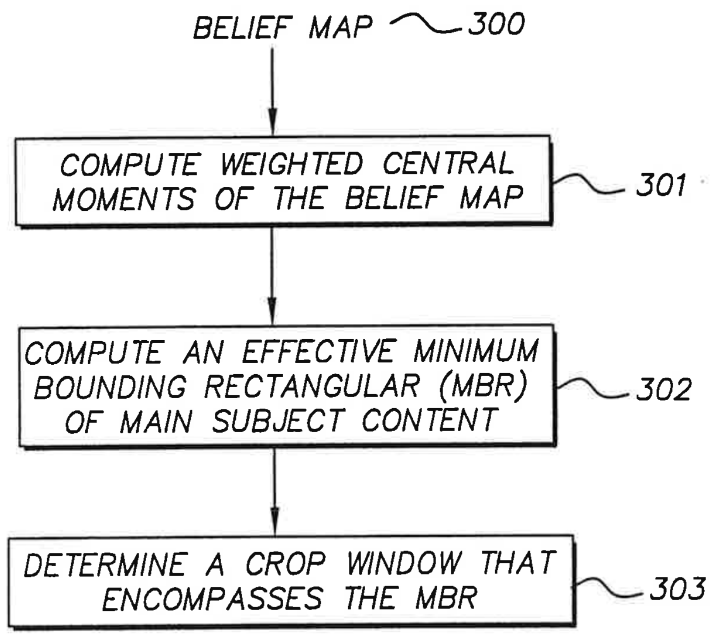

BELIEF MAP
300
COMPUTE WEIGHTED CENTRAL MOMENTS OF THE BELIEF MAP
301
COMPUTE AN EFFECTIVE MINIMUM BOUNDING RECTANGULAR (MBR) OF MAIN SUBJECT CONTENT
302
DETERMINE A CROP WINDOW THAT ENCOMPASSES THE MBR
303

FIG. 4

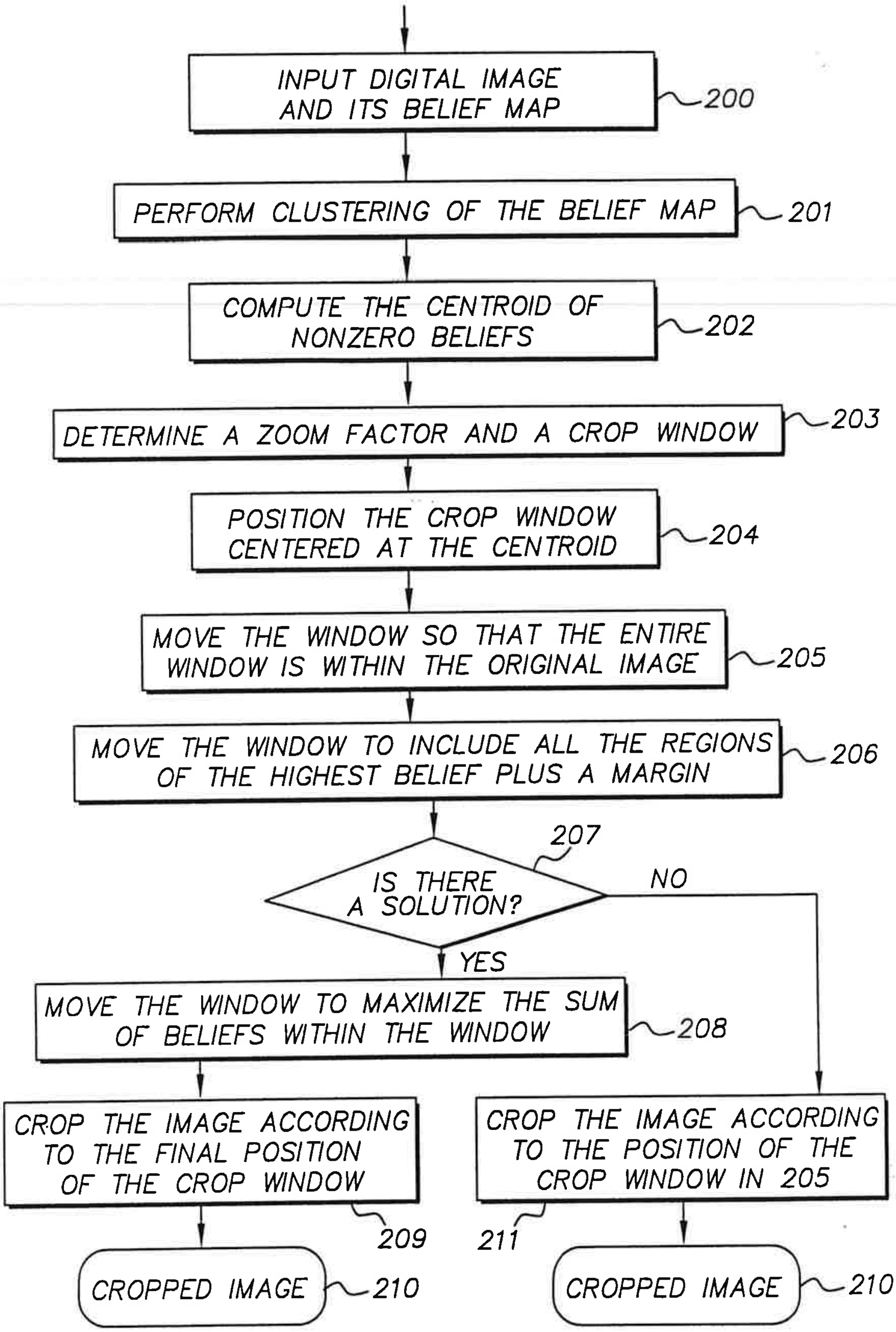

INPUT DIGITAL IMAGE AND ITS BELIEF MAP
200
PERFORM CLUSTERING OF THE BELIEF MAP
201
COMPUTE THE CENTROID OF NONZERO BELIEFS
202
DETERMINE A ZOOM FACTOR AND A CROP WINDOW
203
POSITION THE CROP WINDOW CENTERED AT THE CENTROID
204
MOVE THE WINDOW SO THAT THE ENTIRE WINDOW IS WITHIN THE ORIGINAL IMAGE
205
MOVE THE WINDOW TO INCLUDE ALL THE REGIONS OF THE HIGHEST BELIEF PLUS A MARGIN
206
207
IS THERE A SOLUTION?
NO
YES
MOVE THE WINDOW TO MAXIMIZE THE SUM OF BELIEFS WITHIN THE WINDOW
208
CROP THE IMAGE ACCORDING TO THE FINAL POSITION OF THE CROP WINDOW
209
CROP THE IMAGE ACCORDING TO THE POSITION OF THE CROP WINDOW IN 205
211
CROPPED IMAGE
210
CROPPED IMAGE
210

FIG. 3

Fig. 5

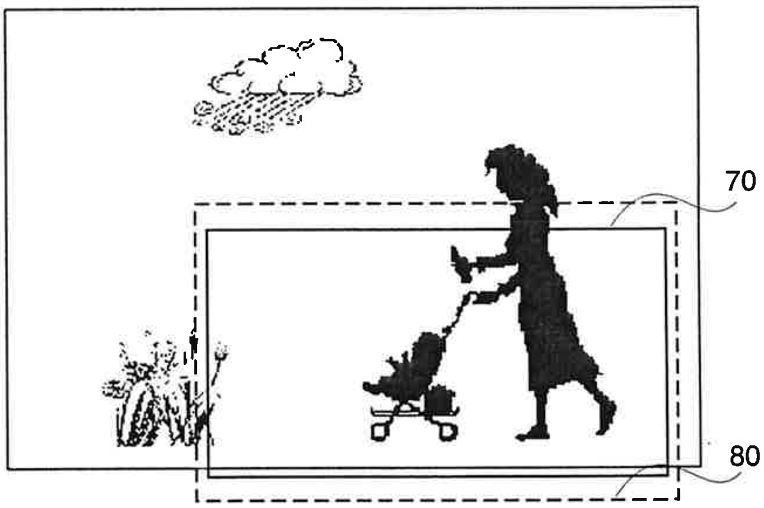


Fig. 6

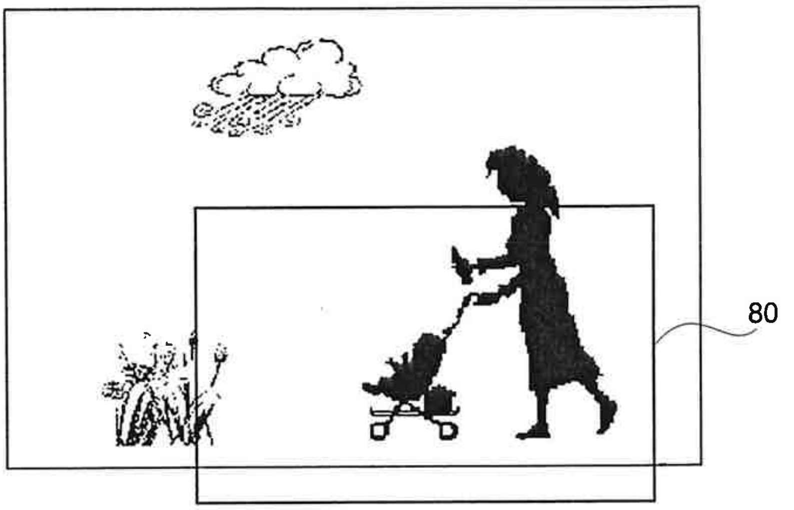


Fig. 7

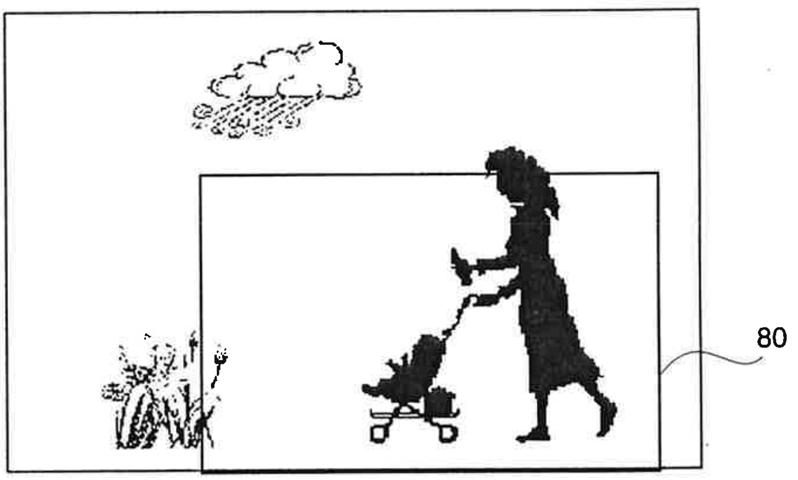


Fig. 8

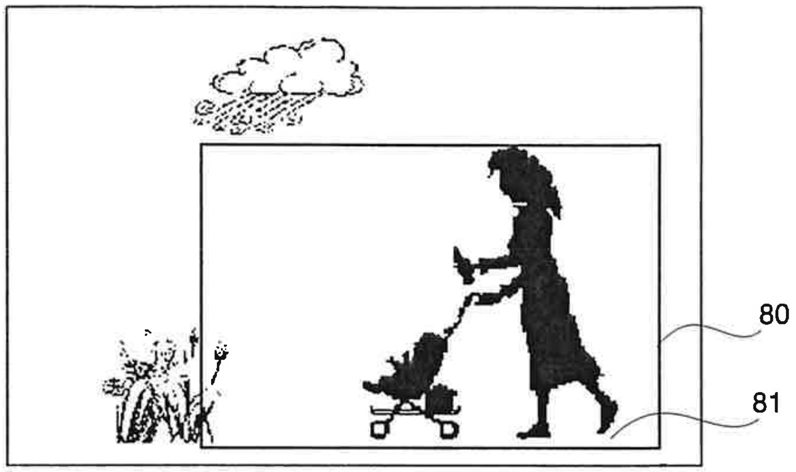


Fig. 9

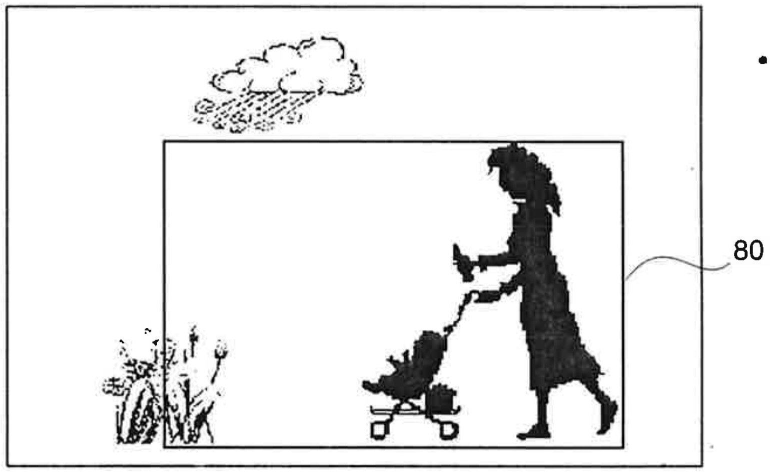


Fig. 10

PRINT OF DRAWINGS
AS ORIGINALLY FILED

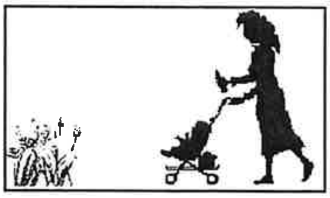


Fig. 11

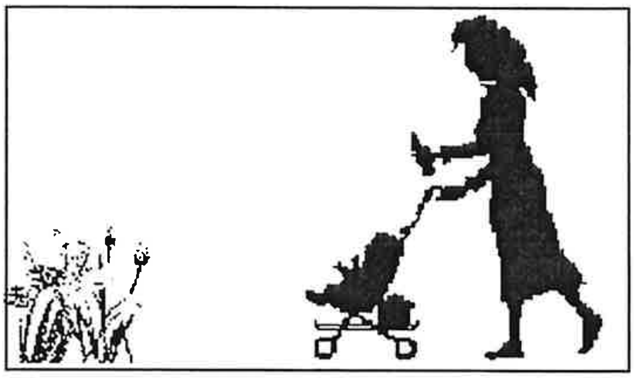


Fig. 12

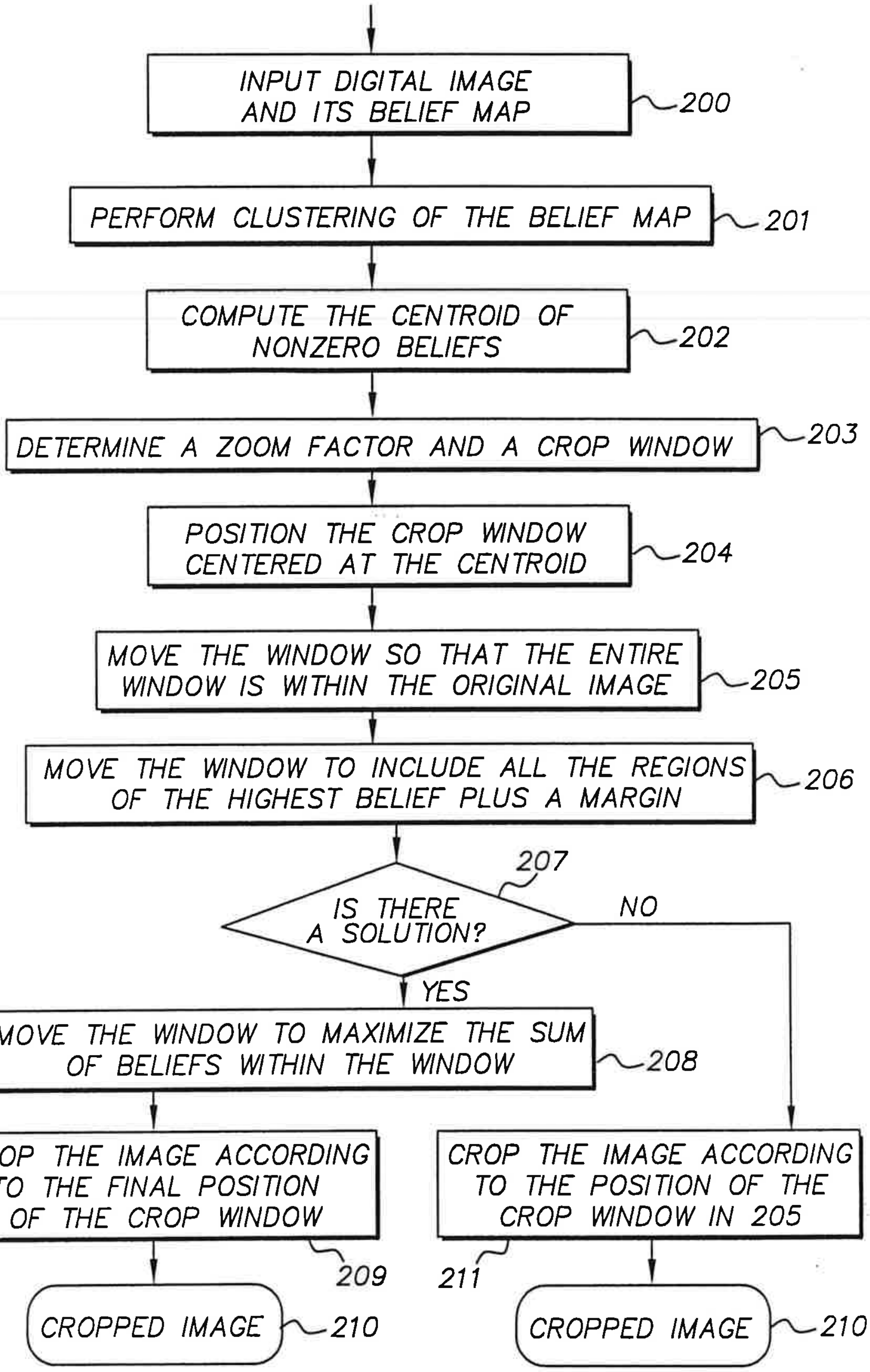

INPUT DIGITAL IMAGE AND ITS BELIEF MAP
200
PERFORM CLUSTERING OF THE BELIEF MAP
201
COMPUTE THE CENTROID OF NONZERO BELIEFS
202
DETERMINE A ZOOM FACTOR AND A CROP WINDOW
203
POSITION THE CROP WINDOW CENTERED AT THE CENTROID
204
MOVE THE WINDOW SO THAT THE ENTIRE WINDOW IS WITHIN THE ORIGINAL IMAGE
205
MOVE THE WINDOW TO INCLUDE ALL THE REGIONS OF THE HIGHEST BELIEF PLUS A MARGIN
206
207
IS THERE A SOLUTION?
NO
YES
MOVE THE WINDOW TO MAXIMIZE THE SUM OF BELIEFS WITHIN THE WINDOW
208
CROP THE IMAGE ACCORDING TO THE FINAL POSITION OF THE CROP WINDOW
209
CROP THE IMAGE ACCORDING TO THE POSITION OF THE CROP WINDOW IN 205
211
CROPPED IMAGE
210
CROPPED IMAGE
210

FIG. 3

PRINT OF DRAWINGS
AS ORIGINALLY FILED

Fig. 5

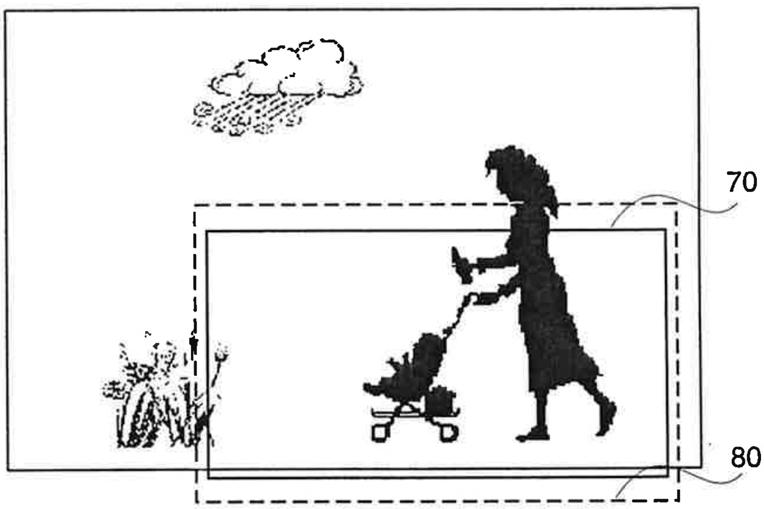


Fig. 6

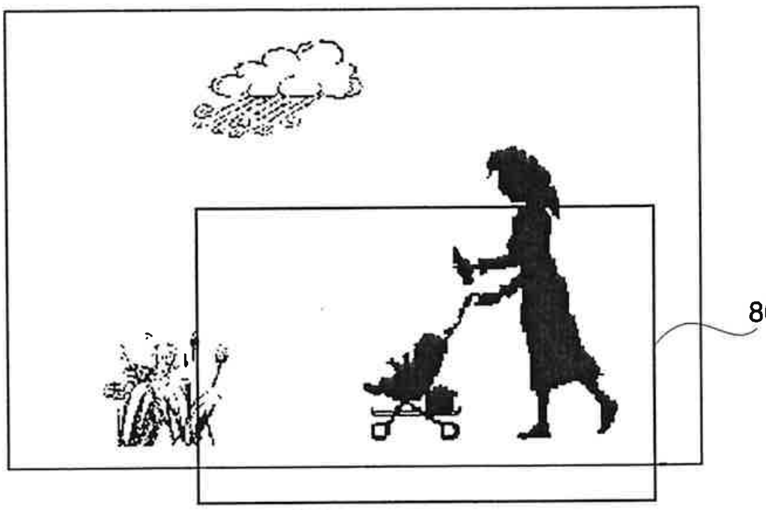


Fig. 7

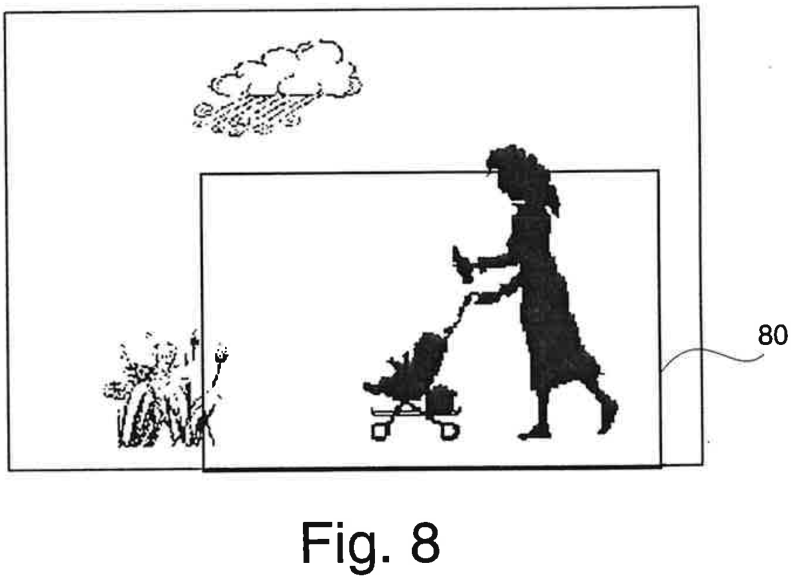


Fig. 8

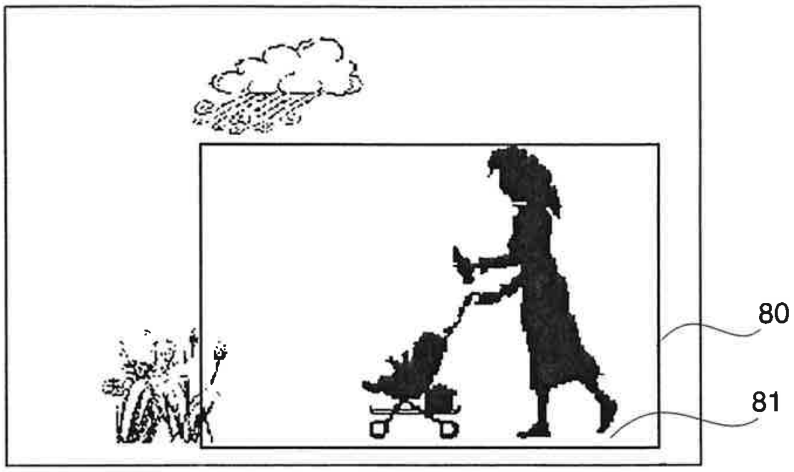


Fig. 9

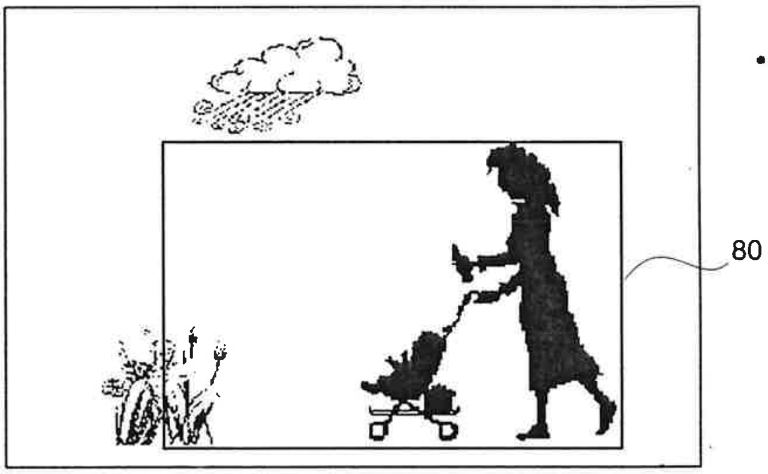


Fig. 10

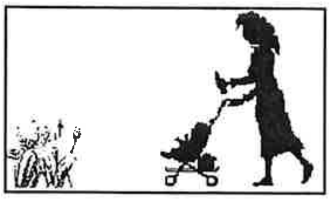


Fig. 11

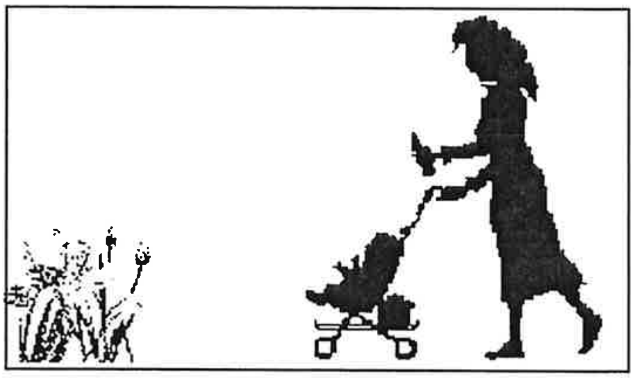


Fig. 12

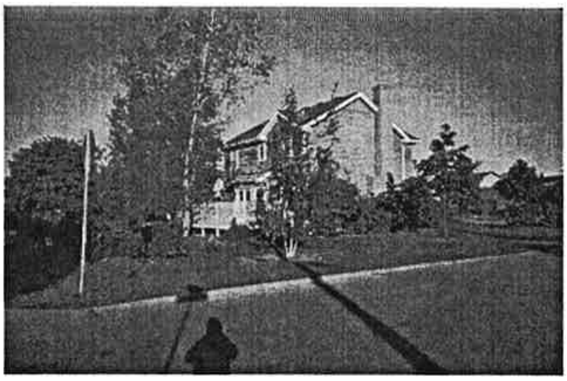


Fig. 13

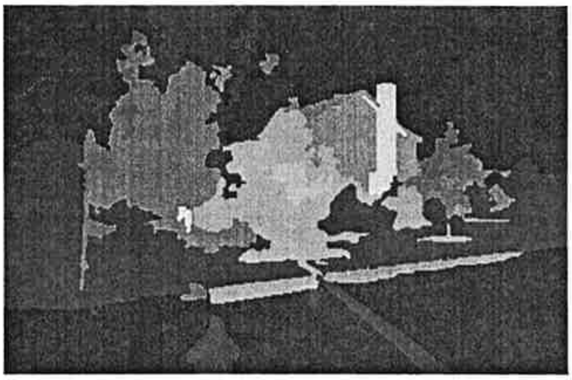


Fig. 14

Fig. 15

PRINT OF DRAWINGS
AS ORIGINALLY FILED

Fig. 16

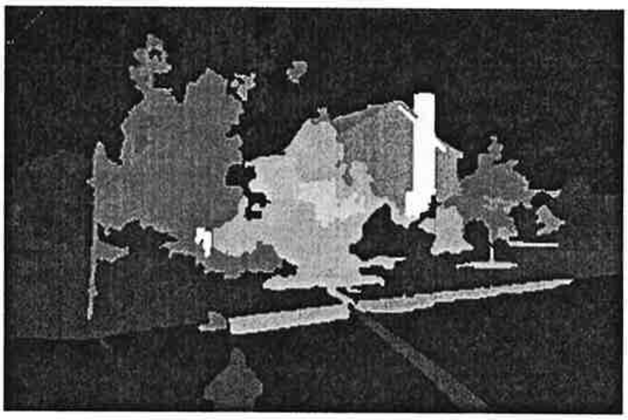


Fig. 17

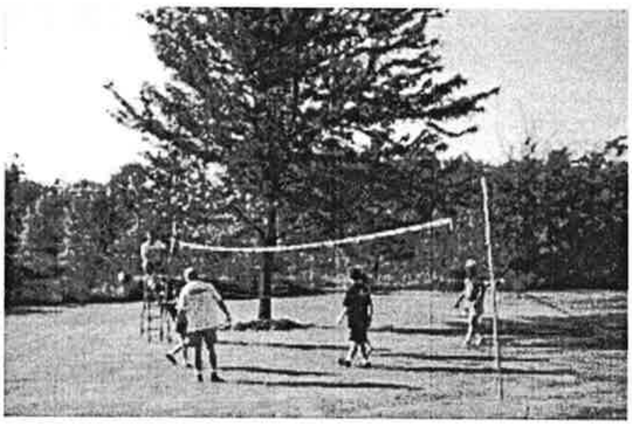


Fig. 18

US006654507B2

(12) **United States Patent**
Luo

(10) **Patent No.: US 6,654,507 B2**
(45) **Date of Patent: *Nov. 25, 2003**

(54) **AUTOMATICALLY PRODUCING AN IMAGE OF A PORTION OF A PHOTOGRAPHIC IMAGE**

(75) Inventor: **Jiebo Luo**, Rochester, NY (US)

(73) Assignee: **Eastman Kodak Company**, Rochester, NY (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 334 days.

This patent is subject to a terminal disclaimer.

(21) Appl. No.: **09/736,825**

(22) Filed: **Dec. 14, 2000**

(65) **Prior Publication Data**

US 2002/0114535 A1 Aug. 22, 2002

(51) **Int. Cl.**^{7} **G06K 9/20**; H04N 1/387

(52) **U.S. Cl.** **382/282**; 382/173; 345/620; 358/453

(58) **Field of Search** 382/291, 173, 382/180, 225, 282; 345/620–628; 358/453

(56) **References Cited**

U.S. PATENT DOCUMENTS

4,809,064 A 2/1989 Amos et al.
5,640,468 A * 6/1997 Hsu 382/190
5,781,665 A * 7/1998 Cullen et al. 382/254
5,872,619 A 2/1999 Stephenson et al.
5,872,643 A 2/1999 Maeda et al.
5,880,858 A * 3/1999 Jin 358/487
5,978,519 A 11/1999 Bollman et al.
5,995,201 A 11/1999 Sakaguchi
6,091,841 A * 7/2000 Rogers et al. 382/132
6,282,317 B1 * 8/2001 Luo et al. 382/203
6,335,985 B1 * 1/2002 Sambonsugi et al. 382/190
6,430,320 B1 * 8/2002 Jia et al. 382/289
6,434,271 B1 * 8/2002 Christian et al. 382/194
6,456,732 B1 * 9/2002 Kimbell et al. 382/112

* cited by examiner

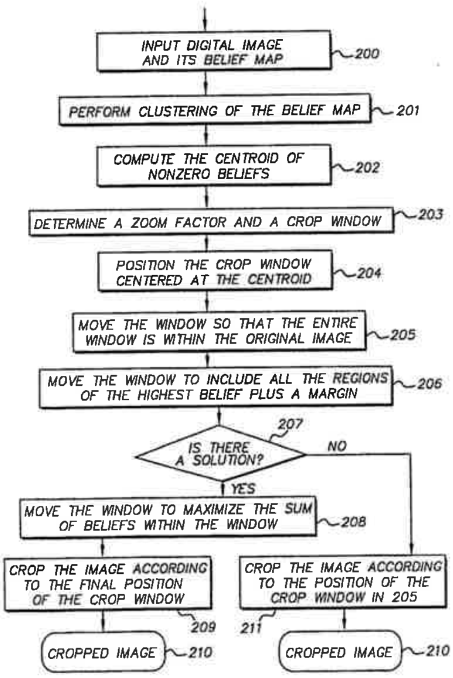
Primary Examiner—Bhavesh M. Mehta
Assistant Examiner—Aaron Carter
(74) *Attorney, Agent, or Firm*—Raymond L. Owens

(57) **ABSTRACT**

A method of producing an image of at least a portion of a digital image that includes pixels includes

computing a belief map of the digital image, by using the pixels of the digital image to determine a series of features and using such features to assign the probability of the location of a main subject of the digital image in the belief map determining a crop window having a shape and a zoom factor, which determine a size of the crop window and cropping the digital image to include a portion of the image of high subject content in response to the belief map and the crop window.

28 Claims, 10 Drawing Sheets



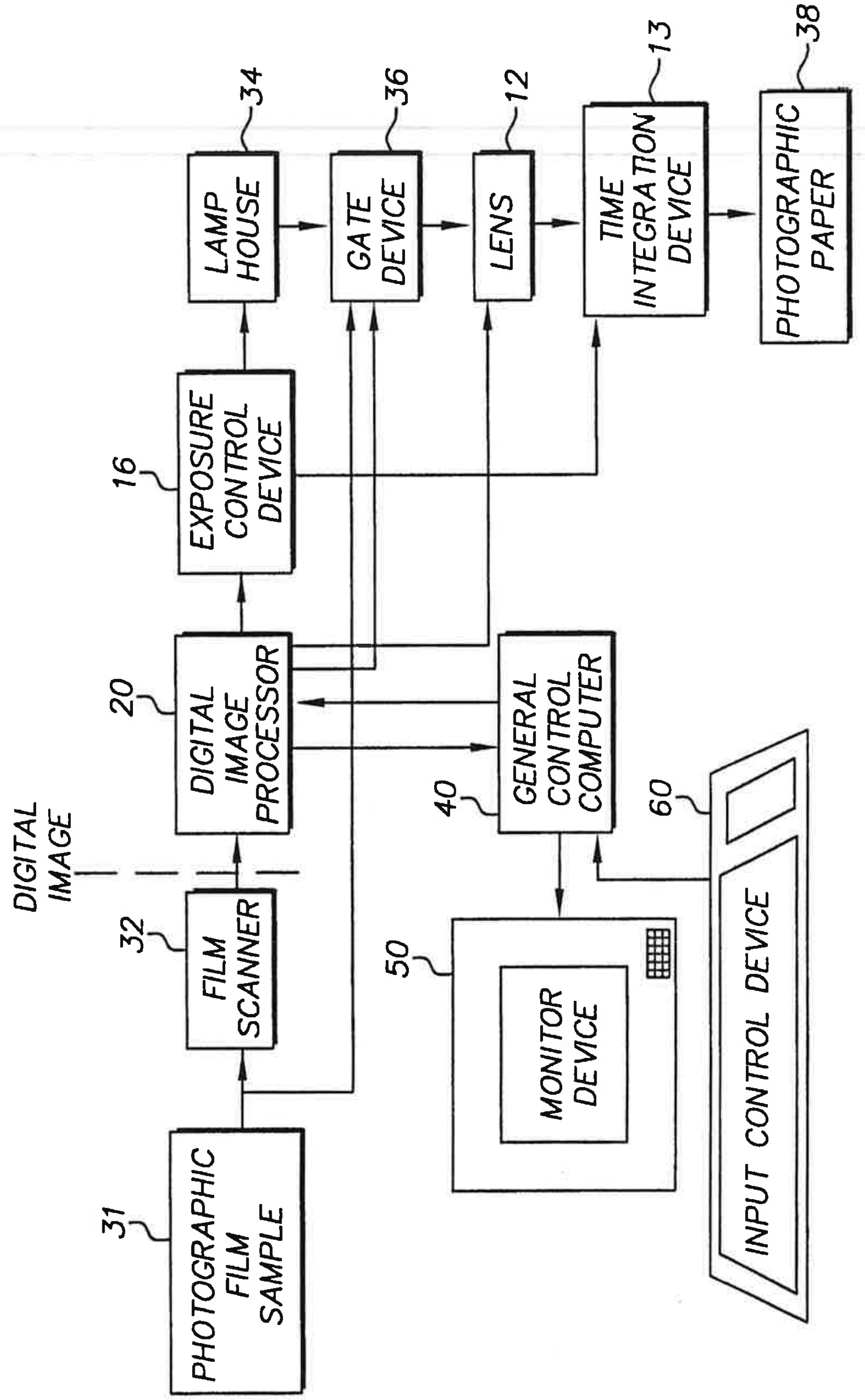

31
PHOTOGRAPHIC FILM SAMPLE
32
FILM SCANNER
DIGITAL IMAGE
20
DIGITAL IMAGE PROCESSOR
16
EXPOSURE CONTROL DEVICE
34
LAMP HOUSE
36
GATE DEVICE
12
LENS
13
TIME INTEGRATION DEVICE
38
PHOTOGRAPHIC PAPER
40
GENERAL CONTROL COMPUTER
50
MONITOR DEVICE
60
INPUT CONTROL DEVICE

FIG. 1A

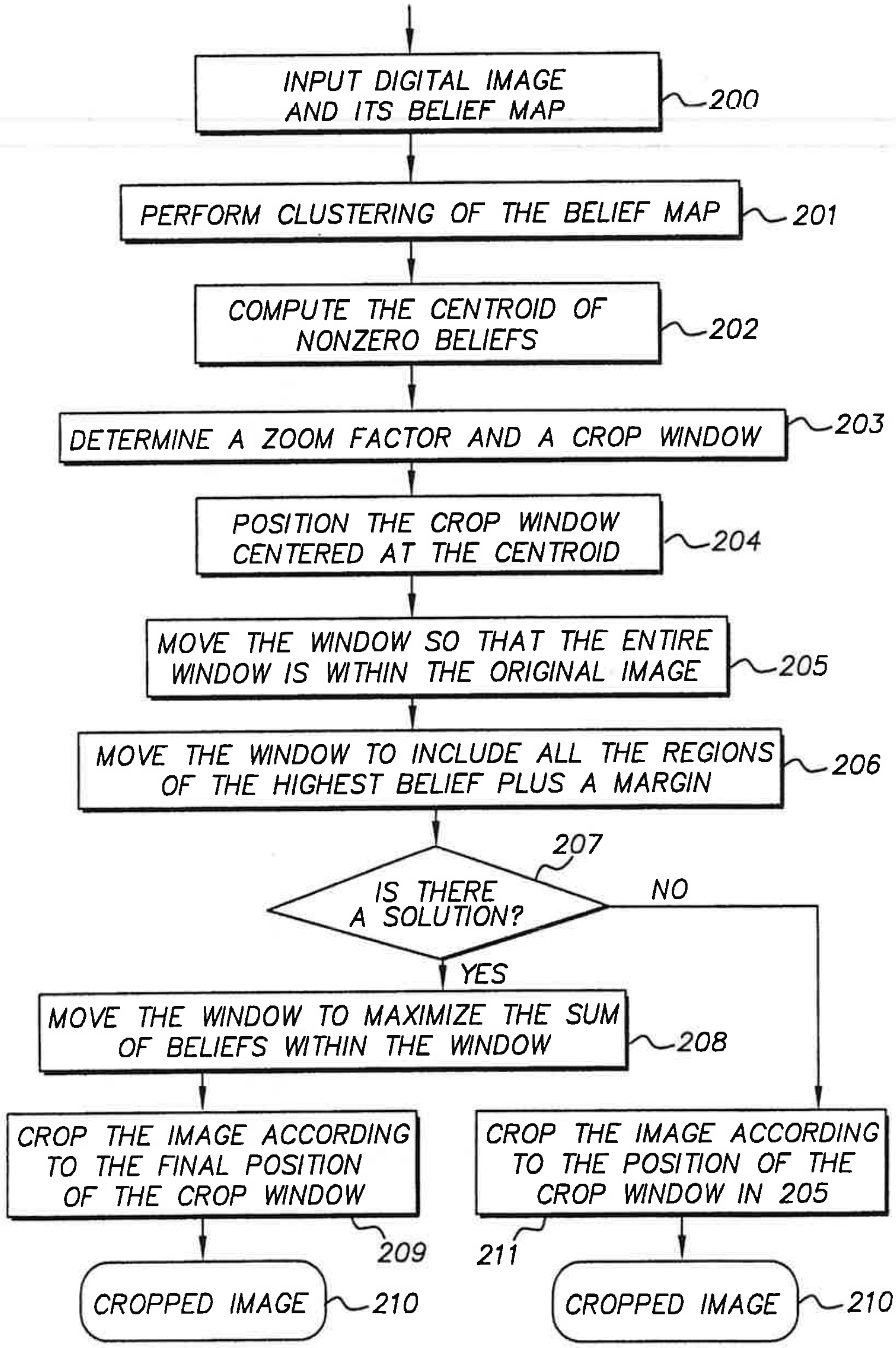

INPUT DIGITAL IMAGE AND ITS BELIEF MAP
200
PERFORM CLUSTERING OF THE BELIEF MAP
201
COMPUTE THE CENTROID OF NONZERO BELIEFS
202
DETERMINE A ZOOM FACTOR AND A CROP WINDOW
203
POSITION THE CROP WINDOW CENTERED AT THE CENTROID
204
MOVE THE WINDOW SO THAT THE ENTIRE WINDOW IS WITHIN THE ORIGINAL IMAGE
205
MOVE THE WINDOW TO INCLUDE ALL THE REGIONS OF THE HIGHEST BELIEF PLUS A MARGIN
206
207
IS THERE A SOLUTION?
NO
YES
MOVE THE WINDOW TO MAXIMIZE THE SUM OF BELIEFS WITHIN THE WINDOW
208
CROP THE IMAGE ACCORDING TO THE FINAL POSITION OF THE CROP WINDOW
209
CROP THE IMAGE ACCORDING TO THE POSITION OF THE CROP WINDOW IN 205
211
CROPPED IMAGE
210
CROPPED IMAGE
210

FIG. 3

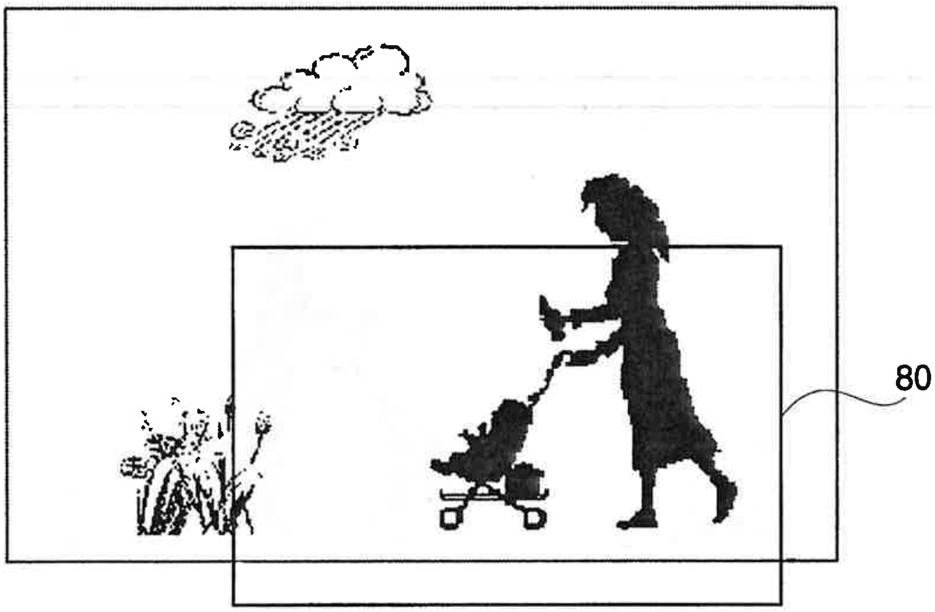


Fig. 7

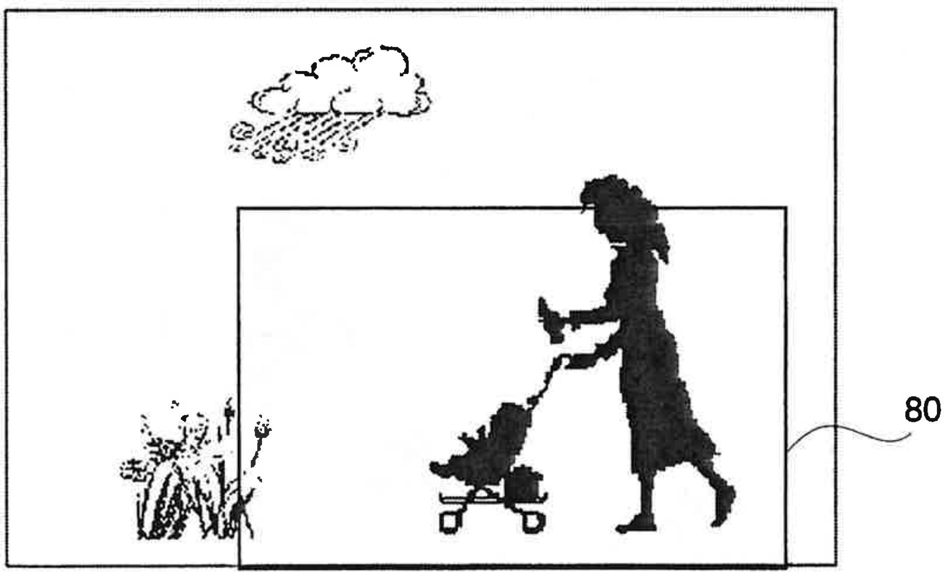


Fig. 8

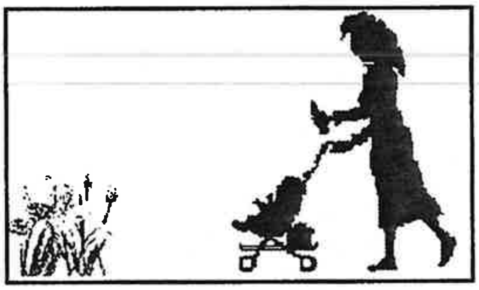


Fig. 11

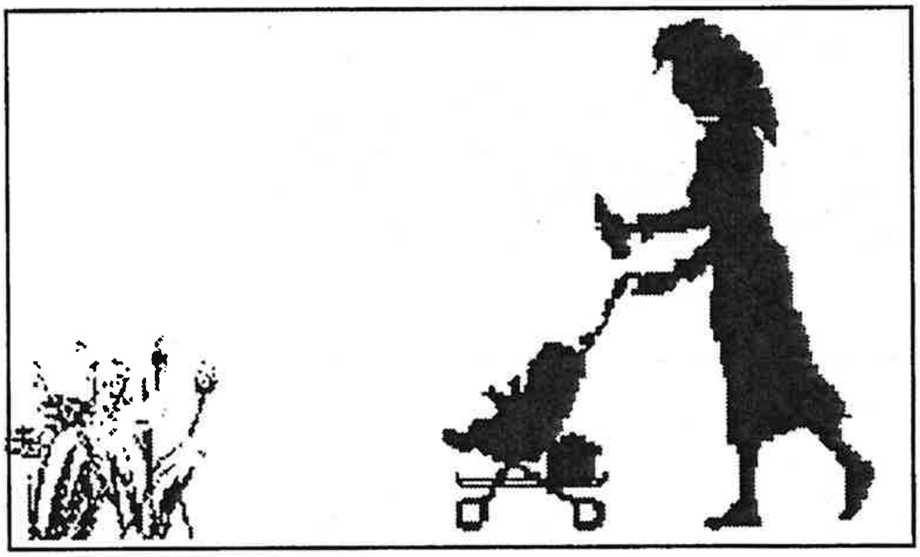


Fig. 12

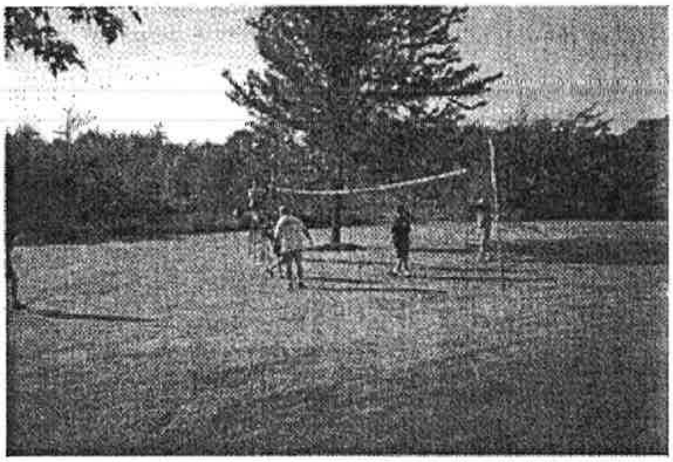


Fig. 16

Fig. 17

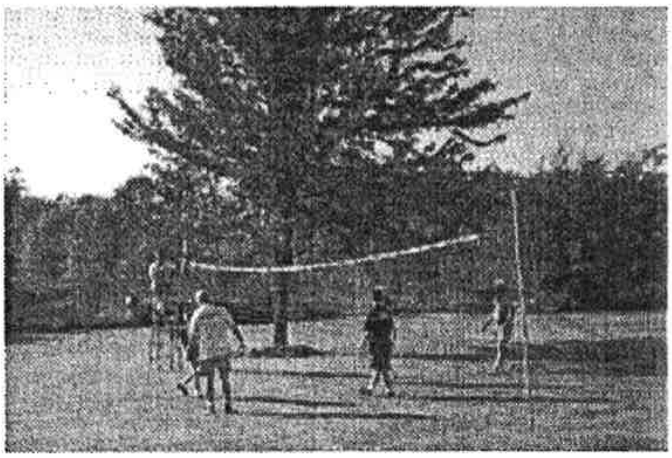


Fig. 18

zoom lens and a printing mask to provide printed images having a selected print width and a selected print height. In U.S. Pat. No. 4,809,064, Amos et al. describe an apparatus for printing a selected region of a photographic negative onto a photosensitive paper to form an enlarged and cropped photographic print. This apparatus includes means for projecting the photographic negative onto first and second zoom lenses, each of the zoom lenses having an adjustable magnification. In U.S. Pat. No. 5,872,643, Maeda et al. describe a film reproducing apparatus that can effectively perform zoom and crop. This apparatus includes an image pick-up device which picks up a film frame image recorded on a film to generate image data, an information reader which reads information about photographing conditions of the film frame image, and a reproducing area designator which designates a reproducing area of the film frame image. However, the reproducing area of the film frame image is determined based on pre-recorded information about the position of the main object, as indicated by which zone of the photograph the automatic focusing (AF) operation in the camera was on-part of the recorded information about photographing conditions. In all the above-mentioned optical printing systems, the position of the photographic film sample and magnification factor of the relay lens are preselected.

SUMMARY OF THE INVENTION

According to the present invention, there is provided a solution to the problems of the prior art. It is an object of the present invention to provide a method for producing a portion of a photographic image by identifying the main subject of the photographic image.

According to a feature of the present invention, there is provided a method of producing an image of at least a portion of a digital image, comprising the steps of:

a) providing a digital image having pixels;

b) computing a belief map of the digital image, by using the pixels of the digital image to determine a series of features, and using such features to assign the probability of the location of a main subject of the digital image in the belief map;

c) determining a crop window having a shape and a zoom factor, the shape and zoom factor determining a size of the crop window; and

d) cropping the digital image to include a portion of the image of high subject content in response to the belief map and the crop window.

ADVANTAGEOUS EFFECT OF THE INVENTION

One advantage of the invention lies in the ability to automatically crop and zoom photographic images based upon the scene contents. The digital image processing steps employed by the present invention includes a step of identifying the main subject within the digital image. The present invention uses the identified main subject of the digital image to automatically zoom and crop the image. Therefore, the present invention produces high-quality zoomed or cropped images automatically, regardless whether the background is uniform or not.

BRIEF DESCRIPTION OF THE DRAWINGS

The foregoing and other objects, aspects and advantages will be better understood from the following detailed description of a preferred embodiment of the invention with reference to the drawings, in which:

FIG. 1 is a schematic diagram of a system embodiment of the invention;

FIG. 2 is a schematic architectural diagram of an embodiment of the invention;

FIG. 3 is a schematic architectural diagram of an embodiment of the invention;

FIG. 4 is a schematic architectural diagram of an embodiment of the invention;

FIG. 5 illustrates the application of the invention to a simulated photograph;

FIG. 6 illustrates the application of the invention to a simulated photograph;

FIG. 7 illustrates the application of the invention to a simulated photograph;

FIG. 8 illustrates the application of the invention to a simulated photograph;

FIG. 9 illustrates the application of the invention to a simulated photograph;

FIG. 10 illustrates the application of the invention to a simulated photograph;

FIG. 11 illustrates the application of the invention to a simulated photograph;

FIG. 12 illustrates the application of the invention to a simulated photograph;

FIG. 13 is an exemplary uncropped photograph;

FIG. 14 is a belief map of the image shown in FIG. 13;

FIG. 15 is a cropped version of the image shown in FIG. 13;

FIG. 17 is a belief map of the image shown in FIG. 16; and

FIG. 18 is a cropped version of the image shown in FIG. 16.

DETAILED DESCRIPTION OF THE INVENTION

The invention automatically zooms and crops digital images according to an analysis of the main subject in the scene. Previously, a system for detecting main subjects (e.g., main subject detection or "MSD") in a consumer-type photographic image from the perspective of a third-party observer has been developed and is described in U.S. patent application Ser. No. 09/223,860, filed Dec. 31, 1998, the disclosure of which is incorporated herein by reference. Main subject detection provides a measure of saliency or relative importance for different regions that are associated with different subjects in an image. Main subject detection enables a discriminative treatment of the scene content for a number of applications related to consumer photographic images, including automatic crop and zoom.

Conventional wisdom in the field of computer vision, which reflects how a human observer would perform such tasks as main subject detection and cropping, calls for a problem-solving path via object recognition and scene content determination according to the semantic meaning of recognized objects. However, generic object recognition remains a largely unsolved problem despite decades of effort from academia and industry.

The MSD system is built upon mostly low-level vision features with semantic information integrated whenever available. This MSD system has a number of sub-tasks, including region segmentation, perceptual grouping, feature extraction, and probabilistic and semantic reasoning. In particular, a large number of features are extracted for each

In item **200**, the belief map is created using MSD. The present invention automatically determines a zoom factor (e.g. 1.5×) and a crop window **80** (as shown in FIG. **7**), as referred to in item **201** of FIG. **3**. This zoom factor is selected by an automatic method based directly on the main subject belief map (e.g., an estimate of the size of the main subject). The crop window is typically a rectangular window with a certain aspect ratio. After the zoom factor is determined by the digital image processor **20**, the value of the zoom factor is used subsequently by the digital image processor **20** shown in FIG. **1**. In FIG. **1***a*, the zoom factor is used to communicate with the lens **12** to adjust the lens magnification setting. This adjustment allows the lens **12** to image the appropriate size of the photographic film sample **31** onto the photographic paper **38**.

In item **201**, regions of the belief map are clustered and the lowest belief cluster (e.g., the background belief) is set to zero using a predefined threshold. As discussed in greater detail below, sections of the image having a belief value below a certain threshold are considered background sections. In item **202** such sections are given a belief of zero for purposes of this embodiment of the invention.

Then, in item **202** the centroid, or center-of-mass (used interchangeably hereon forth), of nonzero beliefs are computed. More specifically, in FIG. **5** the subject having the highest belief in the belief map is the woman and the stroller. FIG. **7** illustrates that the centroid of this subject is approximately the top of the baby's head.

The centroid $(\hat{x},\hat{y})$ of a belief map is calculated using the following procedure:

$$\hat{x}=\sum_i x_i bel(x_i,\ y_i),\ \hat{y}=\sum_i y_i bel(x_i,\ y_i),$$

where x_i and y_i denote that coordinates of a pixel in the belief map and $bel(x_i,y_i)$ represents the belief value at this pixel location.

Before the crop window is placed, a proper crop window is determined in item **203**. Referring to FIG. **4**, there is shown a block diagram of a method that automatically determines a zoom factor in response to the belief map. In item **301**, two second-order central moments, c_{xx} and c_{yy}, with respect to the center-of-mass, are computed using the following procedure:

$$c_{xx}=\frac{\sum_i (x_i-\hat{x})^2\times bel(x_i,\ y_i)}{\sum_i bel(x_i,\ y_i)},\quad c_{yy}=\frac{\sum_i (y_i-\hat{y})^2\times bel(x_i,\ y_i)}{\sum_i bel(x_i,\ y_i)}.$$

Note that these two terms are not the conventional central moments that are computed without any weighting functions. In the preferred embodiment, a linear weighting function of the belief values is used. However, the conventional central moments, or central moments by a nonlinear function of the belief values, can also be used.

An effective bounding rectangle (MBR) of the regions of high subject content can be calculated using the following procedure, where the dimensions of the MBR are calculated by:

$$D_x=2\times\sqrt{3\times c_{xx}},\ D_y=2\times\sqrt{3\times c_{yy}}$$

FIG. **6** illustrates that the effective bounding rectangle **70** is centered at approximately the top of the boy's head and approximately encompasses the region of high subject content. In general, the aspect ratio of the original image is maintained. Therefore, a crop window **80** is determined in item **303** such that it is the smallest rectangle of the original aspect ratio that encompasses the effective MBR **70**.

In item **204**, the initial position of the crop window p **80** is centered at the centroid, as shown in FIG. **7**.

The crop window is **80** then moved so that the entire crop window is within the original image (e.g. item **205**) as shown in FIG. **8**. In item **206**, the crop window **80** is moved again so that all the regions of the highest belief values ("main subject") are included within the crop window and to create a margin **81**, as shown in FIG. **9**. This process (e.g., **206**) captures the entire subject of interest. Therefore, as shown in FIG. **9**, the top of the woman's head is included in the crop window. Compare this to FIG. **8** where the top of the woman's head was outside the crop window.

Decision box **207** determines whether an acceptable solution has been found, i.e., whether it is possible to include at least the regions of the highest belief values in the crop window.

If an acceptable solution exists, the window is again moved, as shown in item **208**, to optimize a subject content index for the crop window. The preferred embodiment of the present invention defines the subject content index as the sum of belief values within the crop window. It should be noted that the present invention specifies higher numerical belief values corresponding to higher main subject probability. Therefore, finding a numerical maximum of the sum of the belief values is equivalent to finding an optimum of the subject content index. This is shown in FIG. **10** where the secondary objects (e.g. flowers) are included within the crop window **80** to increase the sum of beliefs. The sum of beliefs for a crop window is computed as follows.

$$sum(w)=\sum_{(x,y)\in w} bel(x,\ y),$$

where $bel(x, y)$ represents the belief value at a given pixel location (x, y) within the crop window w.

Provided that the primary subjects are included, moving the crop window so that more of the secondary subjects are included would increase the sum of belief values within the crop window. Recall that the primary subjects are indicated by the highest belief values and the secondary subjects are indicated by belief values lower than those of the primary subjects but higher than those of the background subjects. The goal is to find the crop window that has the highest sum of belief values while ensuring that the primary subjects are completely included in the crop window, i.e.,

$$\tilde{w}=\max_{w\in W} sum(w),$$

where W denotes the set of all possible crop windows that satisfy all the aforementioned constraints (e.g., those that are completely within the uncropped image and those that encompass the entire primary subjects).

Then, in item **212** (in place of item **209**, not shown), the position of the center of the crop window is used to calculate the translational component of the film sample position **9**. The gate device **36**, shown in FIG. **1***a*, receives the film sample position **9** and uses this information to control the position of the photographic film sample **31** relative to the lens **12**. Those skilled in the art will recognize that either or both of the lens **12** and the photographic film sample **31** may be moved to achieve the centering of the effective cropped image region on the photographic paper **38**.

as magnetic disk (such as a floppy disk) or magnetic tape; optical storage media such as optical disk, optical tape, or machine readable bar code; solid-state electronic storage devices such as random access memory (RAM), or read-only memory (ROM); or any other physical device or media employed to store a computer program having instructions for practicing a method according to the present invention.

While the invention has been described in terms of preferred embodiments, those skilled in the art will recognize that the invention can be practiced with modification within the spirit and scope of the appended claims.

The subject matter of the present invention relates to digital image understanding technology, which is understood to mean technology that digitally processes a digital image to recognize and thereby assign useful meaning to human understandable objects, attributes or conditions and then to utilize the results obtained in the further processing of the digital image.

The invention has been described in detail with particular reference to certain preferred embodiments thereof, but it will be understood that variations and modifications can be effected within the spirit and scope of the invention.

PARTS LIST

9	film sample position
10	source digital image
11	zoom factor
12	lens
13	time integration device
20	digital image processor
31	photographic film sample
32	film scanner
34	lamp house
36	gate device
38	photographic paper
40	general control computer
50	monitor device
60	input control device
80	crop window
81	margin
99	cropped digital image
200	image and belief map
201	decision box for performing clustering of the belief map
202	decision box for computing the center-of-mass
203	decision box for determining a zoom factor and a crop window
204	decision box for positioning the crop window
205	decision box moving a window
206	decision box for moving a window to contain the highest belief
207	decision box for determining if a solution exists
208	decision box for moving a window to optimize the sum of beliefs
209	decision box for cropping the image
210	cropped image
211	decision box for cropping the image
300	belief map
301	decision box for computing weighted central moments of the belief map with respect to the center-of-mass
302	decision box for computing an effective bounding rectangle (MBR) of the main subject content
303	decision box for determining a zoom factor and a crop window that encompasses the MBR

What is claimed is:

1. A method of producing an image of at least a portion of a digital image, comprising:

a) providing a digital image having pixels;

b) computing a belief map of the digital image by using the pixels of the digital image to determine a series of features and using such features to assign a probability of a location of a main subject of the digital image in the belief map;

c) determining a crop window having a shape factor and a zoom factor, the shape and the zoom factors determining a size of the crop window; and

d) cropping the digital image to include a portion of the image of high subject content in response to the belief map and the crop window.

2. The method of claim 1 wherein determining the crop window includes

i) computing a weighted center-of-mass of the belief map, weighted by belief values of the belief map;

ii) computing weighted central moments of the belief map relative to the center-of-mass and weighted by a weighting function of each belief value of the belief map;

iii) computing an effective rectangular bounding box according to the central moments; and

iv) determining a crop window having a shape and a zoom factor, the shape and the zoom factors determining a size of the crop window.

3. The method of claim 1 wherein cropping the digital image includes

i) selecting an initial position of the crop window at a location which includes a center of mass;

ii) using belief values corresponding to the crop window to select the position of the crop window to include a portion of the image of high subject content in response to the belief map; and

iii) cropping the digital image according to the position of the crop window.

4. The method of claim 2 wherein cropping the digital image includes

i) selecting a crop window of a rectangular shape and of a similar aspect ratio to the digital image; and

ii) selecting a zoom factor to determine the size of the crop window such that the crop window encompasses the effective bounding box.

5. The method of claim 2 wherein the weighting function of computing a belief map is a linear weighting function.

6. The method of claim 2 wherein the weighting function of computing a belief map is a constant function.

7. The method of claim 3 wherein computing a belief map includes

i) calculating a subject content index value for the crop window derived from the belief values;

ii) following a positioning procedure of repeating selecting an initial position of the crop window for at least two positions of the crop window; and

iii) using the subject content index values to select the crop window position.

8. The method of claim 1 wherein the crop window is completely within the digital image.

9. The method of claim 2 wherein computing a belief map includes clustering of the belief map to identify at least a cluster of highest belief values corresponding to the main subject, a cluster of intermediate belief values corresponding to secondary subjects, and a cluster of lowest belief values corresponding to background.

10. The method of claim 9 wherein clustering includes setting background portions to a zero belief value.

11. The method of claim 5 further comprising positioning the crop window such that the subject content index of the crop window is at an optimum.

12. The method of claim 3 further comprising positioning the crop window such that the crop window includes all of the main subject cluster.

13. The method of claim 12 further comprising positioning the crop window to include a buffer around the main subject cluster.

US006504951B1

(12) **United States Patent**
Luo et al.

(10) **Patent No.: US 6,504,951 B1**
(45) **Date of Patent: Jan. 7, 2003**

(54) **METHOD FOR DETECTING SKY IN IMAGES**

(75) Inventors: **Jiebo Luo**, Rochester, NY (US); **Stephen Etz**, Fairport, NY (US)

(73) Assignee: **Eastman Kodak Company**, Rochester, NY (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

(21) Appl. No.: **09/450,190**

(22) Filed: **Nov. 29, 1999**

(51) **Int. Cl.**[7] **G06K 9/00**
(52) **U.S. Cl.** **382/165**
(58) **Field of Search** 382/162–167, 382/305, 224–228; 358/515–527; 345/597–618

(56) **References Cited**

U.S. PATENT DOCUMENTS

5,642,443 A		6/1997	Goodwin et al.	382/289
5,822,453 A	*	10/1998	Lee et al.	382/169
5,889,578 A		3/1999	Jamzadeh	355/41
5,901,245 A		5/1999	Warnick et al.	382/190
6,097,470 A	*	8/2000	Buhr et al.	355/38
6,282,317 B1	*	8/2000	Luo et al.	382/203
6,408,103 B1	*	6/2002	Gallagher et al.	382/260

FOREIGN PATENT DOCUMENTS

EP 0 707 283 A2 4/1996 .... G06T/7/00

OTHER PUBLICATIONS

Azami et al. "Route guidance sign identification using 2–D structural description" Proc. of the 1996 IEEE Intelligent Vehicles Sysposium, 1996, pp. 153–158, Sep. 1996.*

"Neural Network Toolbox for Use with MATLAB" by Howard Demuth and Mark Beale. The MathWorks, Inc., 1998.

"The Nature of Light & Colour in the Open Air" by M. Minnaert. Dover Publications, Inc., 1954.

"Towards Physics–based Segmentation of Photographic Color Images" by Jiebo Luo, Robert T. Gray, Hsien–Che Lee. Proc. IEEE Int. Conf. Image Process., 997.

"On Image Classification: City vs. Landscape" by Aditya Vailaya, Anil Jain, and Hong Jiang Zhang, 1998.

"Decoding Image Semantics Using Composite Region Templates" by John R. Smith and Chung–Sheng Li. Proc. IEEE Intl. workshop on Content–based Access of Image and Video Database, 1998.

"Indoor–Outdoor Image Classification" by Martin Szummer and Rosalind W. Picard. Proc. IEEE Intl. Workshop on Content–based Access of Image and Video Database, 1998.

"Automatic Image Annotation Using Adaptive Color Classification" by Eli Saber, A. Murat Tekalp, Reiner Eschbach and Keith Knox. Graphical Models and Image Processing, vol. 58, No. 2, Mar. pp. 115–126, 1996.

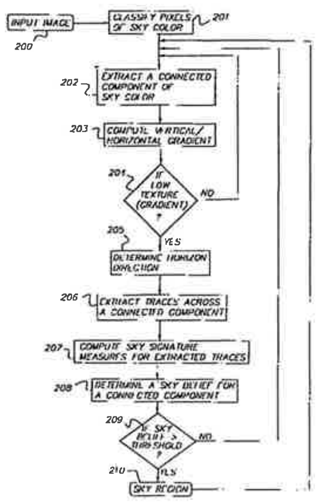
* cited by examiner

Primary Examiner—Jingge Wu
(74) *Attorney, Agent, or Firm*—David M. Woods

(57) **ABSTRACT**

A method, image recognition system, computer program, etc., for detecting sky regions in an image comprise classifying potential sky pixels in the image by color, extracting connected components of the potential sky pixels, eliminating ones of the connected components that have a texture above a predetermined texture threshold, computing desaturation gradients of the connected components, and comparing the desaturation gradients of the connected components with a predetermined desaturation gradient for sky to identify true sky regions in the image.

30 Claims, 13 Drawing Sheets



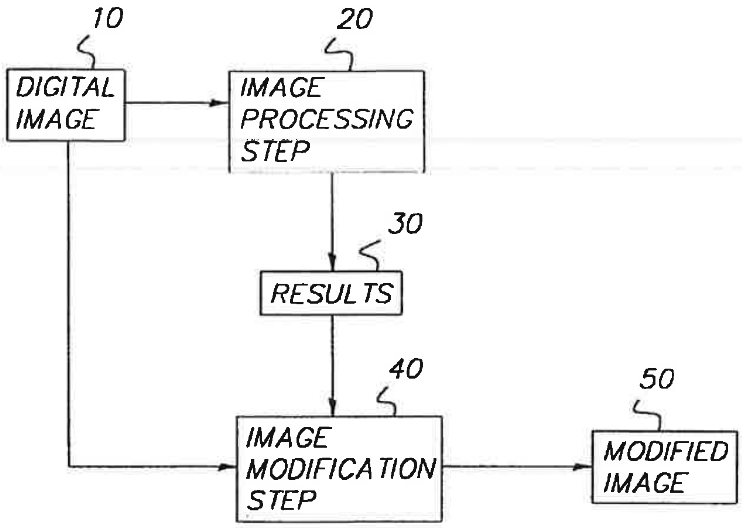


FIG. 1

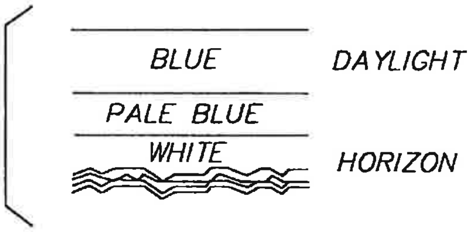


FIG. 3a

FIG. 3b

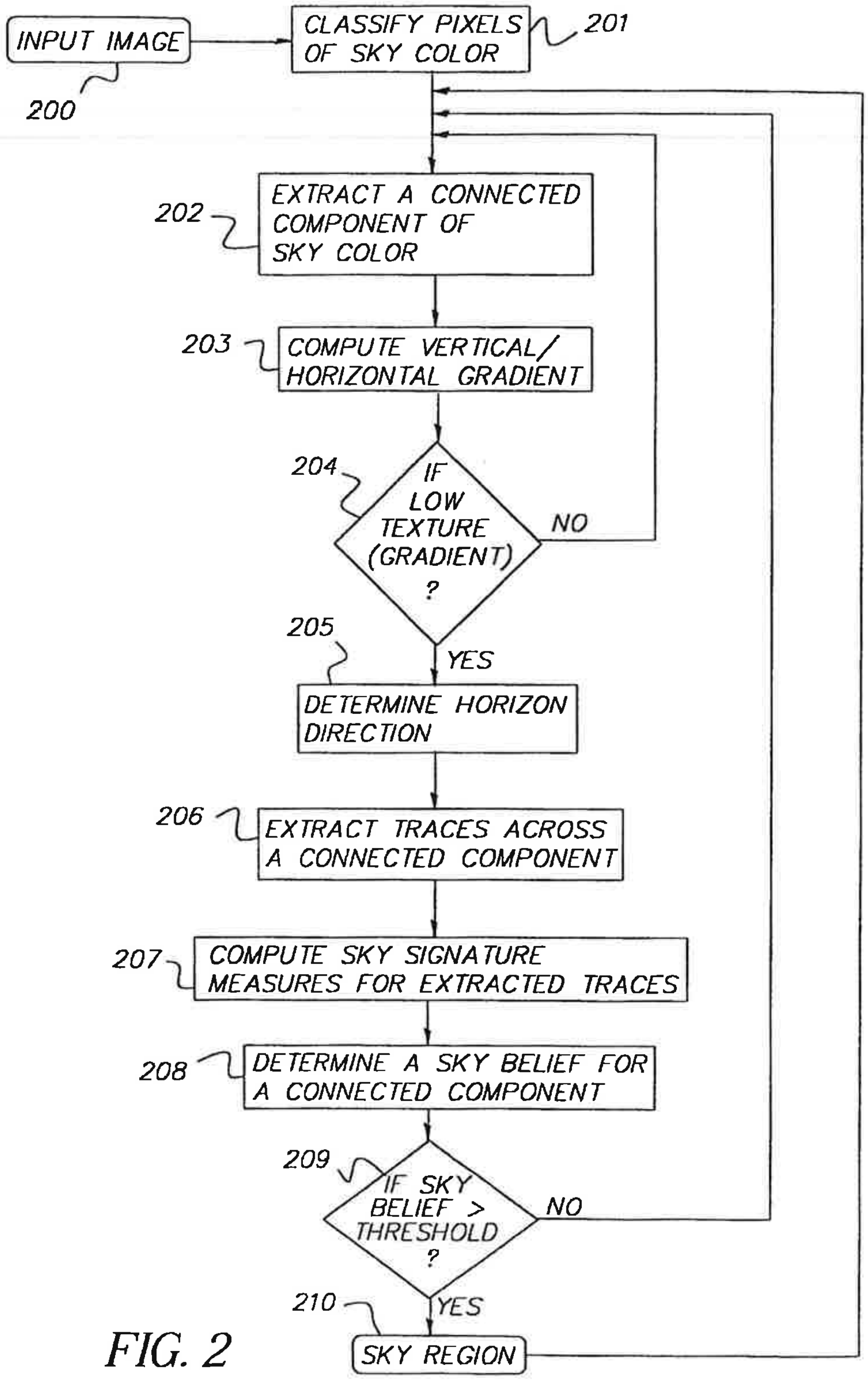

INPUT IMAGE
200
CLASSIFY PIXELS OF SKY COLOR
201
EXTRACT A CONNECTED COMPONENT OF SKY COLOR
202
COMPUTE VERTICAL/ HORIZONTAL GRADIENT
203
IF LOW TEXTURE (GRADIENT) ?
204
NO
YES
205
DETERMINE HORIZON DIRECTION
206
EXTRACT TRACES ACROSS A CONNECTED COMPONENT
207
COMPUTE SKY SIGNATURE MEASURES FOR EXTRACTED TRACES
208
DETERMINE A SKY BELIEF FOR A CONNECTED COMPONENT
209
IF SKY BELIEF > THRESHOLD ?
NO
YES
210
SKY REGION

FIG. 2

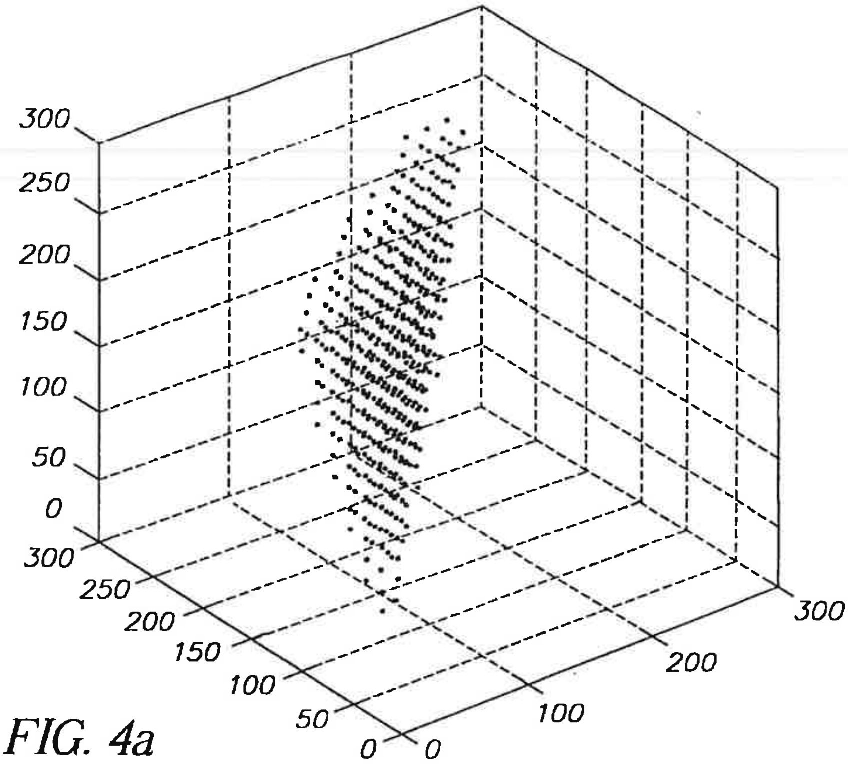

300
250
200
150
100
50
0
300
250
200
150
100
50
0
0
100
200
300

FIG. 4a

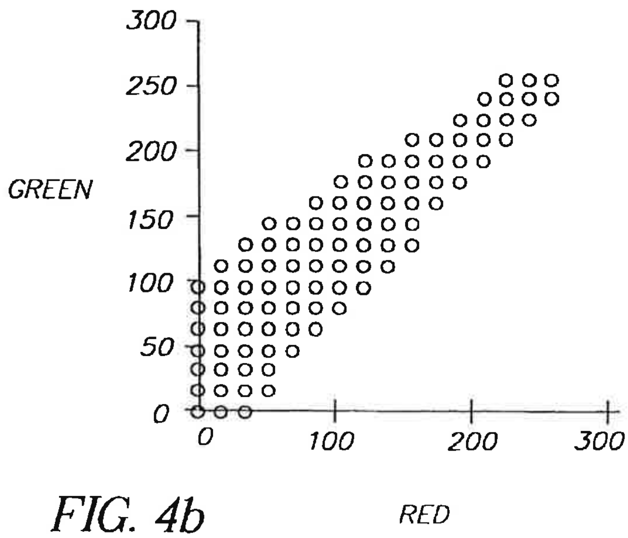

300
250
200
150
100
50
0
GREEN
0
100
200
300
RED

FIG. 4b

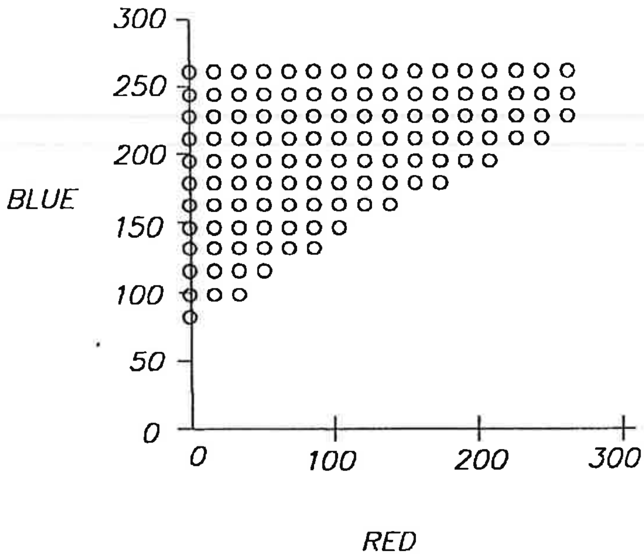

300
250
200
BLUE
150
100
50
0
0
100
200
300
RED

FIG. 4c

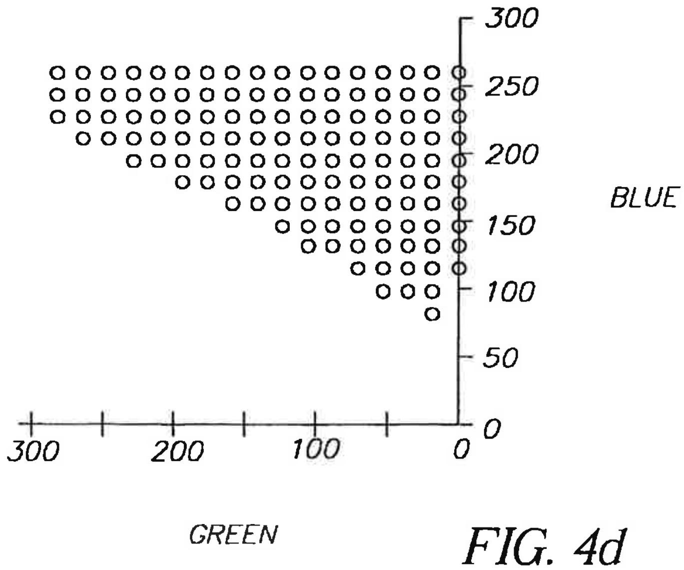

300
250
200
BLUE
150
100
50
0
300
200
100
0
GREEN

FIG. 4d

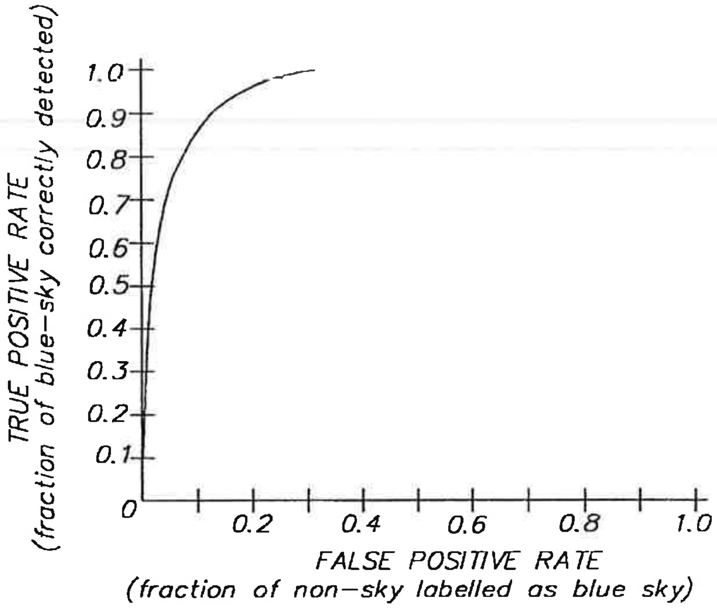

TRUE POSITIVE RATE
(fraction of blue-sky correctly detected)
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0.2
0.4
0.6
0.8
1.0
FALSE POSITIVE RATE
(fraction of non-sky labelled as blue sky)

FIG. 5

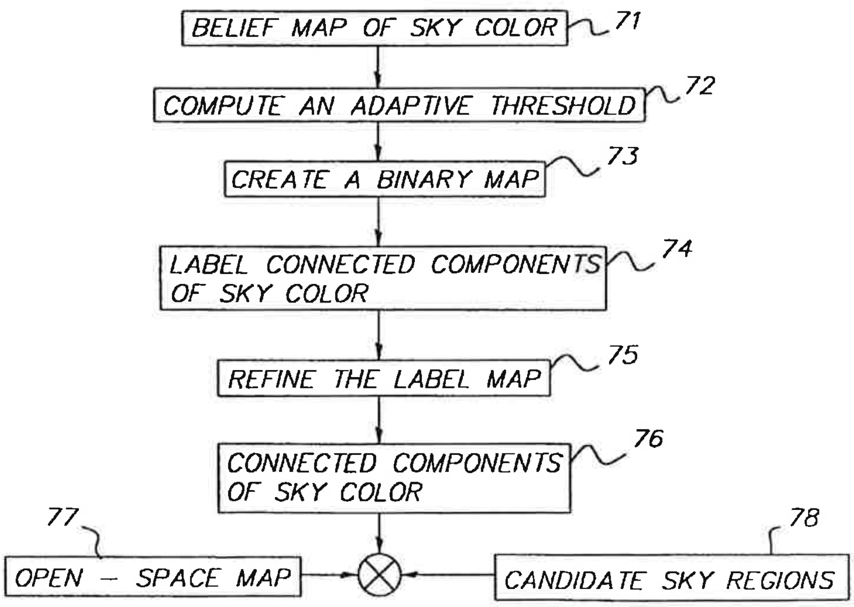

BELIEF MAP OF SKY COLOR
71
COMPUTE AN ADAPTIVE THRESHOLD
72
CREATE A BINARY MAP
73
LABEL CONNECTED COMPONENTS OF SKY COLOR
74
REFINE THE LABEL MAP
75
CONNECTED COMPONENTS OF SKY COLOR
76
77
OPEN - SPACE MAP
78
CANDIDATE SKY REGIONS

FIG. 6

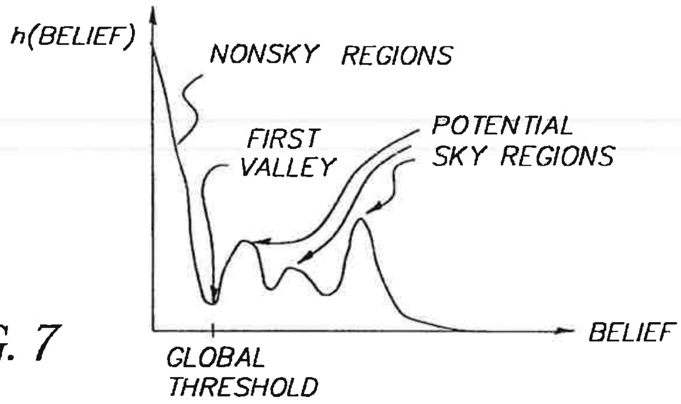


FIG. 7

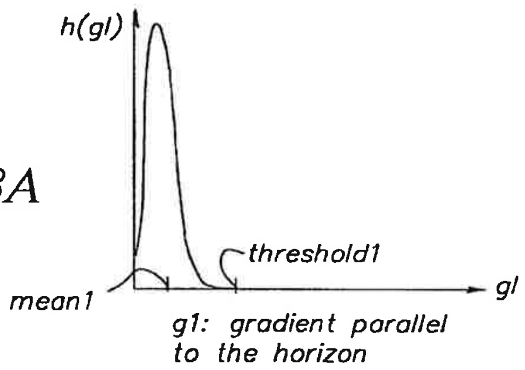


FIG. 8A

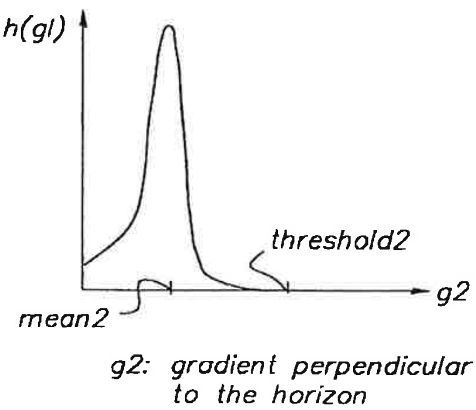


FIG. 8B

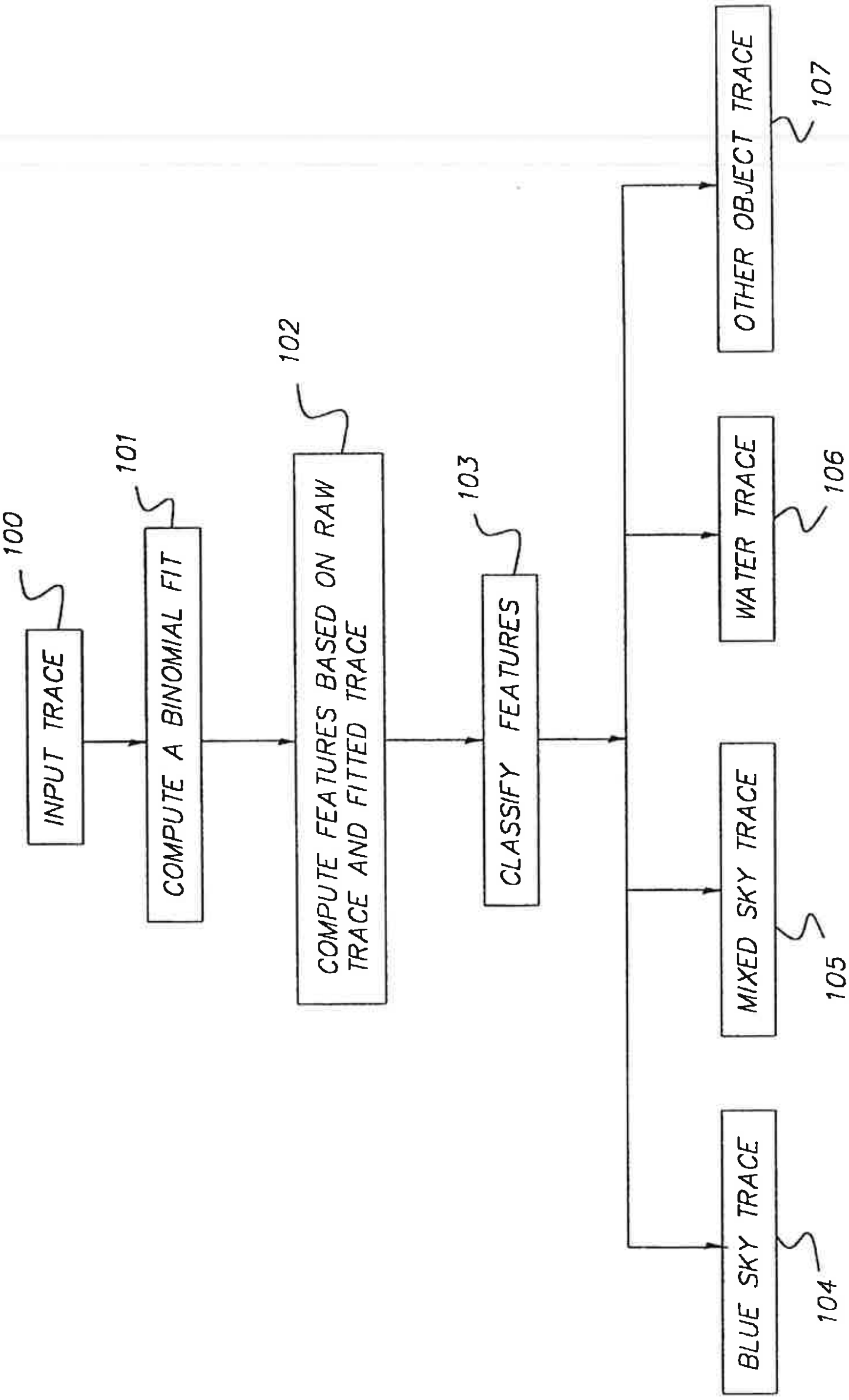

INPUT TRACE
100
COMPUTE A BINOMIAL FIT
101
COMPUTE FEATURES BASED ON RAW TRACE AND FITTED TRACE
102
CLASSIFY FEATURES
103
BLUE SKY TRACE
104
MIXED SKY TRACE
105
WATER TRACE
106
OTHER OBJECT TRACE
107

FIG. 9

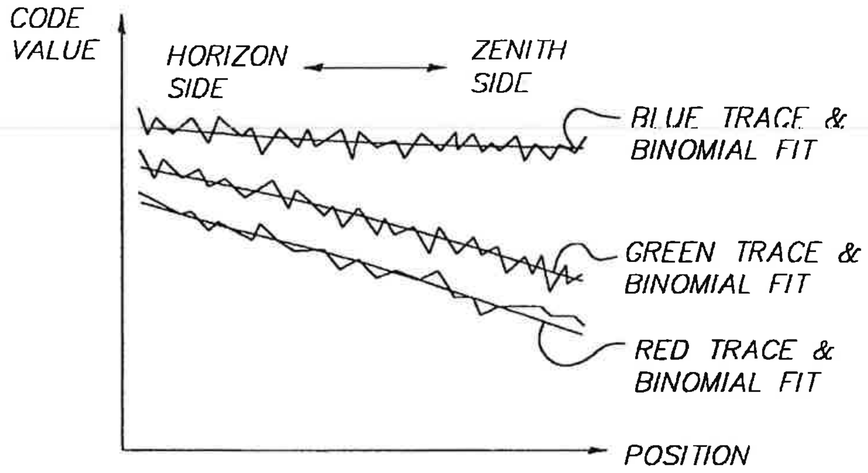


FIG. 10A

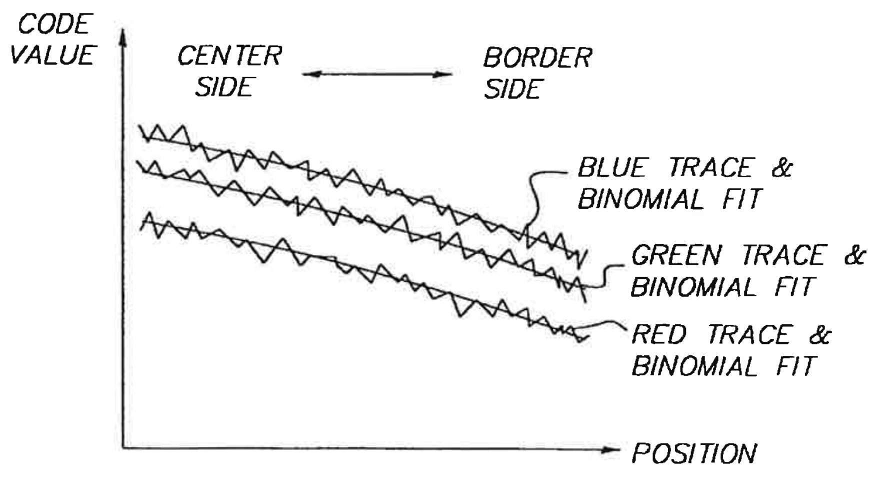


FIG. 10B

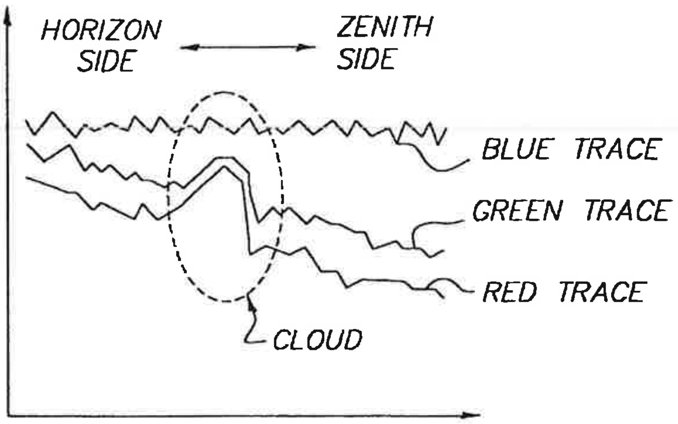


FIG. 11A

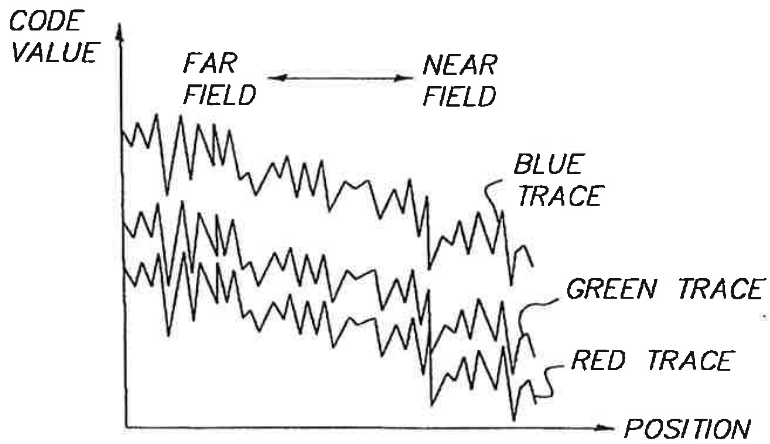


FIG. 11B

Fig. 12A

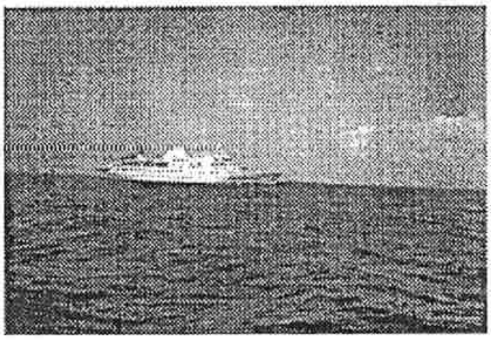


Fig. 12C

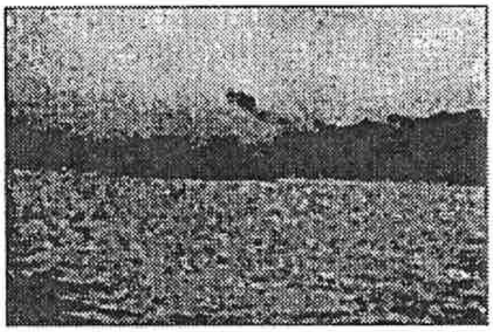


Fig. 12E

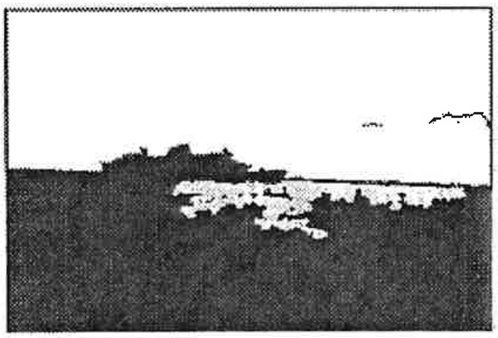


Fig. 12G

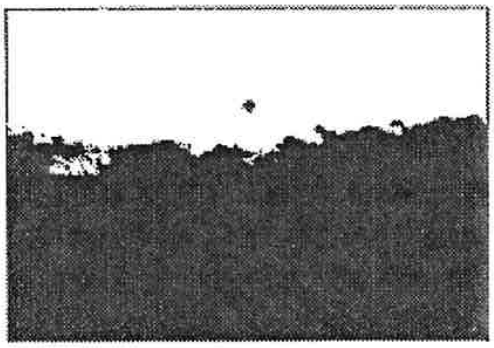


Fig. 12B

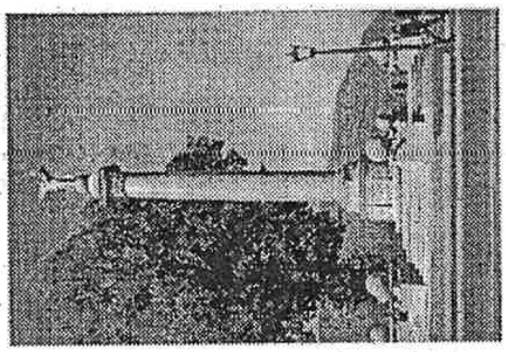


Fig. 12D

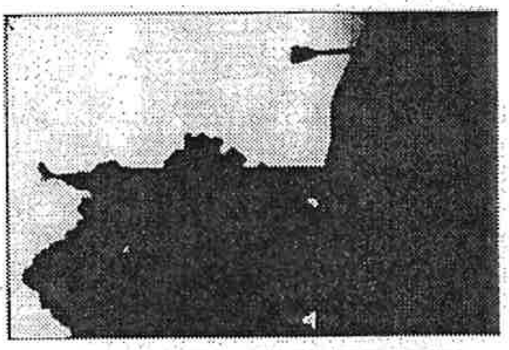


Fig. 12F

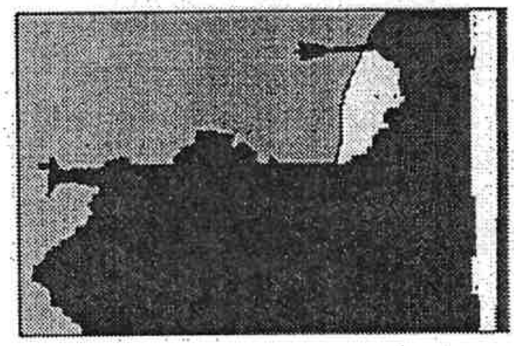


Fig. 12H

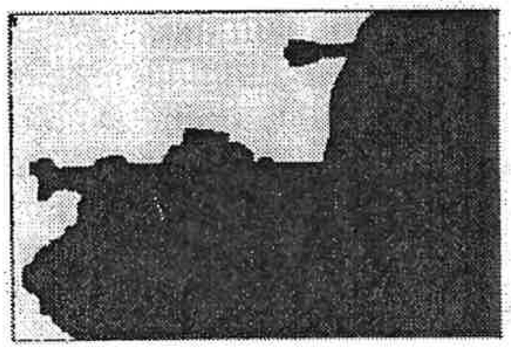


Fig. 13A

Fig. 13C

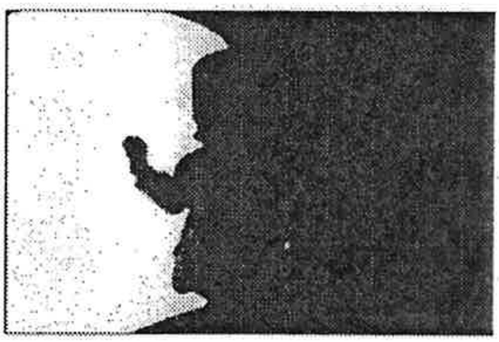


Fig. 13E

Fig. 13G

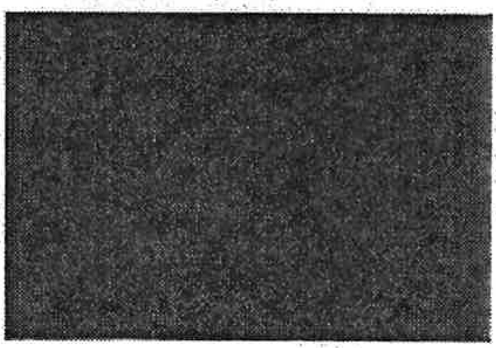


Fig. 13B

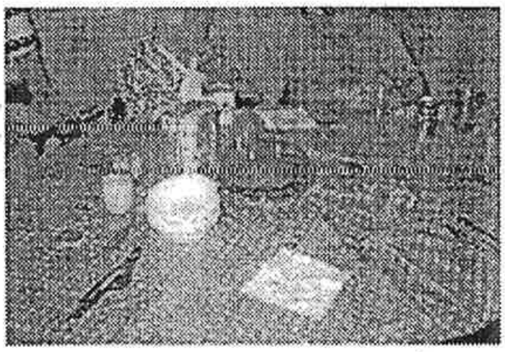


Fig. 13D

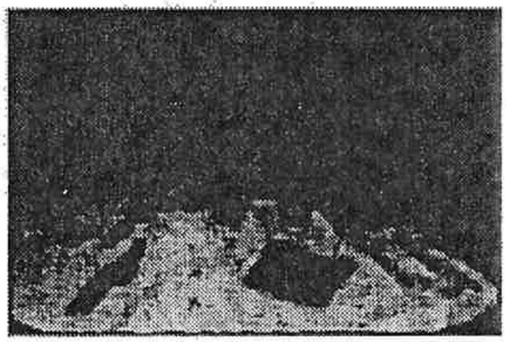


Fig. 13F

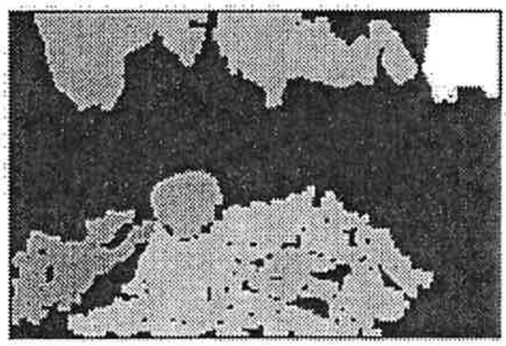
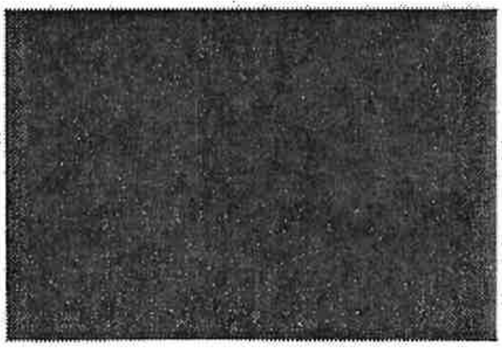


Fig. 13H



METHOD FOR DETECTING SKY IN IMAGES

CROSS-REFERENCE TO RELATED APPLICATION(S)

Reference is made to commonly assigned copending application Ser. No. 09/450,366, entitled "Determining Orientation of Images Containing Blue Sky" and filed in the name of Jiebo Luo, which is assigned to the assignee of this application.

FIELD OF THE INVENTION

The invention relates generally to the field of digital image processing and digital image understanding, and more particular to a system for detecting which regions in photographic and other similar images are of the sky and more particularly to a sky detection system based on color classification, region extraction, and physics-motivated sky signature validation.

BACKGROUND OF THE INVENTION

Sky is among the most important subject matters frequently seen in photographic images. Detection of sky can often facilitate a variety of image understanding, enhancement, and manipulation tasks. Sky is a strong indicator of an outdoor image for scene categorization (e.g., outdoor scenes vs. indoor scenes, picnic scenes vs. meeting scenes, city vs. landscape, etc.). See, for example M. Szummer and R. W. Picard, "Indoor-Outdoor Image Classification," in *Proc. IEEE Intl. Workshop on Content-based Access of Image and Video Database,* 1998 and A. Vailaya, A. Jain, and H. J. Zhang, "On Image Classification: City vs. Landscape," in *Proc. IEEE Intl. Workshop on Content-based Access of Image and Video Database,* 1998 (both of which are incorporated herein by reference). With information about the sky, it is possible to formulate queries such as "outdoor images that contain significant sky" or "sunset images" etc. (e.g., see J. R. Smith and C.-S. Li, "Decoding Image Semantics Using Composite Region Templates," in *Proc. IEEE Intl. Workshop on Content-based Access of Image and Video Database,* 1998, incorporated herein by reference). Thus, sky detection can also lead to more effective content-based image retrieval.

For recognizing the orientation of an image, knowledge of sky and its orientation may indicate the image orientation for outdoor images (contrary to the common belief, a sky region is not always at the top of an image). Further, in detecting main subjects in the image, sky regions can usually be excluded because they are likely to be part of the background.

The most prominent characteristic of sky is its color, which is usually light blue when the sky is clear. Such a characteristic has been used to detect sky in images. For example, U.S. Pat. No. 5,889,578, entitled "Method and Apparatus for Using Film Scanning Information to Determine the Type and Category of an Image" by F. S. Jamzadeh, (which is incorporated herein by reference) mentions the use of color cue ("light blue") to detect sky without providing further description.

U.S. Pat. No. 5,642,443, entitled, "Whole Order Orientation Method and Apparatus" by Robert M. Goodwin, (which is incorporated herein by reference) uses color and (lack of) texture to indicate pixels associated with sky in the image. In particular, partitioning by chromaticity domain into sectors is utilized by Goodwin. Pixels with sampling

zones along the two long sides of a non-oriented image are examined. If an asymmetric distribution of sky colors is found, the orientation of the image is estimated. The orientation of a whole order of photos is determined based on estimates for individual images in the order. For the whole order orientation method in Goodwin to be successful, a sufficiently large group of characteristics (so that one with at least an 80% success rate is found in nearly every image), or a smaller group of characteristics (with greater than a 90% success rate -which characteristics can be found in about 40% of all images) is needed. Therefore, with Goodwin, a very robust sky detection method is not required.

In a work by Saber et al. (E. Saber, A. M. Tekalp, R. Eschbach, and K. Knox, "Automatic Image Annotation Using Adaptive Color Classification", CVGIP: *Graphical Models and Image Processing,* vol. 58, pp. 115–126, 1996, incorporated herein by reference), color classification was used to detect sky. The sky pixels are assumed to follow a 2D Gaussian probability density function (PDF). Therefore, a metric similar to the Mahalonobis distance is used, along with an adaptively determined threshold for a given image, to determine sky pixels. Finally, information regarding the presence of sky, grass, and skin, which are extracted from the image based solely on the above-mentioned color classification, are used to determine the categorization and annotation of an image (e.g., "outdoor", "people").

Recognizing that matching natural images solely based on global similarities can only take things so far. Therefore, Smith, supra, developed a method for decoding image semantics using composite regions templates (CRT) in the context of content-based image retrieval. With the process in Smith, after an image is partitioned using color region segmentation, vertical and horizontal scans are performed on a typical 5×5 grid to create the CRT, which is essentially a 5×5 matrix showing the spatial relationship among regions. Assuming known image orientation, a blue extended patch at the top of an image is likely to represent clear sky, and the regions corresponding to skies and clouds are likely to be above the regions corresponding to grass and trees. Although these assumptions are not always valid, nevertheless it was shown in Smith, supra, that queries performed using CRTs, color histograms and texture were much more effective for such categories as "sunsets" and "nature".

PROBLEMS TO BE SOLVED BY THE INVENTION

The major drawback of conventional techniques is that they cannot differentiate other similarly colored or textured subject matters, such as a blue wall, a body of water, a blue shirt, and so on. Furthermore, some of these techniques have to rely on the knowledge of the image orientation. Failure to reliably detect the presence of sky, in particular false positive detection, may lead to failures in the downstream applications.

SUMMARY OF THE INVENTION

The invention provides a robust sky detection system which is based on color hue classification, texture analysis, and physics-motivated sky trace analysis. The invention utilizes hue color information to select bright, sky colored pixels and utilizes connected component analysis to find potential sky regions. The invention also utilizes gradient to confirm that sky regions are low in texture content and segments open space, defined as smooth expanses, to break up adjacent regions with similar sky color beliefs but dissimilar sky colors. The invention also utilizes gradient to determine the zenith-horizon direction and uses a physics-motivated sky trace signature to determine if a candidate region fits a sky model.

More specifically, the invention can take the form of a method, image recognition system, computer program, etc., for detecting sky regions in an image and comprises classifying potential sky pixels in the image by color, extracting connected components of the potential sky pixels, eliminating ones of the connected components that have a texture above a predetermined texture threshold, computing desaturation gradients of the connected components, and comparing the desaturation gradients of the connected components with a predetermined desaturation gradient for sky to identify true sky regions in the image.

The desaturation gradients comprise desaturation gradients for red, green and blue trace components of the image and the predetermined desaturation gradient for sky comprises, from horizon to zenith, a decrease in red and green light trace components and a substantially constant blue light trace component.

The color classifying includes forming a belief map of pixels in the image using a pixel classifier, computing an adaptive threshold of sky color, and classifying ones of the pixels that exceed the threshold comprises identifying a first valley in a belief histogram derived from the belief map. The belief map and the belief histogram are unique to the image.

The invention also determines a horizontal direction of a scene within the image by identifying a first gradient parallel to a width direction of the image, identifying a second gradient perpendicular to the width direction of the image and comparing the first gradient and the second gradient. The horizontal direction of the scene is identified by the smaller of the first gradient and the second gradient.

ADVANTAGES OF THE INVENTION

One advantage of the invention lies in the utilization of a physical model of the sky based on the scattering of light by small particles in the air. By using a physical model (as opposed to a color or texture model), the invention is not likely to be fooled by other similarly colored subject matters such as bodies of water, walls, toys, and clothing. Further, the inventive region extraction process automatically determines an appropriate threshold for the sky color belief map. By utilizing the physical model in combination with color and texture filters, the invention produces results which are superior to conventional systems.

The invention works very well on 8-bit images from sources including film and digital cameras after pre-balancing and proper dynamic range adjustment. The sky regions detected by the invention show excellent spatial alignment with perceived sky boundaries.

BRIEF DESCRIPTION OF THE DRAWINGS

The foregoing and other objects, aspects and advantages will be better understood from the following detailed description of a preferred embodiment of the invention with reference to the drawings, in which:

FIG. 1 is a schematic architectural diagram of one embodiment of the invention;

FIG. 2 is a schematic architectural diagram of one embodiment of the invention;

FIGS. 3A–3B are schematic diagrams illustrating the colors of daylight and twilight, respectively, in a clear sky;

FIGS. 4A–4D show is a three-dimensional graphical illustration of the cluster of blue sky in color space and the respective color planes that produce the cluster and the respective color planes thar produce the cluster;

FIG. 5 is a graphical illustration of the receiver operating characteristic (ROC) of sky color classification;

FIG. 6 is a schematic architectural diagram of the region extraction portion of the invention;

FIG. 7 is a graphical illustration of the threshold determination for sky color beliefs according to the invention;

FIGS. 8A–8B are graphical illustrations of typical distributions of gradient magnitudes in a sky region;

FIG. 9 is a schematic architectural diagram of the trace analysis performed by the invention;

FIG. 10A is a graph showing a typical trace of clear sky;

FIG. 10B is a graph showing a typical trace of a blue wall;

FIG. 11A is a graph showing a typical trace of mixed sky and clouds;

FIG. 11B is a graph showing a typical trace of water;

FIGS. 12A–12H illustrate different stages of images processed by the invention; and

FIGS. 13A–13H illustrate different stages of images processed by the invention.

DETAILED DESCRIPTION OF THE INVENTION

As shown above, a robust sky detection process needs to go beyond color and texture. Specifically, a physical model of the sky is desirable, if possible, to differentiate true sky regions from other similarly colored and textured subject matters. The invention described below provides a robust sky detection process that differentiates true sky regions from other similarly colored and textured subject matters.

In this application, sky detection comprises identifying all the pixels in an image that correspond to the unoccluded part of the sky. Furthermore, sky detection assigns each individual segmented regions a probability that it contains sky. It is left to the subsequent conventional processing of the image understanding system to either utilize the probability representation or convert it into a crisp decision. Some important features of the invention include a robust sky detection process based on color hue classification, texture analysis, and physics-motivated sky trace analysis; utilization of color hue information to select bright, sky colored pixels; utilization of connected component analysis to find potential sky regions; utilization of gradient to confirm that sky regions are low in texture content (i.e., open space); utilization of open space segmentation to break up adjacent regions with similar sky color beliefs and dissimilar sky colors; utilization of gradient to determine the zenith-horizon direction; and utilization of a physics-motivated sky trace signature to determine if a candidate region fits a sky model.

The subject matter of the present invention relates to digital image understanding technology, which is understood to mean technology that digitally processes a digital image to recognize and thereby assign useful meaning to human understandable objects, attributes or conditions and then to utilize the results obtained in the further processing of the digital image.

A block diagram of the overall sky detection system (e.g., the digital image understanding technology) is shown in FIG. 1. First, a digital image **10** is digitally processed **20**. The results **30** obtained from processing step **20** are used along with the original digital image **10** in an image modification step **40** to produce a modified image **50**.

A more specific block diagram of the inventive sky detection process is shown in FIG. 2. The inventive method comprises three main stages. In the first main stage (e.g., item **201**), color classification is performed by a multi-layer back-propagation neural network trained in a bootstrapping fashion using positive and negative examples, that is discussed in detail below. The output of the color classification is a map of continuous "belief" values, which is preferable over a binary decision map.

In the next main stage, a region extraction process (e.g., item **202**) automatically determines an appropriate threshold for the sky color belief map by finding the first valley point encountered moving from lower beliefs to high beliefs in the belief histogram, and performs a connected component analysis. In addition, open space detection (e.g., item **204**) is incorporated to (1) rule out highly textured regions and (2) separate sky from other blue-colored regions such as bodies of water. Taking the intersection between pixels with supra-threshold belief values, and the connected components in the open-space map creates seed regions. For pixels with sub-threshold belief values, the continuity in belief values as well as continuity in color values guide region growing from the seed regions.

Finally, in the third main stage, the sky signature validation process (e.g., items **205**–**209**) estimates the orientation of sky by examining vertical/horizontal gradients for each extracted region, extracting 1D traces within the region along the estimated horizon-to-zenith direction, determining (by a set of rules discussed below) whether a trace resembles a trace from the sky, and finally computing the sky belief of the region based on the percentage of traces that fit the physics-based sky trace model. In one embodiment, the invention identifies the horizontal direction of a scene within the image by identifying a first gradient parallel to a width direction of the image and a second gradient perpendicular to the width direction of said image, where the smaller of the first gradient and the second gradient indicate the horizontal direction of the scene.

More specifically, in FIG. 2, an input image is received in digital form **200**. The pixels are then classified into sky-colored and non sky-colored pixels **201**, using the inventive color classification process, as discussed below. Using the connected component analysis also discussed below, a spatially contiguous region of sky-colored pixels is extracted **202**. Gradient operators are overlaid on every interior pixel of the connected component (or "region") to compute horizontal and vertical gradient values **203**. The pixels near the boundary of the connected component are preferably excluded in one embodiment because they often represent the large-magnitude transition between the sky and other subject matters, for example, at the horizon.

The average horizontal and vertical gradient values, Gx and Gy, are computed using all the interior pixels of the region. A number of tests will disqualify a candidate region based on excessive texture. Thus, if either gradient value is above a pre-determined high threshold T_{high}, indicating that the region is highly textured, the region is not considered a sky region. If $|G_x|$ and $|G_y|$ are almost identical, the region is also not considered a sky region. Furthermore, if the color (hue) distribution of all the pixels in the candidate region does not fit the expected characteristic of a sky region, the region is also not considered a sky region.

The invention recognizes that, the 3D shape of the sky color distribution should resemble a tilted ellipsoid with its long axis approximately along the luminance direction, which is partially a result of the desaturation effect, as discussed in detail below.

If the region passes the low texture test **204**, the possible direction of zenith to horizon orientation is determined **205**. If not, processing returns to item **202** to analyze the next potential region of pixels that has sky color. In particular, the gradient in the red channel is examined. If $|Gx|>|Gy|$, there is an indication of a landscape image. Otherwise, the image is most likely a portrait image. Furthermore, for a landscape image, if Gx<0, there is an indication of an upright image, otherwise it is most likely an upside-down image. For a portrait image, if Gy<0, there is an indication of a leftside-up image, otherwise it is most likely a rightside-up image.

Traces are then extracted across a candidate sky region along the horizon-zenith direction **206**. For each trace, a plurality of sky-trace signatures **207** are measured to determine whether each trace likely comes from a sky region. The likelihood **208**, or belief that a candidate region is sky, is determined by the voting from all the extracted sky traces. If the overall belief of a candidate region is above a pre-determined threshold **209**, the candidate region is declared a sky region **210**. Processing then returns to analyze all candidate regions in the same fashion (e.g., processing returns to item **202**). In the case where detected sky regions disagree on the sky orientation, the overall orientation of the image is decided by the results from larger, higher belief sky regions. Regions with conflicting sky orientations are rejected.

It is almost axiomatic that, to the human visual system, the sky is blue, grass is green, dirt is gray/red/brown, and water is blue/green. However, what is actually recorded in a digital image is somewhat different. This is true not only for sky regions that contain warm colors associated with sunrise and sunset, but also for sky regions that appear more blue than their color records indicate. To confound even more the problem, color balance of the whole image can be off due to the error introduced during image capture and in other stages of the imaging chain.

The blue appearance of the sky in a color image is the results of human physiology and psychology, as well as physics—the red and green component at a blue-appearing sky pixel can be more intense (by a small percentage) than the blue component. In addition, clear, unclouded sky is usually the brightest subject matter in an image, although the sun itself, illuminated clouds, snow, ice or some man-made objects can be brightened than the blue sky. The sun radiates most brightly in the orange-yellow wavelength. The wavelength selective scattering of air particles disperses the blue light component of the sun ray's much more strongly than the longer wavelength according to Rayleigh's law, which states that scattering is inversely proportional to the fourth power of the wavelength (e.g., see C. F. Bobren and D. R. Huffman, *Absorption and Scattering of Light by Small Particles*, New York, John Wiley and Sons, 1983, incorporated herein by reference). The color of the sky is, indeed, largely composed of violet (to which our eyes are not very sensitive) and further a fair amount of blue, a little green and very little yellow and red—the sum of all these components is sky-blue (e.g., see M. Minnaert, *The Nature of Light and Color in the Open Air.* New York: 1954, incorporated herein by reference).

However, the blue appearance of the sky is not uniform. Sky often appears desaturated toward the horizon. When one looks at the clear sky directly overhead with the sun off to the side, the scattered blue light dominates and the sky appears as deep blue. As one shifts the gaze towards a distant horizon, the various selective factors tend to equalize and the sky appears desaturated to almost white.

There are a number of interesting effects regarding the distribution of light in the sky, e.g., halos, mirages, and rainbows. Among them, the light intensity increases from the zenith to the horizon while at the same time the color changes from deep blue to white. This effect arises primarily from the great thickness of the layer of air between our eyes and the horizon. Although the small particles of the air scatter the blue rays by preference, the scattered rays are weakened most in their long path from the scattering particles to our eyes. Because of a very thick stratum of air, the scattering and attenuation effects counteract each other.

Suppose a small particle at a distance s from a given spot scatters the fraction sds (where s is the color-dependent scattering factor and ds is the size of the particle). The amount of light is weakened in the ratio e^{-sx} before reaching that given spot. The light received from an infinitely thick layer of air (a reasonable approximation) would consist of the sum of contributions from all the particles ds, that is,

$$\int_0^{\infty} se^{-sx}dx,$$

which is equal to one. Evidently, the amount of received light is then independent of s, and thus the color of the light.

Therefore, the sky close to the horizon shows the same brightness and color as a white screen illuminated by the sun. Moreover, the layers of air close to the ground may contain more floating large particles of dust, which scatter light of all colors equally intensely and make the color of the light whiter (even when the layer of air cannot be considered to be of infinite thickness).

If the observer is facing away from the sun, when the sunshines behind the observer of laterally, the concentric distribution of the light can be approximately parallel to the horizon because of the position of the sun (high above the horizon) as well as the observer's limited view. If the observer looks in the direction of the sun (one should stand in the shadow of a building near the edge of the shadow), the brightness of the sky increases rapidly close to the sun and even becomes dazzling, its color becoming more and more white. In photographic images, it is extremely unlikely that one would take a picture of the direct sun light, except at sunrise or sunset, when the sun is on the horizon and the intensity of the light is much weaker.

While the blue sky can be considered as the finest example of a uniform gradation of color, twilight's exhibit much more dramatic color gradation in a similar form of concentric distribution of constant brightness and color, as illustrated in FIGS. 3A–B. More specifically, FIGS. 3A–B illustrate the different colors, which are seen at the eastern horizon as the sun sets (e.g., daylight vs. twilight) in the western horizon. Although it is not the focus of this invention to detect twilight sky, these unique signatures of the twilight sky can be exploited in a more general sky detection process. In fact, when one of the features used in the invention was turned off, the process successfully detected the twilight sky in FIG. 3B, as discussed below.

It is also important to look at the factors determining the color of the water, which is often indistinguishable from that of the sky. Part of the light our eye receives from water is reflected by the surface; it acts like a mirror when it is smooth, and the color of the water is blue, gray according to the color of the sky. The color of the sea (or any large open body of water) in the distance is about the same as that of the sky at the height of 20° to 30°, and darker than the sky immediately above the horizon. This is because only part of the light is reflected when our gaze falls on the slopes of distant wavelets (e.g., see Minnaert, supra).

Apart from reflection, deep water has a "color of its own"—the color of the light scattered back from below. The depth of the deep water and similar deep water can be considered so great that practically no light returns form the bottom of it. The "color of its own" is to be attributed to the combined effects of scattering and absorption in the water. The color of deep, almost pure water is blue due to the absorption by the water in the orange and red parts of the spectrum, after the light penetrates the water and is scattered back again.

For the purpose of sky detection, one important issue is to differentiate bodies of blue (usually deep) water, whether they co-appear with the sky or not, from the sky. The factors of great concern are the absorption of orange and red components of the light by the water. The waves and undulations of such deep water bodies create small surfaces of various slopes. In general, the color is darker when our gaze falls on a surface more perpendicular to the gaze or closer to us. However, the changes are primarily in brightness instead of hue.

Turning now to color classification, mentioned briefly above (e.g., item **201** in FIG. **2**), the invention first trains a color classifier specifically for clear, light-blue sky seen at daytime for simplicity and clarity. Sky regions which contain the warm colors associated with sunrise and sunset are not be lumped in with the blue-sky and gray-sky regions that form the background in many outdoor scenes. In the context of the invention, the color-based detection identifies all candidate blue sky pixels, which are then screen as regions for spatial signatures consistent with clear sky.

Neutral network training is then utilized to complete the training of the color classifier. The initial training set includes images having ideal blue sky characteristics, gray sky images, and non-sky (primarily indoor) images. All blue sky pixels were included as positive examples, and negative examples were included by sampling from among all pixels that are neither blue sky nor water.

A feedforward neural network was constructed with two hidden layers, containing 3 or 2 neurons, and a single output neuron (e.g., see Howard Demuth and Mark Beale, *Matlab Neural Network Toolbox*, The Math Works, Inc., 1998). The hidden layer neurons had tangent-sigmoidal transfer functions, while the output neuron's transfer function was log-sigmoidal. The network was trained using Levenberg-Marquardt backpropagation to classify pixel values as ideal blue sky or non-sky (e.g., see Howard Demuth and Mark Beale). The target responses are a=1 for ideal blue sky pixels and a=0 for non-sky.

The color classifier, so trained, outputs a belief value between 0 and 1 for each pixel processed, 1 indicating a pixel highly likely to be blue sky and 0 indicating a pixel not very likely to be blue sky. To help visualize the invention's response to points in the (r,g,b) input space, a regularly-spaced grid of (r,g,b) triplets from example images processed with the invention is shown in FIG. 4A, with each color plane shown separately in FIGS. 4A–4D.

Points producing a blue-sky belief great that 0.1 are marked by "." in FIG. 4A. The projections of this distribution onto the three planes are also shown (marked by "o"). Note that the distribution is highly elongated along the direction of luminance, and starts to diverge a bit towards lower luminance. For a specific input image, each pixel is classified independently, and a belief map is created by setting the brightness of each pixel proportional to its belief value. Examples of such belief maps are shown in FIGS. **12**E–F and **13**E–F.

A pixel-level receiver operating characteristic (ROC) of the inventive color classifier is shown in FIG. **5**. This curve shows the true positive and false positive performance if the processing in the color classifier was immediately followed by a hard threshold at a variety of levels.

Conventionally, the global threshold is not dynamic and is found by locating the position on the curve closest to the upper left-hand corner of the graph shown in FIG. 5. For example, using a threshold of 0.0125 gives correct detection of 90.4% of bluesky pixels, but also detects (incorrectly) 13% of non-blue-sky pixels. Among those detected non-blue-sky pixels, water accounts for a significant portion. To the contrary, the invention does not employ a predefined "hard" threshold, but instead performs a region-extraction process before validating each region against a set of sky signatures. This process is discussed in detail below with respect to FIG. 7.

More specifically, the inventive region extraction process (e.g., item **202** discussed above) automatically determines an appropriate threshold for the sky color belief map by finding the first valley point encountered moving from lower beliefs to higher beliefs in the belief histogram, and then performs a connected component analysis, as shown in FIG. 7. In addition, with the invention, the connected components are refined to produce a region-level representation of the sky segments, which facilitates sky signature validation that is otherwise impossible at the pixel level.

In FIG. 6, more detail is given for the region extraction process **202** (in FIG. 2). For a belief map **71**, where the value of each pixel is proportional to the belief of that pixel having a sky color, a global threshold **72** is determined in an adaptive fashion, as discussed below with respect to FIG. 7. A binary map **73** is created using this threshold, whereas a "1" pixel is considered as a candidate sky pixel and a "0" pixel is considered as a non-sky pixel. Connected components, which are regions of spatially contiguous "1" pixels, are uniquely labeled **74** to produce spatially separated nonzero regions of sky color. Note that non-sky pixels are labeled to "0" (referred to herein as "unlabeled") regardless of their connectivity. Each connected component of sky color is refined **75** using two operations, which are discussed in greater detail below, to produce the connected components of sky color **76**. An open space map **77** (which is also discussed below) is combined with the connected components to produce the candidate sky regions that are output by item **202** in FIG. 2.

FIG. 7 illustrates the inventive process for dynamically determining the global threshold. First, a histogram of the belief values is obtained form the belief map of sky color. Next, the histogram is smoothed to remove noise (e.g., producing the cart shown in FIG. 7). The first significant valley (e.g., "First valley" in FIG. 7) is found in the smoothed histogram. In a simple image where there is a distinctive sky region and everything else is distinctively non-sky, the histogram has only two peaks and one valley in between. In complex images there are sky, water and other blue regions. Therefore, the invention utilizes a different histogram for each image, which permits a dynamic threshold to be created for each individual image processed by the invention.

In Saber, supra, the last valley in the smoothed histogram was used to adjust a universal threshold in a maximum likelihood estimation (MLE) scheme based on the assumption that the true sky region in an image always has the highest probability. However, in some cases, a blue-colored non-sky region may have higher sky belief in terms of color. Therefore, the invention retains all sky-colored regions for further analysis and rules out non-sky regions that happen to have sky colors in the alter stages of the sky detection process. Therefore, the belief value at which the first valley is located is chosen as the global threshold. As mentioned above, this threshold is determined adaptively for each individual image to accommodate different shades of sky as well as the image capturing conditions.

The first of the two refinement operations, discussed in item **75** above, is region splitting. Region splitting is used to split spatially connected bluish (potential sky) regions that belong to different objects but otherwise have similar belief values in terms of having sky color. For example, such a region could be blue cloth against blue sky. Such regions may have similar beliefs (in being typical colors of sky) and thus are not separable in the belief map.

However, such regions have different shades of blue colors and thus are separable using a general-purpose color segmentation processes, such as an adaptive k-means processing (e.g., see J. Luo, R. T. Gray, and H.-C. Le "Towards a Physics-Based Segmentation of Photographic Color Images," in *Proc. IEEE Int. Conf. Image* Process., 1997, incorporated herein by reference). The invention utilizes this process and splits a labeled region of sky color into two or more regions (with unique new labels) if the region is a conglomerate of multiple regions indicated by the color segmentation process.

In another embodiment of the invention, an open-space detection process **77** (described in J. Warnick, R. Mehrotra and R. Senn, U.S. Pat. No. 5,901,245, "Method and system for detection and characterization of open space in digital images," incorporated herein by reference) can be used instead of a general-purpose color segmentation process. Open space is defined a smooth and contiguous region in an image. It is very useful for placing a desired caption or figurative element in an image.

The automatic open-space detection process mentioned above (Warnick, supra) is based on two separate stages of operation. First, after a proper color space transformation is performed, a gradient-based activity map is computed and a proper threshold is determined according to a multi-region histogram analysis. In the second stage, a connected component analysis is performed on the binary activity map to fill voids and small regions are discarded. The open-space process as implemented in Warnick, supra, is both effective and efficient. Its speed is only a fraction of that required for the color segmentation process. In addition, open-space detection provides additional confirmation of the smoothness of the candidate regions. Therefore, in this preferred embodiment, the invention utilizes the open-space detection process. Thus, open space detection is incorporated to (1) rule out highly textured regions and (2) separate sky from other blue-colored regions such as bodies of water.

The second refinement operation performed in item **75** of FIG. 6 comprises region growing. The inventive region growing process is used to fill in holes and extend boundaries. This is especially useful where "marginal" pixels may have sky-color belief values that barely fail the global threshold but are close enough to the belief values of the neighboring pixels that have passed the initial global threshold.

With the invention a "growing threshold" is used to rebel such marginal pixels to a connected component if the difference in belief values between an "unlabeled" pixel and its neighboring "labeled" pixel is smaller than a second threshold for region growing. More specifically, seed regions are created by taking the intersection between pixels with supra-threshold belief values and the connected components in the open-space map. For pixels with sub-threshold belief values, region growing is guided by the continuity in belief values as well as continuity in color values. Small, isolated sky regions are ignored.

In the sky signature measures, which are discussed above in item **207** in FIG. **2**, one-dimensional traces are extracted within the region along the horizon-to-zenith direction. The invention automatically determines the sky orientation based on the distribution of both vertical-horizontal gradients in each extracted region.

More specifically, the invention uses the red signal to determine the sky orientation, because of the physics-motivated model of sky. As discussed above, with the physics-motivated model of sky, the amount of light scattering depends on the wavelength of the light and the scattering angle. In general, the desaturation effect towards the horizon is caused by the increase in red light and green light relative to blue light. Furthermore, the present inventors have determined that blue light stays relatively unchanged along the horizon-zenith direction. The change in the green signal may not be as pronounced as in the red signal. Therefore, the red signal provides the most reliable indication of the desaturation effect. Consequently, the uneven gradient distribution is most observable in the red signal.

Because of the desaturation effect, sky has low gradient in the horizon-zenith direction, but is essentially constant in the perpendicular direction. When the position of the sun is high above the horizon, the concentric distribution of the scattering light can be approximated by horizontal strips of different color regions (e.g., see FIG. **3**, barring lens falloff effect). Therefore, the distribution of gradient has different characteristics in horizontal and vertical directions, as shown by FIGS. **8**A and **8**B (which are parallel and perpendicular to the horizon, respectively), where mean1<<mean2.

After regions extraction **202** and orientation determination **205**, the sky signature validation process extracts one-dimensional traces within the region along the determined horizon-to-zenith direction **206**, determines by a set of rules whether the trace resembles a trace from the sky **207**, and finally computes the sky belief of the region by the percentage of traces that fit the physics-based sky trace model **208**, as discussed above.

Based on the analysis of numerous one-dimensional traces from sky as well as a few other typical sky-colored subject matters in images, the invention includes models to quantify these traces. In particular, traces extracted along the horizon-zenith direction reveal a signature of sky traces shown in FIG. **10**A. The blue signal of a key trace tends to be constant across the sky; the green signal and red signal gradually decrease away from the horizon; the red signal decreases faster than the green signal. More specifically, all the three signals can be approximated by lower-order polynomials (e.g., quadratic polynomials). The micro-variations in the three signals are not correlated. In comparison, a few other blue-colored subject matters do not exhibit such a signature. To the contrary, in FIG. **10**B, there is shown a typical trace of a blue wall in a flash-fired picture, where the three signals change smoothly in parallel.

Similarly, FIG. **11**B shows a typical trace through a body of water, where the three signals are highly correlated in local variations. Both of these two cases indicate that the changes are mostly in luminance. Furthermore, as illustrated in FIG. **11**A, in mixed sky where (white) clouds are present together with clear blue sky, the red and green signals jump high in the clouds while the blue signal stays the same to create a neutral cloud region. Typically, the red signal jumps up by a larger amount than the green signal in the clouds.

FIG. **9** is a flowchart illustrating the processing of the input trace. More specifically, in item **100**, an extracted trace is analyzed with respect to the trace models shown in FIGS. **10**A–**11**B. First a quadratic polynomial fit **102** is computed for the three signals: red, green and blue, respectively. The quadratic polynomial is given as $y=f(x)=c1+c2*x+c3*x^2$ where x denotes the index of the one-dimensional trace and y is the code value of the corresponding signal.

Next, a plurality of features ("signatures") are computed based on either the raw trace or the fitted trace **102**. Features are classified **103** so that a trace can be characterized as a blue sky trace **104**, or a non-blue-sky trace (a mixed sky trace **105**, a water trace **106**, or "unknown" **107**). In the example shown below, ten measure ("signatures") are computed for each extracted sky trace. However, one ordinarily skilled in the art could prepare any number of such signatures in light of this disclosure.

The first signature regards the offset of the fitting quadratic polynomial. The offsets are related to the mean values in red, green, and blue channels. This signature feature requires the average blue component to be above the average red and green components. Due to the specific way a trace is extracted, this features actually translates into the requirement that the blue component is the strongest at the most blue side of the trace. C-language-like pseudo code for such a logical statement follows:

```
If (cb[1]>cr[1]+BR_OFFSET&&cb[1]>cg[1]-BF_OFFSET&&cg
   [1]>cr[1]-RG_OFFSET) sig1=1
```

where

# define BR_OFFSET	10
# define BG_OFFSET	5
# define RG_OFFSET	5

Instead of using the above crisp rule, it may be advantageous to use a trapezoidal fuzzy scoring function of continuous values with a cutoff point with a certain HUGEPENALTY if this condition is violated.

The second exemplary signature regards the slope of the fitting quadratic polynomial. In general, due to the specific way a trace is extracted, the slopes of RGB signals are negative. This feature requires that the blue signal decreases (if so) slower than the red and green signals. On the other hand, monotonic increase (positive slope) is also allowed by this feature. C-language-like pseudo code for such a logical statement follows.

```
if (cb[2]>cg[2]&&cb[2]>cr[2]) sig2=1;

if (!sig2&&sig2bg&&sig2br) sig2=1;
```

This is implemented as a crisp rule. Exception is granted to relax the strict condition of sig2 to two more loosely defined conditions sig2bg and sig2br when sig2 is not satisfied.

The third signature regards the similarity or parallelism among the fitted signals. Pseudo code for such a logical statement follows.

```
if ((rgdist<brdist&&bgdist<brdist)||(rgdist<bgdist&&rgdist<brdist))
    sig3=1;
```

Note that rgdist is used to indicate the difference ("distance") between two fitted red and green signals. It is determined in the following way. First, one of the two signals is shifted appropriately such that the shifted signal has the same value at the starting point as the unshifted signal. Let the fitted red and green signals be

$$r(x)=c^r_1+c^r_2+c^r_3x^2 g(x)=c_1^g+c_2^gx+c^g_3x^2 \quad (7)$$

then

$$\hat{r}(x)=r(x)+(c_1^g-c_1^r)=c_1^g+c^r_2x+c^r_2x+c^r_2x+c^r_3x^2 \quad (8)$$

such that $\hat{r}(0)=g(0)$.

Next, the difference or distance between the fitted red and green signals is given by

$$rg\text{dist}=|r(L/2)-g(L/2)| \quad (9)$$

where L is the total length of the trace. In other words, this feature measures the difference between two fitted signals by the distance at two midpoints when one of them is shifted so that both signals have the same starting value. The other two terms, bgdist and brdist, are defined in a similar fashion. One possibility here is not to use the absolute values such that sign information is used in conjunction with the absolute difference.

The fourth signature regards red-green similarity. The red and green signals should be reasonably similar. Pseudo code for such a logical statement follows

```
if (rgdist<MAXRGDIST) sig4=1;
```

where

define MAXRGDIST 15

The fifth signature regards low nonlinearity. All the three signals should have low nonlinearity. Pseudo code for such a logical statements follows.

```
If (fabs (cb[3])<MAXNONLINEARITY&&fabs
    (cg[3])<MAXNONLINEARITY&&fabs (cr[3])<MAXNON-
    LINEARITY) sig5=1;
```

where

#define MAXNONLINEARITY 0.05

Instead of using the above crisp rule, it may be advantageous to use a sigmoid fuzzy scoring function of continuous values with a cutoff point with a certain HUGEPENALTY if this condition is violated. The sixth signature regards red-green-blue correlation for large modulation. Pseudo code for such a logical statement follows.

```
if (largesignal&&corr_rg>0.5&&corr_br<0.3&&corr_bg<0.3)
    sig6=1; //red-grn-blue correlation for large modulation

else if (!largesignal&&corr_rg>0.2&&corr_br<0.4&&corr_
    bg<0.4) sig6-1; //red-grn-blue correlation for small modulation

else if (largesignal==-1) sig6=1; //red-grn-blue correlation for
    micro modulation

if (largesignal!=-1&&corr_rg>0.9&&corr_rg>5*corr_
    br&&corr_rg>5*corr_bg) sig6=-1; //significantly higher red-
    grn correlation
```

where corr-xy denotes the correlation coefficient between signal x and y. Again, instead of using the above crisp rule, it may be advantageous to use a sigmoid fuzzy scoring function of continuous values with a cutoff point with a certain HUGEPENALTY if this condition is violated (s>0.95). The seventh signature regards red-green-blue similarity or near parallelism. Pseudo code for such a logical statement follows.

```
if (rgdist<MINDIST&&bgdist<MINDIST&&brdist<MINDIST)
    sig7=0;
```

where

#define MINDIST 1.5

As before, instead of using the above crisp rule, it may be advantageous to use a sigmoid fuzzy scoring function of continuous values with a cutoff point with a certain HUGEPENALTY if this condition is violated (s>0.95).

The eighth signature regards negative red/green slope. Pseudo code for such a logical statement follows.

```
if (cr[2]>0ANDcg[2]>0) sig8=0;
```

This is implemented as a crisp rule.

The ninth signature regards goodness of the fit. Pseudo code for such a logical statement follows.

```
if (rchisq>MAXCHISQ_R&&gchisq>MAXCHISQ_
    G&&bchisq>MAXCHISQ_B) sig9=0;
```

where

#define MAXCHISQ_R 50

#define MAXCHISQ_G 25

#define MAXCHISQ_B 100

where CHISQ denotes a χ-square fitting error.

Also, instead of using the above crisp rule, it may be advantageous to use a sigmoid fuzzy scoring function of continuous values with a cutoff point where a certain HUGEPENALTY if this condition is violated (s<0.1).

Signature ten regards the decrease in red and green signals. Pseudo code for such a logical statement follows.

```
sigA=rdec*gdec;
```

where rdec indicates whether the red signal decreases (monotonically). In particular, rdec is determined using the fitted red signal by taking two samples first x1 at ¼th point and second x2 at ¾th point of the total length, respectively Pseudo code for such a logical statement follows.

```
if (x2<x1) rdec=1;
```

else

```
rdec=0
```

the other term gdec is determined in a similar fashion for the green signal. This is implemented as a crisp rule. Note that sigA=1 if and only if rdec=1 and gdec=1.

These ten features are integrated in the current rule-based process as a crisp decision; a given trace is only declared a sky trace when all the condition are satisfied, i.e.,

```
if (sig1&&sig2&&sig3&&sig4&&sig5&&sig6!=
    0&&sig7&&sig8&&sig9&&sigA) skysignature=1;
```

Or, in a fuzzy logic-based algorithm,

```
if (sig1>EFFECTIVEZERO&&sig2>EFFECTIVEZERO&&
    sig3>EFFECTIVEZERO&&sig4>EFFECTIVEZERO&&
    sig5>EFFECTIVEZERO&&fabs(sig6)>EFFECTIVEZERO)
    &&sig7>EFFECTIVEZERO&&sig8>EFFECTIVEZERO&&
    sig9>EFFECTIVEZERO&&sigA>EFFECTIVEZERO) skysig-
    nature=(sig1+sig2+ . . . +sig9+sigA)/10
```

where

#define EFFECTIVEZERO 0.1

Upon examination of all candidate traces, which are mostly (e.g., 95%) of sky-colored pixels, the sky belief of the region is computed as the percentage of traces that satisfy the physics-based sky trace model. A sky-colored region is declared as non-sky if the sky belief is below a threshold (empirically determined at 0.25 in this example for general purposes).

FIGS. 12A–13H illustrate the invention's performance on various images. More specifically, FIGS. 12A–B and 13A–B illustrate original images to be processed. FIGS. 12C–D and 13C–D represent the results of the color classification process of the invention shown in item 201 in FIG. 2, discussed above. FIGS. 12E–F and 13E–F illustrate the result of the open space map (item 77 in FIG. 6) produced by the invention. FIGS. 12G–H and 13G–H illustrate the invention's determination of the sky regions as white portions and non-sky regions as black portions. The brightness level in FIGS. 12C–D and 13C–D is proportional to the sky color beliefs, however the brightness level in 12E–F and 13E–F merely indicates separated label regions.

The invention works well on RGB images produced by such sources as film and digital cameras. The detected sky regions show excellent alignment to perceptual boundaries. The few examples shown in FIGS. 12A–12H demonstrate the performance of the invention. The sky and the sea are correctly separated and the true sky region is detected in FIG. 12G. The image in FIG. 12B is an example where the assumption of sky at the top is invalid but the sky is nevertheless correctly detected by the proposed process based on correct determination of the sky orientation. A smooth blue object in FIG. 13A and a textured table cloth in FIG. 13B are correctly rejected, respectively, by the invention.

Given the effectiveness of the inventive sky signature validation process, it is possible to relax the color classification stage to include other off-blue shades of the sky, such as the shades at sunset or sunrise. In contrast to overcast sky, cloudless sky at sunset or sunrise exhibits similar scattering effect as the counterpart during the day. The main difference is the warm color tint from the rising or setting sun.

A 2D planar fit of a candidate region is an alternative way of conducting sky validation. For regions that have holes, the weighting factor at hole locations can be set to zero so that only the sky-colored pixels contribute to the planar fit. It may be necessary to require that the holes can only be due to bright neutral objects (clouds) to limit the potential increase of false positive detection.

Therefore, the invention comprises a system for sky detection that is based on color classification, region extraction, and physics-motivated sky signature validation. The invention works very well on 8-bit images from sources including film and digital cameras after pre-balancing and proper dynamic range adjustment. The detected sky regions also show excellent spatial alignment with perceived sky boundaries.

As mentioned above, the invention utilizes a physical model of the sky based on the scattering of light by small particles in the air. By using a physical model (as opposed to a color or texture model), the invention is not likely to be fooled by other similarly colored subject matters such as bodies of water, walls, toys, and clothing. Further, the inventive region extraction process automatically determines an appropriate threshold for the sky color belief map. By utilizing the physical model in combination with color and texture filters, the invention produces results that are superior to conventional systems.

While the invention has been described in terms of preferred embodiments, those skilled in the art will recognize that the invention can be practiced with modification within the spirit and scope of the appended claims.

What is claimed is:

1. A method of detecting sky regions in an image comprising:

classifying potential sky pixels in said image by color;

extracting connected components of said potential sky pixels;

eliminating ones of said connected components that have a texture above a predetermined texture threshold;

computing desaturation gradients of said connected components; and

comparing said desaturation gradients of said connected components with a predetermined desaturation gradient for sky to identify true sky regions in said image.

2. The method in claim 1, wherein said desaturation gradients comprise desaturation gradients for red, green and blue trace components of said image.

3. The method in claim 1, wherein said predetermined desaturation gradient for sky comprises, from horizon to zenith, a decrease in red and green light trace components and a substantially constant blue light trace component.

4. The method in claim 1, wherein said classifying comprises;

forming a belief map of pixels in said image using a pixel classifier;

computing an adaptive threshold of sky color; and

classifying ones of said pixels that exceed said threshold as said potential sky pixels.

5. The method in claim 4, wherein said computing of said adaptive threshold comprises identifying a first valley in a belief histogram derived from said belief map.

6. The method in claim 5, wherein said belief map and said belief histogram are unique to said image.

7. The method in claim 1, further comprising determining a horizontal direction of a scene within said image by:

identifying a first gradient parallel to a width direction of said image;

identifying a second gradient perpendicular to said width direction of said image; and

comparing said first gradient and said second gradient, said horizontal direction of said scene being identified by the smaller of said first gradient and said second gradient.

8. A method of detecting sky regions in an image comprising: computing desaturation gradients of regions in said image; and

comparing said desaturation gradients of said regions with a predetermined desaturation gradient for sky to identify true sky regions in said image.

9. The method in claim 8, further comprising classifying potential sky regions in said image by color.

10. The method in claim 8, further comprising eliminating ones of said regions that have a texture above a predetermined texture threshold.

11. The method in claim 8, wherein said desaturation gradients comprise desaturation gradients for red, green and blue trace components of said image.

12. The method in claim 8, wherein said predetermined desaturation gradient for sky comprises, from horizon to zenith, a decrease in red and green light trace components and a substantially constant blue light trace component.

13. The method in claim 9, wherein said classifying comprises:

forming a belief map of pixels in said image using a pixel classifier;

computing an adaptive threshold of sky color; and

classifying ones of said pixels that exceed said threshold as said potential sky pixels.

14. The method in claim 13, wherein said computing of said adaptive threshold comprises identifying a first valley in a belief histogram derived from said belief map.

15. The method in claim 14, wherein said belief map and said belief histogram are unique to said image.

16. The method in claim 8, further comprising determining a horizontal direction of a scene within said image by:

identifying a first gradient parallel to a width direction of said image;

identifying a second gradient perpendicular to said width direction of said image; and

comparing said first gradient and said second gradient, said horizontal direction of said scene being identified by the smaller of said first gradient and said second gradient.

17. An image identification system comprising:

a color classifier classifying potential sky pixels in said image by color;

an extractor identifying connected components of said potential sky pixels;

a first comparator eliminating ones of said connected components that have a texture above a predetermined texture threshold;

a logic unit computing desaturation gradients of said connected components; and

a second comparator comparing said desaturation gradients of said connected components with a predetermined desaturation gradient for sky to identify true sky regions in said image.

18. The image identification system in claim 17, wherein said desaturation gradients comprise desaturation gradients for red, green and blue trace components of said image.

19. The image identification system in claim 17, wherein said predetermined desaturation gradient for sky comprises, from horizon to zenith, a decrease in red and green light trace components and a substantial constant blue light trace component.

20. The image identification system in claim 17, wherein said color classifier includes:

a pixel classifier forming a belief map of pixels in said image;

a second logic unit computing an adaptive threshold of sky color; and

a third comparator classifying ones of said pixels that exceed said threshold as said potential sky pixels.

21. The image identification system in claim 20, wherein said second logic unit identifies a first valley on a belief histogram derived from said belief map to computer said adaptive threshold.

22. The image identification system in claim 21, wherein said belief map and said belief histogram are unique to said image.

23. The image identification system in claim 17, further comprising:

a horizontal identification unit that identifies a first gradient parallel to a width direction of said image and a second gradient perpendicular to said width direction of said image,

a third comparator comparing said first gradient and said second gradient, said horizontal direction of said comparator identifying said scene by identifying the smaller of said first gradient and said second gradient.

24. A method of using and a computer program running on a computer system to detect sky regions in an image, said method comprising:

using said computer program to classify potential sky pixels in said image by color;

using said computer program to extract connected components of said potential sky pixels;

using said computer program to eliminate ones of said connect components that have a texture above a predetermined texture threshold;

using said computer program to computer desaturation gradients of said connected components; and

using said computer program to compare said desaturation gradients of said connected components with a predetermined desaturation gradient for sky to identify true sky regions in said image.

25. The method in claim 24, wherein said desaturation gradients comprise desaturation gradients for red, green and blue trace components of said image.

26. The method in claim 24, wherein said predetermined desaturation gradient for sky comprises, from horizon to zenith, a decease in red and green light trace components and a substantially constant blue light trace component.

27. The method in claim 24, wherein said using of said computer program to classify comprises:

using said computer program to form a belief map of pixels in said image using a pixel classifier;

using said computer program to compute an adaptive threshold of sky color; and

using said computer program to classify ones of said pixels that exceed said threshold as said potential sky pixels.

28. The method in claim 27, wherein said using of said computer program to computer said adaptive threshold comprises using said computer program to identify a first valley in a belief histogram derived from said belief map.

29. The method in claim 28, wherein said belief map and said belief histogram are unique to said image.

30. The method in claim 24, further comprising using said computer program to determine a horizontal direction of a scene within said image by:

using said computer program to identify a first gradient parallel to a width direction of said image;

using said computer program to identify a second gradient perpendicular to said width direction of said image; and

using said computer program to comparing said first gradient and said second gradient, said horizontal direction of said scene being identified by the smaller of said first gradient and said second gradient.

* * * * *

US006483940B1

(12) **United States Patent**
Wang

(10) **Patent No.: US 6,483,940 B1**
(45) **Date of Patent: Nov. 19, 2002**

(54) **METHOD FOR DIVIDING IMAGE**

(75) Inventor: **Tachun Wang**, Taipei (TW)

(73) Assignee: **Institute for Information Industry** (TW)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

(21) Appl. No.: **09/358,680**

(22) Filed: **Jul. 22, 1999**

(51) **Int. Cl.**^{7} **G06K 9/00**
(52) **U.S. Cl.** **382/164**; 382/170; 382/171; 358/515; 358/538
(58) **Field of Search** 382/164, 170, 382/171, 225; 358/538, 515

(56) **References Cited**

U.S. PATENT DOCUMENTS

5,047,842 A * 9/1991 Bouman, Jr. et al. 358/75
5,649,025 A * 7/1997 Revankar 382/171
5,848,181 A * 12/1998 Ogata 382/169
5,889,885 A * 3/1999 Moed et al. 382/172
5,930,385 A * 7/1999 Fujimoto et al. 382/162
5,933,524 A * 8/1999 Schuster et al. 382/168
6,148,102 A * 11/2000 Stolin 382/164
6,175,427 B1 * 1/2001 Lehmbeck et al. 358/1.9
6,181,817 B1 * 1/2001 Zabih et al. 382/170
6,195,458 B1 * 2/2001 Warnick et al. 382/173
6,272,250 B1 * 8/2001 Sun et al. 382/225
6,282,317 B1 * 8/2001 Luo et al. 382/203
6,310,967 B1 * 10/2001 Heine et al. 382/128

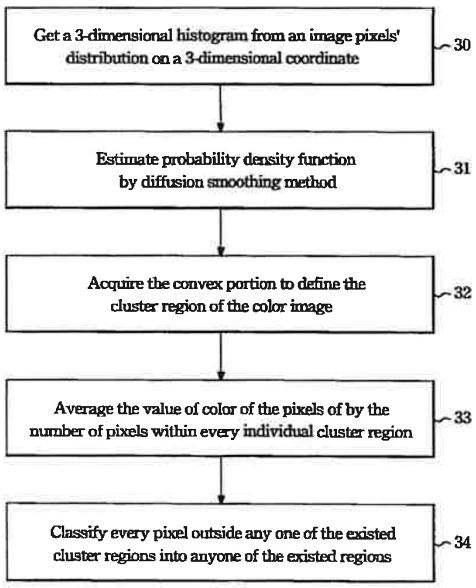
* cited by examiner

Primary Examiner—Joseph Mancuso
Assistant Examiner—Anand Bhatnagar
(74) *Attorney, Agent, or Firm*—Baker Botts LLP

(57) **ABSTRACT**

The method proposed by the present invention divide an image into a plurality of cluster regions, each has its own principle color. The method mentioned above includes the following steps. At first, acquire a three dimensional histogram of the image by counting the number of pixels of the image based on three dimensional color coordinate. Then smooth the three dimensional histogram by a numerical method to obtain a function. Next, define a plurality of cluster regions on the image. Subsequently, assign the color of each pixel of the image as a principle color of a first cluster region among the plurality of cluster regions. The principle color is equal to average of all pixels within the cluster region, the average is equal to A/B, wherein A is summation of color of pixels within the first cluster region, and B is equal to the number of pixels within the first cluster region. Each of the plurality of cluster regions has a corresponding principle color. Finally, joint an unclassified pixel to a cluster region of the plurality of cluster regions according to minimum distance from the unclassified pixel to the plurality of cluster regions. The unclassified pixel is outside the plurality of cluster regions before this joint step, after all pixels of the image have being jointed to one of the plurality of cluster regions, the image is divided into the plurality of cluster regions.

14 Claims, 3 Drawing Sheets



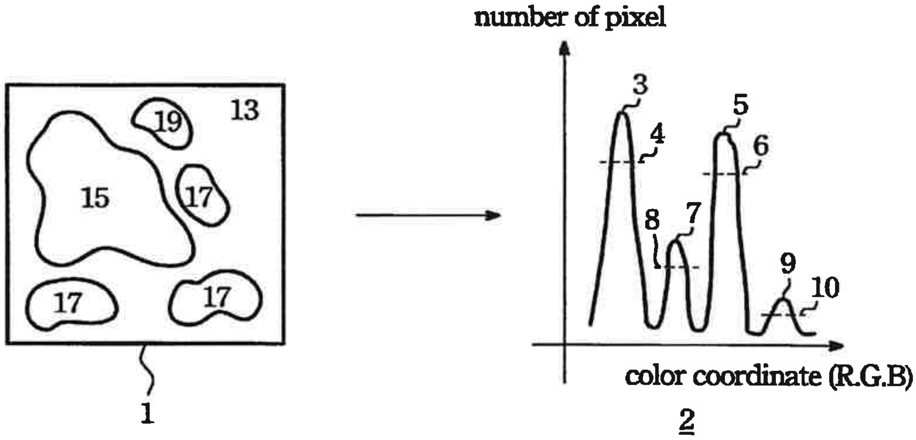

13
19
15
17
17
17
1
number of pixel
3
4
5
6
8
7
9
10
color coordinate (R.G.B)
2

Fig.1

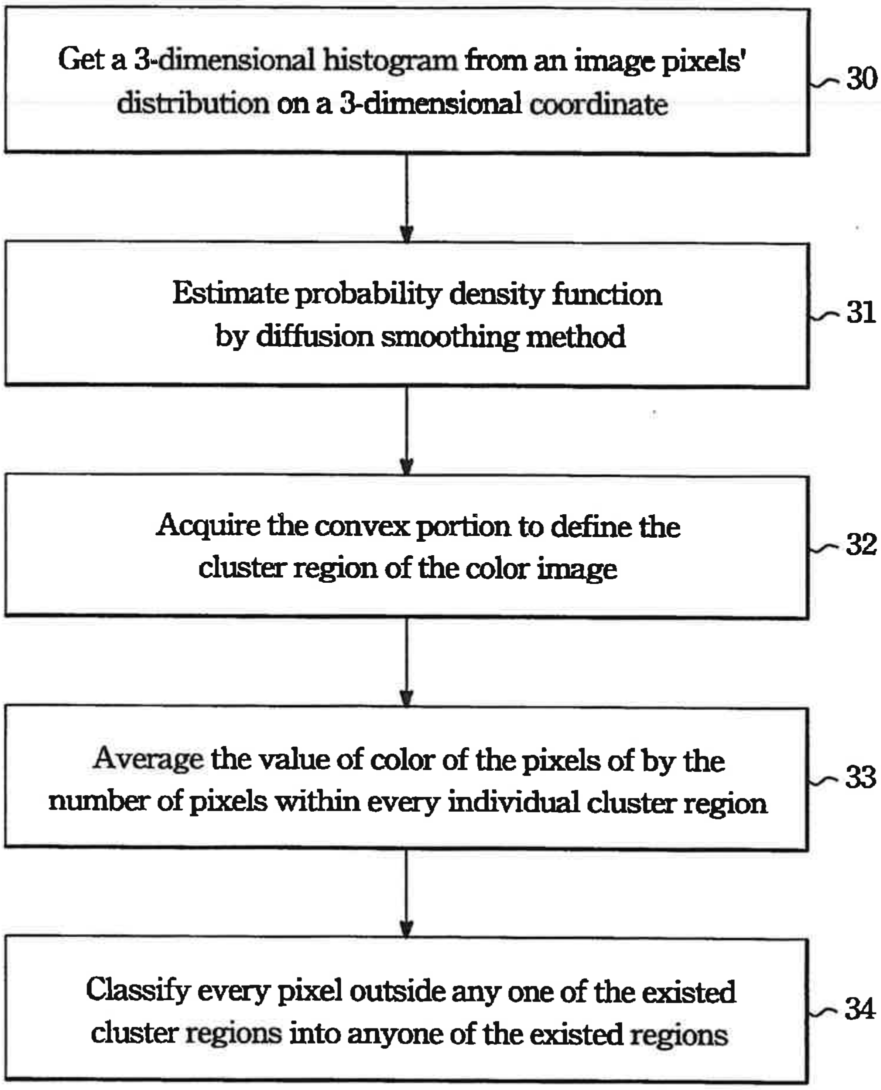

Get a 3-dimensional histogram from an image pixels' distribution on a 3-dimensional coordinate
30
Estimate probability density function by diffusion smoothing method
31
Acquire the convex portion to define the cluster region of the color image
32
Average the value of color of the pixels of by the number of pixels within every individual cluster region
33
Classify every pixel outside any one of the existed cluster regions into anyone of the existed regions
34

Fig.2

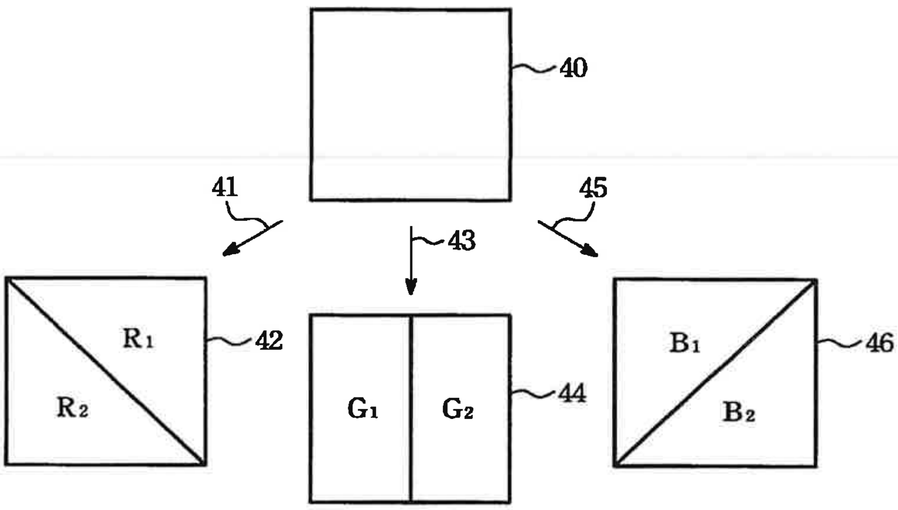

40
41
43
45
R1
R2
42
G1
G2
44
B1
B2
46

Fig.3

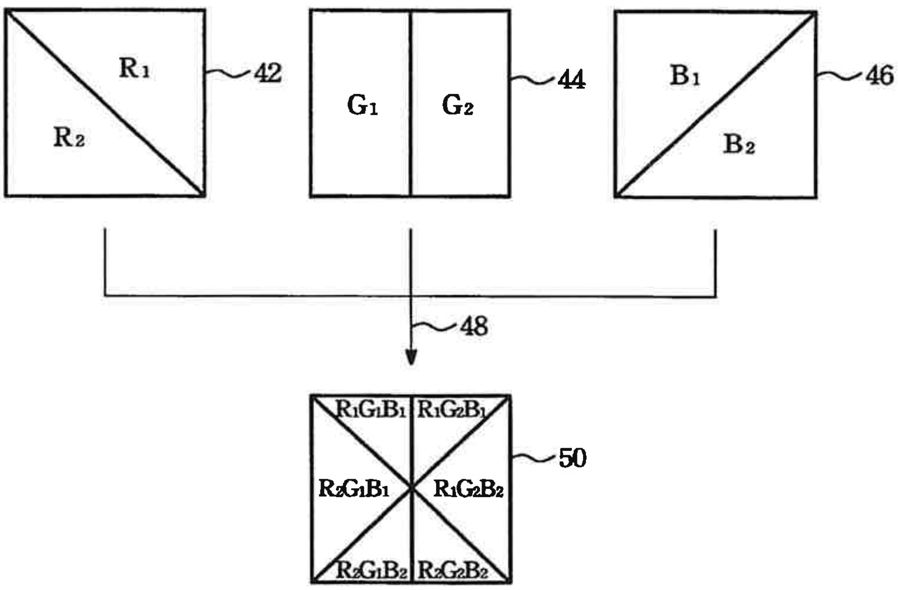

R1
R2
42
G1
G2
44
B1
B2
46
48
R1G1B1
R1G2B1
R2G1B1
R1G2B2
R2G1B2
R2G2B2
50

Fig.4

METHOD FOR DIVIDING IMAGE

BACKGROUND OF THE INVENTION

1. Field of the Invention

This invention relates to a method for dividing image by descending number ot color, and more particularly, to a method for dividing image in accordance with inflection on a distribution curve of image pixel versus color coordinate.

2. Description of the Prior Art

When a traditional technology is used to divide an image, pixels of similar color are classified into a cluster region, so an image can be classified into a plurality of cluster regions. Finally, if any pixel is not classified into one of the existed cluster regions, determine the distance between the color of the pixel and that of every existed cluster region. Then joint the pixel to the existed cluster that has the shortest distance of color to the pixel, in other words, the pixel is classified into the cluster. So the color of the pixel is replaced with the color of the color of the existed cluster that has shortest distance to the pixel mentioned above. So all the pixels of the image are classified into the plurality of cluster regions.

The color histogram of an image is frequently used in dividing an image into a plurality of cluster regions, especially the color of a region having a local maximum within the color histogram is assigned as the principal color of the cluster region. When dividing an image into several cluster regions, the relation between the image mentioned above and its color histogram is shown in FIG. **1**. The image **1** includes a plurality of pixels, and the statistic for the number of pixel having a color on color coordinate acquires its color histogram **2**. Then find a first local maximum value **3**, and set up first threshold value **4**. Next, by the same way, a second local maximum value **5**, its second threshold value **6**, the third local maximum value **7**, its third threshold value **8**, the fourth local maximum value **9**, and its fourth threshold value **10**. Every local maximum value is actually a value on the color coordinate. Subsequently, assign the color of the pixel having the local maximum value to be the color of the pixels of the cluster region enclosing the pixel having the local maximum value. In other words, the principle color of a cluster region is assigned to be the color of the pixel enclosed by the cluster region, which having the local maximum value. Next, calculate the distance from the color of the pixel beyond the plurality of cluster regions to the plurality of local maximum values. So a plurality of distance is acquired, then select a minimum distance from the plurality of distance. The principle color of the cluster region having the minimum distance to the pixel beyond the plurality of cluster regions is assigned as the color of the pixel beyond the plurality of cluster regions. Then it is defined that the pixel beyond the plurality of cluster regions is joint to a cluster region which having minimum distance to the pixel. Also, it is defined that, the pixel is classified to a cluster region. Proceeding with the steps mentioned above continuously until all the pixels of the image is classified into one of the existed cluster regions. Through the steps mentioned above, thus obtain a plurality of cluster regions which including all pixels of image **1**. For example, as shown in FIG. **1**, the first cluster region **13** in image **1** includes the pixels having the color the same as the principle color of the first local maximum value **3** illustrated in the color histogram **2**. Similarly, the second cluster region **15** in image **1** includes the pixels having the color the same as the principle color of the second local maximum value **5** illustrated in the color histogram **2**. The third cluster region **17** in image **1** includes the pixels having the color the same as the principle color of the first local maximum value **7** illustrated in the color histogram **2**. The fourth cluster region **19** in image **1** includes the pixels having the color the same as the principle color of the fourth local maximum value **9** illustrated in the color histogram **2**.

The colorful image dividing method mentioned above is a technology utilizing a moving window to look for a local maximum value in the color histogram. Because the colorful-image dividing method mentioned above does not analysis the color histogram according to the feature of color distribution, thus the obtained principle colors does not fully match the color distribution of the image. Furthermore, the size of the moving window as well as the threshold values are the fixed values set up by the user, i.e., the window size and the threshold values are not set up according to the actual color distribution. So the exactly color distribution can not be obtained, and the resulted color division does not actually fit the color distribution of the colors on the image. Also, the division does not fit the actual color distribution, so the resulted principle color of the cluster region is not correct. In conclusion, the traditional method mentioned above can not properly reduce the number of color of the image.

SUMMARY OF THE INVENTION

According to the prior art mentioned above, the traditional method utilized to divide a pre-processed image utilize the manually determined threshold value as the criterion when dividing the image, so it is impossible to divide the image according to the characteristic of color distribution of the pre-processed image. Thus the traditional method can not divide the image correctly.

A purpose of the present invention is to avoid using the fixed threshold value to divide the image, instead, the present invention uses the convex portion of a function to obtain the plurality of cluster regions. The function mentioned above is obtained by a smoothed histogram of the image, so the cluster regions are acquired according to the characteristic of the distribution of color of the pre-processed image. The present invention is used to divide a pre-processed image into a plurality of cluster regions, such that all the pixels of the pre-processed image are classified into any one of the plurality of cluster regions. In other words, the pre-processed image is divided into a plurality of cluster regions.

According to the purpose mentioned above, the method used to divide an image is disclosed herein. The method divide an image into a plurality of cluster regions, each has its own principle color, and every pixel of the image is classified into one of the plurality of cluster regions. The method mentioned above includes the following steps.

At first, acquire a three dimensional histogram of the image by counting the number of pixels of the image based on three dimensional color coordinate. Then smooth the three dimensional histogram by a numerical diffusion method to obtain a function. The function can be a normal distribution probability function. Next, define a plurality of cluster regions on the image, thus the pixels within the plurality of cluster regions correspond to convex portion of the function. Subsequently, assign the color of each pixel of the image as a principle color of a first cluster region among the plurality of cluster regions. The principle color is equal to average of all pixels within the cluster region, the average is equal to A/B, wherein A is summation of color of pixels within the first cluster region, and B is equal to the number of pixels within the first cluster region. The color of the pixel is represented by the three color coordinate, and each of the plurality of cluster regions have a corresponding principle color. Finally, joint an unclassified pixel to a cluster region of the plurality of cluster regions according to minimum distance from the unclassified pixel to the plurality of cluster regions. The unclassified pixel is outside the plurality of cluster regions before this joint step, after all pixels of the image have being jointed to one of the plurality of cluster regions, the image is divided into the plurality of cluster regions.

BRIEF DESCRIPTION OF THE DRAWINGS

The above features of the present invention will be more clearly understood from consideration of the following descriptions in connection with accompanying drawings in which:

FIG. 1 illustrates the relation between the divided image and the histogram utilized to define the cluster regions of the image, furthermore the pixels defined by a local maximum value and its threshold on the histogram related to a corresponding cluster region;

FIG. 2 illustrates the steps utilized to divide an image according to one preferred embodiment of the present invention;

FIG. 3 illustrates the steps utilized to divide an preprocessed image according to the other preferred embodiment of the present invention, in which the pre-processed image is processed according to three color series respectively to acquire color distribution;

FIG. 4 illustrates the steps utilized to add the divided image according to the other preferred embodiment of the present invention, in which the divided images according to the three color coordinates are added, and the pixel of the added image each has 3-dimensional color coordinate.

DESCRIPTION OF THE PREFERRED EMBODIMENT

At first, in the preferred embodiment of the present invention, the statistic method is used to analyze the colorful image, thus a plurality of cluster regions can be defined according to the actual distribution of a plurality of colors. Subsequently, calculate the average value of the color of the pixels within every individual cluster region, and take the average value as the principle color of the individual cluster region. In the present invention, because the window size and the threshold value are not manually set up, the resulted color distribution stands for the exact feature of the color distribution of the image. The method proposed in the preferred embodiment of the present invention mentioned above is detailed below in FIG. 2. Firstly, as shown in step **30** "get a 3-dimensional histogram from an image pixels' distribution on 3-dimensional coordinate", analyze the 3-dimensional distribution of pixels of an image, and get a 3-dimensional histogram of the image. The first step mentioned above analyze every pixel of the image according to the pixel's color composition, and get a corresponding position in a 3-dimensional coordinate (such as color coordinate: RGB, HSV or YUV . . . etc.). After all the pixels of the image of have been processed by the step mentioned above, calculating the number of pixels on every position of the 3-dimensional color coordinate. Thus the color histogram of the image is acquired.

The following step is step **31** "estimate probability density function by diffusion smoothing method", which uses a numerical diffusion method to calculate the probability density function of the 3-dimensional histogram of the image. The obtained probability density function is a function inverse proportional to the nature log of a squared coordinate. For example, the one dimensional probability density function can be expressed as e^{-x^2/σ^2}, which is known as the normal distribution that is a well-known probability density function. Besides, the numerical diffusion method is a method frequently used in thermodynamics, whereas the numerical diffusion method is used to smooth the color histogram of the colorful image in the preferred embodiment of the present invention.

The next step—step **32** "acquire the convex portion to define the cluster region of the colorful image", is used to acquire a plurality of the cluster regions of the colorful image by determining the convex portion of the probability density. The relation between the convex portion (above threshold value) of the histogram and the cluster region of the colorful image had been illustrated in FIG. **1**. Similar to foregoing description about the convex portion and the cluster region, in the preferred embodiment of the present invention, a plurality of cluster regions can be obtained by proceeding with the histogram of the colorful image by step **32**. Thus the colorful image is divided into a plurality of cluster regions. The next step is step **33** "average the value of color of pixels by the number of pixels within every individual cluster region". The aim of step **33** is to determine the principle color of every individual cluster region, and then assign the principle color as the color of every pixel within the cluster region. So the value of color of every pixel within a specific cluster region is assigned as the value of the principle color on the color coordinate. In step **33**, the average is performed as the follow description. Assume a cluster region includes a number (**n1**) of first pixels having first color coordinate (**c1**) and a number (**n2**) of second pixels having second color coordinate (**c2**). Step **33** is to divide the summation of the value of color of all pixels within a cluster region by the summation of the number of the first pixel and the second pixel, in other words, the result can be expressed as **(c1×n1+c2×n2)/(n1+n2)**. The summation of the value of color of all pixels within a cluster region can be obtained by multiplying the value of color of the first pixel by the number of the first pixel plus multiplying the value of color of the second pixel by the number of the second pixel.

The principle color of every existed cluster region of the image can be obtained by performing step **33**, and the pixels belonging to none of the existed cluster region should be processed with step **34**. Step **34** "classify every pixel outside any one of existed cluster regions into any one of the existed cluster region" is to classify the pixels outside any existed cluster region into any one of the existed cluster regions. The method used in step **34** is continuously classify the pixel outside the existed cluster region into any one of the existed cluster regions till all the pixels are classified. In the preferred embodiment of the present invention, a pixel outside any one of the existed cluster regions can be classified into a cluster region as the description below. Determine the minimum distance from the color of the unclassified pixel to the principle color of every existed cluster region; then classify the unclassified pixel into the cluster region that has the minimum distance to the pixel. Continue the classifying step mention above till all the unclassified pixels are classified (joint) to one of the existed cluster regions. The division of the image is finished after step **34** had been performed. Because the number of color of the colorful image having been divided by the method in the present invention is reduced, the present invention can also be used to reduce the number of color of a colorful image when dividing the image.

It is noted that although the preferred embodiment of the present invention mentioned above is to proceed with colorful image, however it can be used to process the monotone image. During processing a monotone image, the variation of the present invention for the monotone image is to process the gray level of the monotone image. In spite of the variation mentioned above, the method utilized to divide a colorful image is basically the same as the method utilized to divide a monotone image. In other words, if the image is composed of a single color, the step **30** "get a 3-dimensional histogram from an image pixels' distribution on 3-dimensional coordinate" should be replaced with the step "get a gray level histogram from an image pixels' distribution on gray level coordinate". In the following steps, the "colorful image" should be replaced with the "single color image", and then the method shown in FIG. **2** that is utilized to process a colorful image can be used to process a monotone image.

Because the method utilized in dividing a monotone image is basically the same as that utilized in dividing a colorful image, the step utilized to divide the monotone image is not shown in figure due to the similarity to FIG. **2**. The method utilized in dividing a monotone image is to divide a monotone image into a plurality of cluster regions such that every pixel of the monotone image is classified into any one of the cluster regions. The method used to divide the monotone image includes basically the following steps. At first, the statistic method is used to analyze the monotone image, thus a plurality of cluster regions can be defined according to the actual distribution of a plurality of colors. Subsequently, calculate the average value of the gray level of the pixels within every individual cluster region, and take the average value as the principle gray level of the individual cluster region. Next, step "get a gray level histogram from an image pixels' distribution on every position of gray level coordinate", analyze the gray level distribution of pixels of an image, and get a gray level histogram of the image. The first step mentioned above analyze every pixel of the image according to the pixel's gray level, and get a corresponding position in a gray level coordinate. After all the pixels of the image have been processed by the step mentioned above, calculate the number of pixels on every position of the gray level coordinate. Thus the gray level histogram of the image is acquired. The following step is "estimate probability density function by diffusion smoothing method", which uses a numerical diffusion method to calculate the probability density function of the gray level histogram of the image. The next step "acquire the convex portion to define the cluster region of the monotone image", is used to acquire a plurality of cluster regions of the monotone image by determining the convex portion of the probability density function. Thus the monotone image is divided into a plurality of cluster regions. The next step is step "average the gray level value of pixels by the number of pixels within every individual cluster region". The aim of the step mentioned above is to determine the principle gray level of every individual cluster region, and then assign the principle gray level as the gray level of every pixel within the cluster region. So the gray level value of every pixel within a specific cluster region is assigned as the gray level value on the gray level coordinate. In foregoing step, the average is performed by the follow description. Assume a cluster region includes a number (n1) of first pixels having first gray level (g1) and a number (n2) of second pixels having second gray level (g2). The foregoing step is to divide the summation of the gray level value of all pixels within a cluster region by the summation of the number of the first pixel and the second pixel, in other words, the result can be expressed as $(g1 \times n1 + g2 \times n2)/(n1 + n2)$. The summation of the gray level value of all pixels within a cluster region can be obtained by multiplying the gray level value of the first pixel by the number of the first pixel plus multiplying the gray level value of the second pixel by the number of the second pixel. The principle gray level of every existed cluster region of the image can be obtained, and the pixels belonging to none of the existed cluster region should be processed with the following step: classify every pixel outside any one of existed cluster regions into any one of the existed cluster regions. The step "classify every pixel outside any one of existed cluster regions into any one of the existed cluster regions" is to classify the pixels outside any existed cluster region into any one of the existed cluster regions. The method mentioned above is continuously classify the pixel outside the existed cluster region into any one of the existed cluster regions till all the pixels are classified. In the preferred embodiment of the present invention, a pixel outside any one of the existed cluster regions can be classified into a cluster region as the description below. The method mentioned above is used to process the monotone image according to the present invention.

The other preferred embodiment of the present invention is used to divide the colorful image by the other way, which uses the method that is utilized to divide a monotone image. The other preferred embodiment of the present invention firstly divide the image based on a first primary color coordinate (such as Red coordinate), subsequently, divide the image based on a second primary color coordinate (such as Green coordinate), next, divide the image based on a third primary color coordinate (such as Blue coordinate). In the following step, acquire the 3-dimensional coordinate of every pixel from the pixel's position relating to the three independent coordinates (such as R coordinate, G coordinate, and B coordinate), so the three-dimensional color coordinate of a pixel is acquired, and every pixel is then colorful. Thus the division of the colorful image relating to the three color coordinates is finished. Because the method utilized to divide a one-color image is the same as that utilized in the other preferred embodiment of the present invention. The only difference is that the other preferred embodiment of the present invention respectively process the colorful image based on the three primary colors so as to acquire the distribution of every pixel of the colorful image in the three primary colors. So the method according to the preferred embodiment of the present invention is not illustrated in figure, instead, the method mentioned in the other preferred embodiment is described.

The method proposed by the other preferred embodiment of the present invention mentioned above is to divide a colorful image into a plurality of cluster regions, such that the pixel of the colorful image is joint to any one of the plurality of cluster regions. The method proposed by the other preferred embodiment of the present invention includes the following steps.

Firstly, divide the pre-processed image (colorful image before processing) according to the first color coordinate by the steps as follows: Firstly, analyze the first color distribution of pixels of an image, and get a first histogram of the image. The step mentioned above analyze every pixel of the image according to the pixel's value on the first color coordinate, and get a corresponding position in the first coordinate. After all the pixels of the image have been processed by the step mentioned above, calculate the number of pixels on every position of the first color coordinate. Thus the first color histogram of the image is acquired. The following step uses a numerical diffusion method to estimate the first function of the first color histogram of the image. Next, by determining the convex portion of the first function, a plurality of cluster regions of the image based on the first color coordinate are acquired. Thus the image is divided into a plurality of the first type cluster regions. The next step is to divide the total value of the pixels on the first color coordinate by the number of pixels within every individual first type cluster region. The aim of the step mentioned above is to determine the first principle color of every individual first type of cluster region, and then assign the first principle color as the value of the first color coordinate of every pixel within the first type cluster region. So the first color coordinate value of every pixel within a specific first type cluster region is assigned as the first color coordinate value on the first color coordinate. In the foregoing step, the average is performed by the follow description. Assume a first type cluster region includes a number (n1) of first pixels having first color coordinate (p1) and a number (n2) of second pixels having second color coordinate (p2). Then divide the summation of the first color coordinate value of all pixels within a cluster region by the summation of the number of the first pixel and the second pixel, in other words, the result can be expressed as $(p1\times n1+p2\times n2)/(n1+n2)$. The summation of the value of the first color coordinate of all pixels within a first type cluster region can be obtained by multiplying the value of the first color coordinate of the first pixel by the number of the first pixel plus multiplying the value of the second color coordinate of the second pixel by the number of the second pixel. The first principle color of every existed cluster region of the image can be obtained, and the pixels belonging to none of the existed cluster region should be processed with the following step. Classify every pixel outside any one of the existed first type cluster regions into any one of the existed first type cluster regions. The step mentioned above is to classify the pixels outside any existed first type cluster region into any one of the existed first type cluster regions. The method mentioned above is to continuously classify the pixel outside the existed first type of the first type cluster region into any one of the existed first type cluster regions till all the pixels are classified. In the other preferred embodiment of the present invention detailed above, a pixel outside all the existed first type cluster regions can be classified into a first type cluster region.

In the following step of the other preferred embodiment of the present invention, the pre-processed image (colorful image before processing) is divided according to the second color coordinate by the steps as follows. Firstly, analyze the second color distribution of pixels of an image, and get a second histogram of the image. The step mentioned above analyze every pixel of the image according to the pixel's value on the second color coordinate, and get a corresponding position in the second coordinate. After all the pixels of the image have been processed by the step mentioned above, calculate the number of pixels on every position of the second color coordinate. Thus the second color histogram of the image is acquired. The following step uses a numerical diffusion method to estimate the second function of the second color histogram of the image. Next, by determining the convex portion of the second function, a plurality of cluster regions of the image based on the second color coordinate are acquired. Thus the image is divided into a plurality of the second type cluster regions. The next step is to divide the total value of the pixels on the second color coordinate by the number of pixels within every individual second type cluster region. The aim of the step mentioned above is to determine the second principle color of every individual second type of cluster region, and then assign the second principle color as the value of the second color coordinate of every pixel within the second type cluster region. So the second color coordinate value of every pixel within a specific second type cluster region is assigned as the second color coordinate value on the second color coordinate. In the foregoing step, the average is performed by the follow description. Assume a second type cluster region includes a number (n1) of first pixels having second color coordinate (q1) and a number (n2) of second pixels having second color coordinate (q2). Then divide the summation of the second color coordinate value of all pixels within a cluster region by the summation of the number of the first pixel and the second pixel, in other words, the result can be expressed as $(q1\times n1+q2\times n2)/(n1+n2)$. The summation of the value of the second color coordinate of all pixels within a second type cluster region can be obtained by multiplying the value of the second color coordinate of the first pixel by the number of the first pixel plus multiplying the value of the second color coordinate of the second pixel by the number of the second pixel. The second principle color of every existed cluster region of the image can be obtained, and the pixels belonging to none of the existed cluster region should be processed with the following step. Classify every pixel outside any one of the existed second type cluster regions into any one of the existed second type cluster regions. The step mentioned above is to classify the pixels outside any existed second type cluster region into any one of the existed second type cluster regions. The method mentioned above is to continuously classify the pixel outside the existed second type of the second type cluster region into any one of the existed second type cluster regions till all the pixels are classified. In the other preferred embodiment of the present invention detailed above, a pixel outside all the existed second type cluster regions can be classified into a second type cluster region.

In the next step of the other preferred embodiment of the present invention, the pre-processed image (colorful image before processing) is divided according to the third color coordinate by the steps as follows. Firstly, analyze the third color distribution of pixels of an image, and get a third histogram of the image. The step mentioned above analyze every pixel of the image according to the pixel's value on the third color coordinate, and get a corresponding position in the third coordinate. After all the pixels of the image have been processed by the step mentioned above, calculate the number of pixels on every position of the third color coordinate. Thus the third color histogram of the image is acquired. The following step uses a numerical diffusion method to estimate the third function of the third color histogram of the image. Next, by determining the convex portion of the third function, a plurality of cluster regions of the image based on the third color coordinate are acquired. Thus the image is divided into a plurality of the third type cluster regions. The next step is to divide the total value of the pixels on the third color coordinate by the number of pixels within every individual third type cluster region. The aim of the step mentioned above is to determine the third principle color of every individual third type of cluster region, and then assign the third principle color as the value of the third color coordinate of every pixel within the third type cluster region. So the third color coordinate value of every pixel within a specific third type cluster region is assigned as the third color coordinate value on the third color coordinate. In the foregoing step, the average is performed by the follow description. Assume a third type cluster region includes a number (n1) of first pixels having third color coordinate (r1) and a number (n2) of third pixels having third color coordinate (r2). Then divide the summation of the third color coordinate value of all pixels within a cluster region by the summation of the number of the first pixel and the second pixel, in other words, the result can be expressed as $(r1\times n1+r2\times n2)/(n1+n2)$. The summation of the value of the third color coordinate of all pixels within a third type cluster region can be obtained by multiplying the value of the third color coordinate of the first pixel by the number of the first pixel plus multiplying the value of the third color coordinate of the second pixel by the number of the second pixel. The third principle color of every existed cluster region of the image can be obtained, and the pixels belonging to none of the existed cluster region should be processed with the following step. Classify every pixel outside any one of the existed third type cluster regions into any one of the existed third type cluster regions. The step mentioned above is to classify the pixels outside any existed third type cluster region into any one of the existed third type cluster regions. The method mentioned above is to continuously classify the pixel outside the existed third type of the third type cluster region into any one of the existed third type cluster regions till all the pixels are classified. In the other preferred embodiment of the present invention detailed above, a pixel outside all the existed third type cluster regions can be classified into a third type cluster region.

After the foregoing three steps utilized to process the first color coordinate, the second color coordinate, and the third color coordinate have been performed, the pre-processed image is divided according to the three color coordinates mentioned above. For example, as shown in FIG. 3, the pre-processed image **40** having processed by the first-color processing step **41** produces the first color distribution **42**. Although the step **41** divide the image according to the first color, different coordinate relates to different color in spite of the same color series. Although the pre-processed image is divided according to the first color series, the cluster regions in the divided image can have different color. In other words, though the image is divided according to a first series, the principle color of each cluster region is different to each other. For example, when the pre-processed image is divided according to the red primary color, there can be a cluster region of heavy red and the other cluster region of light red.

In the foregoing example, assume the result of the division is first color distribution **42** having a first region R1 (of a first principle color) and a second region R2 (of a second principle color R2). In the example of the preferred embodiment of the present invention, assume the pre-processed image **40** is divided according to a second-color processing step **43**, and the result is second color distribution **44** (including region G1 and G2). In addition, assume the pre-processed image **40** is divided according to a second-color processing step **45**, and the result is second color distribution **46** (including region B1 and B2). Though there are only two principle colors in the divided image according to one color series in foregoing example, when it is used in realistic application, the divided image should have a plurality of principle colors.

After the pre-processed image had processed by the steps mentioned above, every pixel of the divided image corresponding to different colors is added. In other words, every pixel corresponding to the first color coordinate (e.g., red color series), the second color coordinate (e.g., green color series), and the third color coordinate (blue color series) is added. Because the three color coordinates are mutual independent, so the color of the added pixel is 3-dimensional such as (R, G, B). In the foregoing operation, the addition of different color coordinates is performed by logic "AND" operation, and the visual result is the mixture of the colors in the three color coordinates.

In the preferred embodiment of the present invention illustrated in FIG. 3, after the pre-processed image has been divided according to the three primary color coordinates, the color distribution in each color coordinate is shown as **42**, **44**, and **46** respectively. Refer to FIG. 4, the three color coordinates of every pixel in the divided image (such as first color distribution **42**, second color distribution **44**, and third color distribution **46**) corresponding to the three prime color coordinates is added by the adding process **48**, then the color of the pixel in the added image is 3-dimensional. So the pixel of the divided image **50** has three color coordinates, and the color of each pixel is the superposition of the three colors.

The divided image **50** is the result of dividing the image according to three-color coordinates, and there are regions of different colors in the divided image **50**, and the regions mentioned above are R2G1B1, R1G1B1, R1G2B1, R1G2B2, R2G2B2, R2G1B2. The color of the pixel within a region is the principle color of the region enclosing the foregoing pixel, such as the color of the pixel of region R2G1B1 is equal to the mixture of the following colors: principle color of region R2, principle color of region G1, and principle color of region B1. In addition, the first function, the second function, and the third function in the preferred embodiment of the present invention can be probability density function, furthermore, can be normal distribution probability density function.

One preferred embodiment of the present invention is used to smooth a three (or one) dimensional histogram to acquire probability density function, whereas, the probability density function that can be used in not confined to the examples illustrated in the embodiment.

As will be understood by persons skilled in the art, the foregoing preferred embodiment of the present invention is illustrative of the present invention rather than limiting the present invention. Having described the invention in connection with a preferred embodiment, as long as the principle used to divide the image is the same as that disclosed herein, the modification will now suggest itself to those skilled in the art. While the preferred embodiment of the invention has been illustrated and described, it will be appreciated that various changes can be made therein without departing from the spirit and scope of the invention.

What is claimed is:

1. A method for dividing an image composed of pixels into a plurality of cluster regions, wherein each pixel is classified into one of said plurality of cluster regions, said method comprising:

acquiring a three dimensional histogram of said image by counting the number of pixels of said image in three dimensional color coordinate;

smoothing said three dimensional histogram by a numerical diffusion method to obtain a function;

defining a plurality of cluster regions on said image, wherein pixels within said plurality of cluster regions correspond to a convex portion of said function;

assigning the color of each pixel of said image as a principle color of a first cluster region among said plurality of cluster regions, said principle color being equal to an average of all pixels within said cluster region, said average being equal to A/B, wherein A is a summation of color of pixels within said first cluster region, and B is equal to the number of said pixels within said first cluster region, the color of said pixel being represented by said three color coordinate, each of said plurality of cluster regions having a corresponding principle color; and

jointing an unclassified pixel to one of said plurality of cluster regions according to minimum distance from said unclassified pixel to said plurality of cluster regions, said unclassified pixel being outside said plurality of cluster regions before said jointing step, after all pixels of said image being jointed to one of said plurality of cluster regions, said image being divided into said plurality of cluster regions.

2. The method as claim **1**, wherein said image is a colorful image.

3. The method as claim **1**, wherein said function is a probability density function.

4. The method according to claim **3**, wherein said probability density function is a normal distribution probability density function.

5. A method for dividing a colorful image composed of pixels into a plurality of cluster regions, wherein each pixel is classified into one of said plurality of cluster regions, said method comprising:

acquiring a three dimensional histogram of said colorful image by counting the number of pixels of said colorful image in three dimensional color coordinate;

smoothing said three dimensional histogram by a numerical diffusion method to obtain a probability density function;

defining a plurality of cluster regions on said image, wherein pixels within said plurality of cluster regions correspond to a convex portion of said probability density;

assigning the color of each pixel of said colorful image as a principle color of a first cluster region among said plurality of cluster regions, said principle color being equal to an average of all pixels within said cluster region, said average being equal to A/B, wherein A is a summation of color of pixels within said first cluster region, and B is equal to the number of pixels within said first cluster region, the color of said pixel being represented by said three color coordinate, each of said plurality of cluster regions having a corresponding principle color; and

jointing an unclassified pixel to one of said plurality of cluster regions according to minimum distance from said unclassified pixel to said plurality of cluster regions, said unclassified pixel being outside said plurality of cluster regions before said jointing step, after all pixels of said colorful image being jointed to one of said plurality of cluster regions, said colorful image being completely divided into said plurality of cluster regions.

6. The method as claim **5**, wherein said probability density function is a normal distribution probability density function.

7. A method for driving an image composed of pixels into a plurality of cluster regions, wherein each pixel of said image is classified into one of said plurality of cluster regions, said method comprising:

acquiring a histogram of said image by counting the number of pixels of said image in a gray level;

smoothing said histogram by a numerical diffusion method to obtain a function;

defining a plurality of cluster regions on said image, wherein pixels within said plurality of cluster regions correspond to a convex portion of said function;

assigning the gray level of each pixel of said image as a principle gray level of a first cluster region among said plurality of cluster regions, said principle gray level being equal to an average of all pixels within said cluster region, said average being equal to A/B, wherein A is a summation of gray level of pixels within said first cluster region, and B is equal to the number of pixels within said first cluster region, the color of said pixel being represented by said three color coordinate, each of said plurality of cluster regions having a corresponding principle gray level; and

jointing an unclassified pixel to one of said plurality of cluster regions according to minimum distance from said unclassified pixel to said plurality of cluster regions, said unclassified pixel being outside said plurality of cluster regions before said jointing step, after all pixels of said image being jointed to one of said plurality of cluster regions, said image being divided into said plurality of cluster regions.

8. The method as claim **7**, wherein said image is a monotone image.

9. The method as claim **7**, wherein said function is a probability density function.

10. The method as claim **9**, wherein said probability density function is a normal distribution probability density function.

11. A method for dividing an image composed of pixels into a plurality of cluster regions, wherein each pixel is classified into one of said plurality of cluster regions, said method comprising:

acquiring a first histogram of said image by counting the number of pixels of said image in a first color coordinate;

smoothing said first histogram by a numerical diffusion method to obtain a first function;

defining a plurality of cluster regions on said image, wherein pixels within said plurality of cluster regions correspond to a convex portion of said first function;

assigning the color of first color coordinate of each pixel of said image as a first principle color of a first cluster region among said plurality of cluster regions, said first principle color being equal to an average of all pixels within said cluster region, said average being equal to A/B, wherein A is a summation of first color coordinate of pixels within said first cluster region, and B is equal to number of said pixels within said first cluster region, the color of said pixel being represented by said three color coordinate, each of said plurality of cluster regions having a corresponding first principle color;

jointing an unclassified pixel to one of said plurality of cluster regions according to minimum distance from said unclassified pixel to said plurality of cluster regions, said unclassified pixel being outside said plurality of cluster regions before said jointing step, after all pixels of said image being jointed to one of said plurality of cluster regions, said image being divided into said plurality of cluster regions;

acquiring a second histogram of said image by counting the number of pixels of said image in a second color coordinate;

smoothing said second histogram by a numerical diffusion method to obtain a second function;

defining a plurality of cluster regions on said image, wherein pixels within said plurality of cluster regions correspond to a convex portion of said second function;

assigning the color of second color coordinate of each pixel of said image as a second principle color of a second cluster region among said plurality of cluster regions, said second principle color being equal to an average of all pixels within said cluster region, said average being equal to A/B, wherein A is a summation of the second color coordinate of pixels within said second cluster region, and B is equal to the number of pixels within said second cluster region, the color of said pixel being represented by said three color coordinate, each of said plurality of cluster regions having a corresponding second principle color;

jointing an unclassified pixel to one of said plurality of cluster regions according to minimum distance from said unclassified pixel to said plurality of cluster regions, said unclassified pixel being outside said plurality of cluster regions before said jointing step, after all pixels of said image being jointed to one of said plurality of cluster regions, said image being divided into said plurality of cluster regions;

acquiring a third histogram of said image by counting the number of pixels of said image in a third color coordinate;

smoothing said third histogram by a numerical diffusion method to obtain a third function;

defining a plurality of cluster regions on said image, wherein pixels within said plurality of cluster regions correspond to a convex portion of said third function;

assigning the color of third color coordinate of each pixel of said image as a third principle color of a third cluster region among said plurality of cluster regions, said third principle color being equal to an average of all pixels within said cluster region, said average being equal to A/B, wherein A is a summation of the third color coordinate of pixels within said third cluster region, B is equal to the number of said pixels within said third cluster region, the color of said pixel being represented by said three color coordinate, each of said plurality of cluster regions having a corresponding third principle color; and

jointing an unclassified pixel to one of said plurality of cluster regions according to minimum distance from said unclassified pixel to said plurality of cluster regions, said unclassified pixel being outside said plurality of cluster regions before said jointing step, after all pixels of said image being jointed to one of said plurality of cluster regions, said image being divided into said plurality of cluster regions; and

transforming every pixel from said first color coordinate, said second color coordinate, and said third color coordinate to three linear independent coordinates respectively to acquire every pixel of said image having a three dimensional coordinate, said image being composed of a first color of said first color region, a second color of said second color region, and a third color of said third color region.

12. The method as claim **11**, wherein said image is a colorful image.

13. The method according to claim **11**, wherein said first function, said second function and said third function are probability density functions.

14. The method according to claim **13**, wherein said probability density functions are normal distribution probability density functions.

* * * * *

US006545743B1

(12) **United States Patent**
Luo et al.

(10) **Patent No.: US 6,545,743 B1**
(45) **Date of Patent: Apr. 8, 2003**

(54) **PRODUCING AN IMAGE OF A PORTION OF A PHOTOGRAPHIC IMAGE ONTO A RECEIVER USING A DIGITAL IMAGE OF THE PHOTOGRAPHIC IMAGE**

(75) Inventors: **Jiebo Luo**, Pittsford, NY (US); **Robert T. Gray**, Rochester, NY (US); **Edward B. Gindele**, Rochester, NY (US)

(73) Assignee: **Eastman Kodak Company**, Rochester, NY (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

(21) Appl. No.: **09/575,784**

(22) Filed: **May 22, 2000**

(51) **Int. Cl.**[7] **G03B 27/00**; G03B 27/52; G03B 27/34; G06K 9/46

(52) **U.S. Cl.** **355/18**; 355/40; 355/56; 382/195

(58) **Field of Search** 355/18, 55, 40, 355/56; 399/45; 358/518; 382/195, 218, 224, 203

(56) **References Cited**

U.S. PATENT DOCUMENTS

4,809,064 A		2/1989	Amos et al.	
5,675,400 A	*	10/1997	Stephenson, III	355/40
5,872,619 A		2/1999	Stephenson et al.	
5,872,643 A		2/1999	Maeda et al.	
5,978,519 A		11/1999	Bollman et al.	
5,995,201 A		11/1999	Sakaguchi	
6,014,461 A	*	1/2000	Hennessey et al.	382/195
6,282,317 B1	*	8/2001	Luo et al.	382/203

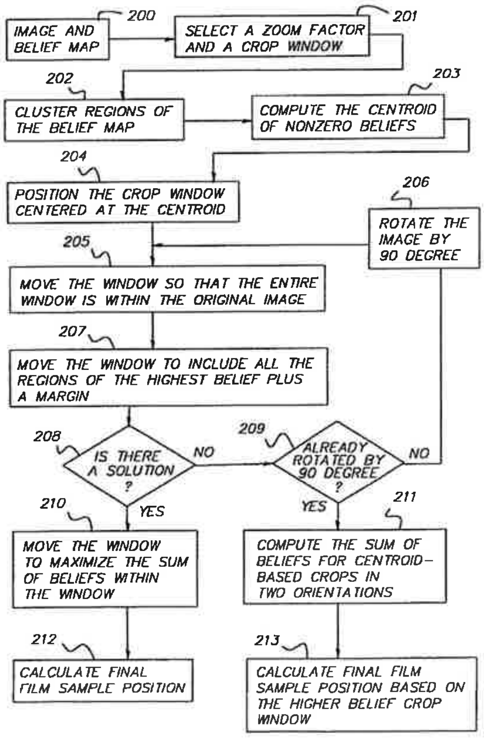
* cited by examiner

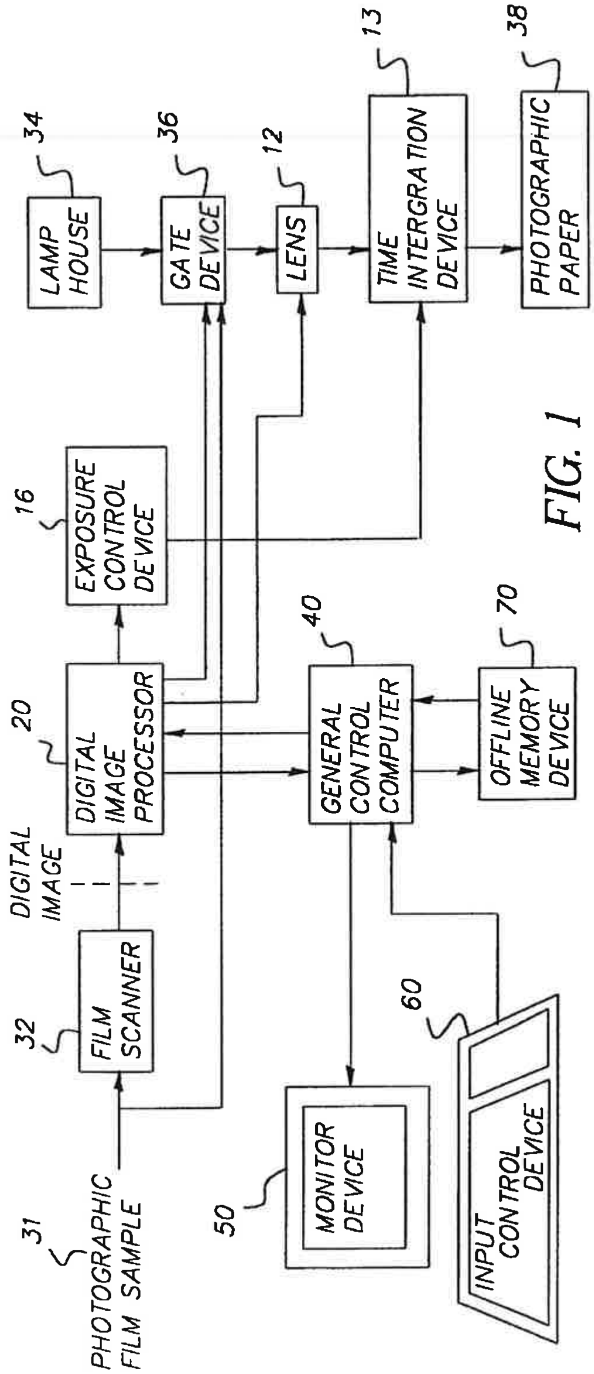
Primary Examiner—Russell Adams
Assistant Examiner—Peter B Kim
(74) *Attorney, Agent, or Firm*—Raymond L. Owens

(57) **ABSTRACT**

A method of producing an image of least a portion of a photographic image onto a photographic receiver including receiving a digital image corresponding to the photographic image, the digital image comprising pixels, and locating the relative optical position of a photographic image, the lens, and the photographic receiver in response to pixels of the digital image and illuminating a portion of the photographic image of high subject content to produce an image of such portion onto the photographic receiver.

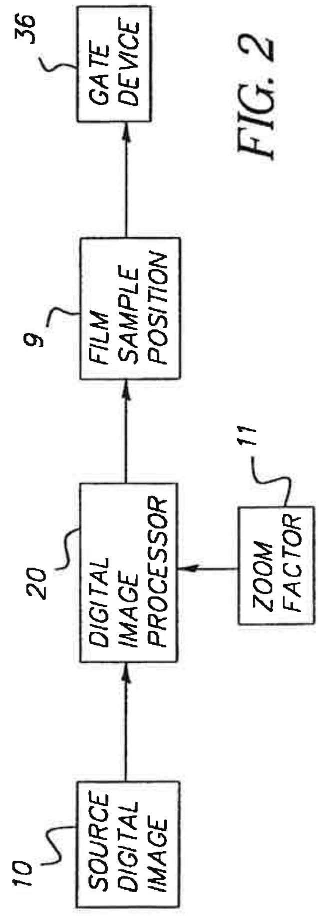
14 Claims, 10 Drawing Sheets





LAMP HOUSE
34
GATE DEVICE
36
LENS
12
TIME INTERGRATION DEVICE
13
PHOTOGRAPHIC PAPER
38
EXPOSURE CONTROL DEVICE
16
DIGITAL IMAGE PROCESSOR
20
GENERAL CONTROL COMPUTER
40
OFFLINE MEMORY DEVICE
70
DIGITAL IMAGE
FILM SCANNER
32
PHOTOGRAPHIC FILM SAMPLE
31
MONITOR DEVICE
50
INPUT CONTROL DEVICE
60

FIG. 1



GATE DEVICE
36
FILM SAMPLE POSITION
9
DIGITAL IMAGE PROCESSOR
20
ZOOM FACTOR
11
SOURCE DIGITAL IMAGE
10

FIG. 2

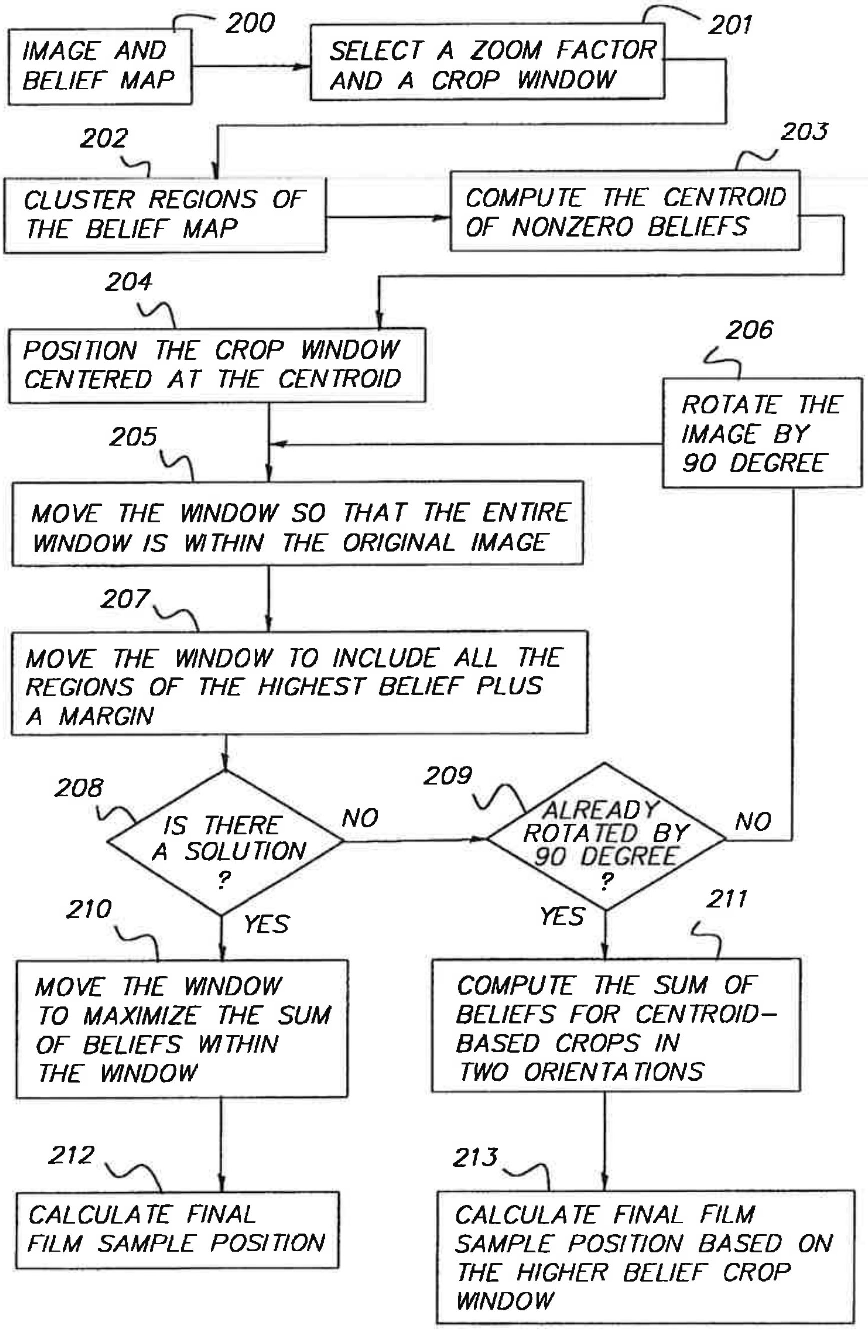

200
IMAGE AND BELIEF MAP
201
SELECT A ZOOM FACTOR AND A CROP WINDOW
202
CLUSTER REGIONS OF THE BELIEF MAP
203
COMPUTE THE CENTROID OF NONZERO BELIEFS
204
POSITION THE CROP WINDOW CENTERED AT THE CENTROID
206
ROTATE THE IMAGE BY 90 DEGREE
205
MOVE THE WINDOW SO THAT THE ENTIRE WINDOW IS WITHIN THE ORIGINAL IMAGE
207
MOVE THE WINDOW TO INCLUDE ALL THE REGIONS OF THE HIGHEST BELIEF PLUS A MARGIN
208
IS THERE A SOLUTION ?
NO
209
ALREADY ROTATED BY 90 DEGREE ?
NO
210
YES
YES
211
MOVE THE WINDOW TO MAXIMIZE THE SUM OF BELIEFS WITHIN THE WINDOW
COMPUTE THE SUM OF BELIEFS FOR CENTROID-BASED CROPS IN TWO ORIENTATIONS
212
CALCULATE FINAL FILM SAMPLE POSITION
213
CALCULATE FINAL FILM SAMPLE POSITION BASED ON THE HIGHER BELIEF CROP WINDOW

FIG. 3

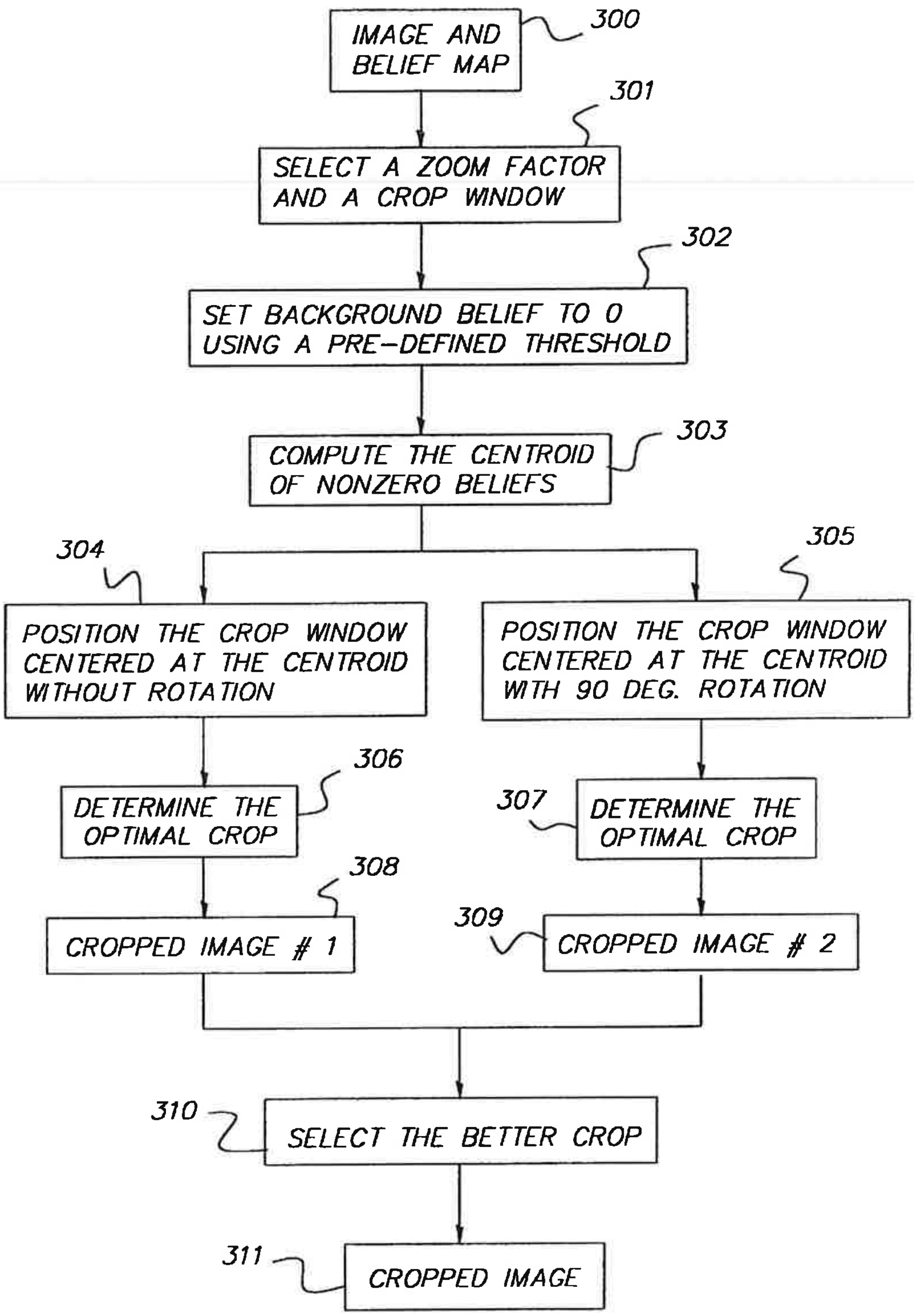


IMAGE AND BELIEF MAP
300
SELECT A ZOOM FACTOR AND A CROP WINDOW
301
SET BACKGROUND BELIEF TO 0 USING A PRE-DEFINED THRESHOLD
302
COMPUTE THE CENTROID OF NONZERO BELIEFS
303
POSITION THE CROP WINDOW CENTERED AT THE CENTROID WITHOUT ROTATION
304
POSITION THE CROP WINDOW CENTERED AT THE CENTROID WITH 90 DEG. ROTATION
305
DETERMINE THE OPTIMAL CROP
306
DETERMINE THE OPTIMAL CROP
307
CROPPED IMAGE # 1
308
CROPPED IMAGE # 2
309
SELECT THE BETTER CROP
310
CROPPED IMAGE
311

FIG. 4

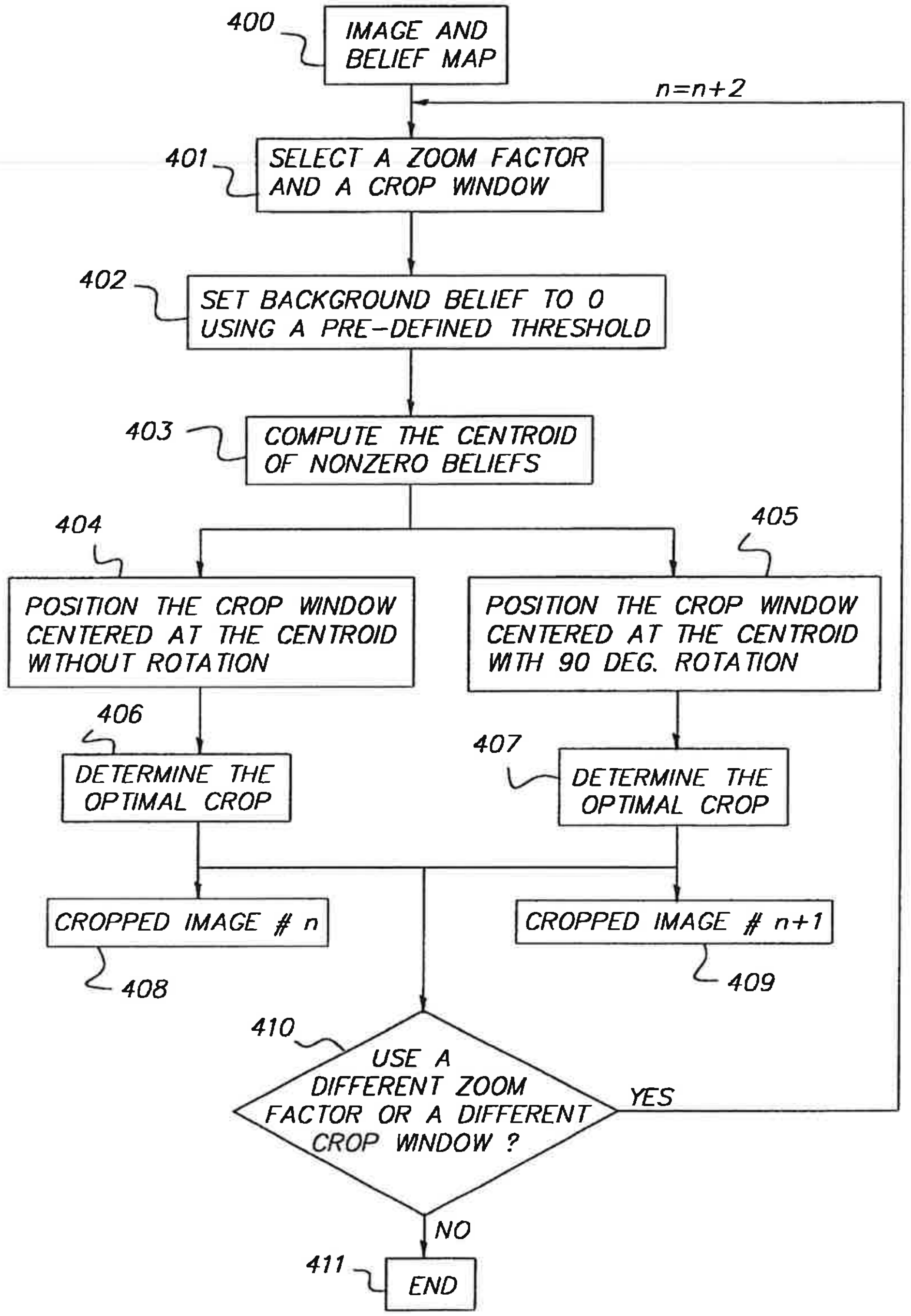

400
IMAGE AND BELIEF MAP
n=n+2
401
SELECT A ZOOM FACTOR AND A CROP WINDOW
402
SET BACKGROUND BELIEF TO 0 USING A PRE-DEFINED THRESHOLD
403
COMPUTE THE CENTROID OF NONZERO BELIEFS
404
POSITION THE CROP WINDOW CENTERED AT THE CENTROID WITHOUT ROTATION
405
POSITION THE CROP WINDOW CENTERED AT THE CENTROID WITH 90 DEG. ROTATION
406
DETERMINE THE OPTIMAL CROP
407
DETERMINE THE OPTIMAL CROP
CROPPED IMAGE # n
408
CROPPED IMAGE # n+1
409
410
USE A DIFFERENT ZOOM FACTOR OR A DIFFERENT CROP WINDOW ?
YES
NO
411
END

FIG. 5

Fig. 6

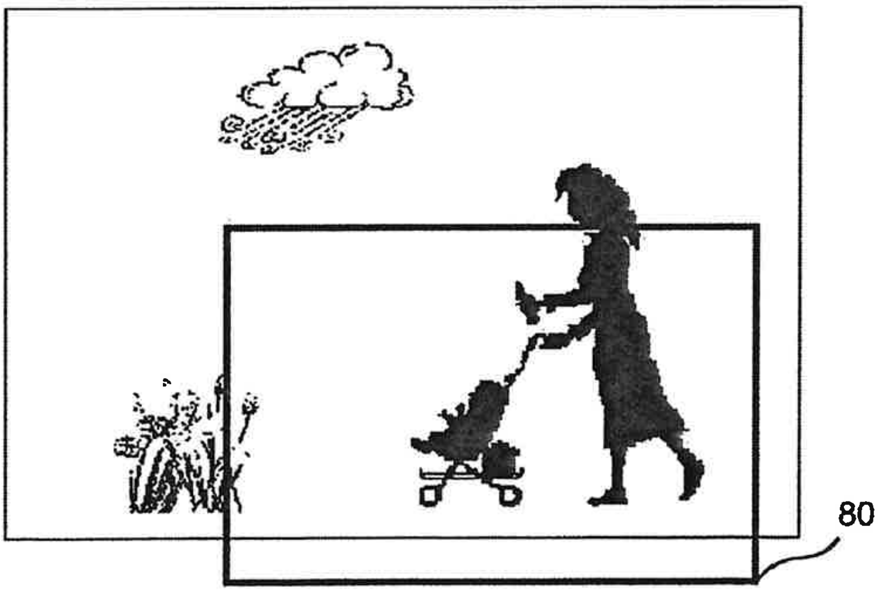


Fig. 7

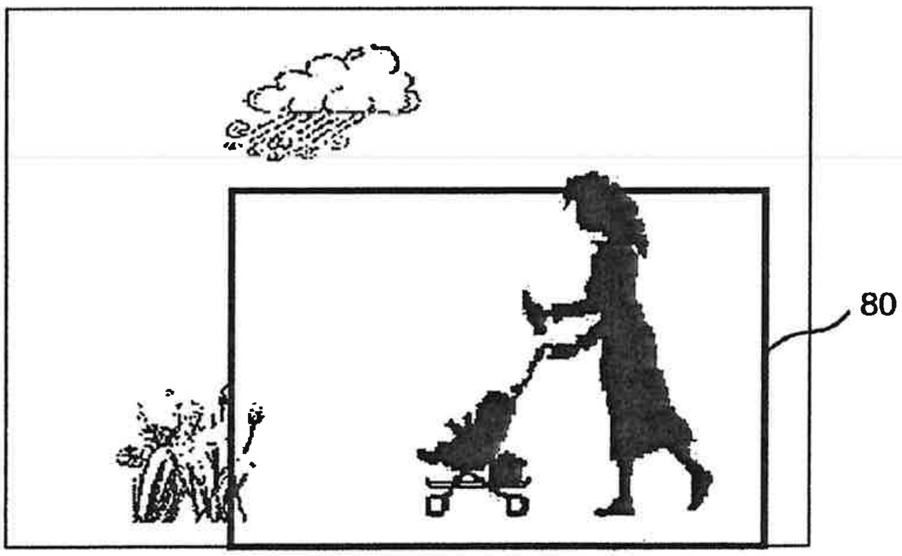


Fig. 8

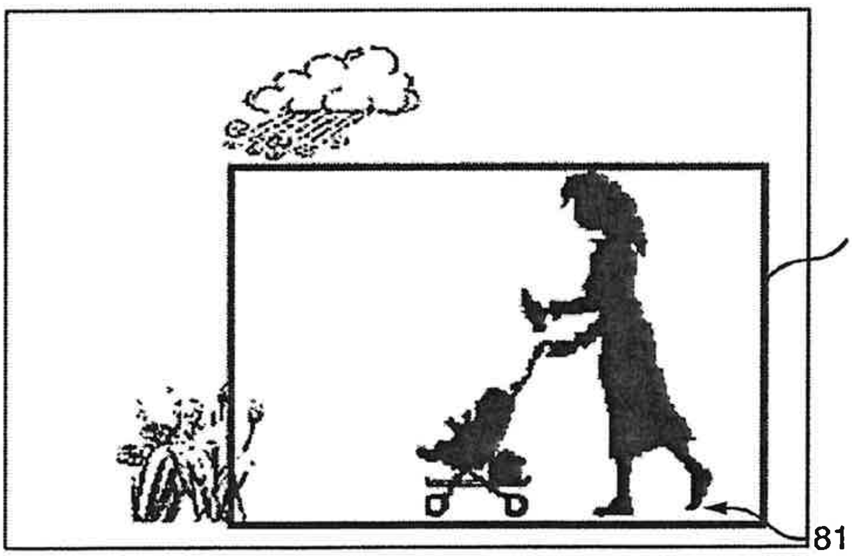


Fig. 9

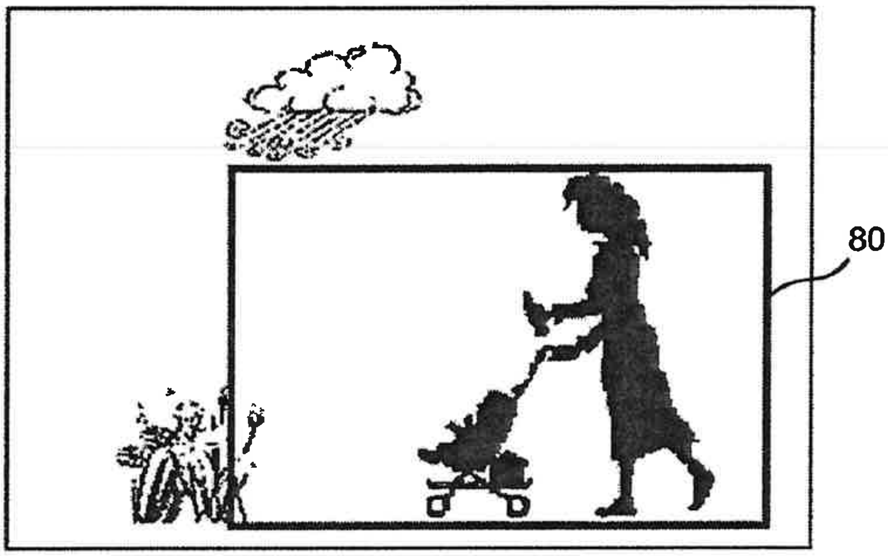


Fig. 10

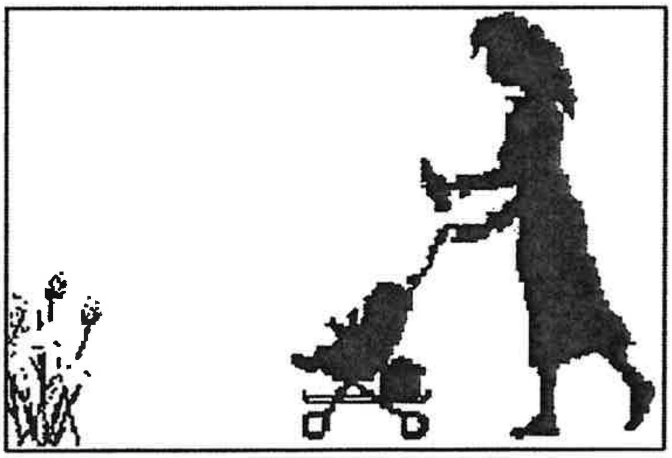


Fig. 11

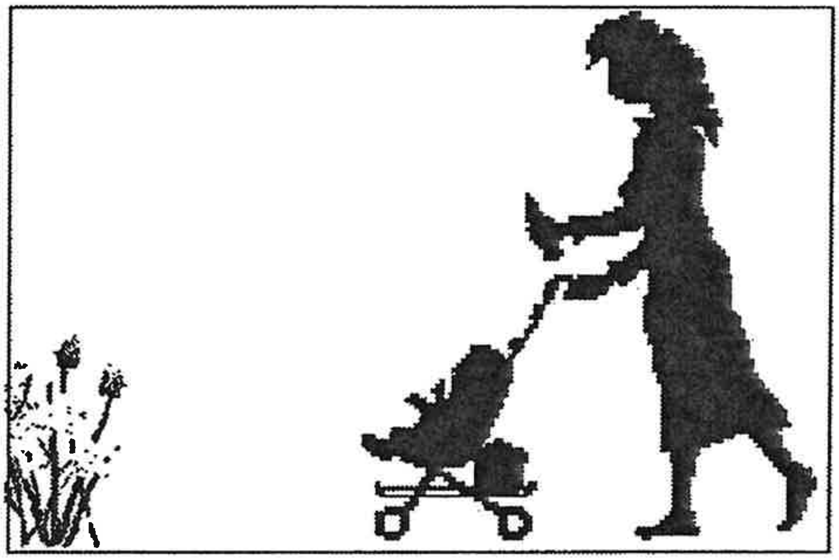


Fig. 12

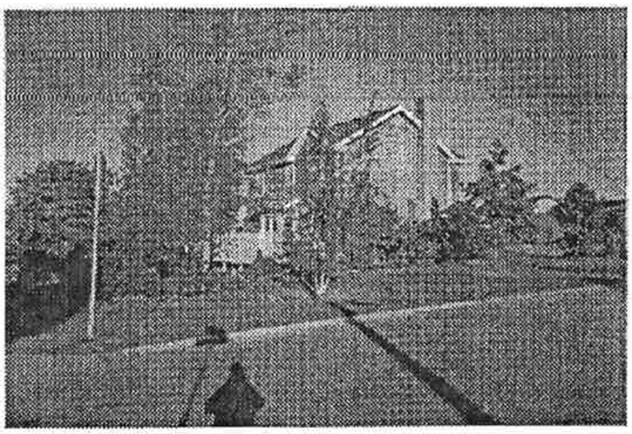


Fig. 13

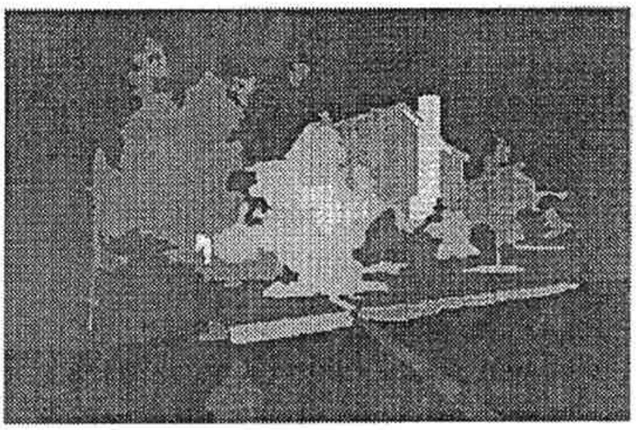


Fig. 14

Fig 15

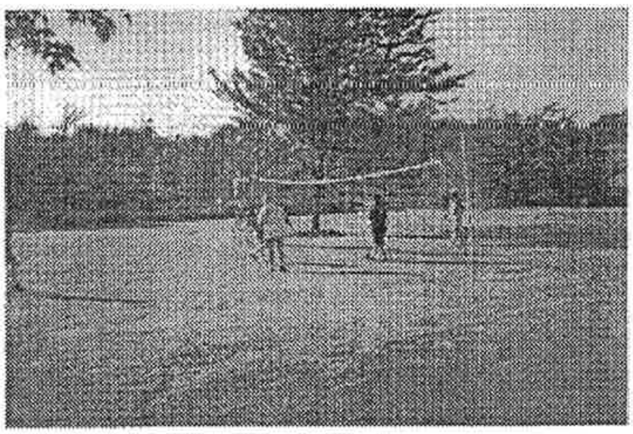


Fig. 16

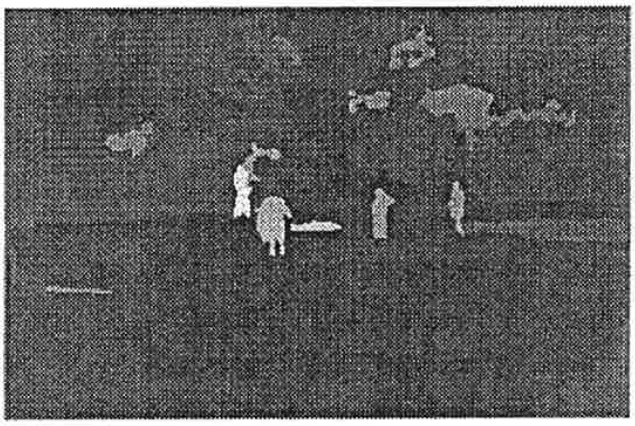


Fig. 17

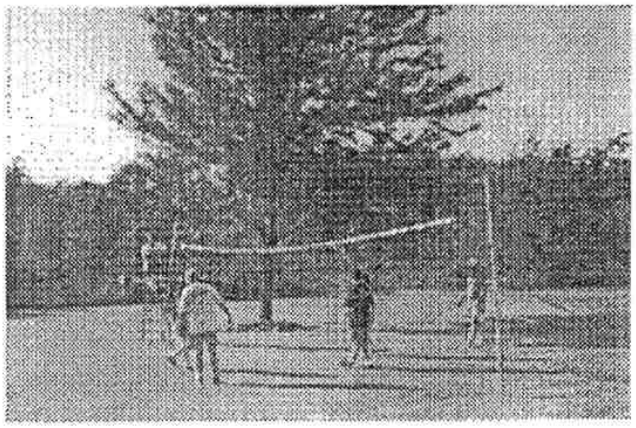


Fig 18

PRODUCING AN IMAGE OF A PORTION OF A PHOTOGRAPHIC IMAGE ONTO A RECEIVER USING A DIGITAL IMAGE OF THE PHOTOGRAPHIC IMAGE

CROSS REFERENCE TO RELATED APPLICATIONS

Reference is made to commonly assigned U.S. patent application Ser. No. 09/490,915 filed Jan. 25, 2000, filed Jan. 25, 2000 entitled "Method for Automatically Creating Cropped and Zoomed Versions of Photographic Images" by Jiebo Luo et al., and assigned U.S. patent application Ser. No. 09/223,860, filed Dec. 31, 1998 entitled "Method for Automatic Determination of Main Subjects in Photographic Images" by Jiebo Luo et al., the disclosures of which are incorporated herein by reference.

FIELD OF THE INVENTION

The invention relates to producing an image of a portion of a photographic image onto a receiver by using digital image processing.

BACKGROUND OF THE INVENTION

For many decades, traditional commercial photofinishing systems have placed limits on the features offered to consumers to promote mass production. Among those features that are unavailable conventionally, zooming and cropping have been identified by both consumers and photofinishers as extremely useful additional features that could potentially improve the quality of the finished photographs and the subsequent picture sharing experiences. With the advent of, and rapid advances in digital imaging, many of the technical barriers that existed in traditional photography no longer stand insurmountable.

Hybrid and digital photography provide the ability to crop undesirable content from a picture, and magnify or zoom the desired content to fill the entire photographic print. In spite of the fact that some traditional cameras with zoom capability provide consumers greater control over composing the desired scene content, studies have found that photographers may still wish to perform a certain amount of cropping and zooming when viewing the finished photograph at a later time. Imprecise viewfinders of many point-and-shoot cameras, as well as simply second-guessing their initial compositions, are factors in the desirability of zoom and crop. In addition, it maybe desirable to use some other regular border templates such as ovals, heart shapes, squares, etc. In another scenario, some people commonly referred to as "scrapbookers" tend to perform more aggressive crop in making a scrapbook, e.g., cutting along the boundary of objects.

There are significant differences in objectives and behaviors between these two types of cropping, namely album-making and scrapbook making, with the latter more difficult to understand and summarize. The invention described below performs automatic zooming and cropping for making photographic prints. One customer focus group study indicated that it would be beneficial to provide customers a double set of prints—one regular and one zoom. Moreover, it is preferred that the cropping and zooming be done automatically. Most customers do not want to think about how the zooming and cropping is being done as long as the content and quality (e.g., sharpness) of the cropped and zoomed pictures is acceptable.

There has been little research on automatic zoom and crop due to the apparent difficulty involved in performing such a

task. None of the known conventional image manipulation software uses scene content in determining the automatic crop amount. For example, a program entitled "XV", a freeware package developed by John Bradley at University of Pennsylvania, USA (Department of Computer and Information Science), provides an "autocrop" function for manipulating images and operates in the following way:

1. The program examines a border line of an image, in all of the four directions, namely from the top, bottom, left and right sides;
2. The program checks the variation within the line. In grayscale images, a line has to be uniform to be cropped. In color images, both the spatial correlation and spectral correlation have to be low, except for a small percentage of pixels, for the line to be qualified for cropping. In other words, a line will not be cropped if it contains a significant amount of variation;
3. If a line along one dimension passes the criterion, the next line (row or column) inward is then examined; and
4. The final cropped image is determined when the above recursive process stops.

This program essentially tries to remove relatively homogeneous margins around the borders of an image. It does not examine the overall content of the image. In practice, the XV program is effective in cropping out the dark border generated due to imprecise alignment during the scanning process. However, disastrous results can often be produced due to the apparent lack of scene understanding. In some extreme cases,-the entire image can be cropped.

Another conventional system, described by Bollman et al. in U.S. Pat. No. 5,978,519 provides a method for cropping images based upon the different intensity levels within the image. With this system, an image to be cropped is scaled down to a grid and divided into non-overlapping blocks. The mean and variance of intensity levels are calculated for each block. Based on the distribution of variances in the blocks, a threshold is selected for the variance. All blocks with a variance higher than the threshold variance are selected as regions of interest. The regions of interest are then cropped to a bounding rectangle. However, such a system is only effective when uncropped images contain regions where intensity levels are uniform and other regions where intensity levels vary considerably. The effectiveness of such a system is expected to be comparable to that of the XV program. The difference is that the XV program examines the image in a line by line fashion to identify uniform areas, while Bollman examines the image in a block by block fashion to identify uniform areas. In summary, both techniques cannot deal with images with non-uniform background.

Some optical printing systems have the capability of changing the optical magnification of the relay lens used in the photographic copying process. In U.S. Pat. No. 5,995,201, Sakaguchi describes a method of varying the effective magnification of prints made from film originals utilizing a fixed optical lens instead of zoom lens. In U.S. Pat. No. 5,872,619, Stephenson et al. describe a method of printing photographs from a processed photographic filmstrip having images of different widths measured longitudinally of the filmstrip and having heights measured transversely of the filmstrip. This method uses a photographic printer having a zoom lens and a printing mask to provide printed images having a selected print width and a selected print height. In U.S. Pat. No. 4,809,064, Amos et al. describe an apparatus for printing a selected region of a photographic negative onto a photosensitive paper to form an enlarged and cropped photographic print. This apparatus includes means for projecting the photographic negative onto first and second zoom lenses, each of the zoom lenses having an adjustable magnification. In U.S. Pat. No. 5,872,643, Maeda et al. describe a film reproducing apparatus that can effectively perform zoom and crop. This apparatus includes an image pick-up device which picks up a film frame image recorded on a film to generate image data, an information reader which reads information about photographing conditions of the film frame image, and a reproducing area designator which designates a reproducing area of the film frame image. However, the reproducing area of the film frame image is determined based on pre-recorded information about the position of the main object, as indicated by which zone of the photograph the automatic focusing (AF) operation in the camera was on—part of the recorded information about photographing conditions. In all the above mentioned optical printing systems, the position of the photographic film sample and magnification factor of the relay lens are preselected.

SUMMARY OF THE INVENTION

It is an object of the present invention to provide a method for producing a portion of an image from a photographic image by identifying the main subject of the photographic image.

This object is achieved by a method of producing an image of at least a portion of a photographic image onto a photographic receiver, comprising the steps of:

a) receiving a digital image corresponding to the photographic image, the digital image comprising pixels; and

b) locating the relative optical position of a photographic image, the lens, and the photographic receiver in response to pixels of the digital image and illuminating a portion of the photographic image of high subject content to produce an image of such portion onto the photographic receiver.

ADVANTAGES

One advantage of the invention lies in the ability to automatically crop and zoom photographic images based upon the scene contents. The digital image processing steps employed by the present invention includes a step of identifying the main subject within the digital image. The present invention uses the identified main subject of the digital image to automatically zoom and crop the image. Therefore, the present invention produces high-quality zoomed or cropped images automatically, regardless whether the background is uniform or not.

BRIEF DESCRIPTION OF THE DRAWINGS

The foregoing and other objects, aspects and advantages will be better understood from the following detailed description of a preferred embodiment of the invention with reference to the drawings, in which:

FIG. 1 is a schematic diagram of a system embodiment of the invention;

FIG. 2 is a schematic architectural diagram of an embodiment of the invention;

FIG. 3 is a schematic architectural diagram of an embodiment of the invention;

FIG. 4 is a schematic architectural diagram of an embodiment of the invention;

FIG. 5 is a schematic architectural diagram of an embodiment of the invention;

FIG. 6 illustrates the application of the invention to a simulated photograph;

FIG. 7 illustrates the application of the invention to a simulated photograph;

FIG. 8 illustrates the application of the invention to a simulated photograph;

FIG. 9 illustrates the application of the invention to a simulated photograph;

FIG. 10 illustrates the application of the invention to a simulated photograph;

FIG. 11 illustrates the application of the invention to a simulated photograph;

FIG. 12 illustrates the application of the invention to a simulated photograph;

FIG. 13 is an exemplary uncropped photograph;

FIG. 14 is a belief map of the image shown in FIG. 13;

FIG. 15 is a cropped version of the image shown in FIG. 13;

FIG. 16 depicts another uncropped photograph;

FIG. 17 is a belief map of the image shown in FIG. 16; and

FIG. 18 is a cropped version of the image shown in FIG. 16.

DETAILED DESCRIPTION OF THE INVENTION

The invention automatically zooms and crops digital images according to an analysis of the main subject in the scene. Previously, a system for detecting main subjects (e.g., main subject detection or "MSD") in a consumer-type photographic image from the perspective of a third-party observer has been developed and is described in U.S. patent application Ser. No. 09/223,860, filed Dec. 31, 1998, the disclosure of which is incorporated herein by reference. Main subject detection provides a measure of saliency or relative importance for different regions that are associated with different subjects in an image. Main subject detection enables a discriminative treatment of the scene content for a number of applications related to consumer photographic images, including automatic crop and zoom.

Conventional wisdom in the field of computer vision, which reflects how a human observer would perform such tasks as main subject detection and cropping, calls for a problem-solving path via object recognition and scene content determination according to the semantic meaning of recognized objects. However, generic object recognition remains a largely unsolved problem despite decades of effort from academia and industry.

The MSD system is built upon mostly low-level vision features with semantic information integrated whenever available. This MSD system has a number of sub-tasks, including region segmentation, perceptual grouping, feature extraction, and probabilistic and semantic reasoning. In particular, a large number of features are extracted for each segmented region in the image to represent a wide variety of visual saliency properties, which are then input into a tunable, extensible probability network to generate a belief map containing a continuum of values.

Using MSD, regions that belong to the main subject are generally differentiated from the background clutter in the image. Thus, automatic zoom and crop becomes possible. Automatic zoom and crop is a nontrivial operation that was considered impossible for unconstrained images, which do not necessarily contain uniform background, without a certain amount of scene understanding. In the absence of content-driven cropping, conventional systems have concentrated on simply using a centered crop at a fixed zoom (magnification) factor, or removing the uniform background touching the image borders. The centered crop has been found unappealing to customers.

The output of MSD used by the invention is a list of segmented regions ranked in descending order of their likelihood (or belief) as potential main subjects for a generic or specific application. This list can be readily converted into a map in which the brightness of a region is proportional to the main subject belief of the region. Therefore, this map can be called a main subject "belief" map. This "belief" map is more than a binary map that only indicates location of the determined main subject. The associated likelihood is also attached to each region so that regions with large values correspond to regions with high confidence or belief of being part of the main subject.

To some extent, this belief map reflects the inherent uncertainty for humans to perform such a task as MSD because different observers may disagree on certain subject matter while agreeing on other subject matter in terms of main subjects. However, a binary decision, when desired, can be readily obtained by using an appropriate threshold on the belief map. Moreover, the belief information may be very useful for downstream applications. For example, different weighting factors can be assigned to different regions (subject matters) in determining the amount of crop.

For determination of crop, the invention uses the main subject belief map instead of a binarized version of the map to avoid making a bad cropping decision that is irreversible. Furthermore, using the continuous values of the main subject beliefs helps trade-off different regions under the constraints encountered in cropping. A binary decision on what to include and what not to include, once made, leaves little room for trade-off. For example, if the main subject region is smaller than the crop window, the only reasonable choice, given a binary main subject map, is to leave equal amounts of margin around the main subject region. On the other hand, secondary main subjects are indicated by lower belief values in the main subject belief map, and can be included according to a descending order of belief values once the main subject of highest belief values are included. Moreover, if an undesirable binary decision on what to include/exclude is made, there is no recourse to correct the mistake. Consequently, the cropping result becomes sensitive to the threshold used to obtain the binary decision. With a continuous-valued main subject belief map, every region or object is associated with a likelihood of being included or a belief value in its being included.

To reduce the degrees of freedom in determining the amount of crop, in particular for making photographic prints, in one embodiment, the invention restricts the set of allowable zoom factors to {1.5×, 2×, 3×, 4×}. This is based on the findings in the customer focus studies. Those skilled in the art would recognize that the present invention could be used with any the zoom factor. In one example, the default zoom factor is set at 1.5 ×.

"Digital image understanding technology" as used herein means technology which is specifically designed to exclusively perform on a digital image representing at least a portion of an original scene captured by an image capture device other than a copier or scanner, the functions of:

(i) digitally processing such digital image (other than by color and/or luminance, i.e., non-spatial analysis of color) to recognize and thereby assign useful meaning to:

a human understandable object other than (a) a human-created graphic object, including without limitation, a diagram, chart, alphanumeric character, mark, symbol, pattern and any combination of the foregoing, (b) the entire original scene, or (c) any part of such original scene completely designated by a human in such original scene;

an object attribute (including without limitation, characteristics, movement or direction, of such human understandable object) in such original scene (other than for using such object attribute to control the functions, including image signal processing functions, of the image capture device); or

a scene condition (other than for using such scene condition to control the functions, including image signal processing functions, of the image capture device) of such original scene; and

(ii) based on the results of step (i) and in accordance with such useful meaning, modifying, transforming, compressing, expanding, enhancing, editing, classifying, formatting, transmitting, or storing the digital image representing at least a portion of the original scene, or causing an operation other than image processing to be performed.

General Description of an Optical Printer System Application

Referring to FIG. 1, the following description relates to an optical printing system. A photographic film sample **31** is received by a film scanner **32** which produces a source digital image **10** relating to the spatial density distribution of the photographic film sample. This source digital image is received by a digital image processor **20**. The digital image processor **20** may be connected to a general control computer **40** under operator control from an input control device **60**. The monitor device **50** displays diagnostic information about the optical printing system. The general control computer **40** keeps track of the lens magnification setting.

Referring to FIG. 2, a zoom factor **11**, which corresponds to the lens magnification setting may also received by the image processor **20** from the general control computer **40** under operator control. The image processor **20** receives the source digital image **10** and uses the zoom factor **11** and the source digital image **10** to calculate the proper position for the photographic film sample in the form of a film sample position **9**. The photographic film sample is positioned in a gate device **36** which holds the film negative in place during the exposure. The gate device **36** receives the film sample position **9** to position the photographic film sample to adjust which portion of the imaging area of the photograph will be printed.

Referring to FIG. 1, a lamp house **34** provides the illumination source which is transmitted through the photographic film sample **31** and focused by a lens **12** onto photographic paper **38**. The time integration device **13** opens and closes a shutter for a variable length of time allowing the focused light from the lamp house **34** to expose the photographic paper **38**. The exposure control device **16** receives a brightness balance value from the digital image processor **20**. The exposure control device **16** uses the brightness balance value to regulate the length of time the shutter of the time integration device stays open.

A block diagram of the inventive cropping process (e.g., the digital image understanding technology) is shown in FIG. 3, which is discussed in relation to FIGS. **6**–**12**. FIGS. **6**–**12** illustrate the inventive process being applied to an original image shown in FIG. **6**.

In item 200, the belief map is created using MSD. The invention selects a zoom factor (e.g. 1.5×) and a crop window 80 (as shown in FIG. 7), as referred to in item 201 of FIG. 3. This zoom factor may be selected by an operator, or by an automatic method based directly on the main subject belief map (e.g., an estimate of the size of the main subject). The crop window is typically a rectangular window with a certain aspect ratio. If the zoom factor is selected by the operator, the value of the zoom factor is sent from the general control computer 40 to the digital image processor 20 shown in FIG. 1. For both the operator selected case or the main subject belief map case, the zoom factor is used to communicate with the lens 12 to adjust the lens magnification setting. This adjustment allows the lens 12 to image the appropriate size of the photographic film sample 31 onto the photographic paper 38.

In item 202, regions of the belief map are clustered and the lowest belief cluster (e.g., the background belief) is set to zero using a predefined threshold. As discussed in greater detail below, sections of the image having a belief value below a certain threshold are considered background sections. In item 202 such sections are given a belief of zero for purposes of this embodiment of the invention.

Then, in item 203 the centroid, or center-of-mass, of nonzero beliefs are computed and, as shown in item 204, the crop window position is centered at the centroid, as shown in FIG. 7. More specifically, in FIG. 6 the subject having the highest belief in the belief map is the woman and the stroller. FIG. 7 illustrates that the centroid of this subject is approximately the top of the baby's head.

The centroid ($\hat{x}$, $\hat{y}$) of a belief map is calculated using the following procedure:

$$\hat{x} = \sum_i x_i bel(x_i, y_i), \quad \hat{y} = \sum_i y_i bel(x_i, y_i),$$

where x_i and y_i denote that coordinates of a pixel in the belief map and bel(x_i, y_i,) represents the belief value at this pixel location.

The crop window is then moved so that the entire crop window is within the original image (e.g. item 205) as shown in FIG. 8. In item 207, the crop window 80 is moved again so that all the regions of the highest belief values ("main subject") are included within the crop window and to create a margin 81, as shown in FIG. 9. This process (e.g., 207) captures the entire subject of interest. Therefore, as shown in FIG. 9, the top of the woman's head is included in the crop window. Compare this to FIG. 8 where the top of the woman's head was outside the crop window.

Decision box 208 determines whether an acceptable solution has been found, i.e., whether it is possible to include at least the regions of the highest belief values in the crop window.

If an acceptable solution exists, the window is again moved, as shown in item 210, to optimize a subject content index for the crop window. The preferred embodiment of the present invention defines the subject content index as the sum of belief values within the crop window. It should be noted that the present invention specifies higher numerical belief values corresponding to higher main subject probability. Therefore, finding a numerical maximum of the sum of the belief values is equivalent to finding an optimum of the subject content index. This is shown in FIG. 10 where the secondary objects (e.g. flowers) are included within the crop window 80 to increase the sum of beliefs. The sum of beliefs for a crop window is computed as follows.

$$\text{sum}(w) = \sum_{(x,y) \in w} bel(x, y),$$

where bel(x, y) represents the belief value at a given pixel location (x, y) within the crop window w.

Provided that the primary subjects are included, moving the crop window so that more of the secondary subjects are included would increase the sum of belief values within the crop window. Recall that the primary subjects are indicated by the highest belief values and the secondary subjects are indicated by belief values lower than those of the primary subjects but higher than those of the background subjects. The goal is to find the crop window that has the highest sum of belief values while ensuring that the primary subjects are completely included in the crop window, i.e.,

$$\tilde{w} = \max_{w \in W} \text{sum}(w),$$

where W denotes the set of all possible crop windows that satisfy all the aforementioned constraints (e.g., those that are completely within the uncropped image and those that encompass the entire primary subjects).

Then, in item 212, the position of the center of the crop window is used to calculate the translational component of the film sample position. The film sample position is then calculated in item of block 213. The gate device 36, shown in FIG. 1, receives the film sample position 9 and uses this information to control the position of the photographic film sample 31 relative to the lens 12. Those skilled in the art will recognize that either or both of the lens 12 and the photographic film sample 31 may be moved to achieve the centering of the effective cropped image region on the photographic paper 38.

Referring to FIG. 3, if decision box 208 does not produce an acceptable solution, the digital image is rotated by 90 degrees, as shown in item 206, and the processing returns to item 205. At the second and greater pass through decision box 209, if the crop window 80 has already been rotated by 90 degrees and there is still not an acceptable solution (from decision box 208), then the sum of the beliefs for the centroid-based crop window 80 are computed for both orientations (e.g., original and 90 degrees), as shown in item 211. Then, the position of the center of the crop window is used to calculate the translational component of the film sample position 9 and the rotational component of the film sample position 9 is set to 90 degrees. The gate device 36, shown in FIG. 1, receives the film sample position 9 and uses this information to control the position of the photographic film sample 31 relative to the lens 12. The gate device 36 then rotates the photographic film sample 31 by 90 degrees.

As would be known by one ordinarily skilled in the art given this disclosure, the photographic film sample may be rotated at an arbitrary angles (e.g. 45°,180°, etc.) in item 206. Further, the decision box 209 can allow multiple passes of the process described in items 205, 207 at different rotation angles in an attempt to find the orientation having the highest sum of beliefs.

The simulated image example shown in FIGS. 6–12 illustrates the progress the invention makes as it moves through the process shown in FIG. 3. One could formulate the problem as a global exhaustive search for the best solution. The procedure used in the invention is considered a "greedy" searching approach and is certainly more efficient than conventional processes.

FIGS. 4 and 5 illustrate variations of the process shown in FIG. 3 of determining the film sample position 9. FIG. 4 compares a landscape crop with a portrait crop, regardless of the initial (not necessarily the correct) picture orientation. The second variation in FIG. 5 creates a series of "ring-around" cropped versions with a zoom factor series (e.g., {1×, 1.5×, 2×, 3×, and 4×}) for presentation to a human operator.

More specifically, in FIG. 4 items **300**–**303** are substantially identical to items **200**–**203** in FIG. 3. In item **304**, the crop window **80** is positioned centered at the centroid without any rotation. Simultaneously, in a parallel processing step, the position of the crop window **80** is centered at the centroid with a rotation (e.g., 90 degree rotation). In items **306** and **307** the optimal crop window **80** is determined, as discussed above with respect to FIG. 3. This produces cropped images **308** and **309**. In item **310** the better (e.g., higher sum of beliefs) of the two cropped images is selected and the corresponding translational and rotational components of the film sample position **9** are recorded.

Referring to FIG. 5, items **400**–**407** are substantially identical to items **300**–**307** discussed above. However, in the process shown in FIG. 5, decision box **410** observes the cropped images **408**, **409** produced by the parallel processing in items **404**–**407** and determines whether an acceptable level has been attained. If the processing in item **410** determines that a different zoom factor or a different crop window **80** should be evaluated, processing returns to item **401**. This process is repeated until an acceptable result is produced, ending in item **411**. In similar fashion, the corresponding translational and rotational components of the film sample position **9** are recorded.

The invention utilizes a built-in "k-means" clustering process to determine proper thresholds of MSD beliefs for each application. The invention also uses clustering, as discussed below to enhance the cropping process. In one preferred embodiment, it is sufficient to use three levels to quantize MSD beliefs, namely "high", "medium", and "low." As would be known by one ordinarily skilled in the art, the invention is not limited to simply three levels of classification, but instead can utilize a reasonable number of classification levels to reduce the (unnecessary) variation in the belief map. These three levels allow for the main subject (high), the background (low), and an intermediate level (medium) to capture secondary subjects, or uncertainty, or salient regions of background. Therefore, the invention can perform a k-means clustering with k=3 on the MSD belief map to "quantize" the beliefs. Consequently, the belief for each region is replaced by the mean belief of the cluster in that region. Note that a k-means clustering with k=2 essentially produces a binary map with two clusters, "high" and "low," which is undesirable for cropping based on earlier discussion.

There are two major advantages in performing such clustering or quantization. First, clustering helps background separation by grouping low-belief background regions together to form a uniformly low-belief (e.g., zero belief) background region. Second, clustering helps remove noise in belief ordering by grouping similar belief levels together. The centroiding operation does not need such quantization (nor should it be affected by the quantization). The main purpose of the quantization used here is to provide a threshold for the background.

The k-means clustering effectively performs a multi-level thresholding operation to the belief map. After clustering, two thresholds can be determined as follows:

$$threshold_{low}=(C_{low}+C_{medium})/2,\ threshold_{high}=(C_{medium}+C_{high})/2$$

where $\{C_{low}, C_{medum}, C_{high}\}$ is the set of centroids (average belief values) for the three clusters, and $threshold_{low}$ and $threshold_{high}$ are the low and high thresholds, respectively.

Regions with belief values below the lower threshold are considered "background" and their belief values are set to zero in items **202**, **302** and **402** discussed above. Regions with belief values above the higher threshold are considered part of the main subject and need to be included in their entirety, whenever possible. Regions with intermediate belief values (e.g., less than or equal to the higher threshold and greater than or equal to the lower threshold) are considered part of the "secondary subject" and will be included as a whole or partially, if possible, to maximize the sum of main subject belief values retained by the crop window. Note that the variance statistics on the three clusters can be used to set the thresholds more accurately to reflect cluster dispersions.

The invention initializes the k-means process by finding the maximum value $bel_{maximum}$ and minimum values $bel_{minimum}$ of the belief map, computing the average value $bel_{average}$ of the maximum and minimum values for item in the belief map, and setting the initial centroids (denoted by a superscript of 0) at these three values, i.e.,

$$C_{low}^0=bel_{minimium},\ C_{medium}^0=bel_{mediumm},\ C_{high}^0=bel_{maximum}$$

Other ways of initialization may apply. For more about the k-means process, see *Sonka, Hlavac, and Boyle, Image Procesing Analysis, and Machine Vision, PWS Publishing*, 1999 pagse 307–308. For typical MSD belief maps, the k-means process usually converges in fewer than 10 iterations.

In applications where a zoom version of the cropped area is desired, there are two scenarios to consider. First, the zoom version effectively requires higher spatial resolution than the highest resolution of the original data. However, a visible loss of image sharpness is likely of concern in the situation. Second, the zoom version effectively requires lower spatial resolution than the highest resolution of the original data. In both cases, the invention uses an interpolation process to resample the data in order to retain a maximum amount of image detail. In general, edge or detail-preserving image interpolation processes such as cubic-spline interpolation are preferred because they tend to preserve the detail and sharpness of the original image better.

Example consumer photographs and their various cropped versions are shown in pictures "house" (e.g., FIGS. **13**–**15**) and "volleyball" (FIGS. **16**–**18**). More specifically, FIGS. **13** and **16** illustrate uncropped original photographic images. FIGS. **14** and **17** illustrate belief maps, with lighter regions indicating higher belief values. As would be known by one ordinarily skilled in the art given this disclosure, the light intensity variations shown in FIGS. **14** and **17** are readily converted into numerical values for calculating the sum of the belief values discussed above. Finally, FIGS. **15** and **18** illustrate images cropped according to the invention.

For the "house" picture, both Bradley and Bollman (U.S. Pat. No. 5,978,519) would keep the entire image and not be able to produce a cropped image because of the shadows at the bottom and the tree extending to the top border of the uncropped image (FIG. **13**). There are no continuous flat background regions extending from the image borders in this picture, as required by U.S. Pat. No. 5,978,519. Similarly, the top of the tree in FIG. **16** would not be cropped in the system disclosed in U.S. Pat. No. 5,978,519.

Secondary subjects can lead to a more balanced cropped picture. For the "volleyball" picture (FIG. **16**), the inclusion of some parts of the tree by the algorithm leads to more interesting cropped pictures than simply placing the main subjects (players) in the center of the cropped image (FIG. **18**). The invention was able to do so because the trees are indicated to be of secondary importance based on the belief map FIG. **17**. It is obvious that the art taught by Bradley and Bollman in U.S. Pat. No. 5,978,519 would not be able to produce such a nicely cropped image. In fact, both Bradley and Bollman (U.S. Pat. No. 5,978,519) would at best remove the entire lower lawn portion of the picture and keep the tree branches in the upper-left of the uncropped image.

A computer program product may include one or more storage medium, for example; magnetic storage media such as magnetic disk (such as a floppy disk) or magnetic tape; optical storage media such as optical disk, optical tape, or machine readable bar code; solid-state electronic storage devices such as random access memory (RAM), or read-only memory (ROM); or any other physical device or media employed to store a computer program having instructions for practicing a method according to the present invention.

While the invention has been described in terms of preferred embodiments, those skilled in the art will recognize that the invention can be practiced with modification within the spirit and scope of the appended claims.

PARTS LIST

9	film sample position
10	source digital image
11	zoom factor
12	lens
13	time integration device
20	digital image processor
31	photographic film sample
32	film scanner
34	lamp house
36	gate device
38	photographic paper
40	general control computer
50	monitor device
60	input control device
80	crop window
81	margin
200	image and belief map
201	decision box for zoom factor selection
202	decision box for clustering regions
203	decision box for computing centroids
204	decision box for positioning the crop window
205	decision box moving a window
206	decision box for rotating the image
207	decision box for moving a window with the highest belief
208	decision box for determining if a solution exists
209	decision box for determining if the image is rotated
210	decision box for computing the sum of beliefs
211	decision box for computing the sum of beliefs
300	image and belief map
301	decision box for selecting a zoom factor and crop window
302	decision box for setting the background belief to 0
303	decision box for computing the centroid of nonzero beliefs
304	decision box for positioning the crop window
305	decision box for positioning the crop window
306	decision box for determining the optimal crop
307	decision box for determining the optimal crop
308	cropped image #1
309	cropped image #2
310	decision box for selecting the better crop
311	cropped image
400	image and belief map
401	decision box for selecting a zoom factor and crop window
402	decision box for setting the background belief to 0
403	decision box for computing the centroid of nonzero beliefs
404	decision box for positioning the crop window
405	decision box for positioning the crop window
406	decision box for determining the optimal crop

-continued

PARTS LIST

407	decision box for determining the optimal crop
408	cropped image #n
409	cropped image #n + 1
410	decision box for selecting a different zoom factor or different crop window
411	decision box for ending the process

What is claimed is:

1. A method of producing an image of a portion of at least a portion of a photographic image onto a photographic receiver, comprising the steps of:

a) receiving a digital image corresponding to the photographic image, the digital image comprising pixels;

b) developing a belief map of the digital image, by using pixels of the digital image to determine a series of features, and using such features to assign the probability of the location of a main subject of the digital image in the belief map;

c) selecting a crop window of a size and shape;

d) selecting an initial position of the crop window;

e) using the belief values corresponding to the crop window to select the position of the crop window; and

f) locating the relative optical position of a photographic image, the lens, and the photographic receiver in response to the position of the crop window within the digital image and illuminating a portion of the photographic image to produce an image onto the photographic receiver.

2. A method as set forth in claim **1** wherein step c) further comprises the steps of:

i) receiving a magnification factor; and

ii) using the magnification factor to select the size of the crop window.

3. A method as set forth in claim **1** wherein step e) further comprising the steps of:

i) calculating a subject content index for the crop window derived from the belief values;

ii) following a positioning procedure of repeating step i) for at least two positions of the crop window; and

iii) use the subject content index values to select the crop window position.

4. A method as set forth in claim **3** wherein step d) further comprises the steps of:

i) using the belief map to determine the center of the mass of the belief map; and

ii) using the center of mass of the belief map to select the initial position of the crop window.

5. The method of claim **4** wherein the crop window is completely within the digital image.

6. The method of claim **3** further comprising repeating said positioning procedure with a rotated image and determining if said rotated image produces a higher subject content index.

7. The method of claim **3** further comprising repeating said positioning procedure with a second zoom factor and a second crop window and determining if said second zoom factor and said second crop window produces a higher subject content index.

8. The method of claim **7** wherein said zoom factor is selected according to a size of said main subject cluster.

9. The method of claim **3** further comprising positioning said crop window such that the subject content index of said crop window is at an optimum.

10. The method of claim 1 wherein step b) further comprises the step of performing a clustering of the belief map to identify at least a cluster of highest belief values corresponding to main subject, a cluster of intermediate belief values corresponding to secondary subjects, and a cluster of lowest belief values corresponding to the background.

11. The method of claim 10 wherein said clustering includes setting said background portions to a zero belief value.

12. The method of claim 10 further comprising positioning said crop window such that said crop window includes all of said main subject cluster.

13. The method of claim 10 further comprising positioning said crop window to include a buffer around said main subject cluster.

14. A computer storage product having at least one computer storage medium having instructions stored therein causing one or more computers to perform the method of claim 1.

* * * * *

US006091841A

United States Patent [19]

Rogers et al.

[11] **Patent Number: 6,091,841**

[45] **Date of Patent: Jul. 18, 2000**

[54] **METHOD AND SYSTEM FOR SEGMENTING DESIRED REGIONS IN DIGITAL MAMMOGRAMS**

[75] Inventors: **Steven K. Rogers**, Beavercreek; **Philip Amburn**, Dayton; **Telford S. Berkey**, London; **Randy P. Broussard**, Huber Heights; **Martin P. DeSimio**, Fairborn; **Jeffrey W. Hoffmeister**, Beavercreek, all of Ohio; **Edward M. Ochoa**, San Antonio, Tex.; **Thomas F. Rathbun**, Beavercreek; **John E. Rosenstengel**, Huber Heights, both of Ohio

[73] Assignee: **Qualia Computing, Inc.**, Beavercreek, Ohio

[21] Appl. No.: **09/418,383**

[22] Filed: **Oct. 14, 1999**

Related U.S. Application Data

[62] Division of application No. 09/141,802, Aug. 28, 1998, Pat. No. 5,999,639.

[60] Provisional application No. 60/057,801, Aug. 28, 1997, provisional application No. 60/066,996, Nov. 28, 1997, and provisional application No. 60/076,760, Mar. 3, 1998.

[51] **Int. Cl.**[7] **G06K 9/36**

[52] **U.S. Cl.** **382/132**; 382/173; 382/256

[58] **Field of Search** 382/128, 132, 382/168, 169, 171, 173, 180, 256, 257; 128/922; 378/37

[56] **References Cited**

U.S. PATENT DOCUMENTS

4,453,266	6/1984	Bacus	382/6
4,723,553	2/1988	Miwa et al.	128/660
4,736,439	4/1988	May	382/54
4,747,156	5/1988	Wahl	382/54
4,907,156	3/1990	Doi et al.	364/413.13
5,133,020	7/1992	Giger et al.	382/6
5,212,637	5/1993	Saxena	364/413.26
5,260,871	11/1993	Goldberg	364/413.02
5,268,967	12/1993	Jang et al.	382/6
5,289,374	2/1994	Doi et al.	364/413.13
5,291,560	3/1994	Daugman	382/2

(List continued on next page.)

OTHER PUBLICATIONS

Winsberg, M.D., F., et al., "Detection of Radiographic Abnormalities in Mammograms by Means of Optical Scanning and Computer Analysis," *Radiology*, vol. 89, Aug. 1967, (pp. 211–215).

Ackerman, L.V., "Computer Classification of Radiographs and Xerograms of the Breast," Ph.D. Dissertation, University of Illinois at the Medical Center, Oct. 1970.

Hall, E.L., et al., "A Survey of Preprocessing and Feature Extraction Techniques for Radiographic Images," *IEEE Transactions on Computers*, vol. C–20, No. 9, Sep. 1971, (pp. 1032–1044).

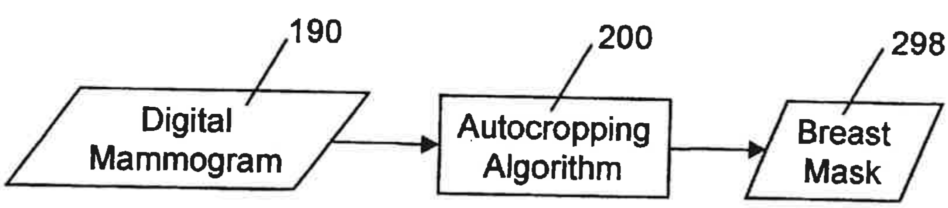
(List continued on next page.)

Primary Examiner—Andrew W. Johns
Attorney, Agent, or Firm—Biebel & French

[57] **ABSTRACT**

A method and system for detecting and displaying clustered microcalcifications in a digital mammogram, wherein a single digital mammogram is first automatically cropped to a breast area sub-image which is then processed by means of an optimized Difference of Gaussians filter to enhance the appearance of potential microcalcifications in the sub-image. The potential microcalcifications are thresholded, clusters are detected, features are computed for the detected clusters, and the clusters are classified as either suspicious or not suspicious by means of a neural network. Thresholding is preferably by sloping local thresholding, but may also be performed by global and dual-local thresholding. The locations in the original digital mammogram of the suspicious detected clustered microcalcifications are indicated. Parameters for use in the detection and thresholding portions of the system are computer-optimized by means of a genetic algorithm. The results of the system are optimally combined with a radiologist's observation of the original mammogram by combining the observations with the results, after the radiologist has first accepted or rejected individual detections reported by the system.

14 Claims, 28 Drawing Sheets



U.S. PATENT DOCUMENTS

5,331,550	7/1994	Stafford et al.	364/413.02
5,343,390	8/1994	Doi et al.	364/413.16
5,359,513	10/1994	Kano et al.	364/413.23
5,365,429	11/1994	Carman	364/413.13
5,388,143	2/1995	MacMahon	378/206
5,452,367	9/1995	Bick et al.	382/128
5,463,548	10/1995	Asada et al.	364/413.02
5,491,627	2/1996	Zhang et al.	364/413.2
5,537,485	7/1996	Nishikawa et al.	382/130
5,572,565	11/1996	Abdel-Mottaleb	378/37
5,574,799	11/1996	Bankman et al.	382/132
5,579,360	11/1996	Abdel-Mottaleb	378/37
5,586,160	12/1996	Mascio	378/37
5,598,481	1/1997	Nishikawa et al.	382/130
5,615,243	3/1997	Chang et al.	378/37
5,622,171	4/1997	Asada et al.	128/653.1
5,625,717	4/1997	Hashimoto et al.	382/260
5,627,907	5/1997	Gur et al.	382/132
5,633,948	5/1997	Kegelmeyer et al.	382/132
5,638,458	6/1997	Giger et al.	382/132
5,657,362	8/1997	Giger et al.	378/37
5,661,820	8/1997	Kegelmeyer, Jr.	382/226
5,666,434	9/1997	Nishikawa et al.	382/128
5,668,888	9/1997	Doi et al.	382/132
5,673,332	9/1997	Nishikawa et al.	382/128
5,729,620	3/1998	Wang	382/128
5,729,662	3/1998	Rozmus	395/23
5,732,697	3/1998	Zhang et al.	128/630
5,740,266	4/1998	Weiss et al.	382/128
5,740,267	4/1998	Echerer et al.	382/132
5,740,268	4/1998	Nishikawa et al.	382/132
5,757,953	5/1998	Jang	382/132
5,761,334	6/1998	Nakajima et al.	382/132
5,768,333	6/1998	Abdel-Mottaleb	378/37
5,768,406	6/1998	Abdel-Mottaleb	382/132
5,769,074	6/1998	Barnhill et al.	128/630
5,799,100	8/1998	Clark et al.	382/132
5,815,591	9/1998	Roehrig et al.	382/130
5,825,910	10/1998	Vafai	382/132
5,825,936	10/1998	Clarke et al.	382/261
5,854,851	12/1998	Bamberger et al.	382/132
5,857,030	1/1999	Gaborski et al.	382/132
5,999,639	12/1999	Rogers et al.	382/132
6,014,452	1/2000	Zhang et al.	382/132

OTHER PUBLICATIONS

Ackerman, L.V., et al., "Breast Lesion Classification by Computer and Xeroradiograph," *Cancer*, vol. 30, No. 4, Oct. 1972, (pp. 1025–1035).

Haralick, R., et al., "Textural Features for Image Classification," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC–3, No. 6, Nov. 1973, (pp. 610–621).

Ballard, D., et al., "Tumor Detection in Radiographs," *Computers and Biomedical Research*, vol. 6, 1973, (pp. 299–321).

Ackerman, L.V., et al., "Classification of Benign and Malignant Breast Tumors on the Basis of 36 Radiographic Properties," *Cancer*, vol. 31, No. 2, 1973, (p. 138).

Chang, N–C., "Computer Characterization of Calcifications for Breast Cancer Detection—A Feasibility Study," Master's Thesis, Department of Electrical Engineering, University of Cincinnati, 1973.

Wes, Ph.D., W.G. "Evaluation of Mammographic Calcifications Using a Computer Program," Work in Progress, *Radiology*, vol. 116, Sep. 1975, (pp. 717–720).

Kimme, C., et al., "Automatic Detection of Suspicious Abnormalities in Breast Radiographs," *Data Structures, Computer Graphics and Pattern Recognition*, 1975, (pp. 427–447).

Ting, Y.C., "A Computer Pattern Recognition System for Breast Cancer Detection," Master's Thesis, University of Cincinnati, 1975.

Millis, R.R., et al., "The Detection and Significance of Calcifications in the Breast: A Radiological and Pathological Study," *British Journal of Radiology*, vol. 49, No. 577, Jan. 1976, (pp. 12–26).

Spiesberger, W., et al., "Outlining of Microcalcifications by Computer–Assisted Mammogram Inspection," *Medicamundi*, vol. 22, No. 3, 1977, (pp. 32–34).

Spiesberger, W., "Mammogram Inspection by Computer," *IEEE Transactions on Biomedical Engineering*, vol. BME–26, No. 4, Apr. 1979, (pp. 213–219).

Kimme–Smith, C., et al., "Toward Reliable Measurements of Breast Parenchymal Patterns," *Proceedings of Sixth Conference on Computer Applications in Radiology and Computer–Aided Analysis of Radiology Images*, Jun. 1979, (pp. 118–121).

Hand, W., et al., "Computer Screening of Xeromammograms: A Technique for Defining Suspicious Areas of the Breast," *Computers and Biomedical Research*, vol. 12, 1979, (pp. 445–460).

Fox, S.H., et al., "A Computer Analysis of Mammographic Microcalcifications: Global Approach," *Proceedings of the IEEE 5th International Conference on Pattern Recognition*, 1980, (pp. 624–631).

Semmlow, J.L., et al., "A Fully Automated System for Screening Xeromammograms," *Computers and Biomedical Research*, vol. 13, 1980, (pp. 350–362).

Dhawan, A.P., et al., "Enhancement of Mammographic Features by Optimal Adaptive Neighborhood Image Processing," *IEEE Transactions on Medical Imaging*, vol. MI–5, No. 1, Mar. 1986, (pp. 8–15).

Dhawan, A.P., et al., Correction to "Enhancement of Mammographic Features by Optimal Adaptive Neighborhood Image Processing," *IEEE Transactions on Medical Imaging*, vol. MI–5, No. 2, Jun. 1986, (p. 120).

Bhahu, B., "Automatic Target Recognition: State of the Art Survey," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES–22, No. 4, Jul. 1986, (pp. 364–379).

Metz, Ph.D., C.E., "ROC Methodology in Radiologic Imaging," *Investigative Radiology*, vol. 21, No. 9, Sep. 1986 (pp. 720–733).

Lippmann, R.P., "An Introduction to Computing with Neural Nets," *IEEE ASSP Magazine*, vol. 96, Apr. 1987 (pp. 4–22).

Chan, H–P, et al., "Image Feature Analysis and Computer–Aided Diagnosis in Digital Radiography. I. Automated Detection of Microcalcifications in Mammography," *Med. Phys.*, vol. 14, No. 4, Jul./Aug. 1987, (pp. 538–548).

Kahn, E., et al., "Computer Analysis of Breast Calcifications in Mammographic Images," *Proceedings of the International Symposium on Computer Assisted Radiology '87*, Lemke, U., et al., editors, 1987, (pp. 729–733).

Chan, H–P, et al., "Original Investigations: Computer–Aided Detection of Microcalcifications in Mammograms—Methodology and Preliminary Clinical Study," *Investigative Radiology*, vol. 23, No. 7, Jul. 1988, (pp. 664–671).

Fam, B.W., et al., "Algorithm for the Detection of Fine Clustered Calcifications on Film Mammograms," *Radiology*, vol. 169, No. 1, Oct. 1988, (pp. 333–337).

Fam, B.W., et al., "The Detection of Calcification Clusters in Film–Screen Mammograms; A Detailed Algorithmic Approach" *Medical Imaging II*, vol. 914, 1988, (pp. 620–634).

Dhawan, A.P., et al., "Mammographic Feature Enhancement by Computerized Image Processing," *Computer Methods and Programs in Biomedicine*, vol. 27, 1988, (pp. 23–35).

Davies, D.H., et al., "Automatic Detection of Microcalcifications in Digital Mammograms Using Local Area Thresholding Techniques," *SPIE's Conference on Medical Imaging III Image Processing*, vol. 1092, Newport Beach, CA, Jan. 31–Feb. 3, 1989, (pp. 153–159).

Lai, S.M., et al., "On Techniques for Detecting Circumscribed Masses in Mammograms," *IEEE Transactions on Medical Imaging*, vol. 8, No. 4, Dec. 1989, (pp. 377–386).

Davies, D.H., et al., "Automatic Detection of Clusters of Calcifications in Digital Mammograms," *Proceedings of the International Symposium on Computer Assisted Radiology*, Lemke, H.U., et al., editors, 1989, (pp. 180–184).

Ayer, K.W. et al., "Forward Looking Infrared Image Segmentation and Pattern Recognition Using Gabor Transform and Joint Transform Correlation Techniques," Wright Laboratories, Air Force Institute of Technology, Technical Report, Mar. 1990.

Davies, D.H., et al., "Automatic Computer Detection of Clustered Calcifications in Digital Mammograms," *Phys. Med. Biol.*, vol. 35, No. 8, Apr. 1990, (pp. 1111–1118).

Boone, J.M., et al., "Neural Networks in Radiologic Diagnosis," *Investigative Radiology*, vol. 25, No. 9 Sep. 1990, (pp. 1012–1016).

Brzakovic, D., et al., "An Approach to Automated Detection of Tumors in Mammograms" *IEEE Transactions on Medical Imaging*, vol. 9, No. 3, Sep. 1990, (pp. 233–241).

Rogers, S.K., et al., *An Introduction to Biological and Artificial Neural Networks*, Oct. 23, 1990, (pp. 47–61).

Chan, Ph.D., H–P, et al., "Improvement in Radiologists' Detection of Clustered Microcalcifications on Mammograms—The Potential of Computer–Aided Diagnosis," *Investigative Radiology*, vol. 25, No. 10, Oct. 1990, (pp. 1102–1110).

Veronin, C.P., et al., "An Optical Image Segmentor Using Wavelet Filtering Techniques as the Front End of a Neural Network Classifier," *SPIE's International Conference on Applications of Artificial Neural Networks*, vol. 1469, Apr. 1991, (pp. 281–291).

Lau, T.K., et al., "Automated Detection of Breast Tumors Using the Asymmetry Approach," *Computers and Biomedical Research*, vol. 24, No. 3, Jun. 1991, (pp. 273–295).

Yin, F.F., et al., "Computerized Detection of Masses in Digital Mammograms: Analysis of Bilateral Subtraction Images," *Med. Phys.*, vol. 18, No. 5, Sep./Oct. 1991, (pp. 955–963).

Zhang, W., et al., "Image Processing of Human Corneal Endothelium Based on a Learning Network," *Applied Optics*, vol. 30, No. 29, Oct. 10, 1991, (pp. 4211–4217).

Laing, J., et al., "Gabor and Multiresolution Wavelet Decomposition Analysis of the Kanisza Triangle Illusion," for Wright Lab Target Recognition Group, Nov. 1991.

Kimme–Smith, Ph.D., C., "New and Future Developments in Screen–Film Mammography Equipment and Techniques," *Radiologic Clinics of North America*, vol. 30, No. 1, Jan. 1992, (pp. 55–66).

Veronin, C.P., et al., "Optical Image Segmentation Using Neural–Based Wavelet Filtering Techniques," *Optical Engineering*, vol. 31, No. 2, Feb. 1992, (pp. 287–294).

Wu, Y., et al., "Computerized Detection of Clustered Microcalcifications in Digital Mammograms: Applications of Artifical Neural Networks," *Medical Physics*, vol. 19, No. 3, May/Jun. 1992, (pp. 555–560).

Ng, S.L., et al., "Automated Detection and Classification of Breast Tumors," *Computers and Biomedical Research*, vol. 25, 1992, (pp. 218–237).

Dhawan, A.P., et al., "Artificial Neural Network Based Classification of Mammographic Microcalcifications Using Image Structure Features," *Proceedings of SPIE's Conference on Biomedical Image Processing and Biomedical Visualization*, San Jose, vol. 1905, Feb. 1–4, 1993, (pp. 820–831).

Wu, Y., et al., "Artificial Neural Networks in Mammography: Applications to Decision Making in the Diagnosis of Breast Cancer," *Radiology*, vol. 187, No. 1, Apr. 1993, (pp. 81–87).

Woods, K.S., et al, "Comparative Evaluation of Pattern Recognition Techniques for Detection of Microcalcifications in Mammography," *Computerized Medical Imaging and Graphics*, vol. 16, No. 5, May 1993, (pp. 1417–1436).

Fletcher, S.W., et al., "Report of the International Workshop on Screeing for Breast Cancer," *Journal of the National Cancer Institute*, vol. 85, No. 20, Oct. 20, 1993, (pp. 1644–1656).

Anand, R., et al., "An Improved Algorithm for Neural Network Classification of Imbalanced Training Sets," *IEEE Transactions on Neural Networks*, vol. 4, No. 6, Nov. 1993, (pp. 962–969).

Dengler, J., et al., "Segmentation of Microcalcifications in Mammograms," *IEEE Transactions on Medical Imaging*, vol. 12, No. 4, Dec. 1993, (pp. 634–642).

Giger, Ph.D., M.L., "Computer–Aided Diagnosis," *RSNA Categorical Course in Physics 1993*, 1993, (pp. 283–298).

Nishikawa, R.M., et al., "Effect of Case Selection on the Performance of Computer–Aided Detection Schemes," *Med. Phys.*, vol. 21, No. 2, Feb. 1994, (pp. 265–269).

Yin, F.F., et al., "Computerized Detection of Masses in Digital Mammograms: Investigation of Feature–Analysis Techniques," *Journal of Digital Imaging*, vol. 7, No. 1, Feb. 1994, (pp. 18–26).

Kegelmeyer, Jr., Ph.D., W. P., et al., "Computer–Aided Mammographic Screening for Spiculated Lesions," *Radiology*, vol. 191, No. 2, May 1994, (pp. 331–337).

Bårman, H., et al. "Feature Extraction For Computer–Aided Analysis of Mammograms," *State of the Art Digital Mammographic Image Analysis*, Boyer, K.W., et al., editors, 1994, (pp. 128–147).

Chitre, Y., et al., "Artificial Neural Network Based Classification of Mammographic Microcalcifications Using Image Structure Features," *State of the Art in Digital Mammographic Image Analysis*, Boyer, K.W., et al., editors, 1994, (pp. 167–197).

Giger, M.L., et al., "Computerized Characterization of Mammographic Masses: Analysis of Spiculation," *Cancer Letters*, vol. 77, 1994, (pp. 201–211).

Kegelmeyer, Jr., W. P., "Evaluation of Stellate Lesion Detection in a Standard Mammogram Data Set," *State of the Art in Digital Mammographic Image Analysis*, Boyer, K.W. et al., editors, 1994, (pp. 262–279).

Lidbrink, E.K., et al., "The General Mammography Screening Program in Stockholm: Organisation and First–Round Results," *Acta Oncologica*, vol. 33, No. 4, 1994, (pp. 353–358).

Nishikawa, R.M., "Computer–Aided Detection and Diagnosis of Masses and Clustered Microcalcifications from Digital Mammograms," *State of the Art in Digital Mammographic Image Analysis*, Boyer, K.W., et al., editors, 1994, (pp. 82–102).

Petrosian, A., et al., "Computer–Aided Diagnosis in Mammography: Classification of Mass and Normal Tissue by Texture Analysis," *Phys. Med. Biol.*, vol. 39, 1994, (pp. 2273–2288).

Shen, L., et al., "Detection and Classification of Mammographic Calcifications," *State of the Art in Digital Mammographic Image Analysis*, Boyer, K.W., et al., editors, 1994, (pp. 198–212).

Wilding, P., et al., "Application of Backpropagation Neural to Diagnosis of Breast and Ovarian Cancer," *Cancer Letters*, vol. 77, 1994, (pp. 145–153).

Woods, K.S., et al., "Comparative Evaluation of Pattern Recognition Techniques for Detection of Microcalcifications in Mammography," *State of the Art in Digital Mammographic Image Analysis*, Boyer, K.W., et al., editors 1994, (pp. 213–231).

Chan, H–P, et al., "Computer–Aided Classification of Mammographic Masses and Normal Tissue: Linear Discriminant Analysis in Texture Feature Space," *Phys. Med. Biol.*, vol. 40, Feb. 1995, (pp. 857–876).

Hojjatoleslami, S.A., et al., "Automatic Detection of Calcification in Mammograms," 5th International Conference on Image Processing and Its Applications, vol. 410, Jul. 1995, (pp. 139–143).

Li, H.D., et al., "Markov Random Field for Tumor Detection in Digital Mammography," *IEEE Transactions on Medical Imaging*, vol. 14, No. 3, Sep. 1995, (pp. 565–576).

Huo Z., et al., "Analysis of Spiculation in the Computerized Classification of Mammographic Masses," *Med. Phys.*, vol. 22, No. 10, Oct. 1995, (pp. 1569–1579).

Bick, U., et al., "Automated Segmentation of Digitized Mammograms," *Academic Radiology*, vol. 2, 1995, (pp. 1–9).

Feig, M.D., S.A., et al., "Digital Mammography, Computer–Aided Diagnosis, and Telemammography," *Radiologic Clinics of North America*, vol. 33, No. 6, Nov. 1995, (pp. 1205–1230).

Zheng, B., et al., "Computerized Detection of Masses from Digitized Mammograms: Comparison of Single–Image Segmentation and Bilateral–Image Subtraction," *Academic Radiology*, vol. 2, No. 12, Dec. 1995, (pp. 1056–1061).

Tahoces, P.G., et al., "Computer–Assisted Diagnosis: The Classification of Mammographic Breast Parenchymal Patterns," *Phys. Med. Biol.*, vol. 40, 1995, (pp. 103–117).

McCandless, D.A., et al., "Wavelet Detection of Clustered Microcalcifications," *SPIE*, vol. 2762 (date unknown), (pp. 388 et seq.).

Braccialarghe, D., et al., "Contrast Enhancement of Mammographic Features: A Comparison of Four Methods," *Optical Engineering*, vol. 35, No. 1, Jan. 1996, (pp. 76–80).

Sahiner, B., et al., "Classification of Mass and Normal Breast Tissue: Feature Selection Using a Genetic Algorithm," *Colloqium on Digital Mammography*, Feb. 1996, (pp. 379–384).

Chang, Y–H, et al., "Computerized Identification of Suspicious Regions for Masses in Digitized Mammograms," *Investigative Radiology*, vol. 31, No. 3, Mar. 1996, (pp. 146–153).

Giger, Ph.D., M., et al., "Image Processing and Computer–Aided Diagnosis," *Radiologic Clinics of North America*, vol. 34, No. 3, May 1996, (pp. 565–596).

Kocur, C.M., et al., "Using Neural Networks to Select Wavelet Features for Breast Cancer Diagnosis," *IEEE Engineering in Medicine and Biology*, May/Jun. 1996, (pp. 95–102).

Dhawan, A.P., et al., "Analysis of Mammographic Microcalcifications Using Gray–Level Image Structure Features," *IEEE Transactions on Medical Imaging*, vol. 15, No. 3, Jun. 1996, (pp. 246–259).

Pohlman, S., et al., "Quantitative Classification of Breast Tumors in Digitized Mammograms," *Med. Phys.*, vol. 23, No. 8, Aug. 1996, (pp. 1337–1345).

Chang, Y–H., "Robustness of Computerized Identification of Masses in Digitized Mammograms," *Investigative Radiology*, vol. 31, No. 9, Sep. 1996, (pp. 563–568).

Cooley, T.R., "An Automated System for the Classification of Mammograms," Ph.D. Dissertation, Rutgers, the State University of New Jersey, Oct. 1996.

Petrick, N., et al., "Automated Detection of Breast Masses on Mammograms Using Adaptive Contrast Enhancement and Texture Classification," *Med. Phys.*, vol. 23, No 10, Oct. 1996, (pp. 1685–1696).

Sahiner, B., et al., "Image Feature Selection by a Genetic Algorithm: Application to Classification of Mass and Normal Breast Tissue," *Med. Phys.*, vol. 23, No. 10, Oct. 1996, (pp. 1671–1684).

Bick, U., et al., "Density Correction of Peripheral Breast Tissue on Digital Mammograms," *RadioGraphics*, vol. 16, No. 6, Nov. 1996, (pp. 1403–1411).

Ochoa, E.M., "Clustered Microcalcification Detection Using Optimized Difference of Gaussians," Master's Thesis, Department of the Air Force, Air University, Air Force Institute of Technology, Dec. 1996.

Zheng, B., et al., "On the Reporting of Mass Contrast in CAD Research," *Med. Phys.*, vol. 23, No. 12, Dec. 1996, (pp. 2007–2009).

Chen, L., et al., "Morphological Filtering and Multiresolution Fusion for Mammographic Microcalcifications Detection," Proceedings of SPIE's Conference on Image Processing, Newport Beach, vol. 3034, Feb. 25–28, 1997, (pp. 938–947).

Freedman, M., et al., "Classification of False Positive Findings on Computer Aided Detection of Breast Microcalcifications," Proceedings of SPIE's Conference on Image Processing, Newport Beach, vol. 3034, Feb. 25–28, 1997, (pp. 853–859).

Gavrielides, M.A., et al., "Automatic Shape Analysis and Classification of Mammographic Calcifications," Proceedings of SPIE's Conference on Image Processing, Newport Beach, vol. 3034, Feb. 25–28, 1997, (pp. 869–876).

Li, H., et al., "Mammographic Mass Detection by Stochastic Modeling and a Multi–Modular Neural Network," Proceedings of SPIE's Conference on Image Processing, Newport Beach, vol. 3034, Feb. 25–28, 1997, (pp. 480–490).

Sahiner, B., et al., "Characterization of Masses on Mammograms: Significance of the Use of the Rubber–Band Straightening Transform," Proceedings of SPIE's Conference on Image Processing, Newport Beach, vol. 3034, Feb. 25–28, 1997 (pp. 491–499).

Wu, C.Y., et al., "Image Feature Analysis for Classification of Microcalcifications in Digital Mammography: Neural Networks and Genetic Algorithms," Proceedings of SPIE's Conference on Image Processing, Newport Beach, vol. 3034, Feb. 25–28, 1997 (pp. 501–509).

Zheng, Y., et al., "Reducing Breast Biopsies by Ultrasonographic Analysis and a Modified Self–Organizing Map," Proceedings of SPIE's Conference on Image Processing, Newport Beach, vol. 3033, Feb. 25–28, 1997 (pp. 384–391).

Rogers, S.K., "Introduction to Artificial Neural Networks," *Fundamentals of Artificial Neural Networks*, Apr. 1997, (pp. 1–41).

Zheng, Y., et al., "Reduction of Breast Biopsies with a Modified Self–Organizing Map," *IEEE Transactions on Neural Networks*, vol. 8, No. 6, Nov. 1997, (pp. 1386–1396).

Anastasio, M.A., et al., "Optimization and FROC Analysis of Rule–Based Detection Schemes Using a Multiobjective Approach," Correspondence in *IEEE Transactions on Medical Imaging*, Aug. 26, 1998.

Fig. 1

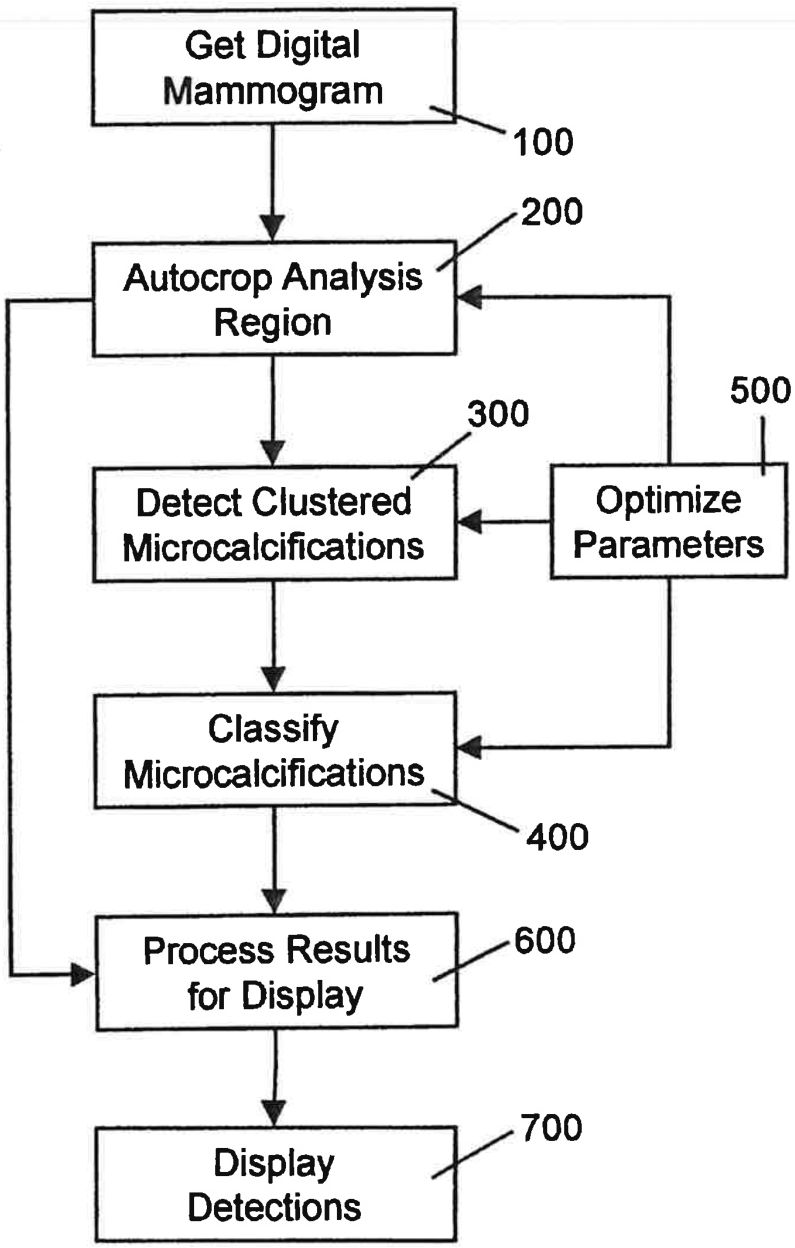

Get Digital Mammogram
100
200
Autocrop Analysis Region
500
300
Detect Clustered Microcalcifications
Optimize Parameters
Classify Microcalcifications
400
Process Results for Display
600
700
Display Detections

Fig. 2

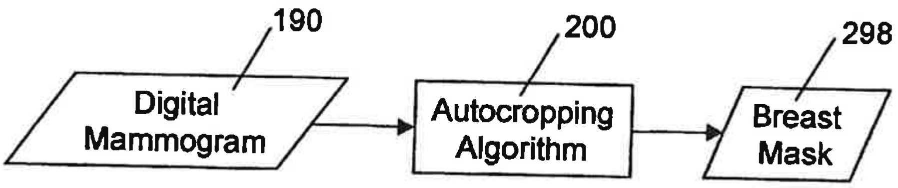

190
Digital Mammogram
200
Autocropping Algorithm
298
Breast Mask

Fig. 3

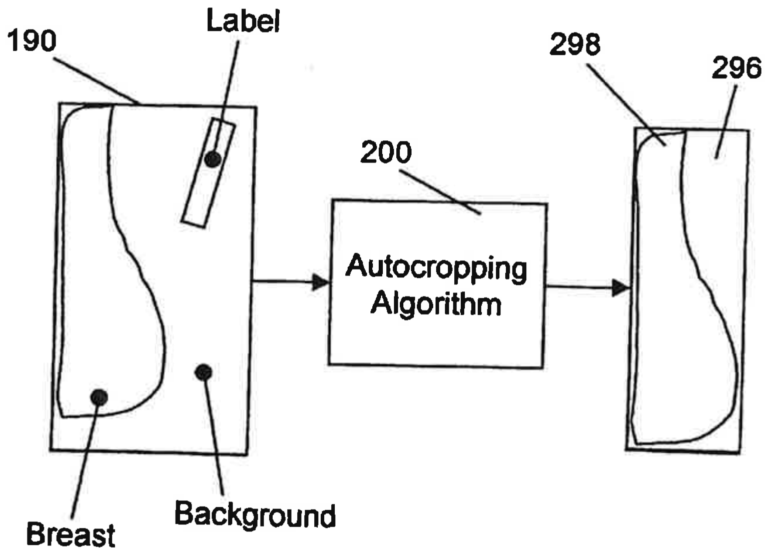

190
Label
Breast
Background
200
Autocropping Algorithm
298
296

Fig. 4

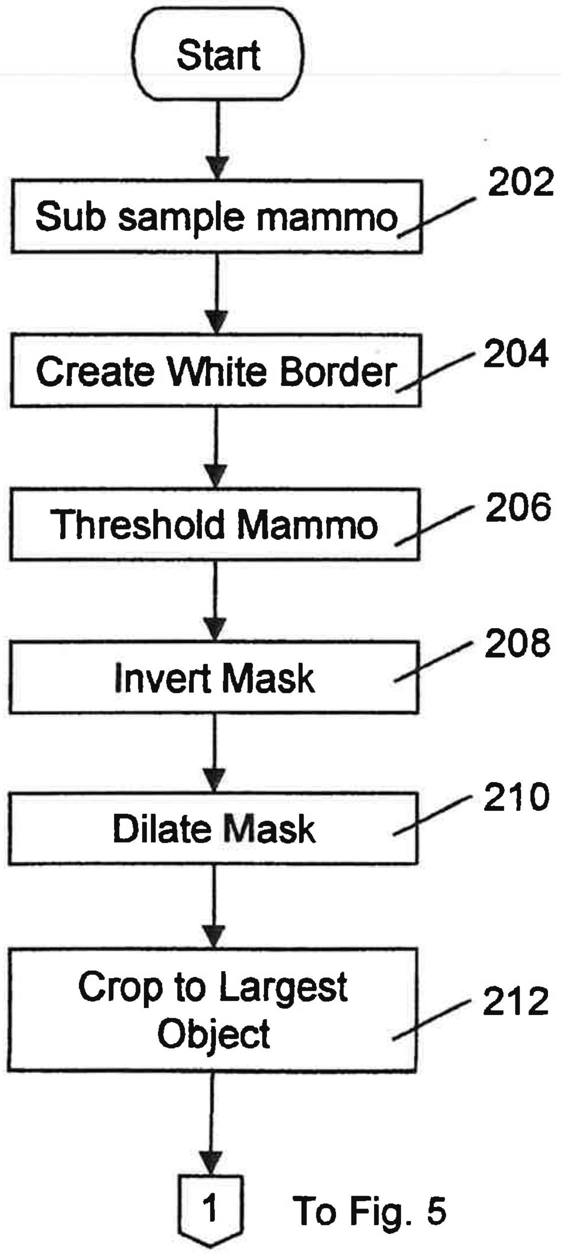

Start
Sub sample mammo
202
Create White Border
204
Threshold Mammo
206
Invert Mask
208
Dilate Mask
210
Crop to Largest Object
212
1
To Fig. 5

Fig. 5

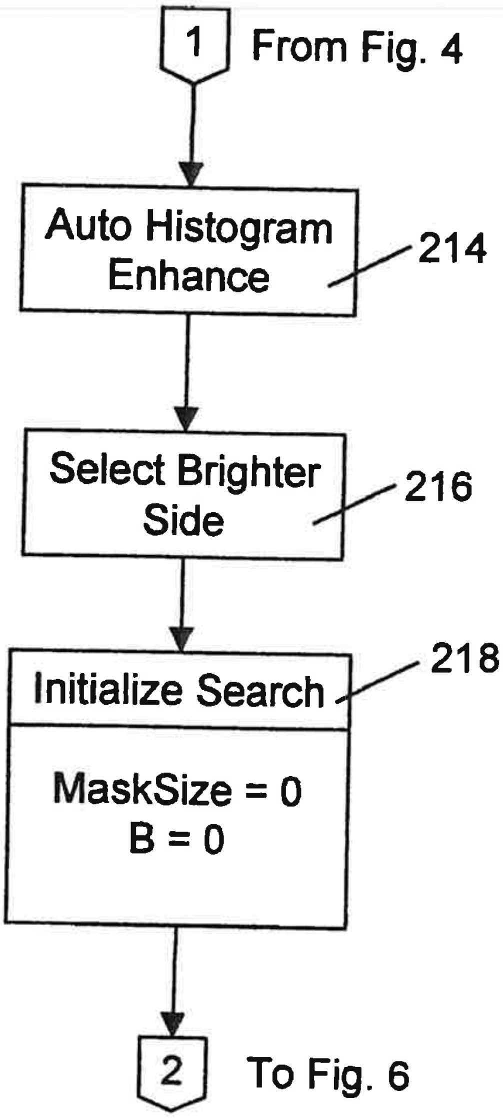

1
From Fig. 4
Auto Histogram Enhance
214
Select Brighter Side
216
Initialize Search
218
MaskSize = 0
B = 0
2
To Fig. 6

Fig. 6

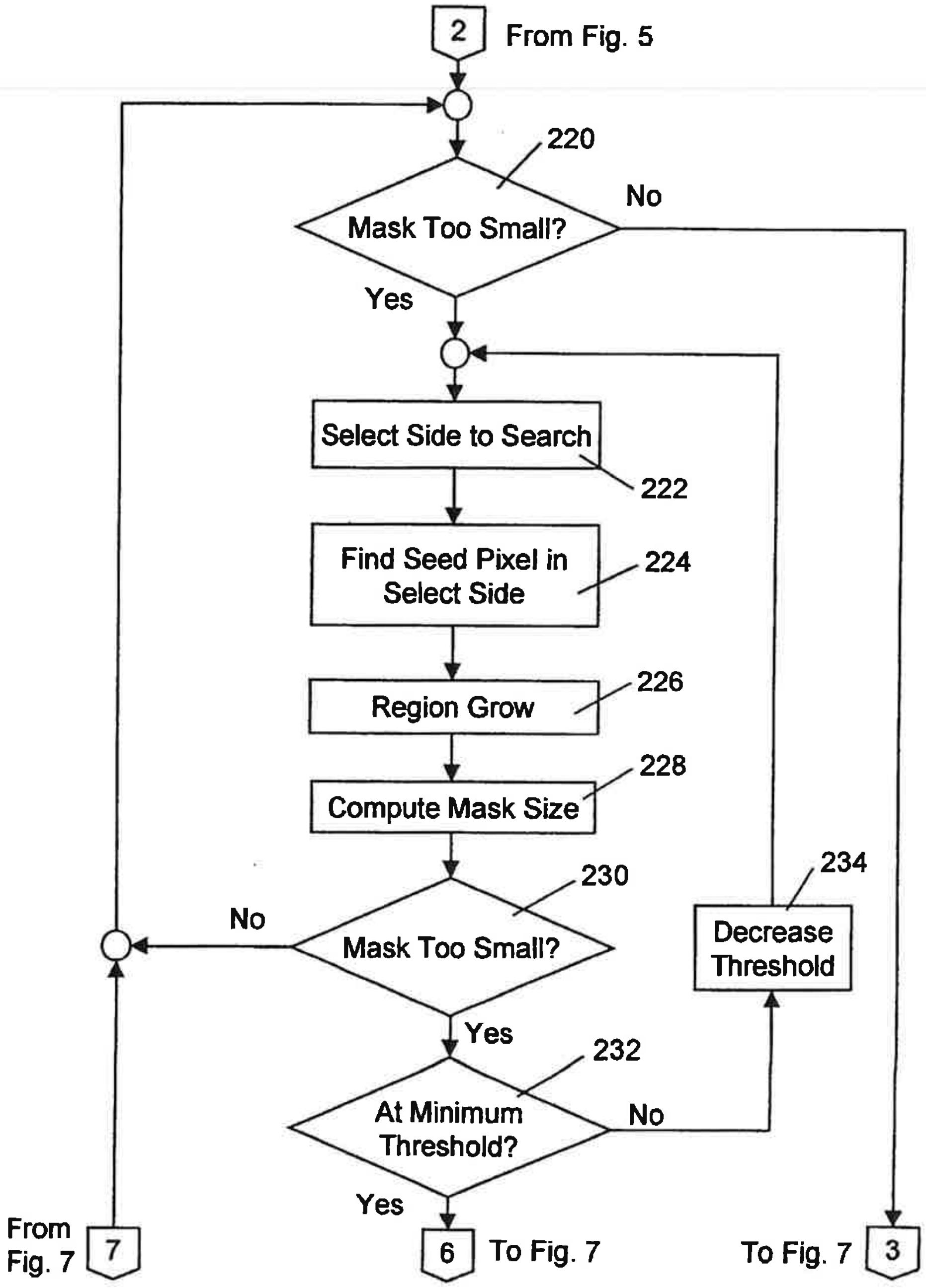

2
From Fig. 5
220
Mask Too Small?
No
Yes
Select Side to Search
222
Find Seed Pixel in Select Side
224
Region Grow
226
Compute Mask Size
228
230
Mask Too Small?
No
Yes
234
Decrease Threshold
232
At Minimum Threshold?
No
Yes
From Fig. 7
7
6
To Fig. 7
To Fig. 7
3

Fig. 7

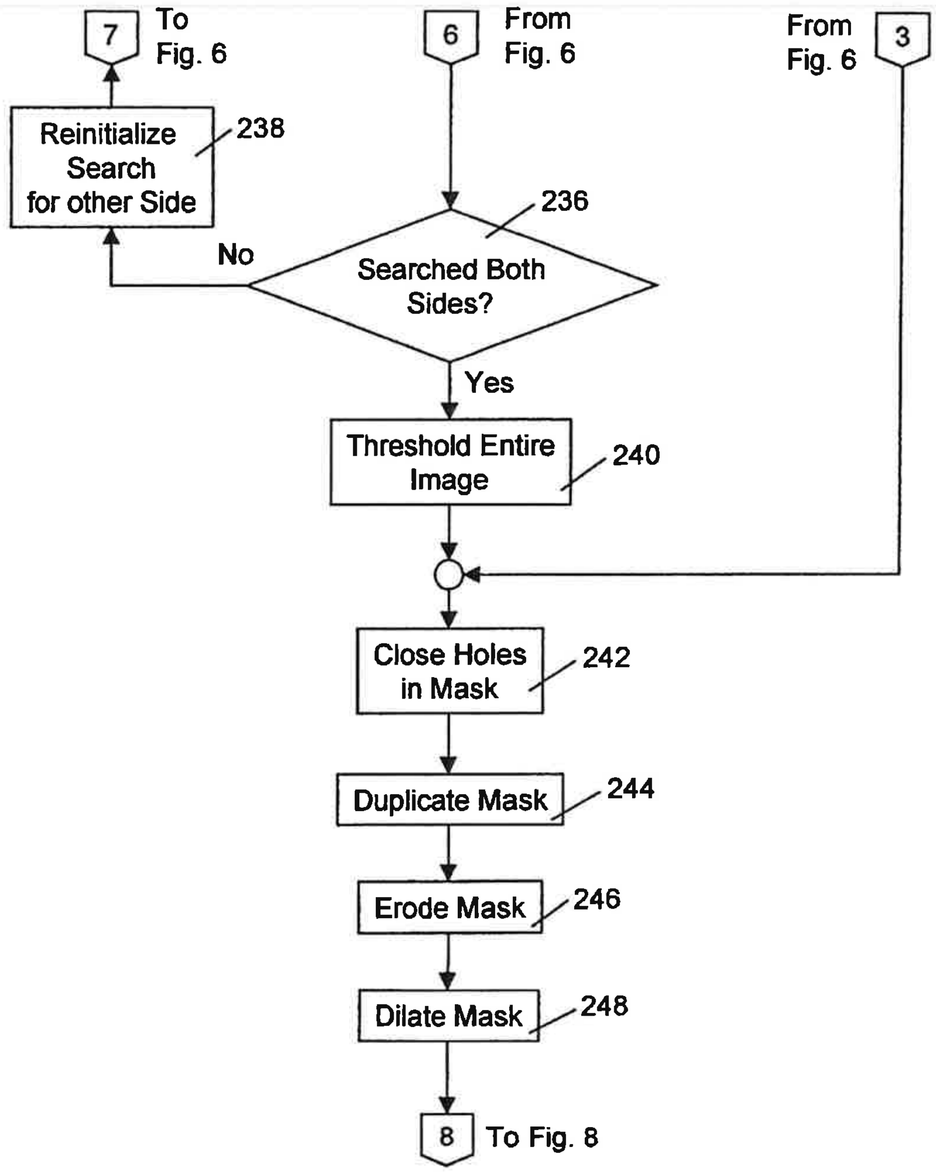

7
To Fig. 6
6
From Fig. 6
From Fig. 6
3
Reinitialize Search for other Side
238
236
No
Searched Both Sides?
Yes
Threshold Entire Image
240
Close Holes in Mask
242
Duplicate Mask
244
Erode Mask
246
Dilate Mask
248
8
To Fig. 8

Fig. 8

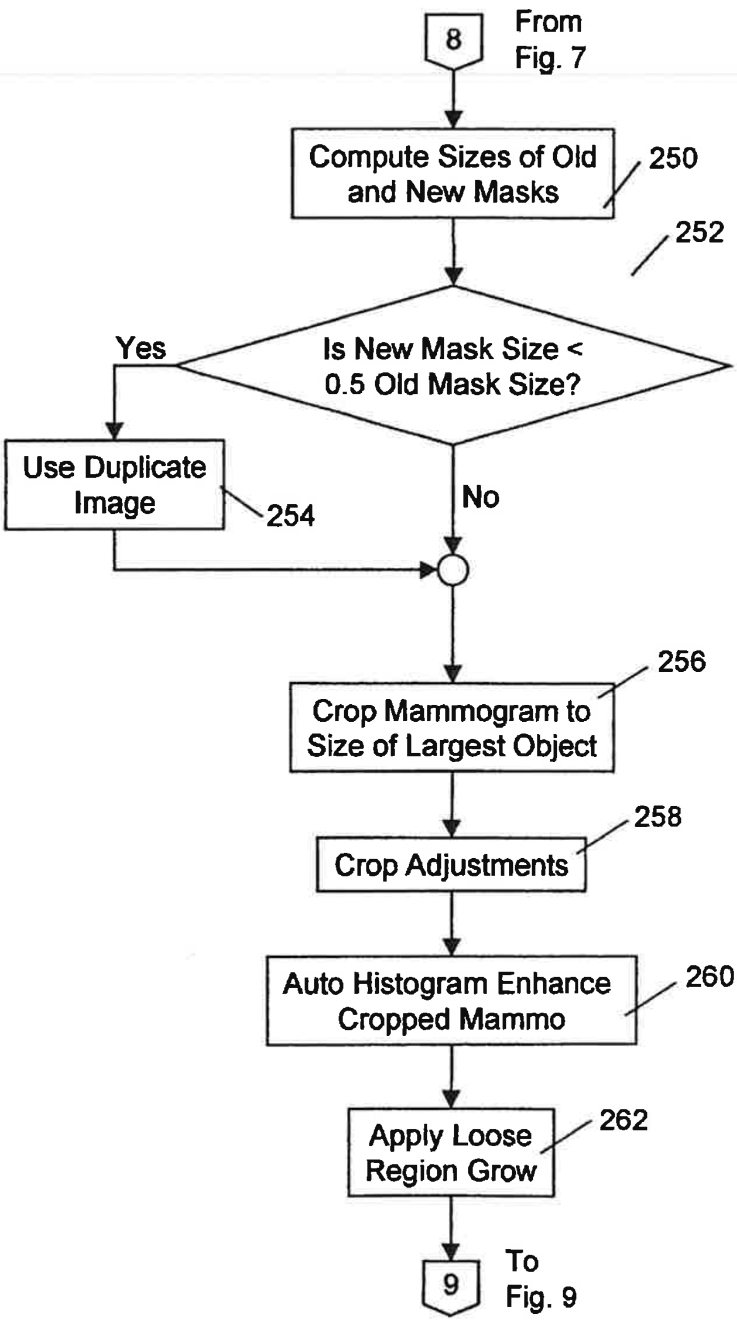

8
From
Fig. 7
Compute Sizes of Old and New Masks
250
252
Is New Mask Size < 0.5 Old Mask Size?
Yes
Use Duplicate Image
254
No
Crop Mammogram to Size of Largest Object
256
Crop Adjustments
258
Auto Histogram Enhance Cropped Mammo
260
Apply Loose Region Grow
262
9
To
Fig. 9

Fig. 9

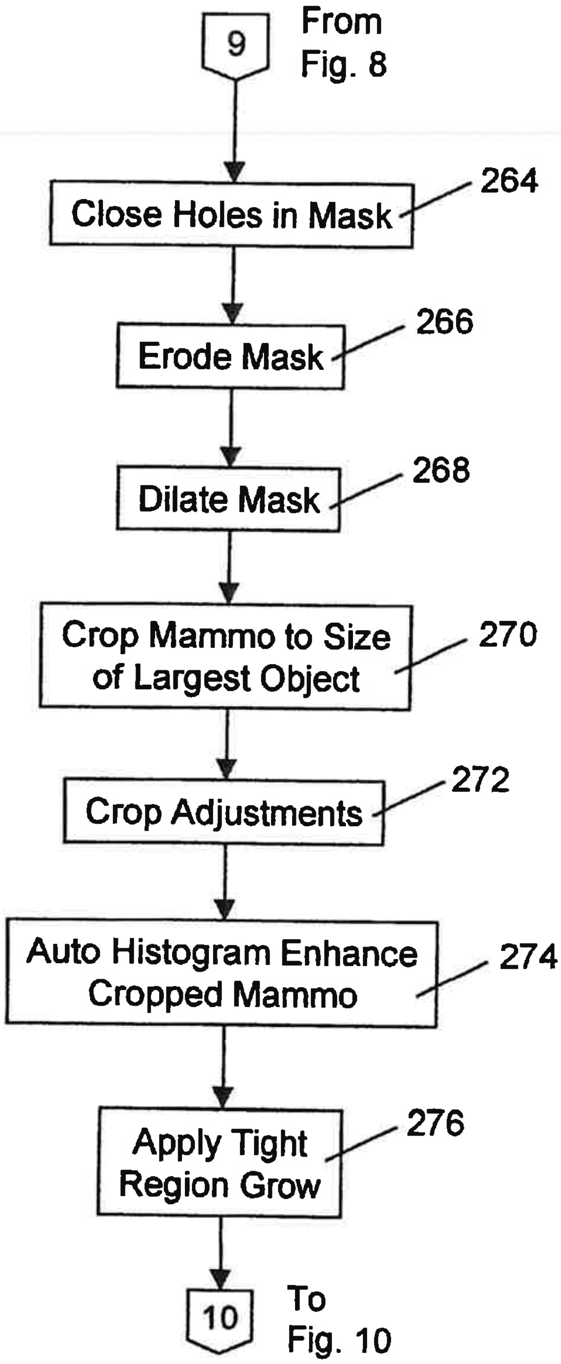

9
From
Fig. 8
Close Holes in Mask
264
Erode Mask
266
Dilate Mask
268
Crop Mammo to Size of Largest Object
270
Crop Adjustments
272
Auto Histogram Enhance Cropped Mammo
274
Apply Tight Region Grow
276
10
To
Fig. 10

Fig. 10

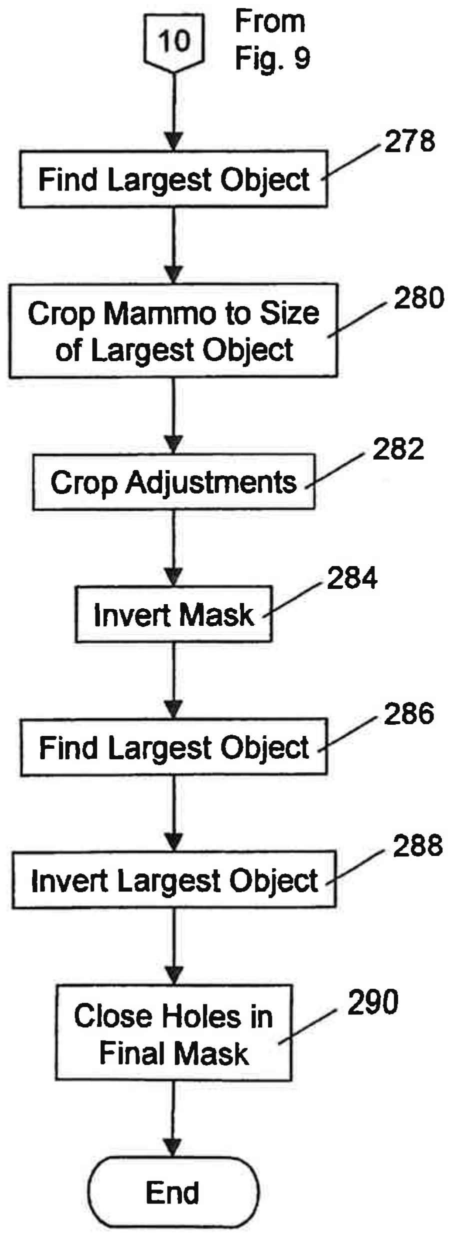

10
From
Fig. 9
Find Largest Object
278
Crop Mammo to Size of Largest Object
280
Crop Adjustments
282
Invert Mask
284
Find Largest Object
286
Invert Largest Object
288
Close Holes in Final Mask
290
End

Fig. 11

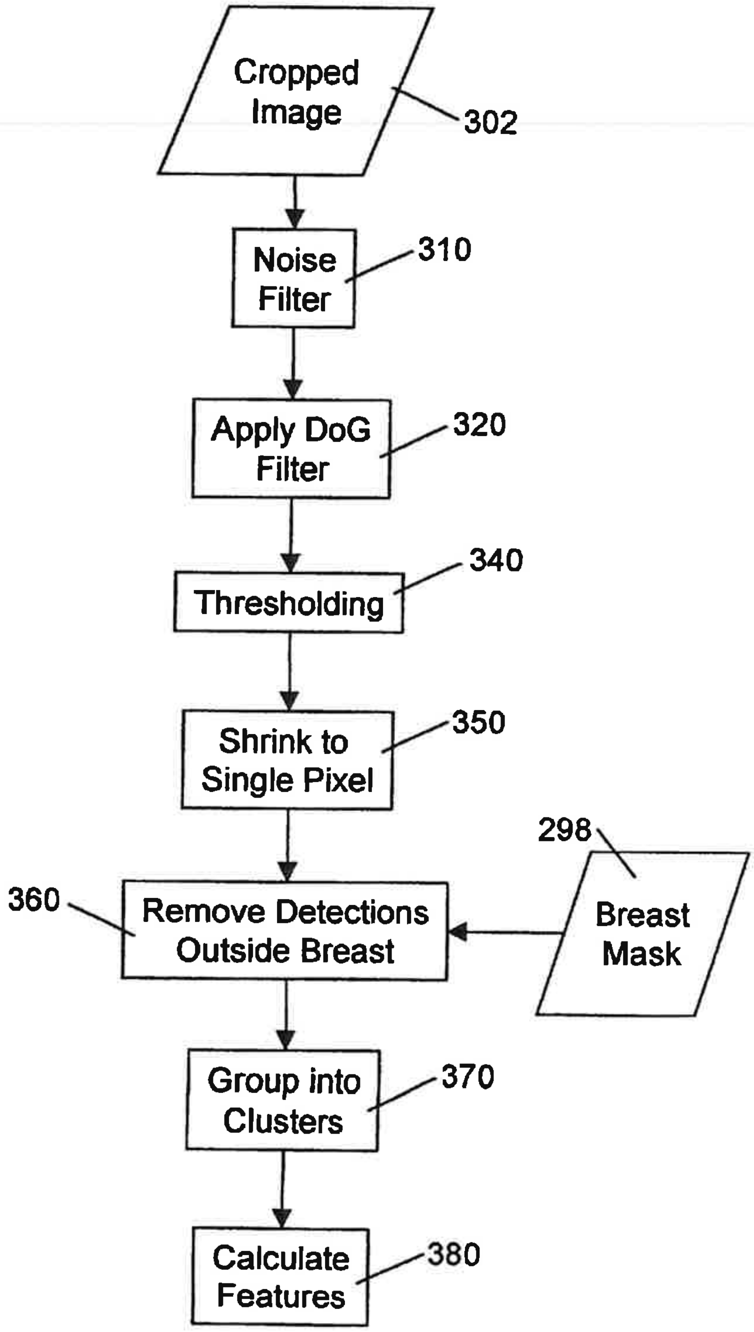

Cropped Image
302
Noise Filter
310
Apply DoG Filter
320
Thresholding
340
Shrink to Single Pixel
350
298
360
Remove Detections Outside Breast
Breast Mask
Group into Clusters
370
Calculate Features
380

Fig. 12

	p(x,y-1)	
p(x-1,y)	p(x,y)	p(x+1,y)
	p(x,y+1)	

Fig. 13

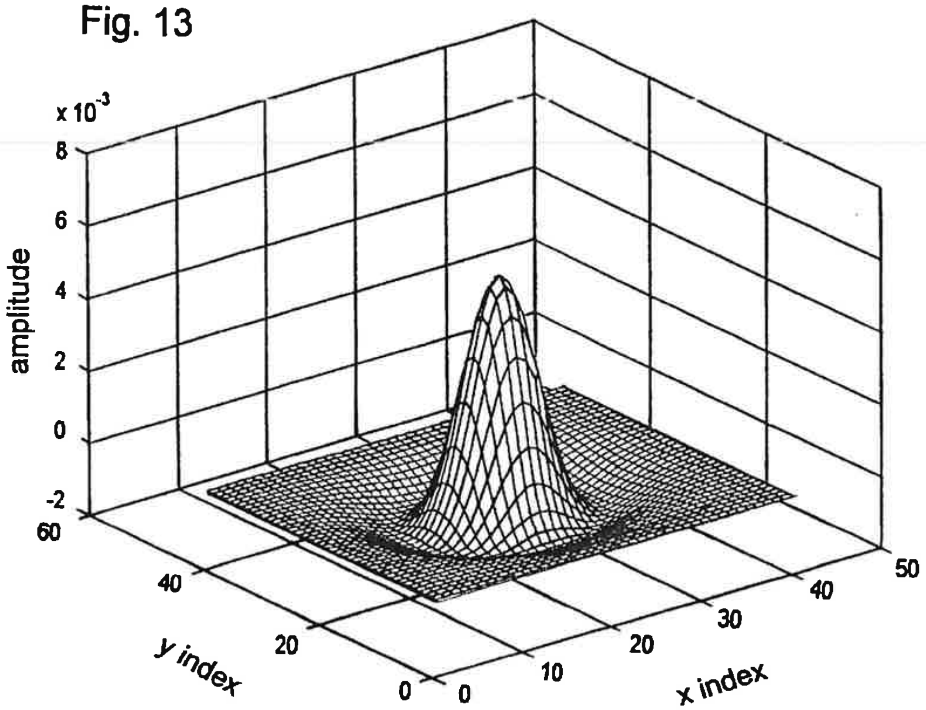

x 10^-3
8
6
4
2
0
-2
amplitude
60
40
20
0
y index
0
10
20
30
40
50
x index

Fig. 14

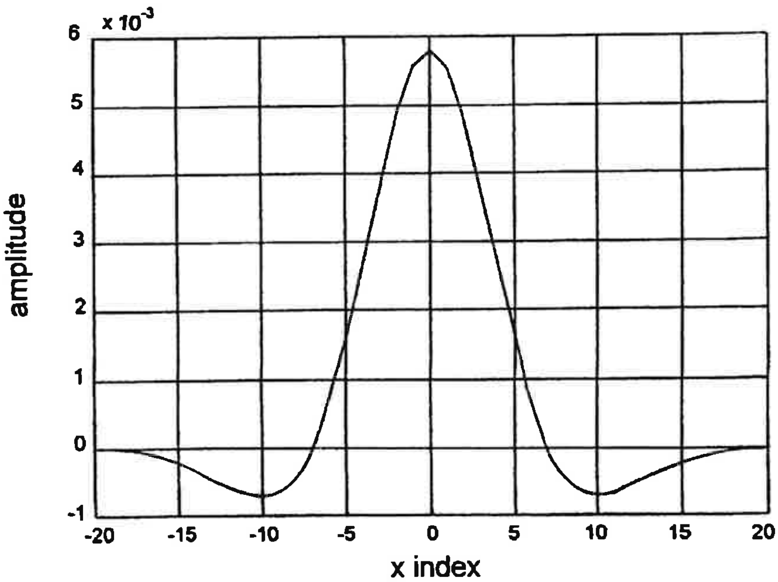

x 10^-3
6
5
4
3
2
1
0
-1
amplitude
-20
-15
-10
-5
0
5
10
15
20
x index

Fig. 15

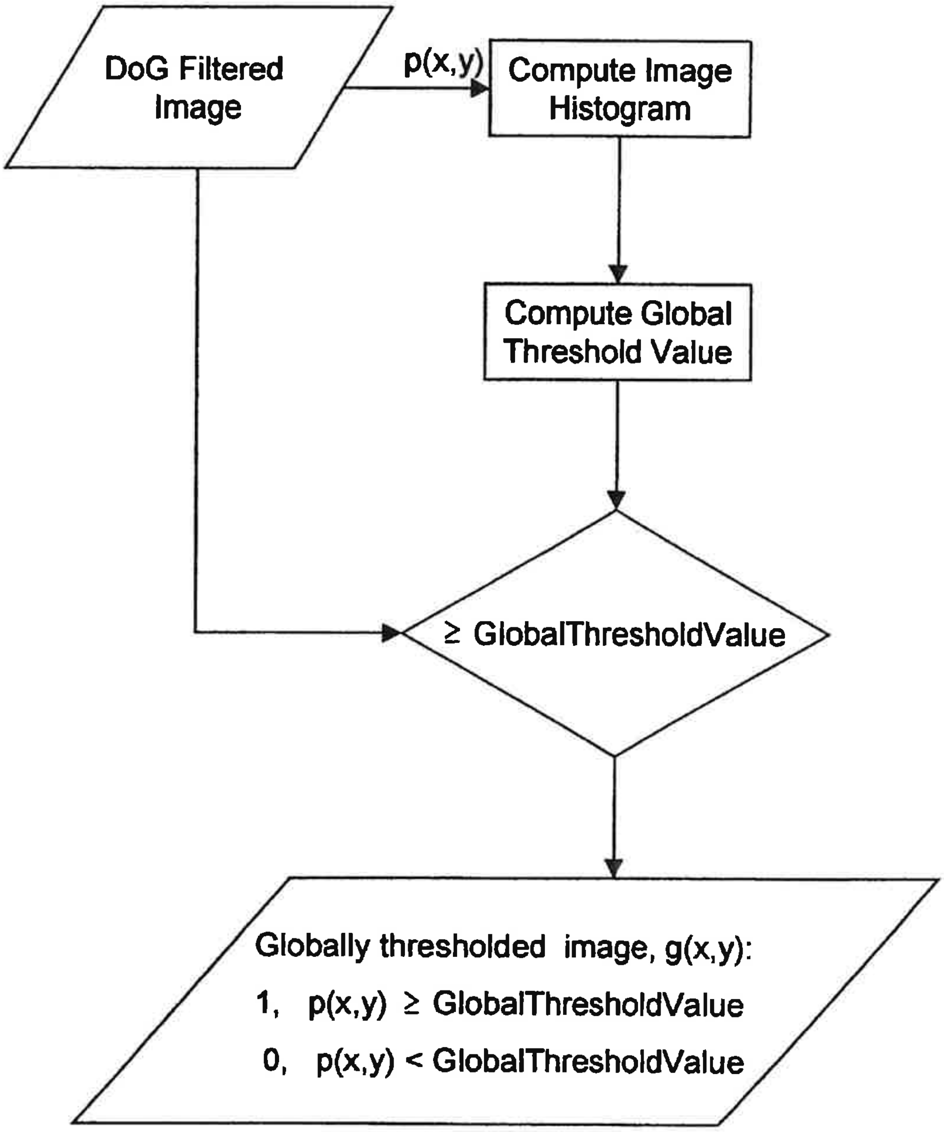

DoG Filtered Image
p(x,y)
Compute Image Histogram
Compute Global Threshold Value
≥ GlobalThresholdValue
Globally thresholded image, g(x,y):
1, p(x,y) ≥ GlobalThresholdValue
0, p(x,y) < GlobalThresholdValue

Fig. 16

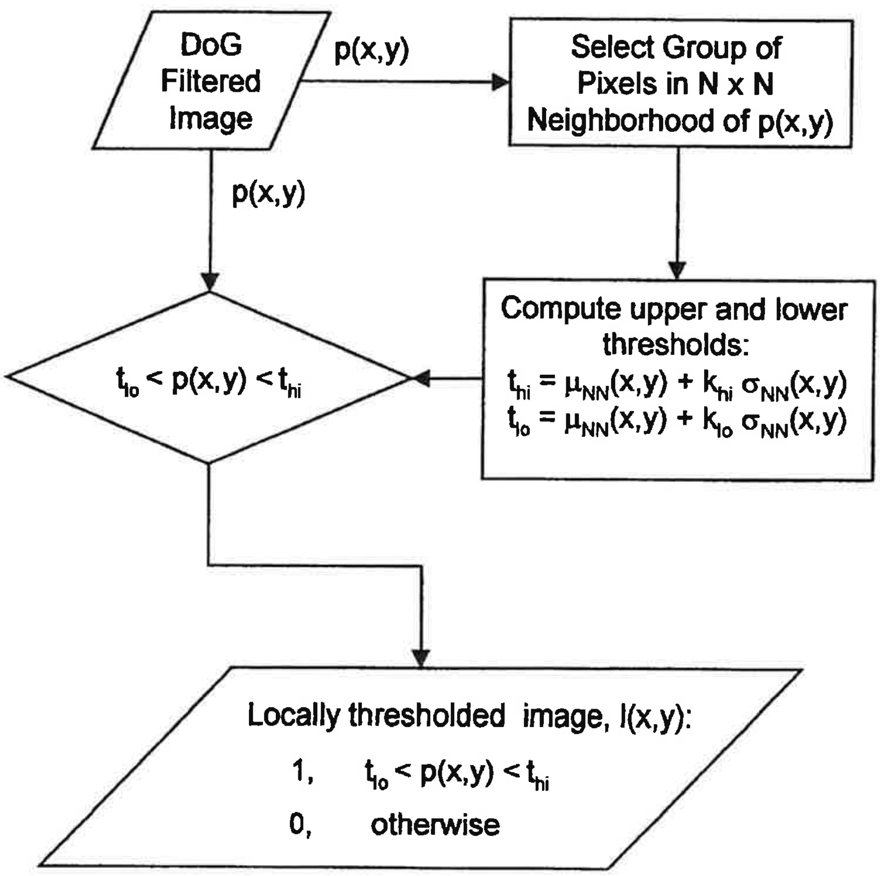

DoG Filtered Image
p(x,y)
Select Group of Pixels in N x N Neighborhood of p(x,y)
p(x,y)
Compute upper and lower thresholds:
$t_{hi} = \mu_{NN}(x,y) + k_{hi}\,\sigma_{NN}(x,y)$
$t_{lo} = \mu_{NN}(x,y) + k_{lo}\,\sigma_{NN}(x,y)$
$t_{lo} < p(x,y) < t_{hi}$
Locally thresholded image, I(x,y):
1, $t_{lo} < p(x,y) < t_{hi}$
0, otherwise

Fig. 17

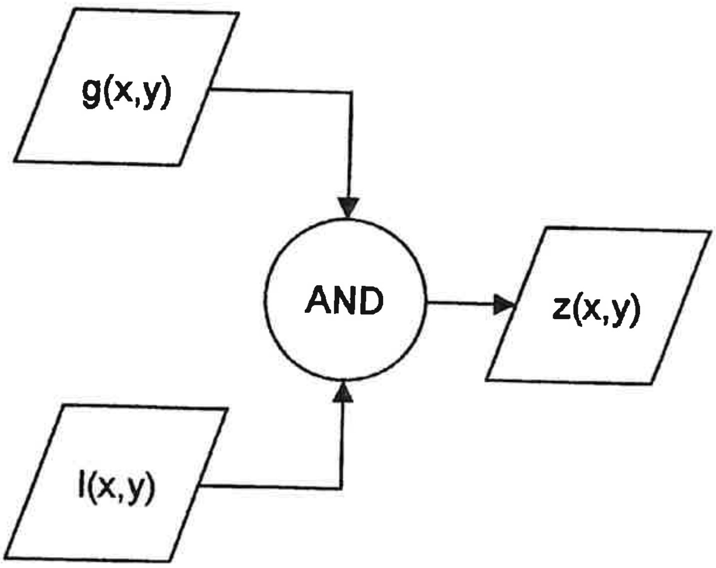


Fig. 18

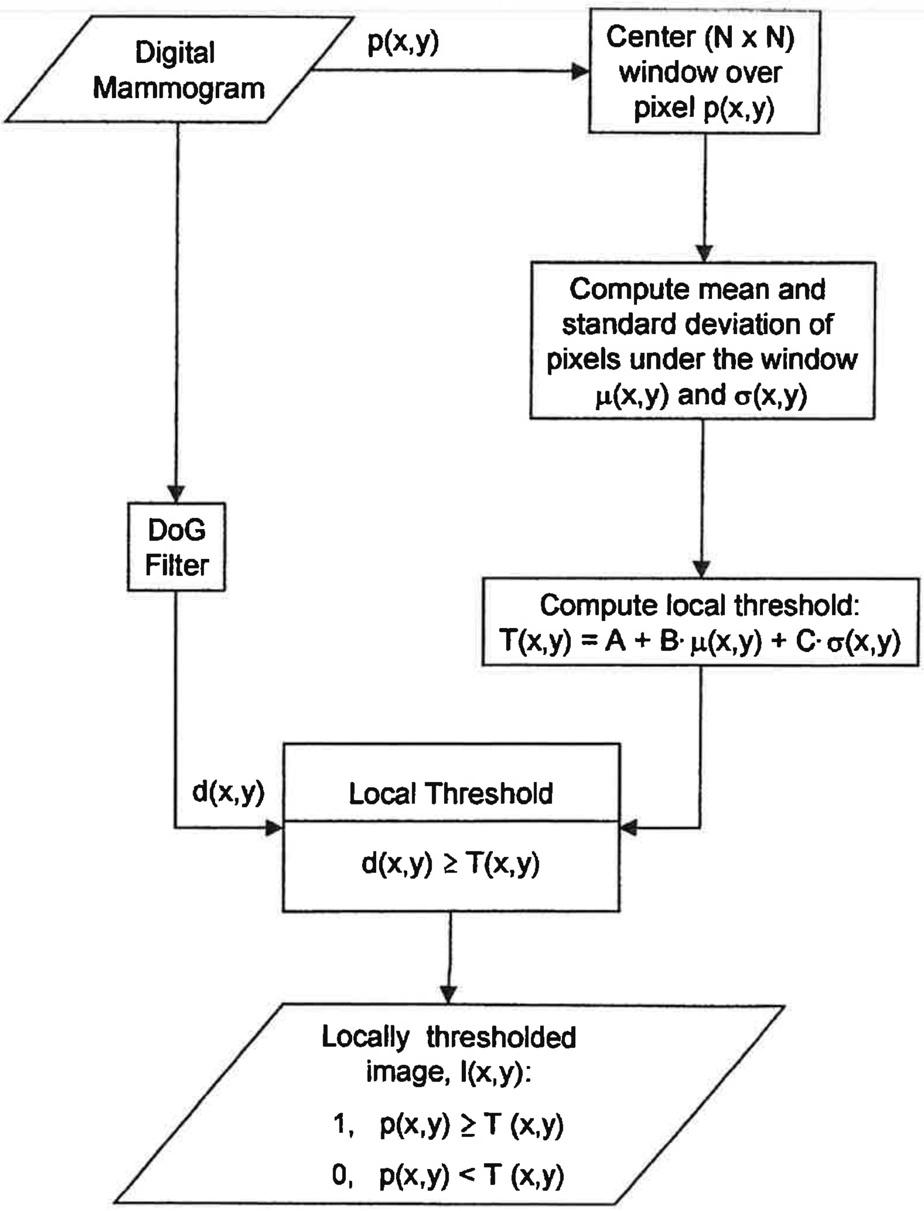

Digital Mammogram
p(x,y)
Center (N x N) window over pixel p(x,y)
Compute mean and standard deviation of pixels under the window μ(x,y) and σ(x,y)
DoG Filter
Compute local threshold: T(x,y) = A + B·μ(x,y) + C·σ(x,y)
d(x,y)
Local Threshold
d(x,y) ≥ T(x,y)
Locally thresholded image, I(x,y):
1, p(x,y) ≥ T (x,y)
0, p(x,y) < T (x,y)

Fig. 19

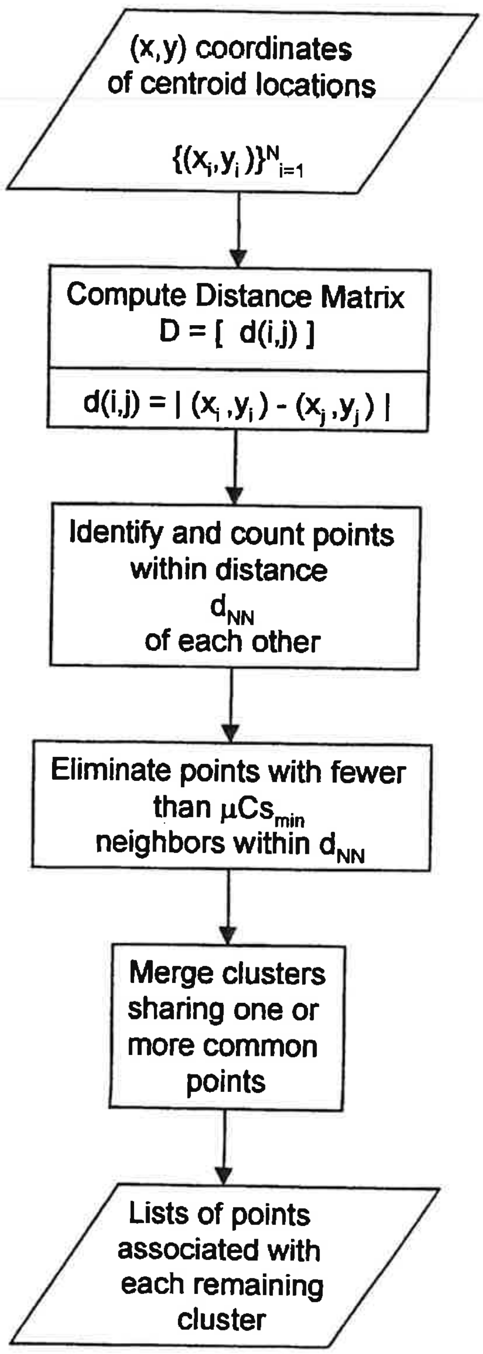

(x,y) coordinates of centroid locations
$\{(x_i, y_i)\}^N_{i=1}$
Compute Distance Matrix
D = [d(i,j)]
$d(i,j) = | (x_i, y_i) - (x_j, y_j) |$
Identify and count points within distance d_{NN} of each other
Eliminate points with fewer than μCs_{min} neighbors within d_{NN}
Merge clusters sharing one or more common points
Lists of points associated with each remaining cluster

Fig. 20

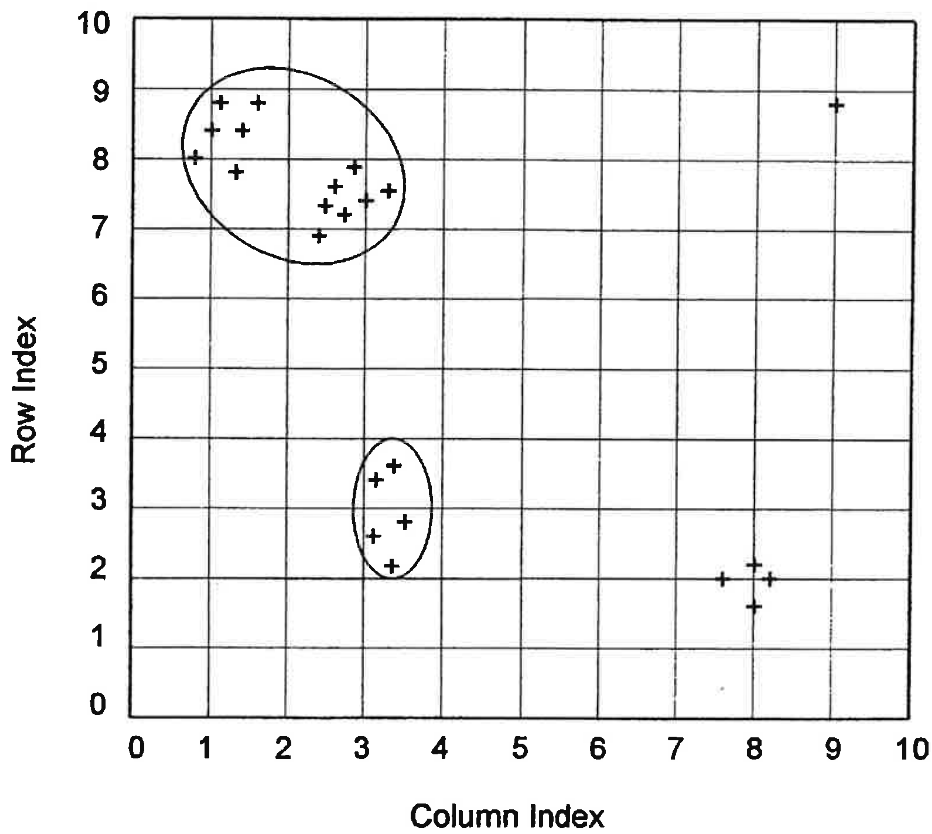


Fig. 21

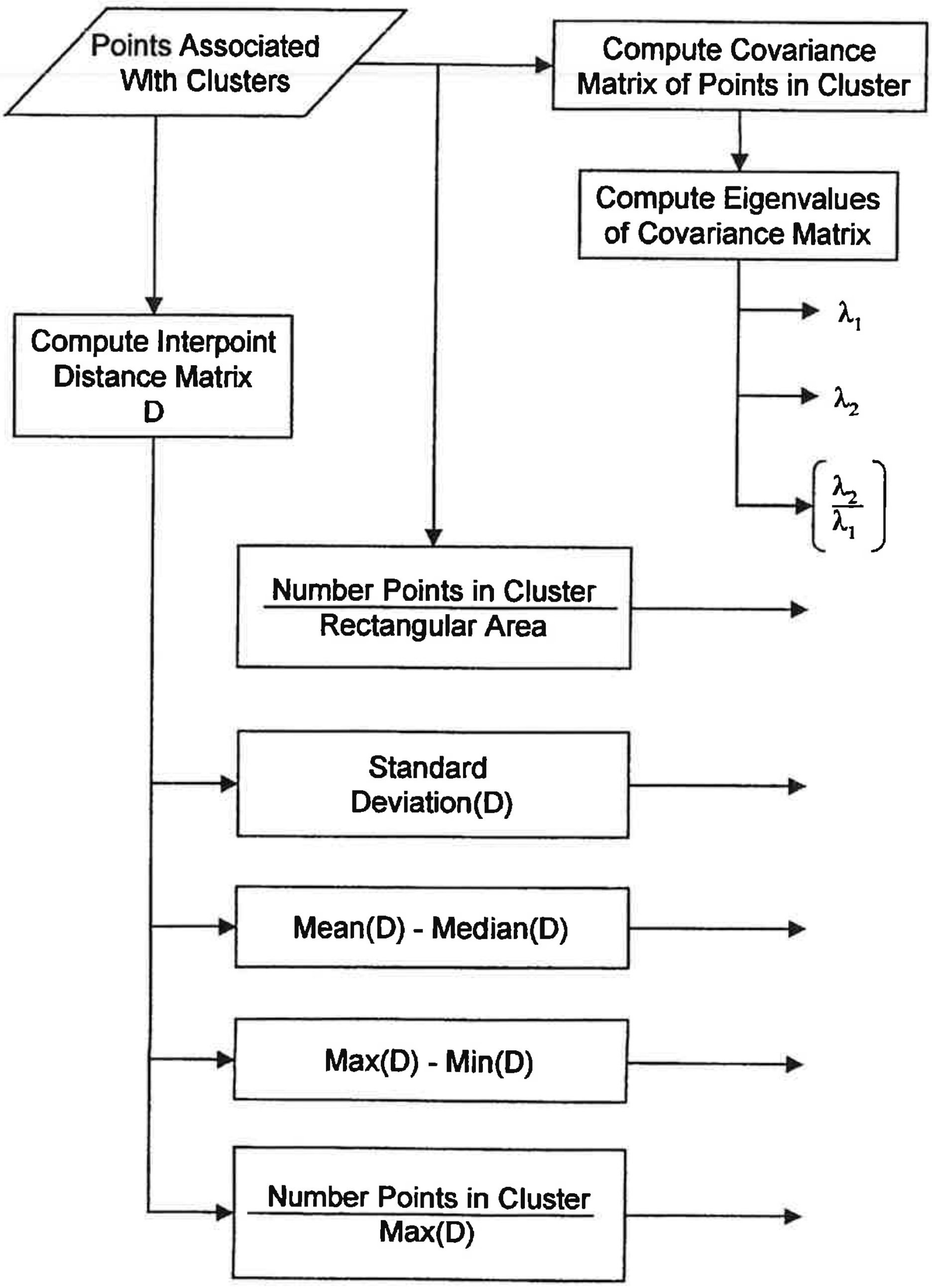

Points Associated With Clusters
Compute Covariance Matrix of Points in Cluster
Compute Eigenvalues of Covariance Matrix
λ_1
λ_2
$\left(\frac{\lambda_2}{\lambda_1}\right)$
Compute Interpoint Distance Matrix D
$\frac{\text{Number Points in Cluster}}{\text{Rectangular Area}}$
Standard Deviation(D)
Mean(D) - Median(D)
Max(D) - Min(D)
$\frac{\text{Number Points in Cluster}}{\text{Max(D)}}$

Fig. 22

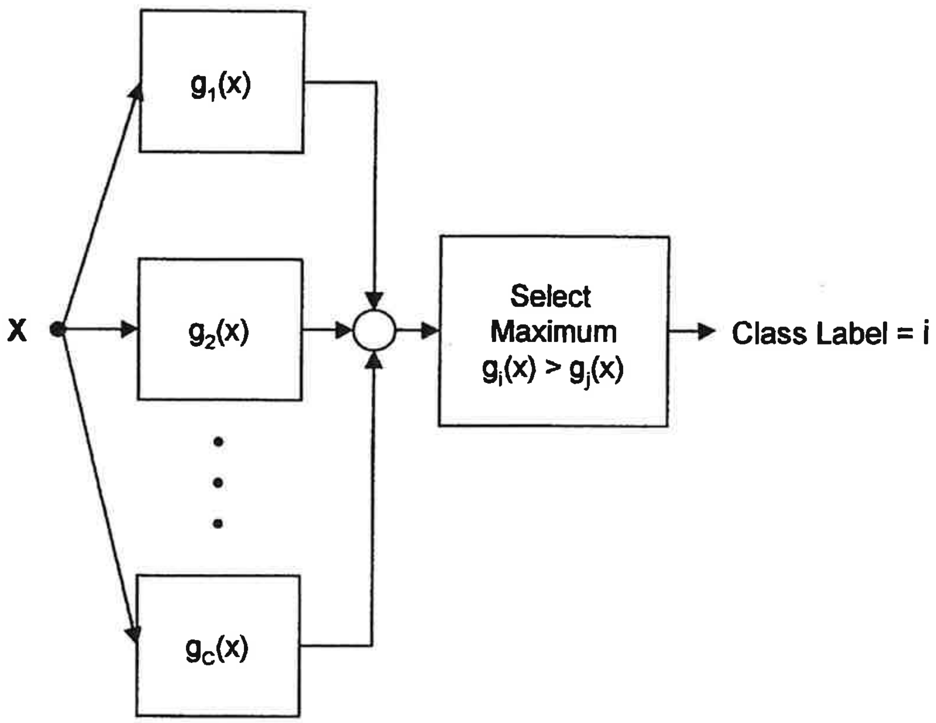


Fig. 23

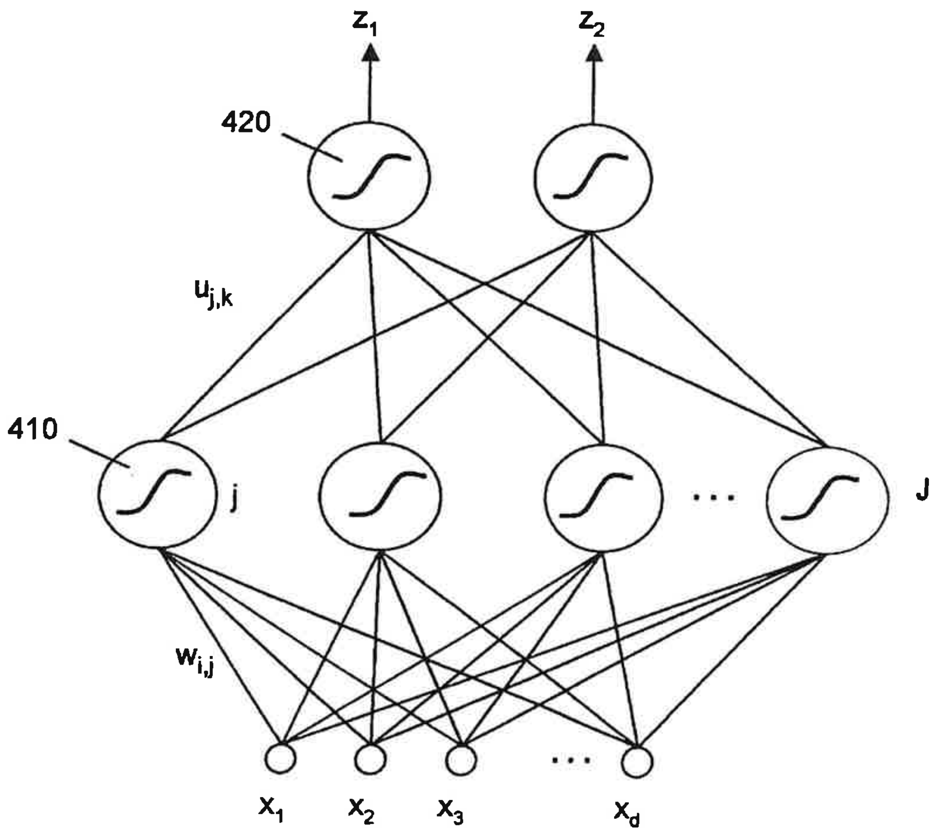


Fig. 24

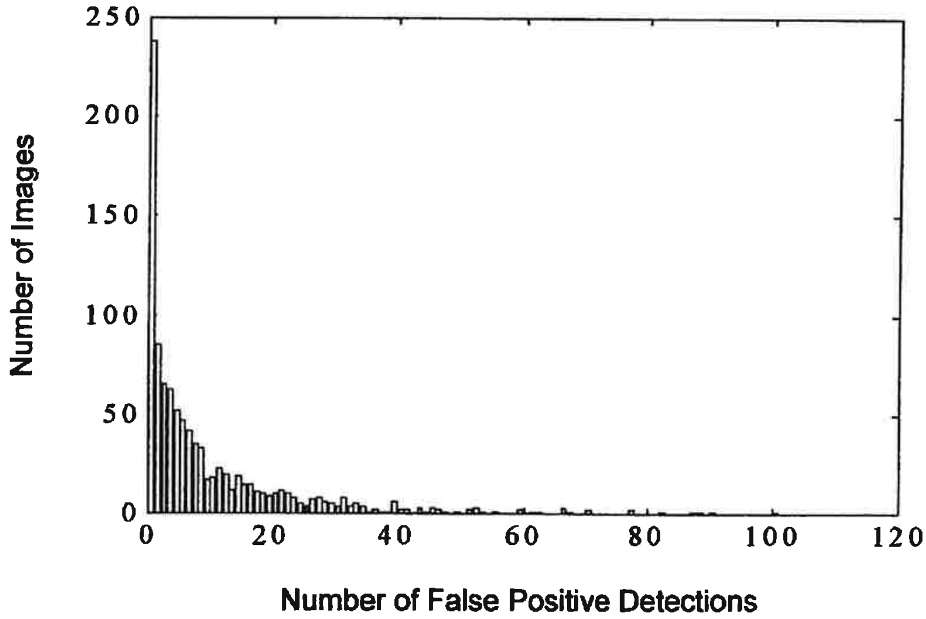

250
200
150
100
50
0
Number of Images
0
20
40
60
80
100
120
Number of False Positive Detections

Fig. 25

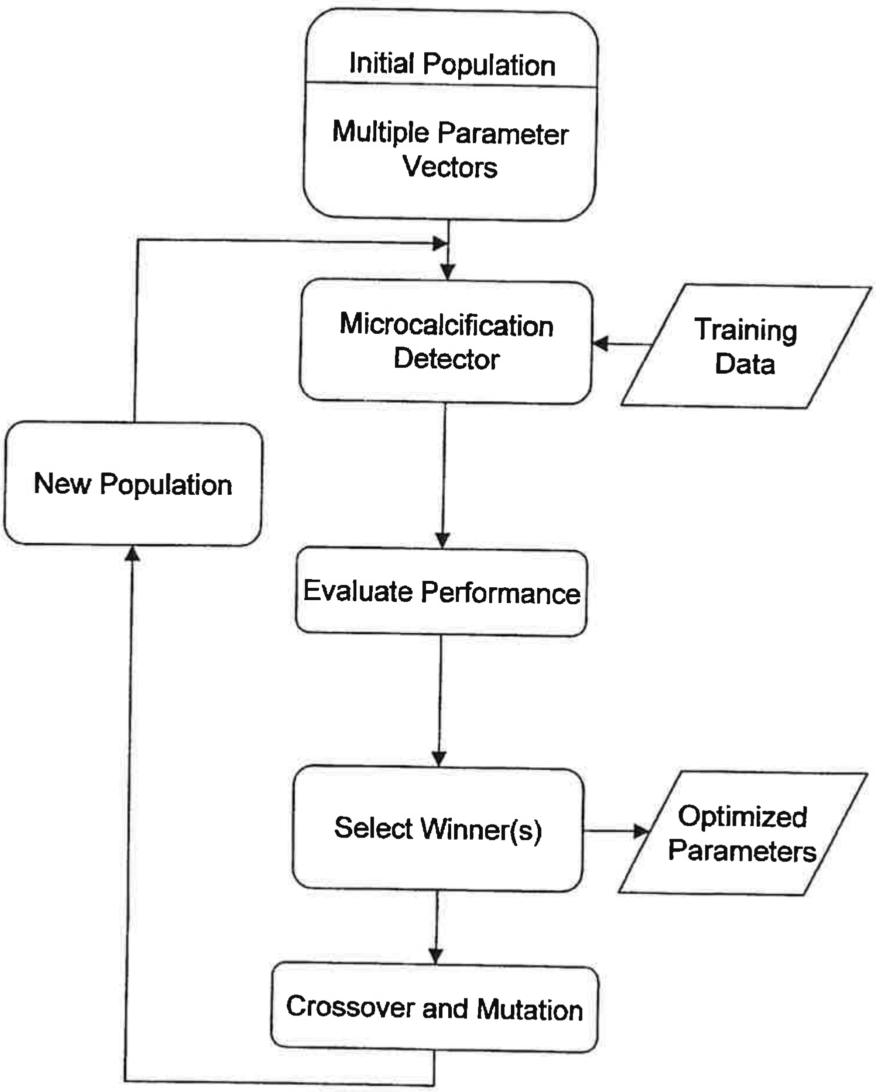

Initial Population
Multiple Parameter Vectors
Microcalcification Detector
Training Data
New Population
Evaluate Performance
Select Winner(s)
Optimized Parameters
Crossover and Mutation

Fig. 26

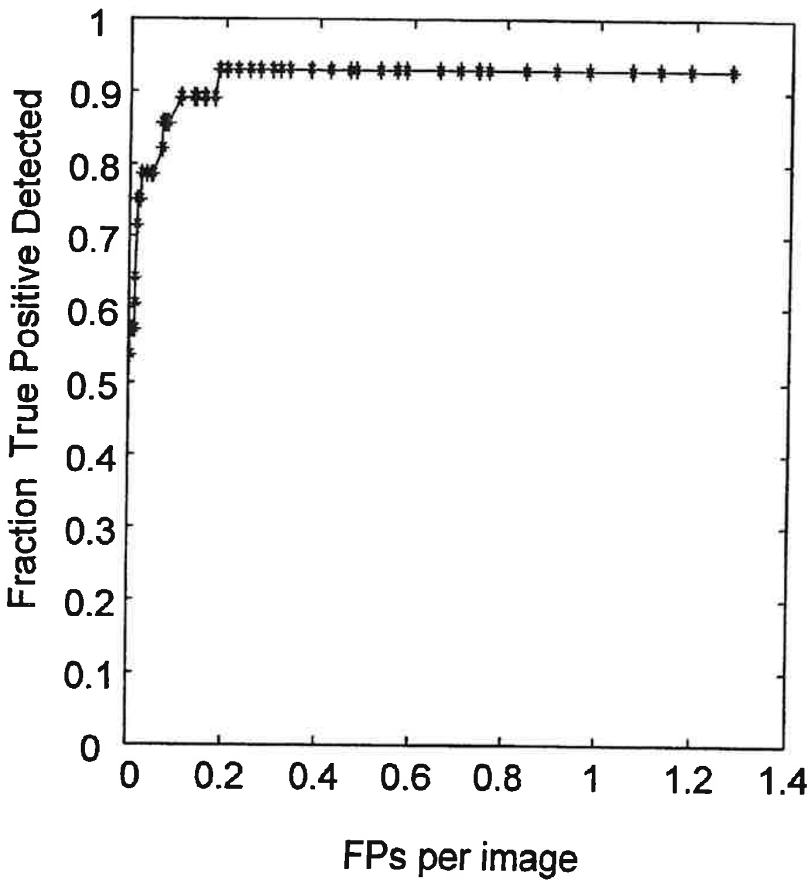

Fraction True Positive Detected
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
0.2
0.4
0.6
0.8
1
1.2
1.4
FPs per image

Fig. 27

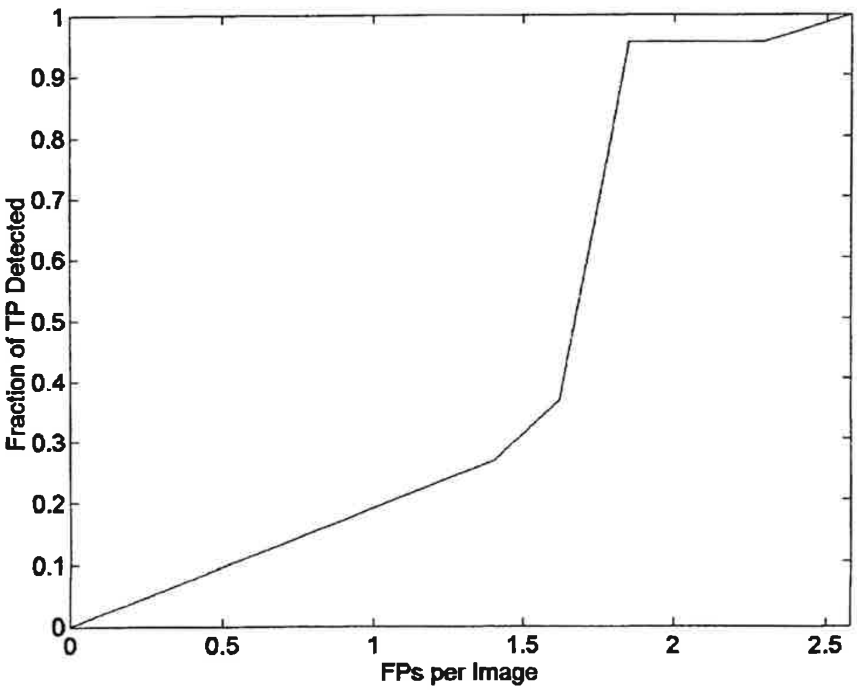

1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Fraction of TP Detected
0
0.5
1
1.5
2
2.5
FPs per Image

Fig. 28

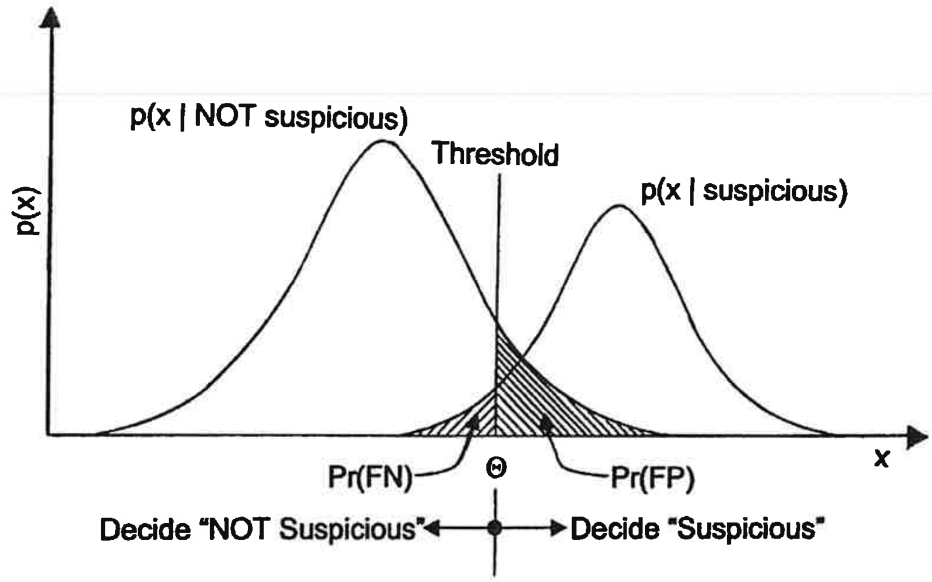

p(x | NOT suspicious)
Threshold
p(x | suspicious)
p(x)
x
Pr(FN)
Θ
Pr(FP)
Decide "NOT Suspicious"
Decide "Suspicious"

Fig. 29

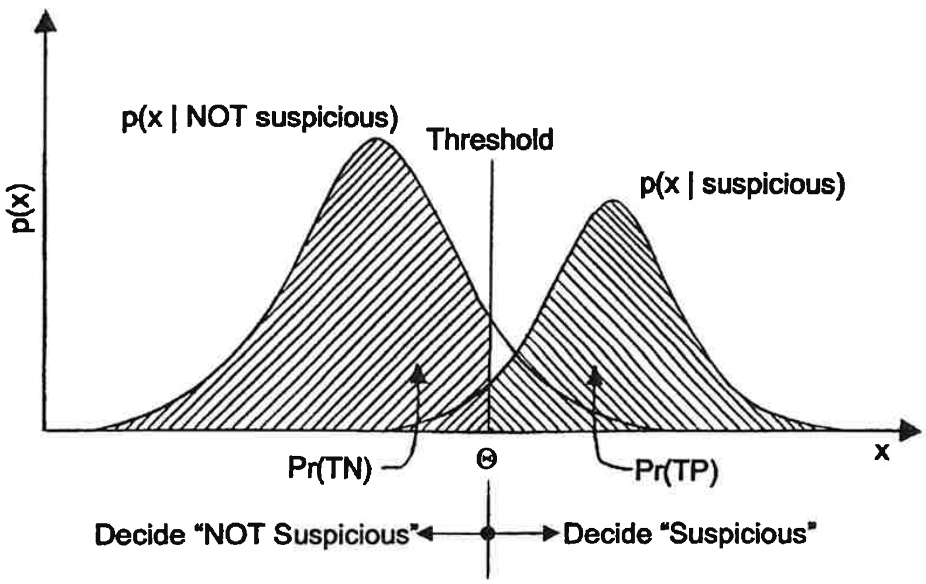

p(x | NOT suspicious)
Threshold
p(x | suspicious)
p(x)
x
Pr(TN)
Θ
Pr(TP)
Decide "NOT Suspicious"
Decide "Suspicious"

Fig. 30

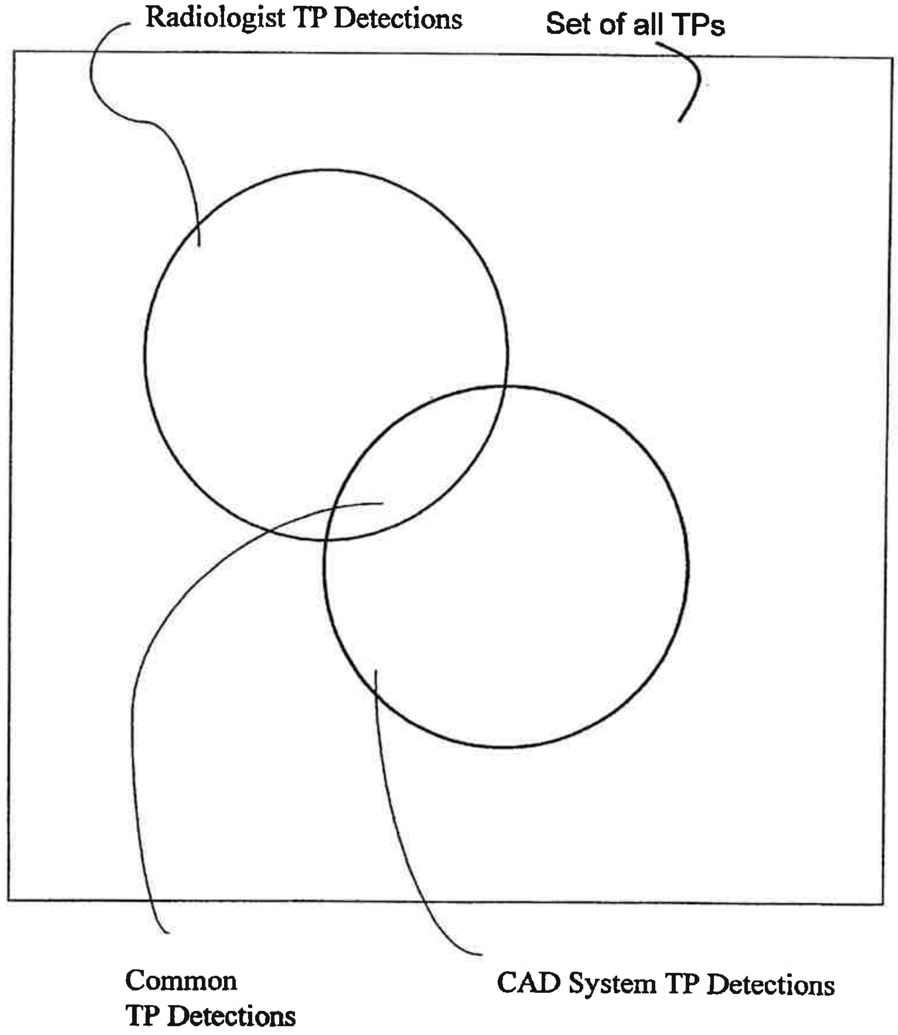
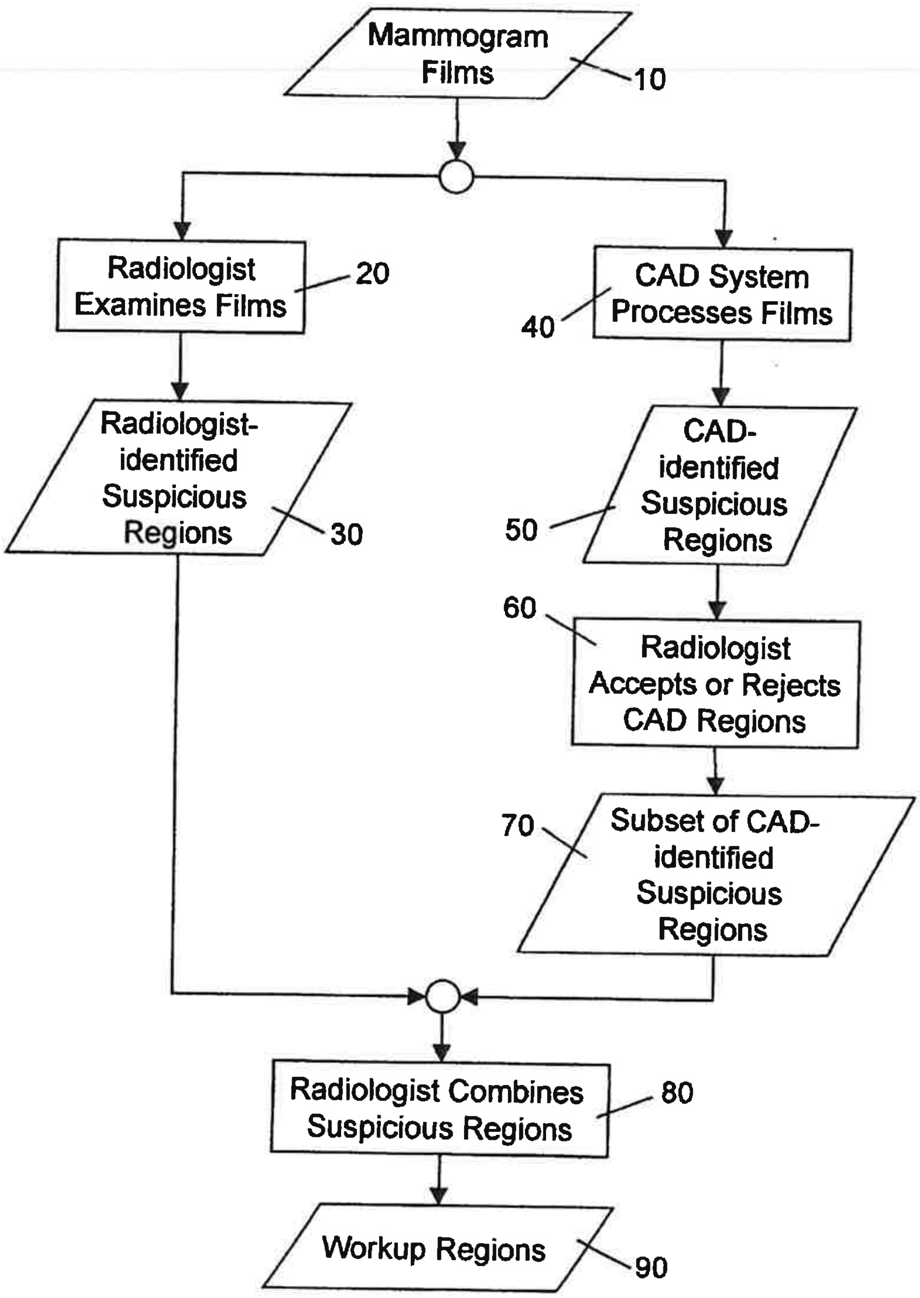

Radiologist TP Detections
Set of all TPs
Common
TP Detections
CAD System TP Detections

Fig. 31


Mammogram Films
10
Radiologist Examines Films
20
CAD System Processes Films
40
Radiologist-identified Suspicious Regions
30
CAD-identified Suspicious Regions
50
60
Radiologist Accepts or Rejects CAD Regions
70
Subset of CAD-identified Suspicious Regions
Radiologist Combines Suspicious Regions
80
Workup Regions
90

METHOD AND SYSTEM FOR SEGMENTING DESIRED REGIONS IN DIGITAL MAMMOGRAMS

CROSS-REFERENCE TO RELATED APPLICATIONS

This application is a division of application Ser. No. 09/141,802, filed Aug. 28, 1998, now U.S. Pat. No. 5,999, 639, which claims the benefit of U.S. Provisional Application Ser. No. 60/057,801, filed Aug. 28, 1997, U.S. Provisional Application Ser. No. 60/066,996, filed Nov. 28, 1997, and U.S. Provisional Application Ser. No. 60/076,760, filed Mar. 3, 1998, the entire disclosures of all of which are incorporated herein by reference.

BACKGROUND OF THE INVENTION

1. Field of the Invention

This invention relates to a method and system for automated detection of clustered microcalcifications from digital images without reduction of radiologist sensitivity.

2. Discussion of Background

Mammography, along with physical examination, is the current procedure of choice for breast cancer screening. Screening mammography has been responsible for an estimated 30 to 35 percent reduction in breast cancer mortality rates. However, in 1996 approximately 185,700 new breast cancer cases were diagnosed and 44,300 women died from this disease. Women have about a 1 in 8 chance of being diagnosed with breast cancer, and 1 in 30 will die of this disease in her lifetime.

Although mammography is a well-studied and standardized methodology, for 10 to 30 percent of women diagnosed with breast cancer, their mammograms were interpreted as negative. Additionally, only 10 to 20 percent of patients referred for biopsy based on mammographic findings prove to have cancer. Further, estimates indicate the malignancies missed by radiologists are evident in two-thirds of the mammograms retrospectively. Missed detections may be attributed to several factors including: poor image quality, improper patient positioning, inaccurate interpretation, fibroglandular tissue obscuration, subtle nature of radiographic findings, eye fatigue, or oversight.

To increase sensitivity, a double reading has been suggested. However, the growing increase in the number of screening mammograms makes this option unlikely. Alternatively, a computer-aided diagnosis (CAD or CADx) system may act as a "second reader" to assist the radiologist in detecting and diagnosing lesions. Several investigators have attempted to analyze mammographic abnormalities with digital computers. However, the known studies are believed to have achieved rates of true-positive detections versus false-positive detections that are undesirably low.

Microcalcifications represent an ideal target for automated detection because subtle microcalcifications are often the first and sometimes the only radiographic findings in early, curable breast cancers, yet individual microcalcifications in a suspicious cluster have a fairly limited range of radiographic appearances. Between 30 and 50 percent of breast carcinomas detected radiographically demonstrate microcalcifications on mammograms, and between 60 and 80 percent of breast carcinomas reveal microcalcifications upon microscopic examination. Any increase in the detection rate of microcalcifications by mammography will lead to further improvements in its efficacy in the detection of early breast cancer.

Although the promise of CAD systems is to increase the ability of physicians to diagnose cancer, the problem is that all CAD systems fail to detect some regions of interest that could be found by a human interpreter. However, human interpreters also miss regions of interest that are subsequently shown to be indicators of cancers. Missing a region that is associated with a cancer is termed a false negative error while associating a normal region with a cancer is termed a false positive error.

It is not yet clear how CAD system outputs are to be incorporated by practicing radiologists into their mammographic analyses. No existing CAD system can claim to find all of the suspicious regions detected by an average radiologist, and they tend to have unacceptably high false positive error rates. However, CAD systems are capable of finding some suspicious regions that may be missed by radiologists.

SUMMARY OF THE INVENTION

Accordingly, an object of this invention is to provide a method and system for automated detection of clustered microcalcifications from digital mammograms.

These and other objects are achieved according to the invention by providing a novel method and system for automated detection of clustered microcalcifications from digital mammograms in which a digital mammogram is obtained, parameters necessary for cropping the digital mammogram image are optimized, the digital mammogram is cropped based on the optimized cropping parameters to select breast tissue for further analysis, parameters necessary for detecting clustered microcalcifications are optimized, and clustered microcalcifications in the cropped digital mammogram are detected based on the optimized clustered microcalcification detection parameters.

The detected clustered microcalcifications are then stored as a detections image, the detections image is processed for display, and a computer-aided detection image is produced for review by a radiologist.

The radiologist first reviews the original mammograms and reports a set of suspicious regions of interest, S1. A CAD system, or more particularly, the CAD system of the invention, operates on the original mammogram and reports a second set of suspicious detections or regions of interest, S2. The radiologist then examines the set S2, accepts or rejects members of S2 as suspicious, thus forming a third set of suspicious detections, S3, that is a subset of set S2. The radiologist then creates a fourth set of suspicious detections, S4, that is the union of sets S1 and S2, for subsequent diagnostic workups. CAD system outputs are thereby incorporated with the radiologist's mammographic analysis in a way that optimizes the overall sensitivity of detecting true positive regions of interest.

Other objects and advantages of the invention will be apparent from the following description, the accompanying drawings and the appended claims.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a flow diagram illustrating the automated system for the detection of clustered microcalcifications in a digital mammogram;

FIGS. 2 and 3 are flow diagrams illustrating the autocropping method and system of the invention;

FIGS. 4–10 are flow diagrams illustrating in more detail the autocropping method and system of the invention;

FIG. 11 is a flow diagram illustrating in greater detail the clustered microcalcification detector of the invention;

FIG. 12 is a schematic diagram illustrating a 3×3 cross-shaped median filter of the invention;

FIG. 13 is a three-dimensional plot of a Difference of Gaussians (DoG) filter kernel;

FIG. 14 is a cross-sectional view through the center of the DoG filter kernel of FIG. 13;

FIG. 15 is a flow diagram illustrating the global thresholding portion of the microcalcification detection system;

FIG. 16 is a flow diagram illustrating the dual local thresholding of the invention;

FIG. 17 is a flow diagram illustrating combining the results of global and dual-local thresholding;

FIG. 18 is a flow diagram illustrating the sloping local thresholding of the invention;

FIG. 19 is a flow diagram illustrating the clustering method of the invention;

FIG. 20 is a schematic diagram illustrating the clustering method of the invention;

FIG. 21 is a flow diagram illustrating the feature computation process of the invention;

FIG. 22 is a flow diagram illustrating a classifier having one discriminant function per class;

FIG. 23 is a schematic diagram illustrating a multi-layer perceptron neural network for a two-class classifier;

FIG. 24 is a histogram of testing results after detection and classification;

FIG. 25 is a flow diagram illustrating the parameter optimization method of the invention;

FIG. 26 is a plot of a free response receiver operating characteristic curve of the invention before classifying detections;

FIG. 27 is a plot of a free response receiver operating characteristic curve of the invention after classifying detections;

FIG. 28 is a plot of probability density functions showing the relationship between the probabilities of false negative and false positive detections;

FIG. 29 is a plot of probability density functions showing the relationship between the probabilities of true negative and true positive detections;

FIG. 30 is a Venn diagram showing the relationship between radiologist and CAD system detections; and

FIG. 31 is a flow diagram illustrating a method for incorporating computer-aided diagnosis detections with those of a human interpreter for optimal sensitivity.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENT

Referring now to the drawings, wherein like reference numerals designate identical or corresponding parts throughout the several views, and more particularly to FIG. **1** thereof, there is shown a flow diagram illustrating a sequence of steps performed in order to detect the locations of clusters of microcalcifications within a digital mammogram.

In a first step **100**, a digital mammogram is obtained using hardware such as digital mammography systems, or by digitizing mammography films using laser or charge-coupled device (CCD) digitizers. In an optimized cropping step **200**, a rectangular analysis region containing breast tissue is segmented from the digital mammogram image and a binary mask corresponding to the breast tissue is created for use in later processing steps to decrease the time required for processing the mammogram image. The binary mask is also used to limit detections to areas of the image containing breast tissue.

Clustered microcalcifications are detected in a clustered microcalcification detection step **300**. After first filtering the cropped image with a median filter to reduce noise, the image is filtered using an optimized difference of Gaussians (DoG) filter to enhance the microcalcifications. The DoG-filtered image is then subjected to optimized threshold tests to detect potential microcalcifications. The detected microcalcifications are shrunk to single-pixel representations and detections outside of the breast area are removed. The remaining microcalcifications are grouped into clusters. Features are then computed for the clusters. Detected clusters are classified as either suspicious or non-suspicious in a classification step **400**.

The parameters used by the autocropping, clustered microcalcification detection, and classification steps **200**, **300**, **400** are optimized in a parameter-optimizing step **500**. The parameters are optimized by parameter-optimizing means that uses a genetic algorithm (GA) so as to maximize the true-positive detection rate while minimizing the false-positive detection rate. Of course, other optimization schemes may be used as well.

The detected clustered microcalcifications are stored in a list of image coordinates. The detection results are processed in a processing step **600** by simply adding an offset to each of the microcalcification coordinates to account for translation of the coordinates incurred as a result of the cropping procedure. Detected clustered microcalcifications are indicated on the digital mammogram by means of rectangles drawn around the clustered microcalcifications in a display step **700**. Other indicators may be used such as, for example, arrows pointing to suspected microcalcifications, or ellipses around suspected microcalcifications.

Acquiring a Digital Representation of a Mammogram

One method of obtaining digital mammograms comprises digitizing radiologic films by means of a laser or charge-coupled device (CCD) scanner. Digital images obtained in this manner typically have a sample spacing of about 100 μm per pixel, with a gray-level resolution of 10 to 12 bits per pixel. In one embodiment of the present invention, radiologic films are scanned using a Model CX812T digitizer manufactured by Radiographic Digital Imaging of Compton, California, to produce digital images having 50 μm spacing per pixel and 12 bits of gray-level resolution per pixel. Another possible input source for digital images is a digital mammography unit from Trex Medical Corporation of Danbury, Connecticut, which has a spatial resolution of about 45 μm per pixel and a gray-level resolution of 14 bits per pixel.

The digital images are stored as digital representations of the original mammogram images on computer-readable storage media. In a preferred embodiment, the digital representations or images are stored on a 2 GB hard drive of a general-purpose computer such as a PC having dual Pentium II® microprocessors running at 200 MHZ, 512 MB of RAM memory, a ViewSonic PT813® monitor, a pointing device, and a Lexmark Optra S1625® printer. The system operates within a Windows NT® operating system.

Autocropping

As may be seen in FIGS. **2** and **3**, a digital mammogram image **190** is first cropped to segment an analysis region **296** from the image and produce a binary mask **298** corresponding to breast tissue in the analysis region. Preferably, the cropping is performed automatically, although it could be cropped manually. The image is cropped as a preliminary step because the breast tissue does not cover the whole radiographic film. Focusing the processing of the image on only that portion of the image which contains breast tissue reduces the time required to process the image. Also, other items appearing on the film, such as labels and patient information, are excluded from consideration, and false-positive indications lying outside of the breast tissue area are eliminated.

Referring to FIGS. **4** through **10**, the autocropping process will be described in detail. The image is first subsampled from 50 μm to 400 μm to reduce the amount of data to be processed in step **202**. Of course, the image may be down-sampled to other resolutions as desired. Not all of the original image data is needed to reliably segment the breast tissue from the remainder of the image. Subsampling every eighth pixel in both the horizontal and vertical directions reduces the amount of data by 64 times. For purposes of segmenting the breast tissue from the rest of the image, the consequent loss of resolution is immaterial.

A white border twenty pixels in width is added around all sides of the subsampled image in step **204**. White corresponds to the maximum pixel value possible given the number of bits used to represent each pixel. For images having 12 bits of gray-scale resolution, the maximum gray-scale value is 4095. The bordered image is then thresholded in step **206** with a relatively high threshold value such that most of the breast tissue is guaranteed to be less than the threshold to produce a binary image. In one embodiment of the invention, the threshold is set equal to a predetermined percentage of the gray-scale value of a pixel near the top middle portion of the image. The thresholded image is then inverted, that is, ones become zeroes and zeroes become ones, in step **208**. The inverted image is then dilated in step **210**. Dilation is a morphological operation in which each pixel in a binary image is turned on, that is, set to a value of one, if any of its neighboring pixels are on. If the pixel is already on, it is left on.

In step **212** the dilated image is cropped to the size of the largest blob. Blobs are contiguous groups of pixels having the value one. This step **212** removes bright borders from the subsampled mammogram representation while ensuring that none of the breast area is reduced. Other techniques that threshold to find the border have a very difficult time dealing with bright areas in the breast adjacent to the border such as, for example, when breast implants are visible in the image. Pixels from the original image, resulting from step **202**, corresponding to the locations of the pixels in the cropped blob, are selected for subsequent processing. Note that this is a simple subset of pixels from the input image.

The image from step **212** is histogram equalized in step **214**. The average brightness of the image will vary widely from mammogram to mammogram. Moreover, different digitizers having different optical density characteristics are an additional source of variability in brightness levels in the digital representation of the mammogram. The breast mask that is the output of the autocropper is mainly defined by means of a region-growing algorithm that requires a single contrast setting to work properly. However, it has been determined experimentally that a single contrast setting will not work for a wide range of image inputs. Therefore, each image is mapped into a normalized image space using an automatic histogram enhancement process, after which a single contrast setting works well.

First, a histogram of the image is obtained. Typically, most of the data in the breast area will be in the lower histogram bins (corresponding to gray-scale values of about 0–1000), with borders and labels being in the higher bins (corresponding to gray-scale values of about 4000–4095) for 12-bit data. The upper and lower bin values that contain the typical breast data are determined. The lower bin value is the first highest peak encountered when going from the lowest gray-scale value toward the highest gray-scale value. The upper bin is the last zero-value bin encountered when going from the highest gray-scale level toward the lowest gray-scale value. Then the data are reduced to an eight-bit representation and linearly stretched over the range of the data type. For example, values in the lower bins are set to zero. Values of data in the upper bins are set to **255**. The rest of the data are then linearly mapped between the lower and upper bins.

After the image has been histogram equalized, the equalized image may be considered to be a matrix. The image matrix is divided into left and right halves, of equal size if possible, and the brighter side is selected in a step **216**. The sums of all the pixels in the left and right halves are computed. The sum values are then compared and the side having the greater sum is the brighter side.

Prior to region growing the brighter side, algorithm variables are initialized in step **218**. The size of the region-grown mask is preliminarily checked in step **220**. If it is large enough, then the mask is acceptable. Otherwise, processing continues to find the mask. The side of the image to be region grown is selected in step **222**. In step **224** this region is searched to find its maximum gray-scale value. This maximum value is used to find a pixel to start a region-growing algorithm. Region growing is the process of grouping connected pixels sharing some like characteristic. The choice of characteristic influences the resultant region. The input to a region growing function is a gray-scale image and a starting point to begin growing. The output is a binary image with ones indicating pixels within the grown region, i.e., blobs. Region growing will create a single blob, but that blob may have within it internal holes, that is, pixels that are off. To grow a blob, each of the four nearest neighbors of a pixel of interest are looked at. The contrast ratio is computed for each nearest neighbor pixel. If the contrast ratio is less than a contrast ratio threshold, then the neighbor pixel is set to a one in a binary mask image. Otherwise, the neighbor pixel is set to zero. The region growing algorithm spirals outwardly from the starting or seed pixel, progressively looking at nearest neighbor pixels until done. To those skilled in the art, it is clear that other region growing algorithms may also be applied.

In step **226**, region growing begins with the pixel identified from the previous step **224** to produce a binary mask. The size of the mask resulting from step **226** is computed in step **228** and checked in step **230**. There may be three points of failure for this approach. First, the brightest point in the search region may be an artifact outside the breast. Therefore, if the resulting mask is not large enough (50 pixels), then the search region is moved closer to the side of the image and searched again. This is repeated three times, each time lowering the contrast value threshold. This corresponds to the path taken through steps **232** and **234**. Second, the side selection approach may be in error. Therefore, if a valid breast mask is not found in the first side searched, then the other side of the equalized image is searched. This corresponds to the path taken through steps **236** and **238**. Third, if a valid breast mask is not found on either side, then the whole breast is thresholded and the largest object is taken to be the breast mask in step **240**.

Since a constant contrast value is used in the region-growing algorithm, some masks will be too large. Typically, there will be "tails" along the edge of the digitized mammogram image where extra light leaked in while the original mammogram film was being digitized. The tails are reduced by applying a series of erodes and then a series of dilates to the image. Erosion is a morphological operation in which each pixel in a binary image is turned off unless all of its neighbors are on. If the pixel is already off, it is left off. But first, the holes in the mask must be filled in or the multiple erodes may break the mask into disjoint sections. Thus, holes in the mask are closed in step **242** by means of a majority operation. The majority operation is a morphological operation in which each pixel in a binary image is turned on if a majority of its neighboring pixels are on. If the pixel is already on, it is left on.

However, another problem is that some smaller breast masks can not undergo as many erodes as can larger breast masks. Therefore, as a fail-safe measure, the sum of the breast mask is taken before and after the erodes and dilates. If the size is reduced too much (i.e., by more than 50%), the original mask before the morphological operators is used. Thus, a duplicate copy of the mask is made in step **244** before the mask is eroded and dilated in steps **246** and **248**, respectively. The size of the resultant mask is then computed in step **250** and compared with the size of the mask from step **242** in step **252**. If the new size is less than half the old size, then the duplicate mask, from step **244**, is selected in step **254** for subsequent processing. Otherwise, the resultant mask from step **248** is used.

The original image (from step **202**) is then cropped to the size of the breast mask just found (either from step **242** or step **248**) in step **256**. In case the resulting mask is too small for subsequent processing, a crop adjustment is always made in step **258**. The adjustment comes in the form of increasing the size of the breast mask bounding box by including additional pixels from the original image in the cropped image.

The cropped image is then automatically histogram enhanced in step **260** as previously described above in connection with step **214**. This enhanced image is passed through a loose region growing step **262** to produce a generous mask. This means that the image is subjected to a lower threshold to yield more "on" pixels. This mask is then subjected to hole-closing, eroding, and dilating in steps **264**, **266**, and **268**, respectively, as above, but to a lesser degree.

The same steps described above are repeated one final time in steps **270** through **276**, but the crop adjustments are less and the contrast value is increased for a tight region growing step **276**. This tight region growing step **276** can afford the higher contrast value since it will be region growing in just the cropped image. This results in a parsimonious estimate of breast tissue. The resulting mask is segmented to find the largest object in step **278** and its bounding box shrunk to just enclose the object in step **280**. There may still be some holes in the breast mask. Therefore, after crop adjustments in step **282**, the mask is inverted in step **284** and the largest object is found in step **286**. This largest object is extracted and then inverted in step **288** to obtain the penultimate mask.

The final mask is obtained by closing holes in the penultimate mask with multiple majority operations and dilations in step **290**. The image is then cropped to the size of the resulting mask and the autocropping is complete. An important result from the autocropper is the offset of the cropped image. This is the pixel location in the original image that corresponds to the pixel in the upper left pixel of the cropped image. Keeping track of all the cropping and crop adjustments determines this offset value.

The output of the autocropping process is a rectangular array of pixels representing a binary mask wherein the pixels corresponding to breast tissue are assigned a value of one while the remainder of the pixels are assigned a value of zero. Put another way, the binary mask is a silhouette of the breast made up of ones while the background is made up of zeroes.

Parameters of the autocropper may be optimized to obtain better breast masks. The procedure is described below in the optimization section.

Detection of Clustered Microcalcifications

Turning now to FIG. **11**, there is seen therein a flow diagram illustrating in greater detail the clustered microcalcification detection system **300** of the invention.

That portion of the digital representation of the mammogram corresponding to the analysis region **296**, designated a cropped sub-image **302**, produced in the cropping step **200**, is first processed to reduce noise in a noise reduction step **310** to reduce digitization noise that contributes to false detections of microcalcifications. The noise-reduced image is then filtered using an optimized target-size-dependent difference of Gaussians (DoG) spatial kernel in step **320** to enhance differences between targets and background, thus creating global and local maxima in the filtered image. The optimized DoG-filtered image is then thresholded in step **340** to segment maxima that represent potential detections of microcalcifications.

The detected maxima are converted to single-pixel coordinate representations in a conversion step **350**. The coordinate representations of the detected maxima are compared with the binary mask of the analysis area in a first false-positive removal step **360** to remove false detections outside the breast mask area. The remaining coordinate representations in the analysis area are clustered in a clustering step **370**. Features are computed for the remaining clusters in a feature computation step **380** and used to remove non-suspicious detections in a classifying step **400** (FIG. **1**). The remaining detections are outputted as detected clustered microcalcifications in an outputting step **600** in the form of cluster coordinates.

Turning now to a more detailed discussion of the steps in the clustered microcalcification detection process, the digital mammogram image is first filtered to reduce noise in the image. Although the main limitation in image quality should be the granularity of the film emulsion, noise is introduced from the process of digitization. This noise may later be detected as a pseudocalcification. In this system, a cross-shaped median filter is used because it is well known to be extremely effective at removing single-pixel noise. The median filter is a non-linear spatial filter that replaces each pixel value with the median of the pixel values within a kernel of chosen size and shape centered at a pixel of interest. Referring to FIG. **12**, it may be seen that the cross shape is formed by the set of pixels which include the center pixel and its four nearest neighbors. The cross shape preserves lines and corners better than typical block-shaped median filters and limits the possible substitution to the four nearest neighbors, thereby reducing the potential for edge displacement.

After noise has been reduced, the image is filtered with an optimized DoG kernel to enhance microcalcifications. Filtering is accomplished by convolving the noise-reduced image with the DoG kernel. In an alternative embodiment, filtering is accomplished by first obtaining the fast Fourier transforms (FFTs) of the noise-reduced image and the DoG kernel, then multiplying the FFTs together, and taking the inverse FFT of the result.

The DoG kernel was chosen because neurophysiological experiments provide evidence that the human visual path-way includes a set of "channels" that are spatial frequency selective. Essentially, at each point in the visual field, there are size-tuned filters or masks analyzing an image. The operation of these spatial receptive fields can be approximated closely by a DoG.

The 2-D Gaussian mask is given as:

$$G(x, y) = ce^{\frac{-(x^2+y^2)}{2\sigma^2}} \quad (1)$$

where c normalizes the sum of mask elements to unity, x and y are horizontal and vertical indices, and σ is the standard deviation. Using Equation 1, the difference of two Gaussians with different σ yields:

$$DoG(x, y) = c_1 e^{\frac{-(x^2+y^2)}{2\sigma_1^2}} - c_2 e^{\frac{-(x^2+y^2)}{2\sigma_2^2}} \quad (2)$$

It has been shown that when $\sigma_2=1.6\sigma_1$, then the DoG filter's response closely matches the response of human spatial receptive filters. Therefore, with motivation from human physiology, let the ratio of the DoG standard deviation constants be 1:1.6. Then, for a target of size (average width) t pixels, use $\sigma_2=t/2$ and, from the rule of thumb, $\sigma_1=\sigma_2/1.6$.

Since microcalcifications typically range from 100 to 300 μm in diameter, potential target sizes for the 50 μm digitized mammograms correspond to 2 to 6 pixels. It has been found that a DoG kernel constructed using an optimization technique for selecting the target size parameter, such as the GA detailed below, has an optimized target size of t=6.01 pixels. The targetsize t will vary depending on such factors as the resolution and scale of the image to be processed. The impulse response of a DoG filter having t=6.01 pixels and $\sigma_2=1.6\sigma_1$ is shown in FIGS. **13** and **14**.

Once the noised-reduced cropped image has been DoG filtered to enhance differences between targets and background, the DoG-filtered subimage contains differences in gray levels between potential microcalcifications and background. Although microcalcifications tend to be among the brightest objects in DoG-filtered subimages, they may exist within regions of high average gray levels and thus prove difficult to reliably segment. The thresholding process used in one embodiment of the invention that generally addresses these concerns involves pair-wise pixel "ANDing" of the results of global histogram and locally adaptive thresholding. However, the preferred embodiment of the invention uses sloping local thresholding.

Since targets tend to exist within an image's higher gray levels, then the global threshold may be approximated by finding the level which segments a preselected percentage of the corresponding higher pixel levels in the image histogram. An embodiment of a global thresholding method is illustrated in FIG. **15**. Locally adaptive thresholding may be implemented by varying the high and low thresholds based on the local pixel value mean and standard deviation. An embodiment of a dual-local thresholding method is illustrated in FIG. **16**.

After computing the image histogram, $p(r_k)$, the gray level threshold, g, used to segment a preselected upper fraction, ƒ, of the histogram, is found using:

$$f = 1 - \sum_{k=0}^{g} p(r_k) \quad (3)$$

where r_k is the k^{th} gray level, $0 \leqq g \leqq g_{max}$, and g_{max} is the maximum gray level in the image.

The locally adaptive thresholds, t_{lo} and t_{hi}, are found using

$$t_{lo}=k_{lo}\sigma_{NN}(x,y)+\mu_{NN}(x,y) \quad (4)$$

and

$$t_{hi}=k_{hi}\sigma_{NN}(x,y)+\mu_{NN}(x,y) \quad (5)$$

where k_{lo} and k_{hi} are used to preselect the multiple of $\sigma_{NN}(x,y)$, the local standard deviation of gray-level intensities, and $\mu_{NN}(x,y)$ is the local gray-level mean of the N×N neighborhood centered on the pixel at (x,y) of the DoG-filtered image. Other neighborhood shapes, such as rectangular, circular, and ellipsoidal, may also be used. Pixels whose brightness or gray-level value falls within the threshold interval, that is, t_{lo}<brightness<t_{hi}, are set equal to one. Optimization of ƒ, k_{lo}, k_{hi}, and N is discussed below in connection with the parameter-optimizing process. The results of the global thresholding process may be combined with the results of the local thresholding step by logically ANDing them as shown in FIG. **17**. Alternatively, either thresholding method may be used alone.

The preferred thresholding means are illustrated in FIG. **18** wherein it may be seen that an N×N window is centered at a pixel x,y in the input image p(x,y). The mean, μ(x,y), and standard deviation, σ(x,y), of the digital mammogram image pixels under the window are computed. A local threshold value, T(x,y), is computed as:

$$T(x,y)=A+B\mu(x,y)+C\sigma(x,y) \quad (6)$$

where values for N, A, B, and C are computed during a parameter optimization stage, discussed below. Values for T(x,y) are computed for every x,y location in the image.

The digital mammogram has also been DoG filtered, producing an image d(x,y). Each pixel of the DoG-filtered image d(x,y) is compared to the threshold value T(x,y). Pixels in the locally thresholded image $l_s(x,y)$ are set to one where values of the DoG-filtered image are greater than the threshold, and set to zero elsewhere.

The advantage of this novel local sloping thresholding method over prior art thresholding methods is that the threshold is computed from the pixels in a pre-DoG-filtered image rather than from a post-DoG-filtered image. This eliminates the need for background trend correction. In conventional local thresholding, the threshold is computed as:

$$T(x,y)=B\mu(x,y)+C(\sigma x,y) \quad (7)$$

from the mean and standard deviation of the DoG-filtered image. The problem of using a local threshold computed from the DoG-filtered image is that DoG-filtered images typically have mean values close to zero and standard deviations significantly affected by the presence of targets.

Local thresholds computed from the statistics of the DoG-filtered image suffer from the following adverse effects. First, since the mean value is close to zero, a degree of freedom is lost in the computation of the threshold, which becomes essentially a function of the standard deviation. Second, the absolute brightness of the input image is lost. To keep many spurious detections from occurring, it is desirable to have high thresholds in bright regions. However, the information about the local mean of the input image is not available in the DoG-filtered image. Finally, the standard deviations of DoG-filtered images are increased by detections of targets. This is so because when local bright spots of proper size exist in the original image, large gray-scale values result in the DoG-filtered image. Thus, the presence of targets in a region increases the local standard deviation thereby raising the threshold of that region. The higher threshold reduces the probability of passing a bright spot to subsequent processing stages.

The novel local thresholding method just described solves the above problems by computing thresholds from the input image, which are then applied to the DoG-filtered image. Additionally, the threshold computed here includes an offset term A, which is independent of the local image mean.

After thresholding, detections are converted to single-pixel representations by computing the centroid or center of gravity of groups of contiguous pixels found by the thresholding process. Detections are thus represented as single pixels having a value of logical one while the remaining pixels have a value of logical zero.

False-positive detections outside of the breast area are removed by logically ANDing the binary mask from the autocropper with the single-pixel representations of the detections.

Calcifications associated with malignancies usually occur in clusters and can be extensive. The cluster detection module identifies clusters based on a clustering algorithm as depicted in FIG. **19**. Specifically, a suspicious cluster is declared when at least μCs_{min} or more detected signals are separated by less than a nearest neighbor distance, d_{nn}. Optimization of μCs_{min} and d_{nn} is discussed below in connection with the parameter optimizing process. FIG. **20** illustrates the clustering process for the case wherein $\mu Cs_{min}=5$ and $d_{nn}=4$.

Additional false-positive clustered microcalcifications are removed by means of a classifier, detailed below. Features are extracted for each of the potential clustered microcalcifications as shown in FIG. **21**. The eight features computed for each of the potential clustered microcalcifications in a preferred embodiment are:

1. The larger eigenvalue (λ_1) of the covariance matrix of the points in a cluster;
2. The smaller eigenvalue (λ_2) of the covariance matrix of the points in a cluster;
3. The ratio of the smaller eigenvalue of the covariance matrix to the larger eigenvalue of the covariance matrix of the points in a cluster. Equivalent to the ratio of the minor axis to the major axis of an ellipse fitted to cover the points in a cluster;
4. Linear density calculated as the number of detected microcalcifications divided by the maximum interpoint distance;
5. Standard deviation of the distances between points in a cluster;
6. Mean minus median of the distances between points in a cluster;
7. Range of points in cluster calculated as maximum interpoint distance minus the minimum interpoint distance; and
8. Density of a cluster calculated as the number of detections divided by the area of a box just large enough to enclose the detections.

Of course, other features could be computed for the potential microcalcification clusters, and the invention is not limited to the number or types of features enumerated herein.

Classifying Detections

The cluster features are provided as inputs to the classifier, which classifies each potential clustered microcalcification as either suspicious or not suspicious. In practice, the clustered microcalcification detector is only able to locate regions of interest in the digital representation of the original mammogram that may be associated with cancer. In any detector, there is a tradeoff between locating as many potentially suspicious regions as possible versus reducing the number of normal regions falsely detected as being potentially suspicious. CAD systems are designed to provide the largest detection rates possible at the expense of detecting potentially significant numbers of regions that are actually normal. Many of these unwanted detections are removed from consideration by applying pattern recognition techniques.

Pattern recognition is the process of making decisions based on measurements. In this system, regions of interest or detections are located by a detector, and then accepted or rejected for display. The first step in the process is to characterize the detected regions. Toward this end, multiple measurements are computed from each of the detected regions. Each measurement is referred to as a feature. A collection of measurements for a detected region is referred to as a feature vector, wherein each element of the vector represents a feature value. The feature vector is input to a discriminant function.

Referring to FIG. **22**, there may be seen therein a classifier having a feature vector x applied to a set of discriminant functions g(x). The classifier shown in FIG. **22** is designed with one discriminant function per class. A discriminant function computes a single value as a function of an input feature vector. Discriminant functions may be learned from training data and implemented in a variety of functional forms. The output of a discriminant function is referred to as a test statistic. Classification is selecting a class according to the discriminant function with the greatest output value. The test statistic is compared to a threshold value. For values of the test statistic above the threshold, the region or detection associated with the feature vector is retained and displayed as potentially suspicious. When the test statistic is below the threshold, the region is not displayed.

Many methods are available for designing discriminant functions. One approach considered for this invention is a class of artificial neural networks. Artificial neural networks require training, whereby the discriminant function is formed with the assistance of labeled training data.

In a preferred embodiment, the classification process is implemented by means of a multi-layer perceptron (MLP) neural network (NN). Of course, other classifier means could be used such as, for example, a statistical quadratric classifier. Only potential clustered microcalcifications classified as suspicious are retained for eventual designation for a radiologist. Alternatively, it may be desirable to iteratively loop between MLP NN analysis of the individual microcalcification detections and the microcalcification clusters.

Referring to FIG. **23**, a schematic diagram of an MLP NN may be seen therein. The MLP NN includes a first layer of J hidden layer nodes or perceptrons **410**, and one output node or perceptron **420** for each class. The preferred embodiment of the invention uses two output nodes, one each for the class of suspicious detections and the class of non-suspicious detections. Of course, more or fewer classes could be used for classifying clusters of microcalcifications. Each computed feature x_i is first multiplied by a weight $w_{i,j}$, where i is an index representing the i^{th} feature vector element, and j is an index representing the j^{th} first layer node. The output y_j of each first layer perceptron **410** is a nonlinear function of the weighted inputs and is given by:

$$y_j = f\left(\sum_{i=1}^{d}(w_{i,j} \times w_i)\right) \quad (8)$$

where d represents the total number of features x_i and f(•) is typically a saturating nonlinearity. In this embodiment, f(•)=tanh(•). The first layer or hidden layer node outputs y_j are then multiplied by a second layer of weights $u_{j,k}$ and applied to the output layer nodes **420**. The output of an output layer node **420** is a nonlinear function of the weighted inputs and is given by:

$$z_k(y) = f\left(\sum_{j=1}^{J}(u_{j,k} \times x_j)\right) \quad (9)$$

where k is an index representing the k^{th} output node.

The hyperbolic tangent function is used in a preferred embodiment of the system because it allows the MLP NN to be trained relatively faster as compared to other functions. However, functions other than the hyperbolic tangent may be used to provide the outputs from the perceptrons. For example, linear functions may be used, as well as smoothly varying nonlinear functions, such as the sigmoid function.

The weight values are obtained by training the network. Training consists of repeatedly presenting feature vectors of known class membership as inputs to the network. Weight values are adjusted with a back propagation algorithm to reduce the mean squared error between actual and desired network outputs. Desired outputs of z_1 and z_2 for a suspicious input are +1 and −1, respectively. Desired outputs of z_1 and Z_2 for non-suspicious inputs are −1 and +1, respectively. Other error metrics and output values may also be used.

In this embodiment of the system, the MLP NN is implemented by means of software running on a general-purpose computer. Alternatively, the MLP NN could also be implemented in a hardware configuration by means readily apparent to those with ordinary skill in the art.

After training, each detected clustered microcalcification is classified as either suspicious or not suspicious by means forming the difference z_1-z_2, then comparing the difference to a threshold, θ. For values of z_1-z_2 greater than or equal to the threshold θ, i.e., $z_1-z_2 \geqq \theta$, the classifier returns a value of +1 for suspicious clustered microcalcifications, and for values of $z_1-z_2<0$, the classifier returns a value of −1 for non-suspicious clustered microcalcifications.

In order to arrive at optimum values for the respective weights, and the number of first layer nodes, the MLP NN was trained with a training set of feature vectors derived from a database of 978 mammogram images.

To develop and test the CAD system of the invention, truth data was first generated. Truth data provides a categorization of the tissue in the digital images as a function of position. Truth data was generated by certified radiologists marking truth boxes over image regions associated with cancer. In addition to the mammogram images, the radiologists also had access to patient histories and pathology reports.

The radiologists identified 57 regions of interest, containing biopsy-confirmed cancers associated with clustered microcalcifications, by means of truth boxes. All 978 images were then processed by the microcalcification detector of the invention to produce a plurality of feature vectors, a subset of which were associated with the 57 truth boxes. Half of the subset feature vectors were randomly chosen, along with about three times as many feature vectors not associated with clustered microcalcifications, to comprise the training set of feature vectors. The MLP NN, having a predetermined number of hidden nodes, was then trained using the training set. The remaining feature vectors were used as a test database to evaluate the performance of the MLP NN after training. Training of the MLP NN was carried out by means of the Levenberg-Marquardt back propagation algorithm.

Alternatively, the MLP NN can be trained with other learning algorithms and may have nonlinearities other than the hyperbolic tangent in either or both layers. In an alternative embodiment with sigmoidal output nodes, the Bayes optimal solution of the problem of classifying clustered microcalcification detections as either suspicious or non-suspicious may be obtained.

In one run of the preferred embodiment during testing, before application of the MLP NN classifier to eliminate false-positive clustered microcalcifications, the detection procedure found about 93% of the true-positive clustered microcalcifications in both the training and test databases while indicating about 10 false-positive clustered microcalcifications per image. It was found that after an MLP NN classifier having 25 first layer nodes was used with the respective optimum weights found during training, 93% of the true-positive detections were retained while 57% of the false-positive detections were successfully removed. Referring to FIG. **24**, there may be seen a histogram of the results of testing on the testing database after classification by the MLP NN. Of course, the MLP NN of the invention may be operated with more or fewer first layer nodes as desired.

Displaying Detections

After the locations of clustered microcalcifications have been determined, they are indicated on the original digitized mammogram image, or a copy of the original image, by drawing rectangular boxes around microcalcifications. Other means for indicating the locations of microcalcifications may be used, such as, for example, placing arrows in the image pointing at detections or drawing ellipses around the detections.

The locations of clustered microcalcifications are passed to the display detections procedure as a list of row and column coordinates of the upper left and lower right pixels bounding each of the clusters. The minimum row and column coordinates and maximum row and column coordinates are computed for each cluster. Bounding boxes defined by the minimum and maximum row and column coordinates are added to the original digitized image, by means well known in the art. The resulting image is then stored as a computer-readable file, displayed on a monitor, or printed as a hard-copy image, as desired.

In one embodiment of the system, the resulting image is saved to a hard disk on a general-purpose computer having dual Pentium II® processors and running a Windows NT® operating system. The resulting image may be viewed on a VGA or SVGA monitor, such as a ViewSonic PT813® monitor, or printed as a hard-copy gray-scale image using a laser printer, such as a Lexmark Optra S1625®. Of course, other hardware elements may be used by those with ordinary skill in the art.

Optimizing the Parameters

Genetic algorithms (GAs) have been successfully applied to many diverse and difficult optimization problems. A preferred embodiment of this invention uses an implementation of a GA developed by Houck, et al. ("A Genetic Algorithm for Function Optimization," Tech. Rep., NCSU-IE TR 95-09, 1995), the entire disclosure of which is incorporated by reference herein, to find promising parameter settings. The parameter optimization process of the invention is shown in FIG. 25. This is a novel application of optimization techniques as compared to current computer-aided diagnosis systems that require hand tuning by experiment.

GAs search the solution space to maximize a fitness (objective) function by use of simulated evolutionary operators such as mutation and sexual recombination. In this embodiment, the fitness function to be maximized reflects the goals of maximizing the number of true-positive detections while minimizing the number of false-positive detections. GA use requires determination of several issues: objective function design, parameter set representation, population initialization, choice of selection function, choice of genetic operators (reproduction mechanisms) for simulated evolution, and identification of termination criteria.

The design of the objective function is a key factor in the performance of any optimization algorithm. The function optimization problem for detecting clustered microcalcifications may be described as follows: given some finite domain, D, a particular set of cluster detection parameters, $x=\{t, f, k_{lo}, k_{hi}, N, \mu Cs_{min}, d_{nn}\}$ where $x \in D$, and an objective function $f_{obj}: D \rightarrow \Re$, where $\Re$ denotes the set of real numbers, find the x in D that maximizes or minimizes f_{obj}. When sloping local thresholding is used in the cluster detector, the parameters N, A, B, and C are optimized. Radiologic imaging systems may be optimized to maximize the TP rate subject to the constraint of minimizing the FP rate. This objective may be recast into the functional form shown in the following equation:

$$f_{obj}(x) = \begin{cases} -FP(x), & TP(x) \geq TP_{min} \\ FP_{penalty}, & \text{otherwise} \end{cases} \quad (10)$$

where maximization is the goal. For a particular set of cluster detection parameters, if the minimum acceptable TP rate, TP_{min}, is exceeded, the objective function returns the negative of the FP rate. Otherwise, if the TP rate falls below TP_{min}, the objective function returns a constant value, $FP_{penalty}=-10$. Other objective functions may also be used.

Since a real-valued GA is an order of magnitude more efficient in CPU time than the binary GA, and provides higher precision with more consistent results across replications, this embodiment of the invention uses a floating-point representation of the GA.

This embodiment also seeds the initial population with some members known beforehand to be in an interesting part of the search space so as to iteratively improve existing solutions. Also, the number of members is limited to twenty so as to reduce the computational cost of evaluating objective functions.

In one embodiment of the invention, normalized geometric ranking is used, as discussed in greater detail in Houck, et al., supra, for the probabilistic selection process used to identify candidates for reproduction. Ranking is less prone to premature convergence caused by individuals that are far above average. The basic idea of ranking is to select solutions for the mating pool based on the relative fitness between solutions. This embodiment also uses the default genetic operation schemes of arithmetic crossover and non-uniform mutation included in Houck, et al.'s GA.

This embodiment continues to search for solutions until the objective function converges. Alternatively, the search could be terminated after a predetermined number of generations. Although termination due to loss of population diversity and/or lack of improvement is efficient when crossover is the primary source of variation in a population, homogeneous populations can be succeeded with better (higher) fitness when using mutation. Crossover refers to generating new members of a population by combining elements from several of the most fit members. This corresponds to keeping solutions in the best part of the search space. Mutation refers to randomly altering elements from the most fit members. This allows the algorithm to exit an area of the search space that may be just a local maximum. Since restarting populations that may have converged proves useful, several iterations of the GA are run until a consistent lack of increase in average fitness is recognized.

Once potentially optimum solutions are found by using the GA, the most fit GA solution may be further optimized by local searches. An alternative embodiment of the invention uses the simplex method to further refine the optimized GA solution.

The autocropping system may also benefit from optimization of its parameters including contrast value, number of erodes, and number of dilates. The method for optimizing the autocropper includes the steps of generating breast masks by hand for some training data, selecting an initial population, and producing breast masks for training data. The method further includes the steps of measuring the percent of overlap of the hand-generated and automatically-generated masks as well as the fraction of autocropped breast tissue outside the hand-generated masks. The method further comprises selecting winning members, generating new members, and iterating in a like manner as described above until a predetermined objective function converges.

In FIGS. 26 and 27, there may be seen therein free response receiver operating characteristic curves for the system of the invention for the outputs of the optimized microcalcification detector and the classifier, respectively. FIG. 26 represents the performance of the optimized detector before classifying detections, while FIG. 27 represents the performance of the system after classifying detections.

Although the GA has been described above in connection with the parameter optimization portion of the preferred embodiment, other optimization techniques are suitable such as, for example, response surface methodology. Of course, processing systems other than those described herein may be optimized by the methods disclosed herein, including the GA.

Incorporating Cad System Outputs for Optimal Sensitivity

Performance metrics for detection of suspicious regions associated with cancer are often reported in terms of sensitivity and specificity. Sensitivity measures how well a system finds suspicious regions and is defined as the percentage of suspicious regions detected from the total number of suspicious regions in the cases reviewed. Sensitivity is defined as:

$$\text{Sensitivity} = \frac{TP}{TP+FN} \quad (11)$$

where TP is the number of regions reported as suspicious by a CAD system that are associated with cancers, and FN is the number of regions that are known to be cancerous that are not reported as suspicious. Specificity measures how well the system reports normal regions as normal. Specificity is defined as:

$$\text{Specificity} = \frac{TN}{FP+TN} \tag{12}$$

where TN represents regions correctly identified as not suspicious and FP represents regions reported as suspicious that are not cancerous.

Current CAD systems increase specificity by reducing FP. However, FP and TP are coupled quantities. That is, a reduction of FP leads to a reduction of TP. This implies that some of the suspicious regions that could have been detected are missed when the objective is to maintain high specificity.

FIGS. **28** and **29** illustrate relationships between the quantities TP, FP, TN, and FN. A measurement from a screening mammography image is represented by test statistic, x. The probability density function of x is represented by p(x) and the decision threshold is represented by θ. If x is greater than θ, a suspicious region is reported. Areas under the probability density functions represent probabilities of events. From FIG. **28** observe that increasing the threshold reduces the probability of FP decisions. However, observe from FIG. **29** that increasing the threshold simultaneously reduces the probability of TP decisions.

Another metric that exists for CAD systems is positive predictive value (PPV), which is defined as the probability that cancer actually exists when a region of interest is labeled as suspicious. PPV can be calculated from the following equation: TP

$$PPV = \frac{TP}{TP+FP} \tag{13}$$

Note that increasing TP or reducing FP increases PPV.

Radiologists and computers find different suspicious regions. FIG. **30** is a Venn diagram depicting a possible distribution of suspicious regions for man and machine detections. Some suspicious regions are found solely by a human interpreter or radiologist, some solely by a CAD system, some are found by both, and some are not found by either.

Referring to FIG. **31**, there may be seen a preferred method for incorporating the outputs of a CAD system, and more particularly for the CAD system of the invention, with the observations of a human interpreter of a screening mammography image **10** for optimal sensitivity, wherein a radiologist examines the screening mammography image **10** in a step **20** and reports a set of suspicious regions **30** designated as S 1. The CAD system then operates on the image **10** in a step **40** and reports a set of suspicious regions **50** designated as S2. The radiologist then examines set S2 and accepts or rejects members of set S2 as suspicious in a step **60**, thereby forming a third set of suspicious regions **70** denoted as S3, which is a subset of S2. The radiologist then creates in a step **80** a set of workup regions **90** denoted as S4 which is the union of sets S1 and S3. The workup regions **90** are then recommended for further examination such as taking additional mammograms with greater resolution, examining the areas of the breast tissue corresponding to the workup regions by means of ultrasound, or performing biopsies of the breast tissue.

While the invention has been described in connection with detecting clustered microcalcifications in mammograms, it should be understood that the methods and systems described herein may also be applicable to other medical images such as chest x-rays.

While the methods herein described, and the forms of apparatus for carrying these methods into effect, constitute

preferred embodiments of this invention, it is to be understood that the invention is not limited to these precise methods and forms of apparatus, and that changes may be made in either without departing from the scope of the invention which is defined in the appended claims.

What is claimed is:

1. A method for automatically segmenting a desired region in a digital mammogram image from undesired regions, comprising the steps of:

normalizing the brightness values of the pixels in said digital mammogram image to produce a normalized image;

selecting a seed pixel in said normalized image;

region growing said normalized image starting from said seed pixel to produce a region-grown mask;

closing holes in the region-grown mask to produce a closed mask;

eroding the closed mask to produce an eroded mask; and

dilating the eroded mask to produce a dilated mask.

2. The method according to claim **1** wherein said step of normalizing comprises the steps of:

obtaining the histogram for the digital mammogram image;

determining the range of histogram bins containing substantially all of the digital mammogram pixels corresponding to said desired region;

setting the pixels below said range to a minimum brightness level;

setting the pixels above said range to a maximum brightness level; and

linearly mapping the brightness values of the pixels in said range between said minimum and maximum brightness levels.

3. The method according to claim **1** wherein said step of selecting comprises the steps of:

dividing said digital mammogram into a plurality of substantially equally-sized regions;

determining which of said equally-sized regions is brightest; and

selecting a pixel in said brightest equally-sized region which has a maximum brightness for that region.

4. The method according to claim **1** wherein said region growing step comprises the steps of:

computing the contrast ratio with respect to said seed pixel of one of the four nearest neighbor pixels of said seed pixel;

comparing said contrast ratio with a threshold value; and

setting the corresponding pixel in said region-grown mask to a predetermined value based on the outcome of said comparing step.

5. The method according to claim **1** further comprising the step of:

cropping the digital mammogram image to the largest rectangle that just encloses the digital mammogram pixels corresponding to the dilated mask.

6. A method for segmenting an area of a digital mammogram image corresponding to breast tissue from the remainder of the image, comprising the steps of:

storing a digital representation of said digital mammogram image;

enhancing the digital representation to produce an enhanced image in which the contrast of the area of the mammogram image corresponding to breast tissue is increased;

thresholding the enhanced image to produce a binary image comprising a seed pixel;

region growing said seed pixel in said binary image to produce a mask;

closing holes in said mask;

eroding said mask; and

dilating said mask.

7. The method according to claim 6 further comprising the step of:

cropping said digital representation to the size of the largest object in said mask.

8. A system for automatically segmenting a desired region in a digital mammogram image from undesired regions, comprising:

normalizing means for normalizing the brightness values of the pixels in said digital mammogram image to produce a normalized image;

seed-selecting means for selecting a seed pixel in said normalized image;

region-growing means for region growing said normalized image starting from said seed pixel to produce a region-grown mask;

hole-closing means for closing holes in the region-grown mask to produce a closed mask;

erosion means for eroding the closed mask to produce an eroded mask; and

dilation means for dilating the eroded mask to produce a dilated mask.

9. The system according to claim 8 wherein said normalizing means comprises:

histogram means for obtaining the histogram for the digital mammogram image;

range determining means for determining the range of histogram bins containing substantially all of the digital mammogram pixels corresponding to said desired region;

means for setting the pixels below said range to a minimum brightness level;

means for setting the pixels above said range to a maximum brightness level; and

linear mapping means for linearly mapping the brightness values of the pixels in said range between said minimum and maximum brightness levels.

10. The system according to claim 8 wherein said seed-selecting means comprises:

means for dividing said digital mammogram into a plurality of substantially equally-sized regions;

means for determining which of said equally-sized regions is brightest; and

means for selecting a pixel in said brightest equally-sized region which has a maximum brightness for that region.

11. The system according to claim 8 wherein said region-growing means comprises:

means for computing the contrast ratio with respect to said seed pixel of one of the four nearest neighbor pixels of said seed pixel;

means for comparing said contrast ratio with a threshold value; and

means for setting the corresponding pixel in said region-grown mask to a predetermined value based on the outcome of said comparing step.

12. The system according to claim 8 further comprising:

cropping means for cropping the digital mammogram image to the largest rectangle that just encloses the digital mammogram pixels corresponding to the dilated mask.

13. A system for segmenting an area of a digital mammogram image corresponding to breast tissue from the remainder of the image, comprising:

storage means for storing a digital representation of said digital mammogram image;

image enhancement means for enhancing the digital representation to produce an enhanced image in which the contrast of the area of the mammogram image corresponding to breast tissue is increased;

thresholding means for thresholding the enhanced image to produce a binary image comprising a seed pixel;

region growing means for region growing said seed pixel in said binary image to produce a mask;

hole closing means for closing holes in said mask;

erosion means for eroding said mask; and

dilation means for dilating said mask.

14. The system according to claim 13 further comprising:

cropping means for cropping said digital representation to the size of the largest object in said mask.

* * * * *

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 0 975 146 A1

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
26.01.2000 Bulletin 2000/04

(51) Int Cl.⁷: H04N 1/387

(21) Application number: 98305918.9

(22) Date of filing: 24.07.1998

(84) Designated Contracting States:
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE
Designated Extension States:
AL LT LV MK RO SI

(71) Applicant: XEROX CORPORATION
Rochester, New York 14644 (US)

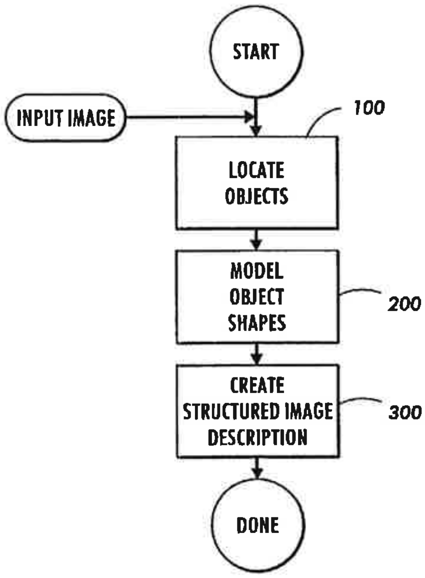
(72) Inventor: Venable, Dennis L.
Marion, New York 14505 (US)

(74) Representative: Skone James, Robert Edmund
GILL JENNINGS & EVERY
Broadgate House
7 Eldon Street
London EC2M 7LH (GB)

(54) Locating the position and orientation of multiple objects with a smart platen

(57) The present invention is an intelligent scanning system (20) for processing a digital input image to automatically characterize a plurality of objects therein. The system then employs the characterizations as the basis for rudimentary image editing operations so as to produce a digital document. In the digital document, the objects may be derotated, shifted, cropped or otherwise aligned in a predetermined fashion in accordance with a template. The scanning apparatus (26) of the present invention not only enables the scanning of a plurality of objects, but does so in an intelligent manner so as to enable further processing and manipulation of the images associated with the objects to create an output document.

FIG. 3



EP 0 975 146 A1

Printed by Jouve, 75001 PARIS (FR)

Description

[0001] This invention relates generally to an image input system, and more particularly to the automatic characterization and editing of a plurality of image objects placed on the platen of an image input device.
[0002] To improve the efficiency of scanning systems and digital copiers, it is desirable to produce a single scanned image composed of several separate objects (e.g., photographs) laid side by side, but preferably not touching, on the platen of the input scanner. To facilitate automated editing of the objects, it is necessary to identify the position, shape and rotation angle of each object. Such a capability can enhance productivity by decreasing the time required for scanning multiple objects and by automating rudimentary editing operations.
[0003] Heretofore, a number of patents and publications have disclosed image segmentation and structured images, the relevant portions of which may be briefly summarized as follows:
[0004] US-A-5,485,568 to Venable et al., issued January 16, 1996, and hereby incorporated by reference, discloses a method and apparatus for representing a complex color raster image as a collection of objects in a structured image format - a hierarchical, device-independent format. A structured image document, generated using the techniques described by Venable, is a representation of data that may be rendered into a raster image. The data includes simple raster images as well as a hierarchical collection of subobjects and raster processing operations. The possible data types for objects in the structured image include a raster image, text, graphics, image processing description, and files containing multiple image representations.
[0005] In "MANAGING AND REPRESENTING IMAGE WORKFLOW IN PREPRESS APPLICATIONS", Technical Association of the Graphic Arts (TAGA) Vol. 1, 1995 Proceedings pp. 373-385, hereby incorporated by reference for its teachings, Venable et al. teach the use of structured images to manage prepress workflow. An operation such as gang scanning is described as a means for capturing several photographs roughly aligned on a scanner platen.
[0006] In accordance with the present invention, there is provided an imaging apparatus, including:

an image input device, said image input device producing a digitized image including representations of each of a plurality of objects imaged by said device;
a programmable computer capable of processing the digitized image, said computer including a first memory for storing at least a portion of the digitized image and program memory for the storage of executable code suitable for causing said computer to execute image processing operations on the digitized image,
said computer, in accordance with preprogrammed instructions, identifying the plurality of objects within the digitized input image, modeling shapes representing boundaries of each of the plurality of objects, and characterizing each of the plurality of objects by parameters including shape, position and orientation; and
said computer automatically composing an output document including a representation of at least one of the plurality of objects.

[0007] In accordance with another aspect of the present invention, there is provided A digital copying apparatus, including:

an image input device having a transparent platen, said image input device producing a digitized image including representations of each of a plurality of objects placed on the platen;
a programmable computer capable of controlling the operation of the digital copying apparatus and of processing the digitized image, said computer including a first memory for storing at least a portion of the digitized image and program memory for the storage of executable code suitable for causing said computer to execute image processing operations on the digitized image, said computer, in accordance with preprogrammed instructions, identifying the plurality of objects within the digitized input image, modeling shapes representing boundaries of each of the plurality of objects, and characterizing each of the plurality of objects by parameters including shape, position and orientation, and said computer automatically composing an output document including a representation of at least one of the plurality of objects;
a user interface, said user interface comprising a display to depict the output document and a plurality of user selectable option, wherein said computer further includes preprogrammed instructions to update the display in accordance with a selection of the user selectable options; and
a printing engine to produce, in accordance with the output document, a substrate bearing marks in accordance with representations contained in the output document.

[0008] The present invention is directed to a system intended to accomplish the automatic determination of independent regions or segments for objects within a scanned image. The invention further provides a user interface and document templates to facilitate the automatic placement of the plurality of objects at positions within a digital document so that the document may be rendered, stored or transmitted. Thus, the present invention combines a number of graphics and image processing techniques into an automated, user-friendly application for productivity enhancement. The application can enhance productivity by decreasing the time required for scanning multiple images, by automating corrections for alignment of multiple images, and even automatically placing multiple images in the document template.

[0009] The present invention accomplishes these objectives by:

1) locating a plurality of independent objects within the image
2) modeling the shape of the identified objects (e.g., rectangle)
3) creating a structured image description identifying the location, shape and orientation of each object within the image.

[0010] One aspect of the invention deals with a basic problem in digital image processing, that of identifying plural objects within a digitized image. The solution to this problem is then leveraged so as to enable the automated recognition and placement of objects to compose a digital document. The techniques described herein enable a user to expediently scan a plurality of distinct documents or objects in a single scanning operation, automatically separate those objects by recognizing them as independent objects within the digitized image, and recomposing a digital document using one or more of the recognized objects. Another aspect of the present invention allows for the automatic creation of a structured image representation of the digitized image so that the image objects may be easily extracted and further processed, independently.

[0011] The techniques described above are advantageous because they improve the efficiency of a scanning process, allowing multiple original documents to be scanned at one time, as well as the document composition process by recomposing the documents into a single document using a predetermined template. In addition, the techniques allow for automatically characterizing physical attributes (e.g., location, shape and orientation) of the objects without user intervention.

[0012] An embodiment of the present invention will now be described with reference to the accompanying drawings, in which:

Figure 1 is an illustration of the equipment that forms an image processing system serving as one embodiment for the present invention;
Figure 2 is a block diagram of the various components comprising the system of Figure 1;
Figure 3 is a flowchart illustrating the general processing steps carried out on the system of Figures 1 and 2 in accordance with the present invention;
Figures 4-7 are detailed flow charts illustrating the processing steps carried out in accordance with various embodiments of the present invention;
Figure 8 is an illustrative example of a portion of a digital document;
Figure 9 is an illustration of the output of the system of Figure 1 when an input image is processed in accordance with the present invention;
Figure 10 is an exemplary user interface screen associated with one embodiment of the present invention;
Figure 11 is an example of an image containing multiple objects obtained by a scanning operation; and
Figure 12 is a comparative illustration of deskewing and derotation operations as applied to an image segment.

[0013] For a general understanding of the present invention, reference is made to the drawings. In the drawings, like reference numerals have been used throughout to designate identical elements. In describing the present invention, the following term(s) have been used in the description.

[0014] The term "data" refers herein to physical signals that indicate or include information. When an item of data can indicate one of a number of possible alternatives, the item of data has one of a number of "values." For example, a binary item of data, also referred to as a "bit," has one of two values, interchangeably referred to as "1" and "0" or "ON" and "OFF" or "high" and "low." A bit is an "inverse" of another bit if the two bits have different values. An N-bit item of data has one of 2N values. A "multi-bit" item of data is an item of data that includes more than one bit.

[0015] "Memory circuitry" or "memory" is any circuitry that can store data, and may include local and remote memory and input/output devices. Examples include semiconductor ROMs, RAMs, and storage medium access devices with data storage media that they can access. A "memory cell" is memory circuitry that can store a single unit of data, such as a bit or other n-ary digit or an analog value.

[0016] A signal "indicates" or "selects" one of a set of alternatives if the signal causes the indicated one of the set of alternatives to occur. For example, a signal can indicate one bit set in a sequence of bit sets to be used in an operation, in which case the signal causes the indicated bit set to be used in the operation.

[0017] An "image" is a pattern of physical light. An image may include characters, words, and text as well as other features such as graphics. A text may be included in a set of one or more images, such as in images of the pages of a document. An image may be processed so as to identify specific "objects" within the image, each of which is itself an image. A object may be of any size and shape and has "physical attributes" or characteristics including, but not limited, to position, shape and orientation.

[0018] An item of data "defines" an image when the item of data includes sufficient information to produce the image. For example, a two-dimensional array can define all or any part of an image, with each item of data in the array providing a value indicating the color of a respective location of the image.

[0019] An item of data "defines" an image set when the item of data includes sufficient information to produce all the images in the set.

[0020] Each location in an image may be called a "pixel". In an array defining an image in which each item of data provides a value, each value indicating the color of a location may be called a "pixel value". Each pixel value is a bit in a "binary form" of an image, a gray scale value in a "gray scale form" of an image, or a set of color space coordinates in a "color coordinate form" of an image, the binary form, gray scale form, and color coordinate form each being a two-dimensional array defining an image.

[0021] An operation performs "image processing" when it operates on an item of data that relates to part of an image.

[0022] Pixels are "neighbors" or "neighboring" within an image when there are no other pixels between them and they meet an appropriate criterion for neighboring. If the pixels are rectangular and appear in rows and columns within a two-dimensional image, each pixel may have 4 or 8 neighboring pixels, depending on the criterion used.

[0023] An "edge" occurs in an image when two neighboring pixels have sufficiently different pixel values according to an appropriate criterion for the occurrence of an edge between them. The terms "edge pixel" or "boundary pixel" may be applied to one or both of two neighboring pixels between which an edge occurs.

[0024] An "image characteristic" or "characteristic" is a measurable attribute of an image. An operation can "measure" a characteristic by producing data indicating the characteristic using data defining an image. A characteristic is measured "for an image" if the characteristic is measured in a manner that is likely to produce approximately the same result each time it occurs.

[0025] A "version" of a first image is a second image produced using an item of data defining the first image. The second image may be identical to the first image, or it may be modified by loss of resolution, by changing the data defining the first image, or by other processes that modify pixel values of the first image.

[0026] An "image input device" is a device that can receive an image and provide an item of data defining a version of the image. A "scanner" is an image input device that receives an image by a scanning operation, such as by scanning a document. A scanner may have a transparent surface (platen) or equivalent means to support a document during scanning. Other well-known image input devices include digital cameras, facsimile machines, and video recorders having the capability to store data signals representative of the intensity of light reflected from the surface of objects at which the device is directed.

[0027] An "image output device" is a device that can receive an item of data defining an image and provide or render the image as output. A "display" is an image output device that provides the output image in human viewable form, and a "printer" is an image output device that renders the output image in a human viewable, hard copy form.

[0028] Referring now to Figures 1 and 2, depicted therein is a smart platen system 20 in which the present invention finds particular use. System 20 includes a computer 22 capable of receiving digital data representing an image of an original document 24 placed upon a platen of scanner 26. Computer 22, initially stores the digital input data from scanner 26 in memory 52 (e.g., RAM or magnetic disk storage) where the image may subsequently be accessed. In addition to the digital data, memory 52 may also include program memory for the storage of object code suitable for directing the processor to execute image processing operations in accordance with the invention described herein. Computer 22 has associated therewith a user interface (U/I) 28 including one or more user input devices 30, such as a keyboard, a keypad, a mouse, trackball, stylus or equivalent pointing device, etc.

[0029] Also part of system 20 is an image output device such as printer 34 which may include a laser-driven, xerographic printing engine as found in a number of commercially available printers. In a preferred embodiment, system 20 is employed to process the digital image data received as input from a scanner 26, utilizing image processing software running in processor 50, so as to produce an output file that may be rendered by printer 34, stored in memory 50, and/or transmitted to another device via network 40.

[0030] Although system 20 is depicted as a plurality of interconnected units, it will be appreciated that the system may also comprise what is commonly referred to as a "digital copier". In such an embodiment, the components of system 20 are integrated within a common housing, and the user interface (including display and input devices) may have a significantly different appearance. An example of such a color digital copying system in which the present invention may find particular use is the Xerox 5775 Digital Color Copier, controlled via a touch-screen color video monitor that shows the progression of operations. It will be appreciated that the document placed upon the scanner platen in system 20 preferably includes a plurality of photographs or other objects represented by marks on a substrate surface, so that such objects may be scanned by a single scanning operation. For example, a particular embodiment to which the following description will be directed is a single scanned image representative of several separate photographs laid side by side on the platen of scanner 26, but not touching or overlapping. In accordance with the present invention it is desired to automatically identify the position, shape and rotation angle of each original photograph.

[0031] Given an input image generated by scanning several separate photographs laid side by side on the scanner platen, the present invention automatically identifies at least the position, shape and orientation angle of each photograph. As shown in the flow chart of Figure 3, the process carried out by computer 22 during the processing of the input image includes three general steps. First, at step 100 the objects within the image are located and boundaries of the object are generally identified. Once the objects are located, the shape of the objects is modeled at step 200. Having located the objects and modeled their shape, a structured image representing the image and objects therein can be created as represented by step 300. The structured image preferably includes data representing not only the image data itself, but data representing the location, shape or orientation of each object, or some combination thereof. Alternatively, the output may be a page description language format or equivalents formats suitable for storing the image information in a retrievable form.

[0032] In a preferred embodiment of the present invention, the scanned input image (or a lower resolution version thereof) is loaded into a memory frame buffer (RAM) where it is analyzed in accordance with the previously described steps. For purposes of the following detailed description, it is assumed that objects do not occlude one another and that the background of the image is contiguous. These simplifying assumptions are intended for purposes of explanation only and are not intended as limitations of the invention. One skilled in the art will appreciate that the invention described herein is extensible so as not to require operation only within the boundaries of such assumptions.

[0033] As depicted by the flow chart of Figure 4, the object location step 100 is performed by first identifying the background region of the input image 102, characterizing the background region 104, and then using the characteristic of the background region as a seed, identifying all the pixels representing the background region with an adaptive seed fill algorithm 106. Background pixels are pixels not associated with any objects, or more simply, they are pixels representative of those regions lying outside of the objects, the values of which are controlled by the "background" against which the objects are placed during scanning (e.g., the underside of the platen cover). One embodiment employs the average color of a small region in the upper left-hand corner of the scanned image as an initial estimate of the background color. Alternatively, other sampling operations may be employed to determine the background color as described, for example, in US-A-5,282,091 for a Programmable Apparatus for Determining Document Background Level by Farrell.

[0034] Once the background color is characterized at step 104, an adaptive algorithm is preferably applied to monitor the background color and to accurately identify the objects. An example of a seed fill algorithm suitable for use in the present invention is described in Graphics Gems I, A. Glassner Ed., Academic Press, pp. 275-277, 1990, hereby incorporated by reference. An adaptive algorithm is required because the background pixels may have significant color variation resulting from a variation in illumination over the platen area. The adaptive seed fill algorithm is applied to the scanned color image data using an initial seed point characterized by the background, for example, the upper-left corner of the image. Generally, the adaptive seed fill algorithm fills a binary frame buffer with a mask indicating all contiguous pixels identified as background pixels. In a simple embodiment, represented by step 112, a pixel is considered to be a background pixel if its color falls within a small distance ε of the current average background pixel value. This distance is calculated as an Euclidean metric in red, green, blue (RGB) color space

$$d = \mathrm{SQRT}\ ((P_r\text{-}AdAvg_r)^2+(P_g\text{-}AdAvg_g)^2+(P_b\text{-}AdAvg_b)^2),$$

where P_k, $AdAvg_k$ are, respectively, the RGB components of the pixel under test and the average background value, and *d* is the distance measurement. The value of ε is fixed and empirically determined in one embodiment. The test conducted at step 112 is:

[0035] if $d<\varepsilon$, then pixel P is a background pixel, else pixel P is a foreground pixel.

[0036] The average background color is adaptively modified at step 114 by taking the average value of the last N pixels that have been classified as background. For efficiency, the system preferably calculates the adaptive average using the equation:

$$AdAvg' = (N*AdAvg-AdAvg+LastVal)/N,$$

where AdAvg' is the modified average, AdAvg is the previous adaptive average, LastVal is the value of the last pixel identified as background, and N is the averaging window. Clearly, this is not a true running average, but it tracks the running average adequately and is more computationally efficient than a strict running average calculation. Alternatively, the value of ε can be adaptively modified. For example, E might be based on the standard deviation of the last several pixels identified as background, etc.

[0037] It will be appreciated that alternative methods for the detection of background regions may be employed, and that the process of obtaining a binary map distinguishing the objects from the background may be accomplished using a simplified thresholding operation based upon the background color. In a preferred embodiment, the background color may be a function of the color of the platen cover, such as a white cover. As a further example, the 5775 Digital Color Copier employs a spectrally reflective platen cover that results in a bluish background color that may be detected and distinguished from the objects.

[0038] Having identified all background pixels and created a binary mask representative of the background regions, the process at step 120 is executed to smooth noisy edges in the background mask using morphological filtering. More specifically, a morphological closure filter is preferably applied to the background mask to eliminate single pixel noise and to smooth object edges. Subsequently, contiguous foreground regions are located, step 122, thereby identifying the objects. Objects are identified by scanning the background mask generated by the adaptive seed fill operation (step 106). Starting with the upper left hand pixel, the mask is searched in a scan line fashion for a pixel not classified in the mask as a background pixel - thus identifying pixels associated with a foreground object. The use of the seed fill algorithm for identifying the background assures that foreground objects are closed, or in other words, complete boundary is formed about the perimeter of the object.

[0039] At step 124, the boundary of an object is identified by tracing its edge. The boundary of the foreground object is traced using a simple 8-connected edge traversal which provides an ordered-set of points tracing the edge of the object. Such an edge traversal operation employs a contour tracing operation to generate a chain code in a manner similar to word or character based recognition systems. An 8-connected process is described, for example, by R. Bozinovic et al. in "Off-Line Cursive Script Word Recognition", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 11, No. 1 (January 1989). Once the edge is traced, all pixels associated with the object in the mask are marked as background so they will not be processed a second time, the object is added to the foreground object list and then the scanning of step 122 is continued as indicated by test step 126. Subsequent to completing the foreground scanning to identify all objects, a review of the identified objects may be completed as represented by step 130. In many cases, the scanned image may contain undesirable foreground objects; such objects can be eliminated from the object list at this step. In one embodiment, the review of the object list may simply eliminate small objects as unlikely images. For example, in a scan of a yearbook page each image has associated with it a text caption that is not to be classified as image data. Such captions consist of many, small perimeter objects, so that by measuring the perimeter length of the traced edges, it is possible to eliminate objects having a perimeter smaller than a specified length, where the threshold length may be predetermined empirically.

[0040] Once the objects have been located, as described with respect to step 100, the next general step, step 200, is to model the shape of the object. For purposes of simplicity, the following description will treat rectangular-shaped objects, however, it will be appreciated that the description is extensible to other polygons and even to shapes having portions thereof represented by curves (e.g., circular or elliptical objects) . The result or output from step 100 is preferably a set of edge traces, in the form of linked lists, that identify bounding pixels about each object within the scanned image. These traces can be used to extract each object, but orientation is not yet determined. To improve the quality of the object extraction, the object traces are fitted to a model shape. Orientation information, etc., may then be extracted from the fitted parameters. In the described embodiment the object traces are fit to a rectangular model, however, other shapes are possible.

[0041] One method of fitting the edge traces to a rectangular shape is a least-squares approach to fit to a rectangle. To accomplish the least-squares fitting, the edge trace is first decomposed into four sets of points, each corresponding to one of the four sides of the rectangular object. The decomposition into four sets of points can be accomplished in several ways as described below.

[0042] The first method has two principal parts, (a) categorizing the edge points into a set of bins associated with a single line, and (b) performing recognition on the bins for rotated shapes. Referring now to Figure 5, where the first decomposition method is depicted in detail, step 204 calculates the slope at each point along the edge trace. Step 204 preferably accomplishes the slope angle calculation by performing a linear regression on a small window of neighboring edge points. For example, 2 points lying on either side of the edge point for which the slope is being determined. The angle of the line passing through the center of each point is determined using linear regression in a small window centered on each point. Each regression requires 4 additions per point in the window, plus 2 subtractions, 2 multiplications, and an arctangent calculation, however, the regression algorithm may be further optimized to remove most of the addition operations. In a preferred embodiment, which reduces the computational complexity, a sample of the edge pixels are employed for slope angle calculations and sorting, thereby reducing the number of calculations necessary to categorize the edge pixels.

[0043] Next, at step 206, the process constructs a list of slope categories or bins. The slope categories are constructed for each edge point by calculating the magnitude of the difference in the slope angle between the current point along the edge (e.g., point B in Figure 8) and the preceding point (e.g., point A in Figure 8). If the difference is less than the value TOLERANCE (determined empirically to be ± 5 degrees in one embodiment), then the point is assigned to the same slope category as the preceding point, otherwise a new slope category is created and the point is assigned to it. Referring to Figure 8, the above-described process would assign points A, B and C to a first slope category, points D, E, F, G and H to a second slope category and points I, J ... to yet another slope category. Finally, if the slope category for the last edge point has approximately the same slope angle as the first slope category, then all points within the first and last slope categories are joined together into a single category.

[0044] Once the slope categories are established at step 206, and stored in a data structure, they are then sorted at step 208 and ordered according to the number of edge points assigned to each category. For rectangular objects, the top four slope categories, those containing the most edge points, should correspond to points along the four edges of the rectangle. The top slope categories are then selected at step 210. It will be appreciated that one would use the top six categories for hexagonal objects, and similarly the top three categories for triangular objects, etc.

[0045] Alternatively, steps 208 and 210 may be replaced by a step that processes the slope angle categories or bins by simple, or even statistical elimination, wherein those categories with few entries are removed. For example, an empirically determined threshold of 5 pixels may be applied so that only bins having more than 5 pixels with a common angle are kept. Subsequently, an average angle for a category may be determined using simple linear regression of all the points assigned to a particular category. With the average angle determined, a further refinement of the categories would be possible, combining those categories having substantially common angles. In particular, each category is checked and if adjacent categories are substantially collinear, the categories are joined. Thus each of the remaining bins or categories represents a set of collinear points lying along an edge. The edge points assigned to each of the remaining slope angle categories represent the edge trace decomposed into the four sides of the rectangle. It will be appreciated that this alternative is broadly directed to the process of "filtering" or refining the categories to identify those representing the actual edge of the objects. Accordingly, equivalent methods of accomplishing the refinement of the categories are contemplated.

[0046] This first method of characterizing the object boundaries is computationally intensive due to the measurement of the average slope at each edge point. In the alternative embodiment mentioned previously, to improve speed, the edge trace may be sampled to reduce the total number of points that must be processed and categorized.

[0047] It will be further appreciated that it may be possible, from an analysis of the ordered categories, to identify the shape. For example, a statistically significant difference in the number of points between a third and fourth category, or the complete lack of a forth category, are indicative of a triangular-shaped object.

[0048] Referring to Figure 6, depicted therein is the second method by which the object shapes may be modeled. After retrieving the edge trace list data at step 202, step 252 calculates the center of mass of the object. Although there are a number of well-known methods for calculating the center of mass of the object, in the case of rectangular objects a straightforward approach would be averaging the (x,y) coordinates of the edge points. Next, the edge point closest to the center of mass would be located at step 254. The closest point will be the approximate center of the long side of the rectangle. Referring again to Figure 8, the angle θ from the center-of-mass (CofM) to the center point ($L_a/2$) is the approximate rotation angle (θ) of the rectangle.

[0049] Once the rotation angle is determined, it is employed in step 256 to determine the approximate length of the minor axis of the rectangle at step 258. In particular, the distance from the center-of-mass to the average position of all edge points that lie in the angular range $\theta-\Delta A$ to $\theta+\Delta A$ is determined. This distance is an approximate measure of one-half the minor axis length L_b of the rectangle. ΔA is an empirically determined value on the order of approximately 5 degrees. Step 260 approximates the length of the major axis (L_a) in much the same manner. The distance from the center-of-mass to the average position $(\theta+90)+\Delta A$ is an approximate measure of one-half the length of the major axis L_a of the rectangle. Having approximated the orientation angle and the lengths of the major and minor axes, step 264 calculates an angular range (as measured with respect to the center-of-mass) for each side of the rectangle that encompasses only those edge points associated with that side:

a)

$\theta'_b = \operatorname{atan2}(L_a, L_b)$ half angle width of major axis;
$\theta'_a = 90-\theta_b$ half angle width of minor axis;
$\theta_b = \theta'_b * TOL$ where TOL=0.95 to avoid corners;
$\theta_a = \theta'_a * TOL$ where TOL=0.95 to avoid corners;

and

b) $Range_1$: $(\theta+\theta_b)$ to $(\theta-\theta_b)$

$Range_2$: $((\theta+90)+\theta_a)$ to $((\theta+90)-\theta_a)$
$Range_3$: $((\theta+180)+\theta_b)$ to $((\theta+180)-\theta_b)$
$Range_4$: $((\theta+270)+\theta_a)$ to $((\theta+270)-\theta_a)$

Once the angular range is determined, step 266 finds all the edge points that lie within each of the four angular ranges (relative to the center-of-mass) determined above, thereby identifying the edge points corresponding to each side of the rectangle. It will be appreciated that this technique is less sensitive to edge-noise than the first method described above.

[0050] Once the edge trace has been decomposed into four sets of points, each set corresponding to one of the four sides of the rectangle, a least squares calculation for fitting the points to rectangle is evaluated at step 280. A rectangle can be described as four mutually perpendicular lines defined by the equations:

$$y=\alpha_0+\beta x,$$

$$y=\alpha_1+Yx,$$

$$y=\alpha_2+\beta x,$$

$$y=\alpha_3+Yx,$$

where $\beta_Y=-1$. A least squares fit yields the fitted parameters:

$$\beta_n=\sum\left(\sum_{i=0}^{n_k} x_{ki}Y_{ki}-\frac{1}{n_k}\sum_{i=0}^{n_k} x_{ki}\sum_{i=0}^{n_k} y_{ki}q\right)$$

$$\beta_d=\left(\sum_{i=0}^{n1} x_{0i}^2-\frac{1}{n_0}\left(\sum_{i=0}^{n1} x_{0i}\right)2\right)+\left(\sum_{i=0}^{n1} y_{1i}^2-\frac{1}{n_1}\left(\sum_{i=0}^{N1} y_{1i}\right)2\right)$$
$$+\left(\sum_{i=0}^{n1} x_{2i}^2-\frac{1}{n_2}\left(\sum_{i=0}^{n1} x_{2i}\right)2\right)+\left(\sum_{i=0}^{n1} y_{3i}^2-\frac{1}{n_3}\left(\sum_{i=0}^{n1} y_{3i}\right)2\right)$$

$$\beta=\beta_n/\beta_d$$

$$\alpha_0=\frac{1}{n_0}\sum_{i=0}^{n_0} y_{0i}-\frac{\beta}{n_0}\sum_{i=0}^{n_0} x_{0i}$$

$$\beta\alpha_1=\frac{1}{n_1}\sum_{i=0}^{n_1} x_{1i}-\frac{\beta}{n_1}\sum_{i=0}^{n_1} y_{1i}$$

$$\alpha_2=\frac{1}{n_2}\sum_{i=0}^{n_2} y_{2i}-\frac{\beta}{n_2}\sum_{i=0}^{n_2} x_{2i}$$

$$\beta\alpha_3 = \frac{1}{n}\sum_{i=0}^{n_3} x_{3i} - \frac{\beta}{n_3}\sum_{i=0}^{n_3} y_{3i}$$

where (x_{ki}, y_{ki}) is the i^{th} edge point of the k^{th} side, and n_k is the number of edge points associated with the k^{th} side. Subsequently, once the least squares fit yields the fitted parameters (β, α_0, α_1, α_2, α_3), they are converted at step 282 the into four coordinate pairs marking the corners of the rectangle. Moreover, the rotation angle of the rectangular object is accurately represented by the slope parameter β.

[0051] Yet another alternative method for fitting the edge traces to a shape is a method employing binary moments for fast image bounding. Using the binary mask generated as described with respect to step 106 (e.g., the adaptive seed algorithm), or alternatively with a simple thresholding operation, the image is rendered in a binary bitmap form where each pixel value is a 0 or 1 indicating background or non-background regions. Once the borders are detected for an object using the binary mask, the alternative embodiment depicted in Figure 7 employs second-order binary moments to fit a shape (e.g., rectangle) to the object.

[0052] Referring to Figure 7, depicted therein is a generalized flowchart representing the steps of the binary moment boundary finding technique. At step 100, the object edges are located and recorded as previously described, thereby providing as an input a linked list of boundary or edge pixels referred to as an edge trace, step 290. Using the boundary list, the second order moments are calculated (step 292) in an efficient manner using the equation:

$$\mu_{kl} = \sum_i \sum_j p(i,j) i^k j^l = \sum_i \left(\sum_j p(i,j) j^l \right) i^k = \sum_i p_i(i) i^k$$

where $p(i,j)$ is the image pixel value at image coordinates (i,j) and $p_1(i)$ is the 1^{th} order moment of the i^{th} scan line. Because the object boundary pixels are previously determined, the process can be simplified and the rightmost and left-most boundary pixels for a particular scanline are used for the 1st order (absolute) moment calculations.

[0053] Subsequently, the 2nd order (central) moments (m_{00}, m_{01}, m_{10}, m_{11}, m_{20}, and m_{02}) are calculated using the 1st order moments and the following equations:

$$x_c = -\mu_{10} / \mu_{00} ,$$

$$y_c = -\mu_{01} / \mu_{00} ,$$

and

$$m_{pq} = \sum_{r=0}^{p} \sum_{s=0}^{q} \binom{p}{r} \binom{q}{s} x_c^{p-r} y_c^{q-s} \mu_{rs} .$$

[0054] Having determined the 2nd order moments, they are employed to characterize an ellipse and from the ellipse the bounding box about the object, step 294. In particular, the center of the ellipse (x, y), the lengths of each axis (a and b) and the rotation angle (θ) are determined. The bounding box for the rectangular object is determined as a rectangle centered at (x,y) with sides of length 2a and 2b, rotated by an angle θ. While this renders a bounding box slightly larger than the object, this is done so as to provide a safety margin for the calculation, and to avoid cropping a portion of the object. If a tighter bounding box is desired, the rectangle would be characterized with sides of length $2\alpha a$ and $2\alpha b$, where α is set equal to $\sqrt{3}/2$ or a slightly smaller value to accomplish edge trimming or cropping (e.g., on the order of one or more pixels).

[0055] After each object has been modeled as a shape (e.g., rectangle), a structured image is created as described, for example, in US-A-5,485,568 to Venable et al. The structured image consists of one "child" structured image for each object detected using one of the methods described above. The structured image definition contains attributes that specify which rectangle of the scanned image contains the object data, and also the rotation angle required to correct for any orientation skew. Figure 9 is an example of a structured image created in accordance with the previously described processes, the structured image containing a pair of rectangular-shaped image objects.

[0056] In one embodiment of the present invention, depicted in Figure 10, the structured image is designed such that when rendered, all objects are de-rotated and laid out in a grid fashion. In particular, Figure 10 illustrates a user interface 400 that may be employed with various aspects of the previously described object shape recognition method to provide an intelligent or "smart" platen scanning system. The smart scanning system represented by Figure 10 preferably provides a means by which a user can interface with a digitizing scanner to efficiently obtain digitized representations of objects placed on platen 24 of a scanner.

[0057] For example, referring to Figure 10 in conjunction with Figure 1, a user may place a number of photographs on the scanner platen. Once placed thereon, the user may then select an operation from region 410 of Figure 10 to cause the computer system 22 to initiate scanning by scanner 26. Although not specifically shown, it will be appreciated that various methods for initiating the digitization of the objects may be employed, including, but not limited to, keyboard entry, touch-sensitive screen selection, depression of a remote switch or even triggering of a platen sensor by a user closing the platen cover. As depicted in Figure 10, after the "Gang & Edit" (412) or equivalent scanning selection is made, system 20 scans the objects placed on platen 24 and temporarily stores the data in the file using the information reflected in region 420 of the user interface screen. For example, the various image objects (A, B, C and D) may be found within an image as illustrated in Figure 11. Once the image is scanned, it is analyzed as described above to identify the image objects. The image objects may then be manipulated by the smart scanning system to automatically orient and position the images, for example they may be automatically placed in a predefined template and rendered, such as the representation depicted in region 430 of the user interface. It will be appreciated that a user may also be given additional edit capability with respect to the template, for example, to add captions to the objects or to include titles 432 and subtitles 434 as illustrated. Input for such text-based editing would be accomplished via the user interface options depicted in region 440.

[0058] It will be appreciated by those skilled in the art that by defining the boundaries of the objects in an image accurately a derotation operation may be employed to more accurately deskew the object. For further illustration of this advantage reference may be had to Figure 12. Depicted therein is an object O within an image segment I. In a typical deskewing operation, the orthogonal boundaries of the image are determined as illustrated by bounding box 500 and the image is deskewed to produce the image indicated as 504 including the object O'. It will be appreciated that in order to deskew the image that is not clearly defined by boundaries, the entire image area must be rotated, leading to both a larger deskewed image area and the need to create pixels to fill in regions 508. The present invention, by accurately defining the boundaries and orientation of the object, avoids the need to incorporate regions larger than the object in the derotation operation. Thus, the pixels representing the object itself are the only pixels that are derotated to produce the O' output image.

[0059] Also enabled by the smart scanning system would be image editing capabilities as illustrated in region 450 of the user interface. Having identified each of the objects within the image, it is possible to isolate the objects, create separate images therefrom, and to then individually process the images. Thus the individual image objects automatically placed within the template of region 430 may be individually selected, manipulated, scaled (button 452), rotated (button 454) or cropped (button 456). It will be appreciated that the scaling, rotation and cropping operations are in addition to those which are preferably automatically applied by the system as the result of the previously described object recognition methods.

[0060] For example, the image scaling button, illustrated with cross-hatching to depict selection, will allow the user to move a cursor (not shown) to select an object (e.g., image object D) and then to drag a side or corner of the object so as to scale the image object. To facilitate the editing of the objects, control points such as those illustrated about the boundary of image object D (436) may be employed in a manner well-known to those who design user interfaces.

[0061] As noted, a predefined template may be used to automatically "place" image objects in relative positions on a document or page thereof. It will be appreciated that such templates may be in the form of a structured image definition, so that the template can be used to specify a different layout for the structured image to be generated. Thus, a family seeking to put its photographs in a "digital photo album" may be able to create a template describing a page similar to that shown in region 430 of the user interface. The template would then be used to automatically organize individual images or plural objects within a larger document image.

[0062] In a preferred embodiment, the output of the smart scanning system would be a structured image document format as described by Venable et al. in US-A-5,485,568. An important characteristic of structured images is the ability to store image processing operations in their description. This means that the structured image can contain image processing operations other than simple object deskewing attributes. For example, automatic image enhancement operations may be included within the structured image such that the objects identified can be individually enhanced. Alternatively, the smart scanning system output may take the form of a digital document in one of many well-known page description languages.

[0063] Once the "page" composed in window 430 in the condition desired by the user, the user may save the image by selecting the "Save Edited Image" button 460. More importantly, a user may then print or otherwise distribute the composed page(s).

[0064] Although the various embodiments of the present invention have been described with respect to the smart scanning system, it will be appreciated that the acquisition of images, and the printing and distribution of the composed pages can be accomplished via networks or on a walk-up digital copier. For example, a user may have photographs automatically scanned by a film processor, and a digitized stamp sheet sent to the user via a network. The stampsheet, being in a structured image format could then be processed using the smart scanning system to produce pages of a digital photo album with one or more objects on each page.

[0065] In recapitulation, the present invention is an intelligent scanning apparatus for processing a digital input image to automatically characterize a plurality of objects therein, and to employ the characterization as the basis for rudimentary image editing operations so as to produce a digital document. In the digital document, the objects may be derotated, shifted, cropped or otherwise aligned in a predetermined fashion in accordance with a template. The scanning apparatus of the present invention not only enables the scanning of a plurality of objects, but does so in an intelligent manner so as to enable further processing and manipulation of the images associated with the objects.

Claims

1. An imaging apparatus, including:

   an image input device (26), said image input device producing a digitized image including representations of each of a plurality of objects imaged by said device;
   a programmable computer (20) capable of processing the digitized image, said computer including a first memory (52) for storing at least a portion of the digitized image and program memory (52) for the storage of executable code suitable for causing said computer to execute image processing operations on the digitized image,
   said computer, in accordance with preprogrammed instructions, identifying the plurality of objects within the digitized input image, modeling shapes representing boundaries of each of the plurality of objects, and characterizing each of the plurality of objects by parameters including shape, position and orientation; and
   said computer automatically composing an output document including a representation of at least one of the plurality of objects.

2. The apparatus of claim 1, further including a user interface (34) for displaying a representation of the output document incorporating the representation of at least one of the plurality of objects wherein one of the parameters of the at least one object has been altered.

3. The apparatus of claim 1 or claim 2, wherein the position of the representation of the object is shifted to a predefined position in the output document.

4. The apparatus of claim 2, wherein only the representation of the object is derotated so as place a derotated representation of the object in the output document.

5. The apparatus of any of the preceding claims, wherein said image input device includes a surface (24) upon which the plurality of objects are placed for digitization.

6. The apparatus of any of the preceding claims, wherein the first memory is a frame buffer.

7. The apparatus of any of the preceding claims, wherein the first memory is a data storage media accessible by said programmable computer.

8. The apparatus of any of the preceding claims, further including template memory for storing at least one template for controlling the position and orientation of the representation of the at least one object during composition of the output document.

9. The apparatus of claim 8, wherein the at least one template is stored in the form of a structured image.

10. The apparatus of any of the preceding claims, further including:

a user interface, said user interface comprising a user input device (30) responsive to a user selection and a display (34) to depict the output document in conjunction with user selectable options;
wherein said computer further includes preprogrammed instructions to update the display in accordance with a selection of the user selectable options.

11. The apparatus of claim 10, wherein the programmable computer, in conjunction with the display, provide user selectable options to:

select at least one object in the output document;
alter at least one parameter of the selected object; and
store, in memory, a version of the output document after a user has altered a parameter of at least one object therein.

12. The apparatus of claim 10 or claim 11, wherein the user selectable options are selected from the group consisting of:

inserting textual content into the output document; defining a storage location for the output document; storing the output document;
scaling the representation of the at least one object;
rotating the representation of the at least one object;
cropping the representation of the at least one object;
altering the colorization of the representation of the at least one object; and
enhancing the representation of the at least one object.

13. A digital copying apparatus, including:

an image input device (26) having a transparent platen (24), said image input device producing a digitized image including representations of each of a plurality of objects placed on the platen;
a programmable computer (20) capable of controlling the operation of the digital copying apparatus and of processing the digitized image, said computer including a first memory (52) for storing at least a portion of the digitized image and program memory (52) for the storage of executable code suitable for causing said computer to execute image processing operations on the digitized image, said computer, in accordance with preprogrammed instructions, identifying the plurality of objects within the digitized input image, modeling shapes representing boundaries of each of the plurality of objects, and characterizing each of the plurality of objects by parameters including shape, position and orientation, and said computer automatically composing an output document including a representation of at least one of the plurality of objects;
a user interface, said user interface comprising a display (34) to depict the output document and a plurality of user selectable option, wherein said computer further includes preprogrammed instructions to update the display in accordance with a selection of the user selectable options; and
a printing engine to produce, in accordance with the output document, a substrate bearing marks in accordance with representations contained in the output document.

14. The digital copying apparatus of claim 13, wherein the digitized image is a color image and wherein the marks placed on the substrate are marks of at least two distinct colors.

15. The apparatus of claim 13 or claim 14, wherein said user interface comprises a user input device (30) responsive to a user selection and a display to depict the output document in conjunction with user selectable options; and

wherein said programmable computer further includes preprogrammed instructions to update the display in accordance with a selection of the user selectable options.

16. The apparatus of claim 15, wherein the programmable computer, in conjunction with the display, provide user selectable options to:

select at least one object in the output document;
alter at least one parameter of the selected object; and
store, in memory, a version of the output document after a user has altered a parameter of at least one object

therein.

17. The apparatus of claim 15 or claim 16, wherein the user selectable options are selected from the group consisting of:

inserting textual content into the output document;
defining a storage location for the output document;
storing the output document;
scaling the representation of the at least one object;
rotating the representation of the at least one object;
cropping the representation of the at least one object;
altering the colorization of the representation of the at least one object; and
enhancing the representation of the at least one object.

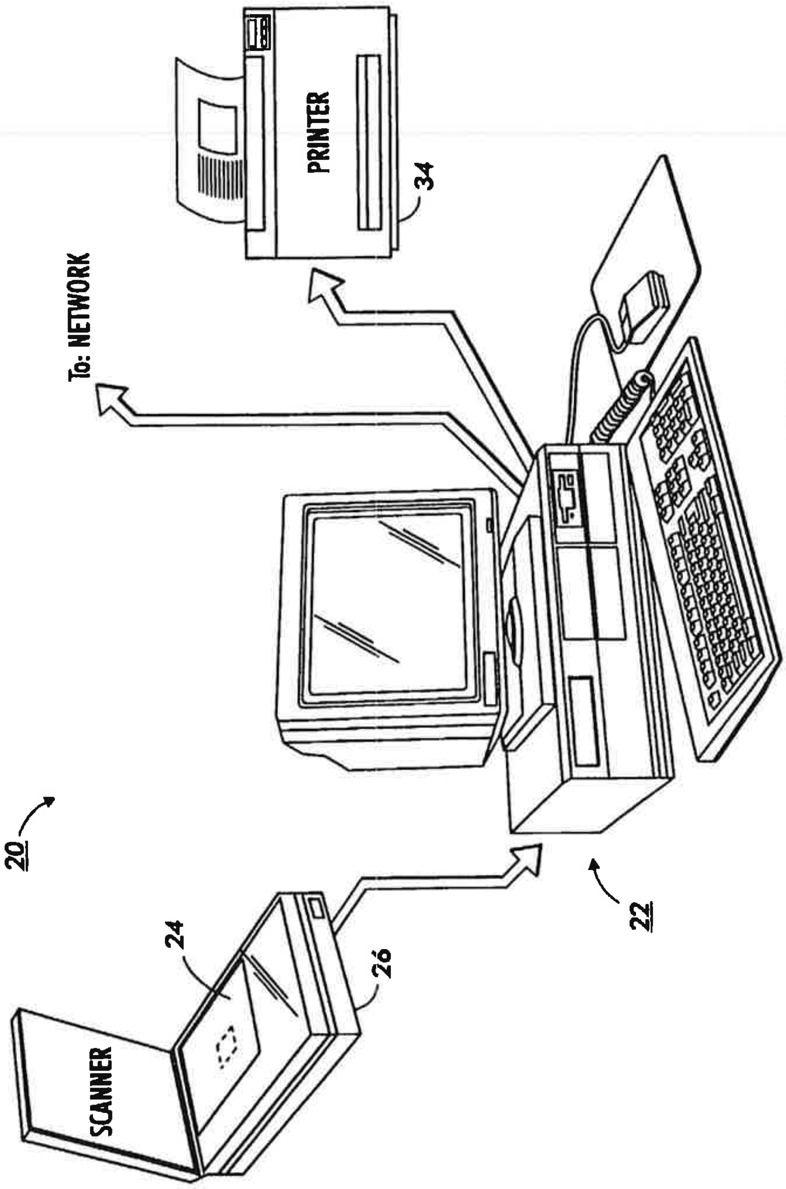


FIG. 1

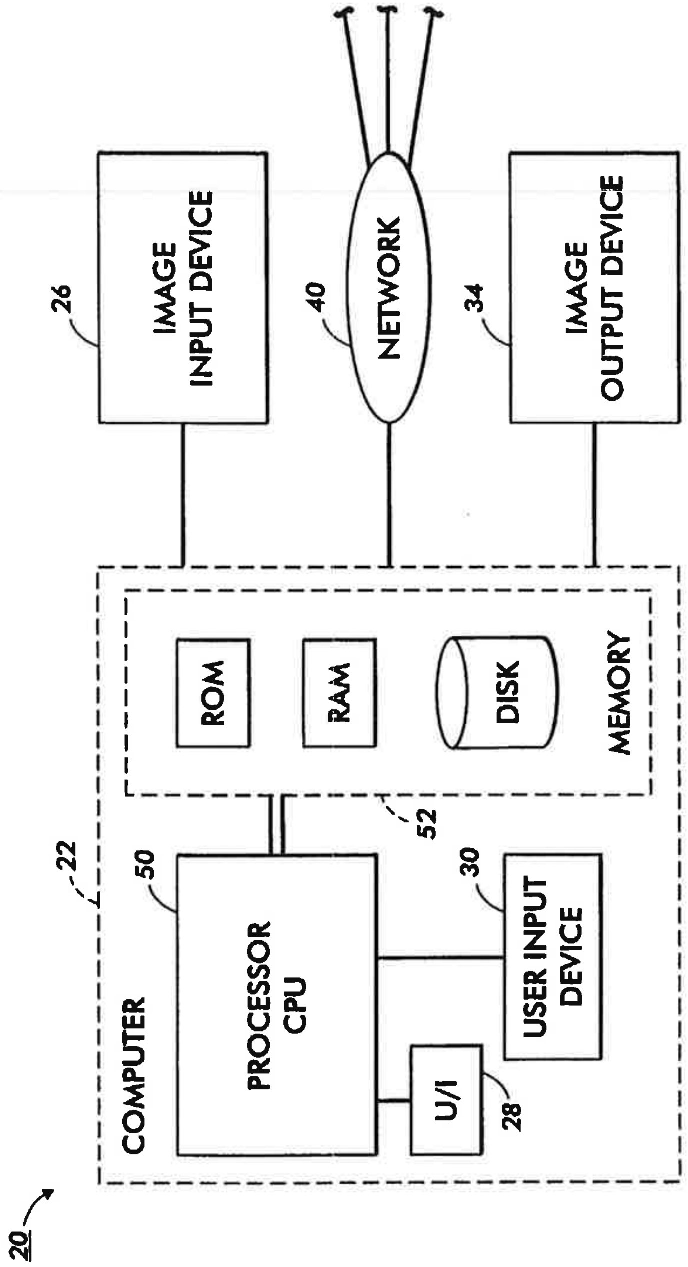

26
IMAGE
INPUT DEVICE
40
NETWORK
34
IMAGE
OUTPUT DEVICE
ROM
RAM
DISK
MEMORY
52
22
50
30
COMPUTER
PROCESSOR
CPU
USER INPUT
DEVICE
U/I
28
20

FIG. 2

FIG. 3

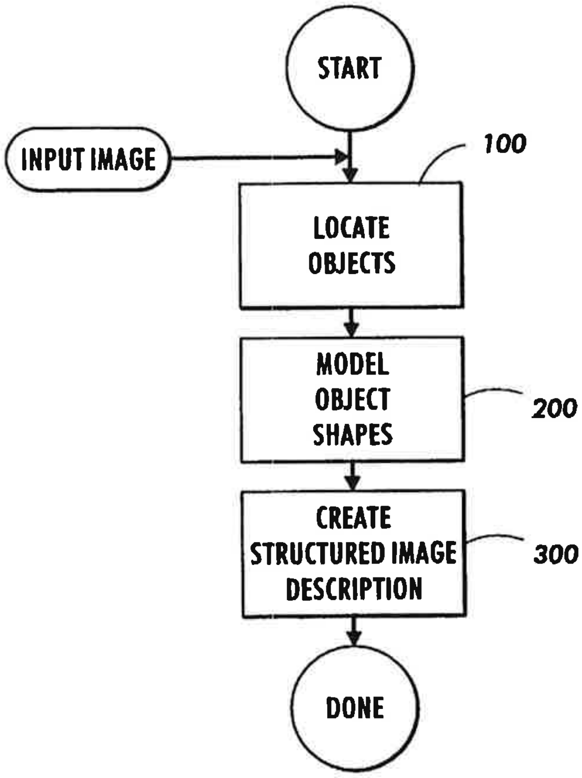

START
INPUT IMAGE
100
LOCATE OBJECTS
MODEL OBJECT SHAPES
200
CREATE STRUCTURED IMAGE DESCRIPTION
300
DONE

FIG. 4

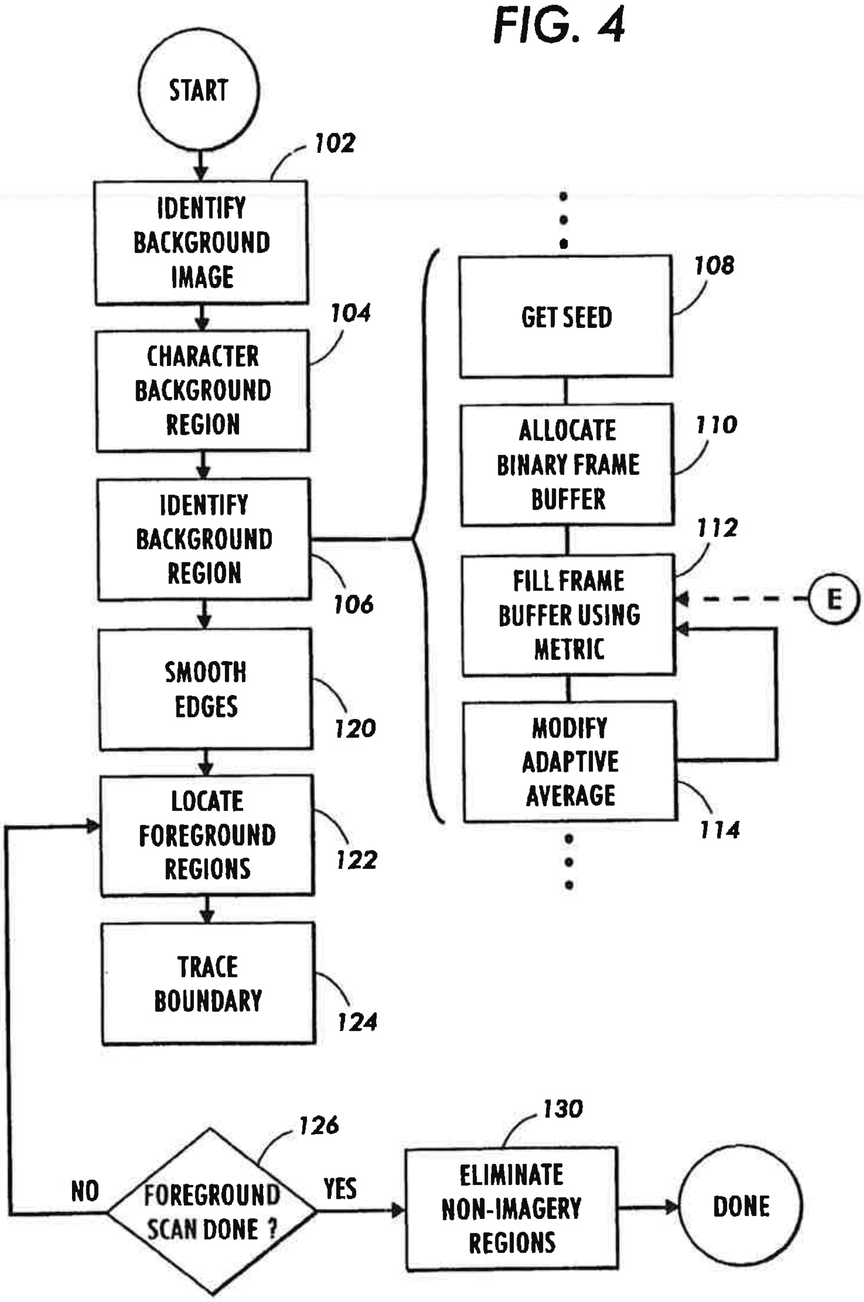
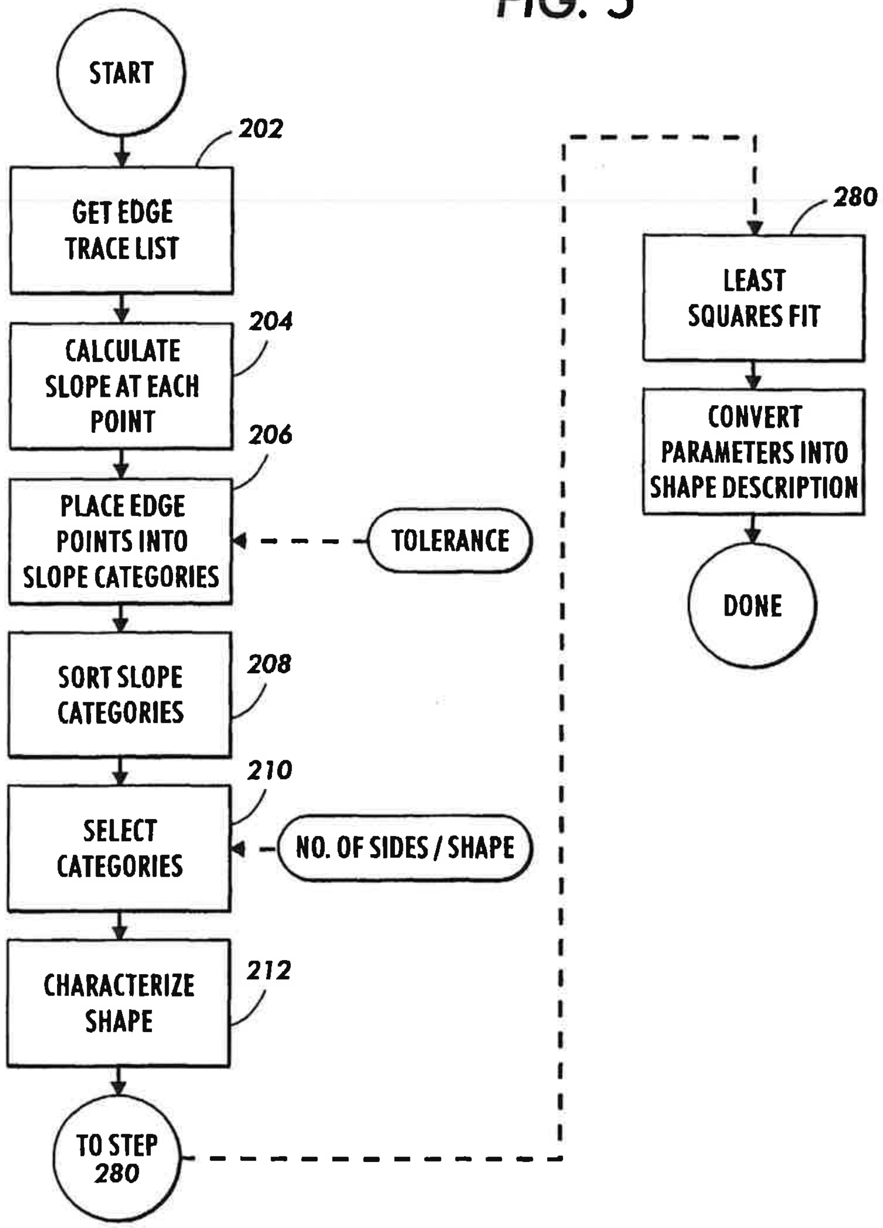

START
102
IDENTIFY BACKGROUND IMAGE
104
CHARACTER BACKGROUND REGION
IDENTIFY BACKGROUND REGION
106
SMOOTH EDGES
120
LOCATE FOREGROUND REGIONS
122
TRACE BOUNDARY
124
GET SEED
108
ALLOCATE BINARY FRAME BUFFER
110
FILL FRAME BUFFER USING METRIC
112
E
MODIFY ADAPTIVE AVERAGE
114
126
NO
FOREGROUND SCAN DONE ?
YES
130
ELIMINATE NON-IMAGERY REGIONS
DONE

FIG. 5


START
202
GET EDGE TRACE LIST
204
CALCULATE SLOPE AT EACH POINT
206
PLACE EDGE POINTS INTO SLOPE CATEGORIES
TOLERANCE
208
SORT SLOPE CATEGORIES
210
SELECT CATEGORIES
NO. OF SIDES / SHAPE
212
CHARACTERIZE SHAPE
TO STEP 280
280
LEAST SQUARES FIT
CONVERT PARAMETERS INTO SHAPE DESCRIPTION
DONE

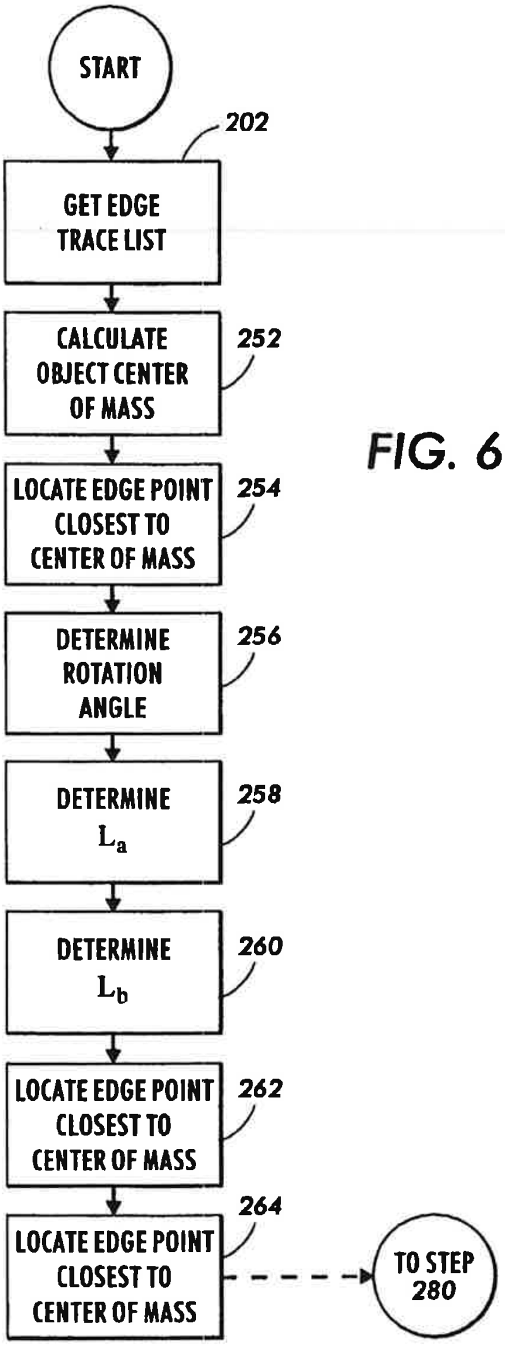
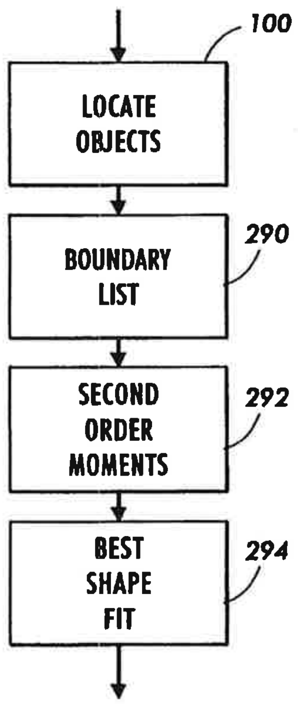

START
202
GET EDGE TRACE LIST
CALCULATE OBJECT CENTER OF MASS
252
LOCATE EDGE POINT CLOSEST TO CENTER OF MASS
254
DETERMINE ROTATION ANGLE
256
DETERMINE L_a
258
DETERMINE L_b
260
LOCATE EDGE POINT CLOSEST TO CENTER OF MASS
262
LOCATE EDGE POINT CLOSEST TO CENTER OF MASS
264
TO STEP 280

FIG. 6

FIG. 7


100
LOCATE OBJECTS
BOUNDARY LIST
290
SECOND ORDER MOMENTS
292
BEST SHAPE FIT
294

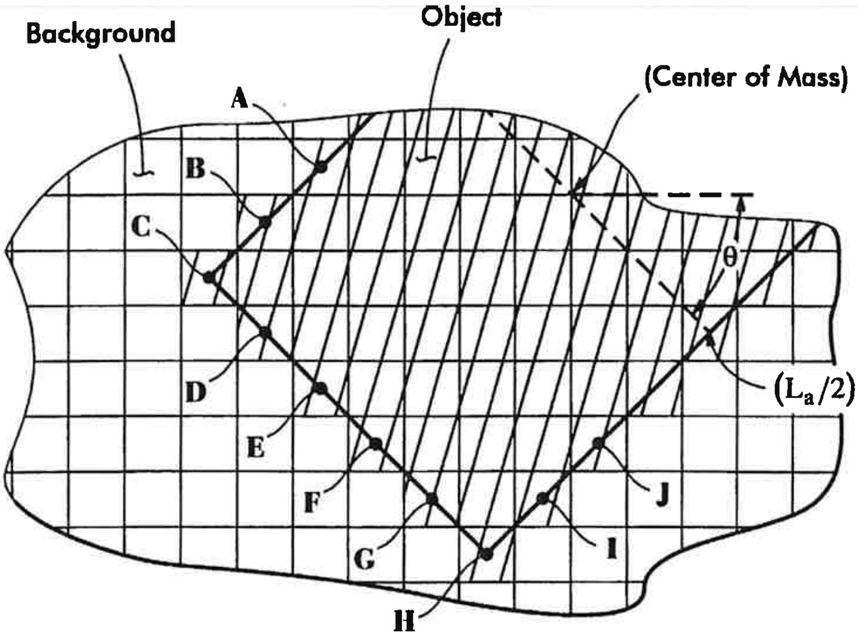


FIG. 8

```
#  Structured Image ver1.0
sid : smp = {
  aspect_ratio = 1.0 ;
  representation = {
    format = ipd ;
    data = [
      merge = {
        xy = 0.000000 0.132377 ;
        path = {
          object = {                      #first detected object
            size = 0.500000 0.735247 ;
            sid : image1 = {
              aspect_ratio = 1.470494 ;
              representation = {
                format = raster ;
                data = "smp.int" ;        # scanned image
                attribute = {
                  selection = $sel1 ;     # object rectangle
                  derotate = 14.445258 ;  # derotation angle
                } ;
              } ;
            } ;
          } ;
        } ;
      } ;
      merge = {
        xy = 0.500000 0.180773 ;
        path = {
          object = {                      #second detected object
            size = 0.500000 0.638454 ;
            sid : image0 = {
              aspect_ratio = 1.276908 ;
              representation = {
                format = raster ;
                data = "smp.int" ;        # scanned image
                attribute = {
                  selection = $sel0 ;     # object rectangle
                  derotate = 12.238364 ;  # derotation angle
                } ;
              } ;
            } ;
          } ;
        } ;
      } ;
    ] ;
  } ;
} ;

selection : sel1 = Selection: In          # rectangle of second object
rect: Include
0.858490 0.958967
```

FIG. 9

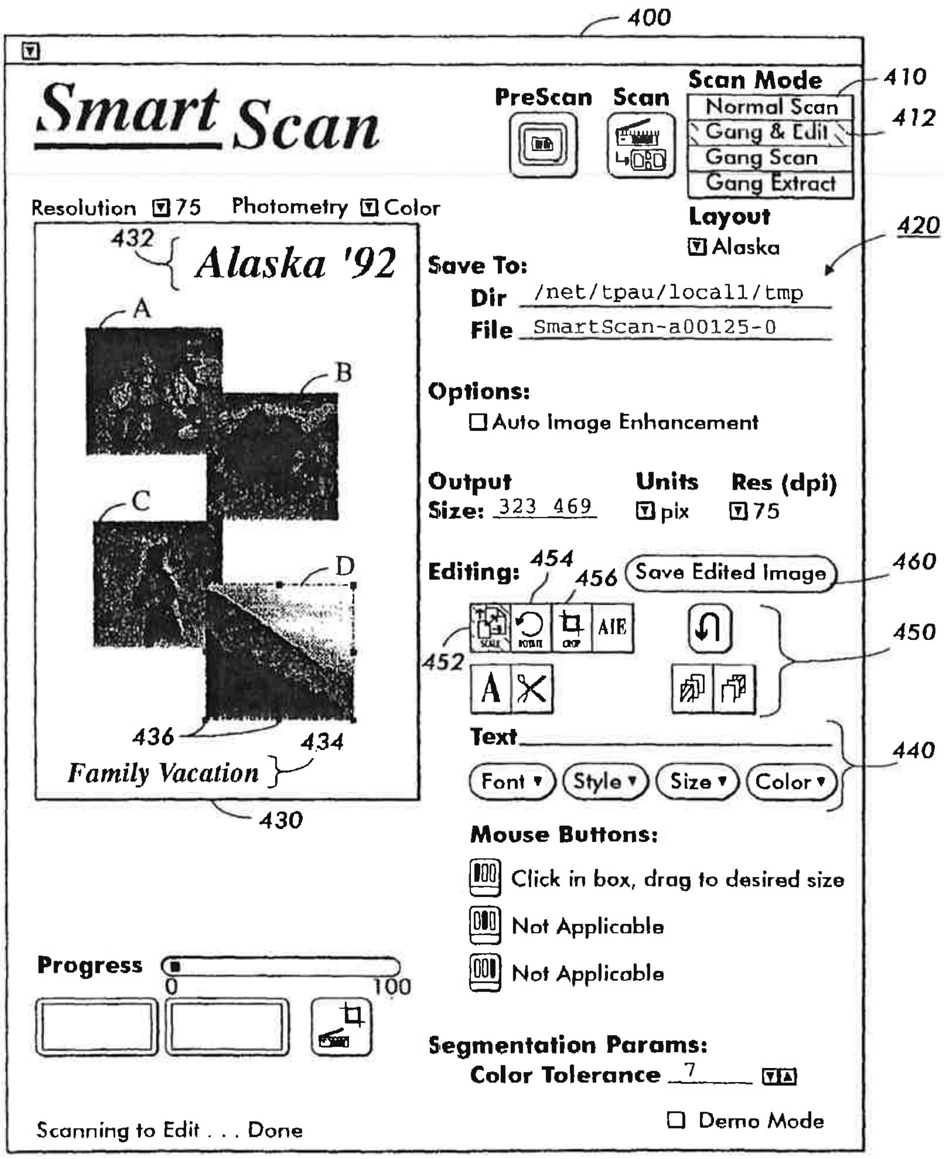

400
Smart Scan
PreScan
Scan
Scan Mode
Normal Scan
Gang & Edit
Gang Scan
Gang Extract
410
412
Resolution 75
Photometry Color
Layout
Alaska
420
432
Alaska '92
A
B
C
D
436
434
Family Vacation
430
Save To:
Dir /net/tpau/local1/tmp
File SmartScan-a00125-0
Options:
Auto Image Enhancement
Output Size: 323 469
Units pix
Res (dpi) 75
Editing:
454
456
Save Edited Image
460
452
AIE
A
450
Text
Font
Style
Size
Color
440
Mouse Buttons:
Click in box, drag to desired size
Not Applicable
Not Applicable
Progress
0
100
Segmentation Params:
Color Tolerance 7
Scanning to Edit . . . Done
Demo Mode

FIG. 10

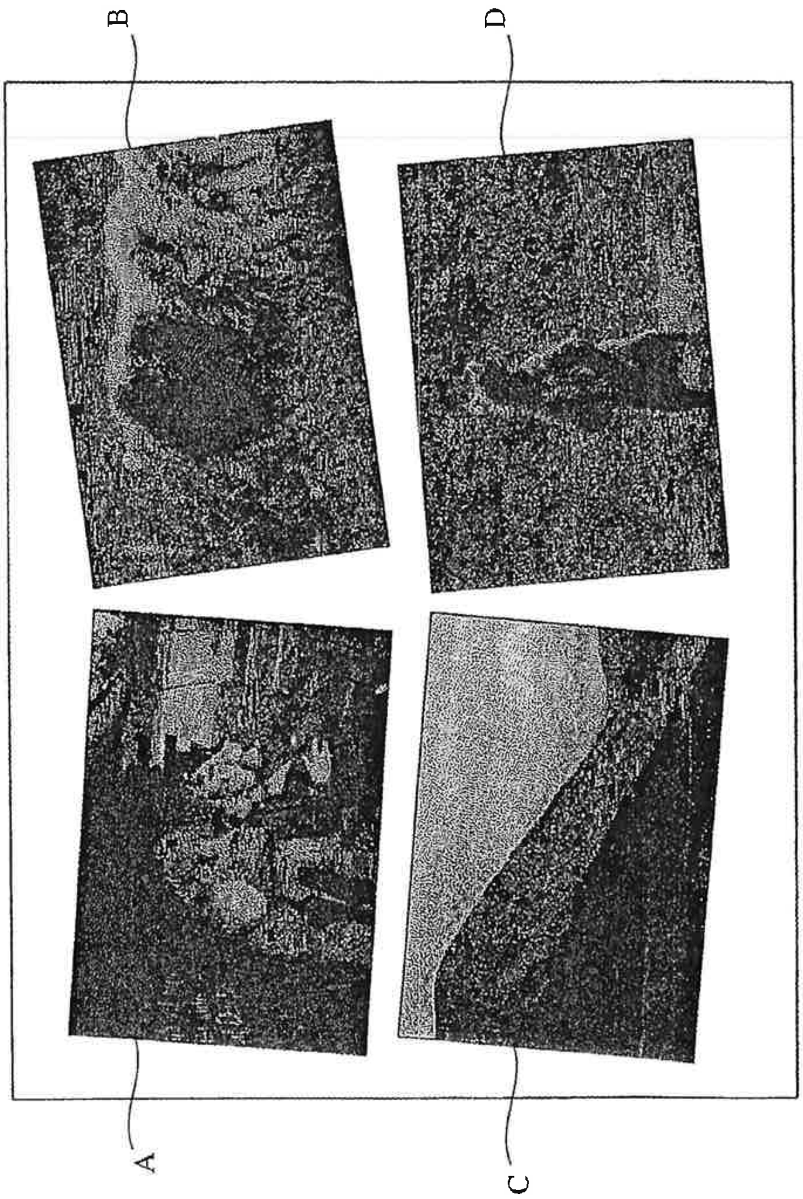


FIG. 11

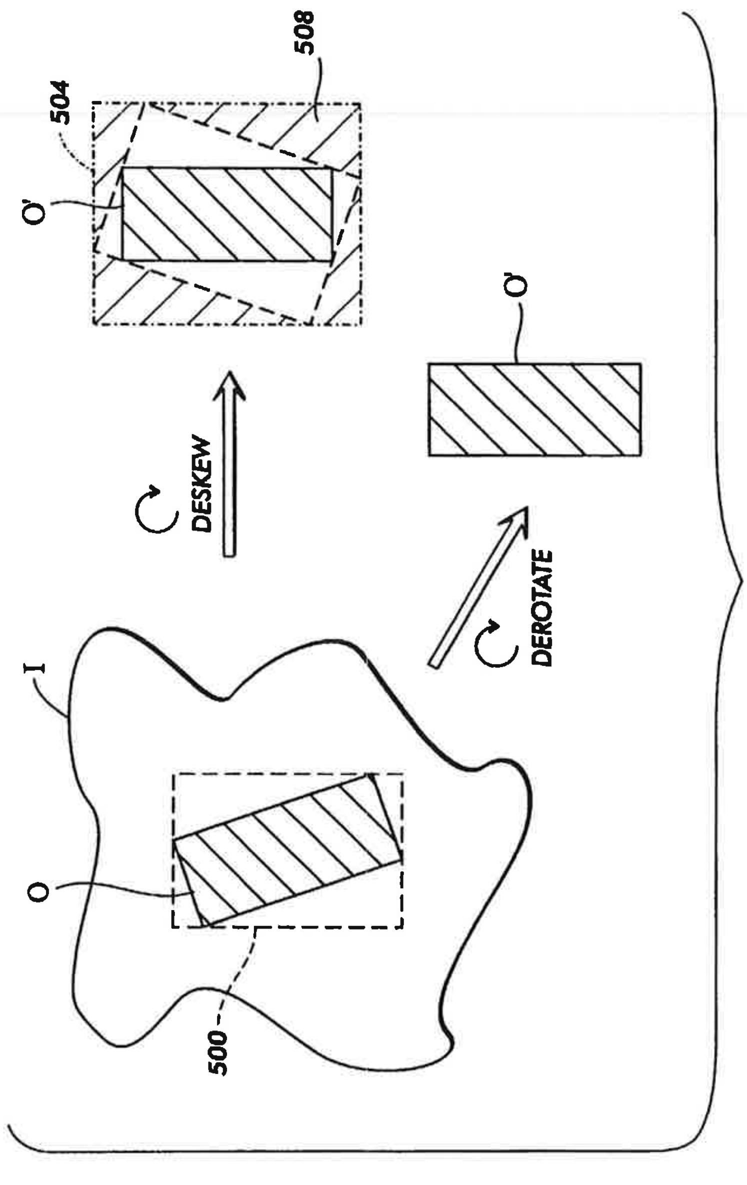


FIG. 12

European Patent Office

EUROPEAN SEARCH REPORT

Application Number
EP 98 30 5918

DOCUMENTS CONSIDERED TO BE RELEVANT

	Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.6)
4	Y	US 5 440 403 A (YOSHIYUKI HASHIMOTO ET AL.) 8 August 1995 (1995-08-08) * column 6, line 13 - column 11, line 23 * * column 13, line 31 - line 37 *	1-17	H04N1/387
2	Y	EP 0 431 962 A (XEROX CORPORATION) 12 June 1991 (1991-06-12) * column 6, line 29 - column 7, line 9 *	1-17	
4	A	EP 0 478 334 A (XEROX CORPORATION) 1 April 1992 (1992-04-01) * column 1, line 57 - column 4, line 24 *	2,4, 10-17	
1	D,A	US 5 485 568 A (D. L. VENABLE ET AL.) 16 January 1996 (1996-01-16) * column 11, line 17 - line 43 *	8,9	

TECHNICAL FIELDS SEARCHED (Int.Cl.6)
H04N

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
THE HAGUE	8 January 1999	DE ROECK, A

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)

ANNEX TO THE EUROPEAN SEARCH REPORT ON EUROPEAN PATENT APPLICATION NO. EP 98 30 5918

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report. The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

08-01-1999

Patent document cited in search report		Publication date	Patent family member(s)			Publication date
US 5440403	A	08-08-1995	JP	6046242	A	18-02-1994
			JP	6152874	A	31-05-1994
			JP	6152930	A	31-05-1994
EP 431962	A	12-06-1991	US	5187753	A	16-02-1993
			DE	69032421	D	23-07-1998
			DE	69032421	T	03-12-1998
			JP	3268189	A	28-11-1991
			US	5355420	A	11-10-1994
EP 478334	A	01-04-1992	AU	8256691	A	02-04-1992
			CA	2048525	A	29-03-1992
			JP	4299465	A	22-10-1992
US 5485568	A	16-01-1996	CA	2131439	A	09-04-1995
			EP	0647921	A	12-04-1995
			JP	7282284	A	27-10-1995

EPO FORM P0459

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 1 017 019 A2

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
05.07.2000 Bulletin 2000/27

(51) Int. Cl.^{7}: G06T 7/00

(21) Application number: 99204116.0

(22) Date of filing: 03.12.1999

(84) Designated Contracting States:
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE
Designated Extension States:
AL LT LV MK RO SI

(30) Priority: 31.12.1998 US 223860

(71) Applicant: EASTMAN KODAK COMPANY
Rochester, New York 14650 (US)

(72) Inventors:
- Luo, Jiebo,
  c/o Eastman Kodak Company
  Rochester, New York 14650-2201 (US)
- Etz, Stephen,
  c/o Eastman Kodak Company
  Rochester, New York 14650-2201 (US)
- Singhal, Amit,
  c/o Eastman Kodak Company
  Rochester, New York 14650-2201 (US)

(74) Representative:
Parent, Yves et al
KODAK INDUSTRIE
Département Brevets - CRT
Zone Industrielle
B.P. 21
71102 Chalon-sur-Saône Cédex (FR)

(54) **Method for automatic determination of main subjects in photographic images**

(57) A method for detecting a main subject in an image, the method comprises: receiving a digital image; extracting regions of arbitrary shape and size defined by actual objects from the digital image; grouping the regions into larger segments corresponding to physically coherent objects; extracting for each of the regions at least one structural saliency feature and at least one semantic saliency feature; and integrating saliency features using a probabilistic reasoning engine into an estimate of a belief that each region is the main subject.

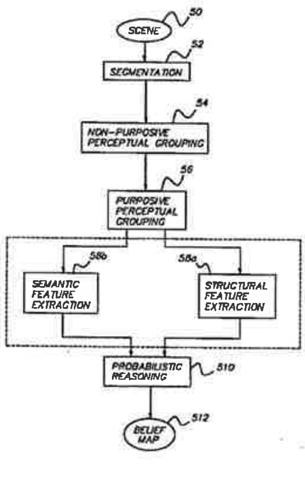


FIG. 2

EP 1 017 019 A2

Printed by Xerox (UK) Business Services
2.16.7 (HRS)/3.6

Description

FIELD OF THE INVENTION

[0001] The invention relates generally to the field of digital image processing and, more particularly, to locating main subjects, or equivalently, regions of photographic interest in a digital image.

BACKGROUND OF THE INVENTION

[0002] In photographic pictures, a main subject is defined as what the photographer tries to capture in the scene. The first-party truth is defined as the opinion of the photographer and the third-party truth is defined as the opinion from an observer other than the photographer and the subject (if applicable). In general, the first-party truth typically is not available due to the lack of specific knowledge that the photographer may have about the people, setting, event, and the like. On the other hand, there is, in general, good agreement among third-party observers if the photographer has successfully used the picture to communicate his or her interest in the main subject to the viewers. Therefore, it is possible to design a method to automatically perform the task of detecting main subjects in images.

[0003] Main subject detection provides a measure of saliency or relative importance for different regions that are associated with different subjects in an image. It enables a discriminative treatment of the scene contents for a number of applications. The output of the overall system can be modified versions of the image, semantic information, and action.

[0004] The methods disclosed by the prior art can be put in two major categories. The first category is considered "pixel-based" because such methods were designed to locate interesting pixels or "spots" or "blocks", which usually do not correspond to entities of objects or subjects in an image. The second category is considered "region-based" because such methods were designed to locate interesting regions, which correspond to entities of objects or subjects in an image.

[0005] Most pixel-based approaches to region-of-interest detection are essentially edge detectors. V. D. Gesu, et al., "Local operators to detect regions of interest," *Pattern Recognition Letters*, vol. 18, pp. 1077-1081, 1997, used two local operators based on the computation of local moments and symmetries to derive the selection. Arguing that the performance of a visual system is strongly influenced by information processing done at early vision stage, two transforms named the discrete moment transform (DMT) and discrete symmetry transform (DST) are computed to measure local central moments about each pixel and local radial symmetry. In order to exclude trivial symmetry cases, nonuniform region selection is needed. The specific DMT operator acts like a detector of prominent edges (occlusion boundaries) and the DST operator acts like a detector of symmetric blobs. The results from the two operators are combined via logic "AND" operation. Some morphological operations are needed to dilate the edge-like raw output map generated by the DMT operator.

[0006] R. Milanese, *Detecting salient regions in an image: From biology to implementation*, PhD thesis, University of Geneva, Switzerland, 1993, developed a computational model of visual attention, which combines knowledge about the human visual system with computer vision techniques. The model is structured into three major stages. First, multiple feature maps are extracted from the input image (for examples, orientation, curvature, color contrast and the like). Second, a corresponding number of "conspicuity" maps are computed using a derivative of Gaussian model, which enhance regions of interest in each feature map. Finally, a nonlinear relaxation process is used to integrate the conspicuity maps into a single representation by finding a compromise among inter-map and intra-map inconsistencies. The effectiveness of the approach was demonstrated using a few relatively simple images with remarkable regions of interest.

[0007] To determine an optimal tonal reproduction, J. R. Boyack, et al., U.S. Patent No. 5,724,456, developed a system that partitions the image into blocks, combines certain blocks into sectors, and then determines a difference between the maximum and minimum average block values for each sector. A sector is labeled an active sector if the difference exceeds a pre-determined threshold value. All weighted counts of active sectors are plotted versus the average luminance sector values in a histogram, which is then shifted via some predetermined criterion so that the average luminance sector value of interest will fall within a destination window corresponding to the tonal reproduction capability of a destination application.

[0008] In summary, this type of pixel-based approach does not explicitly detect region of interest corresponding to semantically meaningful subjects in the scene. Rather, these methods attempt to detect regions where certain changes occur in order to direct attention or gather statistics about the scene.

[0009] X. Marichal, et al., "Automatic detection of interest areas of an image or of a sequence of images," in *Proc. IEEE Int. Conf. Image Process*., 1996, developed a fuzzy logic-based system to detect interesting areas in a video sequence. A number of subjective knowledge-based interest criteria were evaluated for segmented regions in an image. These criteria include: (1) an interaction criterion (a window predefined by a human operator); (2) a border criterion (rejecting of regions having large number of pixels along the picture borders); (3) a face texture criterion (de-emphasizing regions whose texture does not correspond to skin samples); (4) a motion criterion (rejecting regions with no motion and low gradient or regions with very large motion and high gradient); and (5) a continuity criterion (temporal stability in motion). The main application of this method is for directing the resources in video coding, in particular for videophone or videoconference. It is clear that motion is the most effective criterion for this technique targeted at video instead of still images. Moreover, the fuzzy logic functions were designed in an ad hoc fashion. Lastly, this method requires a window predefined by a human operator, and therefore is not fully automatic.

[0010] W. Osberger, et al., "Automatic identification of perceptually important regions in an image," in *Proc. IEEE Int. Conf Pattern Recognition*, 1998, evaluated several features known to influence human visual attention for each region of a segmented image to produce an importance value for each feature in each region. The features mentioned include low-level factors (contrast, size, shape, color, motion) and higher level factors (location, foreground/background, people, context), but only contrast, size, shape, location and foreground/background (determining background by determining the proportion of total image border that is contained in each region) were implemented. Moreover, this method chose to treat each factor as being of equal importance by arguing that (1) there is little quantitative data which indicates the relative importance of these different factors and (2) the relative importance is likely to change from one image to another. Note that segmentation was obtained using the split-and-merge method based on 8 x 8 image blocks and this segmentation method often results in over-segmentation and blotchiness around actual objects.

[0011] Q. Huang, et al., "Foreground/background segmentation of color images by integration of multiple cues," in *Proc. IEEE Int. Conf. Image Process.*, 1995, addressed automatic segmentation of color images into foreground and background with the assumption that background regions are relatively smooth but may have gradually varying colors or be lightly textured. A multi-level segmentation scheme was devised that included color clustering, unsupervised segmentation based on MDL (Minimum Description Length) principle, edge-based foreground/background separation, and integration of both region and edge-based segmentation. In particular, the MDL-based segmentation algorithm was used to further group the regions from the initial color clustering, and the four corners of the image were used to adaptively determine an estimate of the background gradient magnitude. The method was tested on around 100 well-composed images with prominent main subject centered in the image against large area of the assumed type of uncluttered background.

[0012] T. F. Syeda-Mahmood, "Data and model-driven selection using color regions," *Int. J. Comput. Vision*, vol. 21, no. 1, pp. 9-36, 1997, proposed a data-driven region selection method using color region segmentation and region-based saliency measurement. A collection of 220 primary color categories was pre-defined in the form of a color LUT (look-up-table). Pixels are mapped to one of the color categories, grouped together through connected component analysis, and further merged according to compatible color categories. Two types of saliency measures, namely self saliency and relative saliency, are linearly combined using heuristic weighting factors to determine the overall saliency. In particular, self-saliency included color saturation, brightness and size while relative saliency included color contrast (defined by CIE distance) and size contrast between the concerned region and the surrounding region that is ranked highest among neighbors by size, extent and contrast in successive order.

[0013] In summary, almost all of these reported methods have been developed for targeted types of images: video-conferencing or TV news broadcasting images, where the main subject is a talking person against a relatively simple static background (Osberg, Marichal); museum images, where there is a prominent main subject centered in the image against large area of relatively clean background (Huang); and toy-world images, where the main subject are a few distinctively colored and shaped objects (Milanese, Syeda). These methods were either not designed for unconstrained photographic images, or even if designed with generic principles were only demonstrated for their effectiveness on rather simple images. The criteria and reasoning processes used were somewhat inadequate for less constrained images, such as photographic images.

SUMMARY OF THE INVENTION

[0014] It is an object of this invention to provide a method for detecting the location of main subjects within a digitally captured image and thereby overcoming one or more problems set forth above.

[0015] It is also an object of this invention to provide a measure of belief for the location of main subjects within a digitally captured image and thereby capturing the intrinsic degree of uncertainty in determining the relative importance of different subjects in an image. The output of the algorithm is in the form of a list of segmented regions ranked in a descending order of their likelihood as potential main subjects for a generic or specific application. Furthermore, this list can be converted into a map in which the brightness of a region is proportional to the main subject belief of the region.

[0016] It is also an object of this invention to use ground truth data. Ground truth, defined as human outlined main subjects, is used to feature selection and training the reasoning engine.

[0017] It is also an object of this invention to provide a method of finding main subjects in an image in an automatic manner.

[0018] It is also an object of this invention to provide a method of finding main subjects in an image with no constraints or assumptions on scene contents.

[0019] It is further an object of the invention to use the main subject location and main subject belief to obtain estimates of the scene characteristics.

[0020] The present invention comprises the steps of:

a) receiving a digital image;
b) extracting regions of arbitrary shape and size defined by actual objects from the digital image;
c) grouping the regions into larger segments corresponding to physically coherent objects;
d) extracting for each of the regions at least one structural saliency feature and at least one semantic saliency feature; and,
e) integrating saliency features using a probabilistic reasoning engine into an estimate of a belief that each region is the main subject.

[0021] The above and other objects of the present invention will become more apparent when taken in conjunction with the following description and drawings wherein identical reference numerals have been used, where possible, to designate identical elements that are common to the figures.

ADVANTAGEOUS EFFECT OF THE INVENTION

[0022] The present invention has the following advantages of:

a robust image segmentation method capable of identifying object regions of arbitrary shapes and sizes, based on physics-motivated adaptive Bayesian clustering and non-purposive grouping;
emphasis on perceptual grouping capable of organizing regions corresponding to different parts of physically coherent subjects;
utilization of a non-binary representation of the ground truth, which capture the inherent uncertainty in determining the belief of main subject, to guide the design of the system;
a rigorous, systematic statistical training mechanism to determine the relative importance of different features through ground truth collection and contingency table building;
extensive, robust feature extraction and evidence collection;
combination of structural saliency and semantic saliency, the latter facilitated by explicit identification of key foreground- and background- subject matters;
combination of self and relative saliency measures for structural saliency features; and,
a robust Bayes net-based probabilistic inference engine suitable for integrating incomplete information.

BRIEF DESCRIPTION OF THE DRAWINGS

[0023]

Fig. 1 is a perspective view of a computer system for implementing the present invention;
Fig. 2 is a block diagram illustrating a software program of the present invention;
Fig. 3 is an illustration of the sensitivity characteristic of a belief sensor with sigmoidal shape used in the present invention;
Fig. 4 is an illustration of the location PDF with unknown-orientation, Fig. 4(a) is an illustration of the PDF in the form of a 2D function, Fig. 4(b) is an illustration of the PDF in the form of its projection along the width direction, and Fig. 4(c) is an illustration of the PDF in the form of its projection along the height direction;
Fig. 5 is an illustration of the location PDF with known-orientation, Fig. 5(a) is an illustration of the PDF in the form of a 2D function, Fig. 5(b) is an illustration of the PDF in the form of its projection along the width direction, and Fig. 5(c) is an illustration of the PDF in the form of its projection along the height direction;
Fig. 6 is an illustration of the computation of relative saliency for the central circular region using an extended neighborhood as marked by the box of dotted line;
Fig. 7 is an illustration of a two-level Bayes net used in the present invention; and,
Fig. 8 is block diagram of a preferred segmentation method.

DETAILED DESCRIPTION OF THE INVENTION

[0024] In the following description, the present invention will be described in the preferred embodiment as a software program. Those skilled in the art will readily recognize that the equivalent of such software may also be constructed in hardware.

[0025] Still further, as used herein, computer readable storage medium may comprise, for example; magnetic storage media such as a magnetic disk (such as a floppy disk) or magnetic tape; optical storage media such as an optical disc, optical tape, or machine readable bar code; solid state electronic storage devices such as random access memory (RAM), or read only memory (ROM); or any other physical device or medium employed to store a computer program.

[0026] Referring to Fig. 1, there is illustrated a computer system 10 for implementing the present invention. Although the computer system 10 is shown for the purpose of illustrating a preferred embodiment, the present invention is not limited to the computer system 10 shown, but may be used on any electronic processing system. The computer system 10 includes a microprocessor based unit 20 for receiving and processing software programs and for performing other processing functions. A touch screen display 30 is electrically connected to the microprocessor based unit 20 for displaying user related information associated with the software, and for receiving user input via touching the screen. A keyboard 40 is also connected to the microprocessor based unit 20 for permitting a user to input information to the software. As an alternative to using the keyboard 40 for input, a mouse 50 may be used for moving a selector 52 on the display 30 and for selecting an item on which the selector 52 overlays, as is well known in the art.

[0027] A compact disk-read only memory (CD-ROM) 55 is connected to the microprocessor based unit 20 for receiving software programs and for providing a means of inputting the software programs and other information to the microprocessor based unit 20 via a compact disk 57, which typically includes a software program. In addition, a floppy disk 61 may also include a software program, and is inserted into the microprocessor based unit 20 for inputting the software program. Still further, the microprocessor based unit 20 may be programmed, as is well know in the art, for storing the software program internally. A printer 56 is connected to the microprocessor based unit 20 for printing a hardcopy of the output of the computer system 10.

[0028] Images may also be displayed on the display 30 via a personal computer card (PC card) 62 or, as it was formerly known, a personal computer memory card international association card (PCMCIA card) which contains digitized images electronically embodied the card 62. The PC card 62 is ultimately inserted into the microprocessor based unit 20 for permitting visual display of the image on the display 30.

[0029] Referring to Fig. 2, there is shown a block diagram of an overview of the present invention. First, an input image of a natural scene is acquired and stored S0 in a digital form. Then, the image is segmented S2 into a few regions of homogeneous properties. Next, the region segments are grouped into larger regions based on similarity measures S4 through non-purposive perceptual grouping, and further grouped into larger regions corresponding to perceptually coherent objects S6 though purposive grouping (purposive grouping concerns specific objects). The regions are evaluated for their saliency S8 using two independent yet complementary types of saliency features - structural saliency features and semantic saliency features. The structural saliency features, including a set of low-level early vision features and a set of geometric features, are extracted S8a, which are further processed to generate a set of self saliency features and a set of relative saliency features. Semantic saliency features in the forms of key subject matters, which are likely to be part of either foreground (for example, people) or background (for example, sky, grass), are detected S8b to provide semantic cues as well as scene context cues. The evidences of both types are integrated S10 using a reasoning engine based on a Bayes net to yield the final belief map of the main subject S12.

[0030] To the end of semantic interpretation of images, a single criterion is clearly insufficient. The human brain, furnished with its *a priori* knowledge and enormous memory of real world subjects and scenarios, combines different subjective criteria in order to give an assessment of the interesting or primary subject(s) in a scene. The following extensive list of features are believed to have influences on the human brain in performing such a somewhat intangible task as main subject detection: location, size, brightness, colorfulness, texturefulness, key subject matter, shape, symmetry, spatial relationship (surroundedness/occlusion), borderness, indoor/outdoor, orientation, depth (when applicable), and motion (when applicable for video sequence).

[0031] In the present invention, the low-level early vision features include color, brightness, and texture. The geometric features include location (centrality), spatial relationship (borderness, adjacency, surroundedness, and occlusion), size, shape, and symmetry. The semantic features include flesh, face, sky, grass, and other green vegetation. Those skilled in the art can define more features without departing from the scope of the present invention.

S2: Region Segmentation

[0032] The adaptive Bayesian color segmentation algorithm (Luo et al., "Towards physics-based segmentation of photographic color images," Proceedings of the IEEE International Conference on Image Processing, 1997) is used to generate a tractable number of physically coherent regions of arbitrary shape. Although this segmentation method is preferred, it will be appreciated that a person of ordinary skill in the art can use a different segmentation method to obtain object regions of arbitrary shape without departing from the scope of the present invention. Segmentation of arbitrarily shaped regions provides the advantages of: (1) accurate measure of the size, shape, location of and spatial relationship among objects; (2) accurate measure of the color and texture of objects; and (3) accurate classification of key subject matters.

[0033] Referring to Fig. 8, there is shown a block diagram of the preferred segmentation algorithm. First, an initial segmentation of the image into regions is obtained S50. A color histogram of the image is computed and then partitioned into a plurality of clusters that correspond to distinctive, prominent colors in the image. Each pixel of the image is classified to the closest cluster in the color space according to a preferred physics-based color distance metric with respect to the mean values of the color clusters (Luo et al., "Towards physics-based segmentation of photographic color images," Proceedings of the IEEE International Conference on Image Processing, 1997). This classification process results in an initial segmentation of the image. A neighborhood window is placed at each pixel in order to determine what neighborhood pixels are used to compute the local color histogram for this pixel. The window size is initially set at the size of the entire image S52, so that the local color histogram is the same as the one for the entire image and does not need to be recomputed. Next, an iterative procedure is performed between two alternating processes: re-computing S54 the local mean values of each color class based on the current segmentation, and re-classifying the pixels according to the updated local mean values of color classes S56. This iterative procedure is performed until a convergence is reached S60. During this iterative procedure, the strength of the spatial constraints can be adjusted in a gradual manner S58 (for example, the value of β, which indicates the strength of the spatial constraints, is increased linearly with each iteration). After the convergence is reached for a particular window size, the window used to estimate the local mean values for color classes is reduced by half in size S62. The iterative procedure is repeated for the reduced window size to allow more accurate estimation of the local mean values for color classes. This mechanism introduces spatial adaptivity into the segmentation process. Finally, segmentation of the image is obtained when the iterative procedure reaches convergence for the minimum window size S64.

S4 & S6: Perceptual Grouping

[0034] The segmented regions may be grouped into larger segments that consist of regions that belong to the same object. Perceptual grouping can be non-purposive and purposive. Referring to Fig. 2, non-purposive perceptual grouping S4 can eliminate over-segmentation due to large illumination differences, for example, a table or wall with remarkable illumination falloff over a distance. Purposive perceptual grouping S6 is generally based on smooth, noncoincidental connection of joints between parts of the same object, and in certain cases models of typical objects (for example, a person has head, torso and limbs).

[0035] Perceptual grouping facilitates the recognition of high-level vision features. Without proper perceptual grouping, it is difficult to perform object recognition and proper assessment of such properties as size and shape. Perceptual grouping includes: merging small regions into large regions based on similarity in properties and compactness of the would-be merged region (non-purposive grouping); and grouping parts that belong to the same object based on commonly shared background, compactness of the would-be merged region, smoothness in contour connection between regions, and model of specific object (purposive grouping).

S8: Feature Extraction

[0036] For each region, an extensive set of features, which are shown to contribute to visual attention, are extracted and associated evidences are then computed. The list of features consists of three categories - low-level vision features, geometric features, and semantic features. For each feature, either or both of a self-saliency feature and a relative saliency feature are computed. The self-saliency is used to capture subjects that stand out by themselves (for example, in color, texture, location and the like), while the relative saliency is used to capture subjects that are in high contrast to their surrounding (for example, shape). Furthermore, raw measurements of features, self salient or relatively salient, are converted into evidences, whose values are normalized to be within [0, 1.0], by belief sensor functions with appropriate nonlinearity characteristics. Referring to Fig. 3, there is shown a sigmoid-shaped belief sensor function used in the present invention. A raw feature measurement that has a value between a minimum value and a maximum value is mapped to a belief value within [0, 1]. A Gaussian-shaped belief sensor function (not shown) is also used for some features, as will be described hereinbelow.

Structural saliency features

[0037] Structural saliency features include individually or in combination self saliency features and relative saliency features.

[0038] Referring to Fig. 6, an extended neighborhood is used to compute relative saliency features. First, a minimum bounding rectangle (MBR) 14 of a region of concern 10 (shown by the central circular region) is determined. Next, this MBR is extended in all four directions (stopping at the image borders wherever applicable) of the region using an appropriate factor (for example, 2). All regions intersecting this stretched MBR 12, which is indicated by the dotted lines, are considered neighbors of the region. This extended neighborhood ensures adequate context as well natural scalability for computing the relative saliency features.

[0039] The following structural saliency features are computed:

• contrast in hue (a relative saliency feature)

[0040] In terms of color, the contrast in hue between an object and its surrounding is a good indication of the saliency in color.

$$contrast_{color} = \sum_{neighborhood} \frac{\|hue - hue_{surrounding}\|}{hue_{surrounding}} \tag{1}$$

where the neighborhood refers to the context previously defined and henceforth.

• colorfulness (a self-saliency feature) and contrast in colorfulness (a relative saliency feature)

[0041] In terms of colorfulness, the contrast between a colorful object and a dull surrounding is almost as good an indicator as the contrast between a dull object and a colorful surrounding. Therefore, the contrast in colorfulness should always be positive. In general, it is advantageous to treat a self saliency and the corresponding relative saliency as separate features rather than combining them using certain heuristics. The influence of each feature will be determined separately by the training process, which will be described later.

$$colorfulness = saturation \tag{2}$$

$$contrast_{colorfulness} = \frac{\|saturation - saturation_{surrounding}\|}{saturation_{surrounding}} \tag{3}$$

• brightness (a self-saliency feature) and contrast in brightness (a relative saliency feature)

[0042] In terms of brightness, the contrast between a bright object and a dark surrounding is almost as good as the contrast between a dark object and a bright surrounding. In particular, the main subject tends to be lit up in flash scenes.

$$brightness = luminance \tag{4}$$

$$contrast_{brightness} = \frac{|brightness - brightness_{surrounding}|}{brightness_{surrounding}} \tag{5}$$

• texturefulness (a self-saliency feature) and contrast in texturefulness (a relative saliency feature)

[0043] In terms of texturefulness, in general, a large uniform region with very little texture tends to be the background. On the other hand, the contrast between a highly textured object and a nontextured or less textured surrounding is a good indication of main subjects. The same holds for a non-textured or less textured object and a highly textured surrounding.

$$texturefulness = texture\_energy \tag{6}$$

$$contrast_{texturefulness} = \frac{|texturefulness - texturefulness_{surrounding}|}{texturefulness_{surrounding}} \tag{7}$$

• location (a self-saliency feature)

[0044] In terms of location, the main subject tends to be located near the center instead of the peripheral of the image, though not necessarily right in the center of the image. In fact, professional photographers tend to position the main subject at the horizontal gold partition positions.

[0045] The centroid of a region alone is usually not sufficient to indicate the location of the region without any indication of its size and shape. A centrality measure is defined by computing the integral of a probability density function (PDF) over the area of a given region. The PDF is derived from a set of training images, in which the main subject regions are manually outlined, by summing up the ground truth maps over the entire training set. In other words, the PDF represents the distribution of main subjects in terms of location. A more important advantage of this centrality measure is that every pixel of a given region, not just the centroid, contributes to the centrality measure of the region to a varying degree depending on its location.

$$centrality = \frac{1}{N_R} \sum_{(x,y) \in R} PDF_{MSD\_location}(x, y) \tag{8}$$

where (x,y) denotes a pixel in the region R, N_R is the number of pixels in region R, and $PDF_{MSD\_location}$ denotes a 2D probability density function (PDF) of main subject location. If the orientation is unknown, the PDF is symmetric about the center of the image in both vertical and horizontal directions, which results in an orientation-independent centrality measure. An orientation-unaware PDF is shown in Fig. 4(a) and the projection in the width and height directions are also shown in Fig. 4(b) and Fig. 4(c), respectively. If the orientation is known, the PDF is symmetric about the center of the image in the horizontal direction but not in the vertical direction, which results in an orientation-aware centrality measure. An orientation-aware PDF is shown in Fig. 5(a) and the projection in the horizontal and vertical directions are also shown in Fig. 5(b) and Fig. 5(c), respectively.

• size (a self saliency feature)

[0046] Main subjects should have considerable but reasonable sizes. However, in most cases, very large regions or regions that span at least one spatial direction (for example, the horizontal direction) are most likely to be background regions, such as sky, grass, wall, snow, or water. In general, both very small and very large regions should be discounted.

$$size = \begin{cases} 0 & if\ s > s4 \\ 1 - \dfrac{s - s2}{s3 - s2} & if\ s > s3\ and\ s < s4 \\ 1 & if\ s > s2\ and\ s < s3 \\ \dfrac{s - s1}{s2 - s1} & if\ s > s1\ and\ s < s2 \\ 0 & if\ s < s1 \end{cases} \tag{9}$$

where s1, s2, s3, and s4 are predefined threshold (s1 < s2 < s3 < s4).

[0047] In practice, the size of a region is measured as a fraction of the entire image size to achieve invariance to scaling.

$$size = \frac{region_{pixels}}{image_{pixels}} \tag{10}$$

[0048] In this invention, the region size is classified into one of three bins, labeled "small," "medium" and "large" using two thresholds s2 and s3, where s2 < s3.

• shape (a self-saliency feature) and contrast in shape (a relative saliency feature)

[0049] In general, objects that have distinctive geometry and smooth contour tend to be man-made and thus have high likelihood to be main subjects. For example, square, round, elliptic, or triangle shaped objects. In some cases, the contrast in shape indicates conspicuity (for example, a child among a pool of bubble balls).

[0050] The shape features are divided into two categories, self salient and relatively salient. Self salient features characterize the shape properties of the regions themselves and relatively salient features characterize the shape properties of the regions in comparison to those of neighboring regions.

[0051] The aspect ratio of a region is the major axis/minor axis of the region. A Gaussian belief function maps the aspect ratio to a belief value. This feature detector is used to discount long narrow shapes from being part of the main subject.

[0052] Three different measures are used to characterize the convexity of a region: (1) perimeter-based - perimeter of the convex hull divided by the perimeter of region; (2) area-based - area of region divided by the area of the convex hull; and (3) hyperconvexity - the ratio of the perimeter-based convexity and area-based convexity. In general, an object of complicated shape has a hyperconvexity greater than 1.0. The three convexity features measure the compactness of the region. Sigmoid belief functions are used to map the convexity measures to beliefs.

[0053] The rectangularity is the area of the MBR of a region divided by the area of the region. A sigmoid belief function maps the rectangularity to a belief value. The circularity is the square of the perimeter of the region divided by the area of region. A sigmoid belief function maps the circularity to a belief value.

[0054] Relative shape-saliency features include relative rectangularity, relative circularity and relative convexity. In particular, each of these relative shape features is defined as the average difference between the corresponding self salient shape feature of the region and those of the neighborhood regions, respectively. Finally, a Gaussian function is used to map the relative measures to beliefs.

- symmetry (a self-saliency feature)

[0055] Objects of striking symmetry, natural or artificial, are also likely to be of great interest. Local symmetry can be computed using the method described by V. D. Gesu, et al., "Local operators to detect regions of interest," *Pattern Recognition Letters*, vol. 18, pp. 1077-1081, 1997.

- spatial relationship (a relative saliency feature)

[0056] In general, main subjects tend to be in the foreground. Consequently, main subjects tend to share boundaries with a lot of background regions (background clutter), or be enclosed by large background regions such as sky, grass, snow, wall and water, or occlude other regions. These characteristics in terms of spatial relationship may reveal the region of attention. Adjacency, surroundedness and occlusion are the main features in terms of spatial relationship. In many cases, occlusion can be inferred from T-junctions (L. R. Williams, "Perceptual organization of occluding contours," in *Proc. IEEE Int. Conf. Computer Vision*, 1990) and fragments can be grouped based on the principle of perceptual occlusion (J. August, et al., "Fragment grouping via the principle of perceptual occlusion," in *Proc. IEEE Int. Conf. Pattern Recognition*, 1996).

[0057] In particular, a region that is nearly completely surrounded by a single other region is more likely to be the main subject. Surroundedness is measured as the maximum fraction of the region's perimeter that is shared with any one neighboring region. A region that is totally surrounded by a single other region has the highest possible surroundedness value of 1.0.

$$surroundedness = \max_{neighbors} \frac{length\_of\_common\_border}{region\_perimeter} \quad (11)$$

- borderness (a self-saliency feature)

[0058] Many background regions tend to contact one or more of the image borders. In other words, a region that has significant amount of its contour on the image borders tends to belong to the background. The percentage of the contour points on the image borders and the number of image borders shared (at most four) can be good indications of the background.

[0059] In the case where the orientation is unknown, one borderness feature places each region in one of six categories determined by the number and configuration of image borders the region is "in contact" with. A region is "in contact" with a border when at least one pixel in the region falls within a fixed distance of the border of the image. Distance is expressed as a fraction of the shorter dimension of the image. The six categories for *borderness_a* are defined in Table 1.

Table 1

Categories for orientation-independent *borderness_a*.	
Category	**The region is in contact with...**
0	none of the image borders
1	exactly one of the image borders
2	exactly two of the image borders, adjacent to one another
3	exactly two of the image borders, opposite to one another
4	exactly three of the image borders
5	exactly four (all) of the image borders

[0060] Knowing the proper orientation of the image allows us to refine the borderness feature to account for the fact that regions in contact with the top border are much more likely to be background than regions in contact with the bottom. This feature places each region in one of 12 categories determined by the number and configuration of image borders the region is "in contact" with, using the definition of "in contact with" from above. The four borders of the image are labeled as "Top", "Bottom", "Left", and "Right" according to their position when the image is oriented with objects in the scene standing upright. In this case, the twelve categories for *borderness_b* are defined in Table 2, which lists each possible combination of borders a region may be in contact with, and gives the category assignment for that combination.

Table 2

Categories for orientation-dependent *borderness_a*.				
The region is in contact with...				**Category**
Top	Bottom	Left	Right	Category
N	N	N	N	0
N	Y	N	N	1
Y	N	N	N	2
N	N	Y	N	3
N	N	N	Y	3
N	Y	Y	N	4
N	Y	N	Y	4
Y	N	N	N	5
Y	N	N	N	5
Y	Y	N	N	6
N	N	Y	Y	7
N	Y	Y	Y	8
Y	Y	Y	N	9
Y	Y	N	Y	9
Y	N	Y	Y	10
Y	Y	Y	Y	11

[0061] Regions that include a large fraction of the image border are also likely to be background regions. This feature indicates what fraction of the image border is in contact with the given region.

$$borderness\_b = \frac{perimeter\_pixels\_in\_this\_region}{2*(image\_height+image\_width-2)} \quad (12)$$

[0062] When a large fraction of the region perimeter is on the image border, a region is also likely to be background. Such a ratio is unlikely to exceed 0.5, so a value in the range [0,1] is obtained by scaling the ratio by a factor of 2 and saturating the ratio at the value of 1.0.

$$borderness\_c = \frac{min(1,2*num\_region\_perimeter\_pixels\_on\_border)}{region\_perimeter} \quad (13)$$

[0063] Again, note that instead of a composite borderness measure based on heuristics, all the above three borderness measures are separately trained and used in the main subject detection.

Semantic saliency features

• flesh/face/people (foreground, self saliency features)

[0064] A majority of photographic images have people and about the same number of images have sizable faces in them. In conjunction with certain shape analysis and pattern analysis, some detected flesh regions can be identified as faces. Subsequently, using models of human figures, flesh detection and face detection can lead to clothing detection and eventually people detection.

[0065] The current flesh detection algorithm utilizes color image segmentation and a pre-determined flesh distribution in a chrominance space (Lee, "Color image quantization based on physics and psychophysics," Journal of Society of Photographic Science and Technology of Japan, Vol. 59, No. 1, pp. 212-225, 1996). The flesh region classification is based on Maximum Likelihood Estimation (MLE) according to the average color of a segmented region. The conditional probabilities are mapped to a belief value via a sigmoid belief function.

[0066] A primitive face detection algorithm is used in the present invention. It combines the flesh map output by the flesh detection algorithm with other face heuristics to output a belief in the location of faces in an image. Each region in an image that is identified as a flesh region is fitted with an ellipse. The major and minor axes of the ellipse are calculated as also the number of pixels in the region outside the ellipse and the number of pixels in the ellipse not part of the region. The aspect ratio is computed as a ratio of the major axis to the minor axis. The belief for the face is a function of the aspect ratio of the fitted ellipse, the area of the region outside the ellipse, and the area of the ellipse not part of the region. A Gaussian belief sensor function is used to scale the raw function outputs to beliefs.

[0067] It will be appreciated that a person of ordinary skill in the art can use a different face detection method without departing from the present invention.

• key background subject matters (self saliency features)

[0068] There are a number of objects that frequently appear in photographic images, such as sky, cloud, grass, tree, foliage, vegetation, water body (river, lake, pond), wood, metal, and the like. Most of them have high likelihood to be background objects. Therefore, such objects can be ruled out while they also serve as precursors for main subjects as well as scene types.

[0069] Among these background subject matters, sky and grass (may include other green vegetation) are detected with relatively high confidence due to the amount of constancy in terms of their color, texture, spatial extent, and spatial location.

Probabilistic Reasoning

[0070] All the saliency features are integrated by a Bayes net to yield the likelihood of main subjects. On one hand, different evidences may compete with or contradict each other. On the other hand, different evidences may mutually reinforce each other according to prior models or knowledge of typical photographic scenes. Both competition and reinforcement are resolved by the Bayes net-based inference engine.

[0071] A Bayes net (J. Pearl, *Probabilistic Reasoning in Intelligent Systems*, San Francisco, CA: Morgan

Kaufmann, 1988) is a directed acyclic graph that represents causality relationships between various entities in the graph. The direction of links represents causality. It is an evaluation means knowing joint Probability Distribution Function (PDF) among various entities. Its advantages include explicit uncertainty characterization, fast and efficient computation, quick training, high adaptivity and ease of building, and representing contextual knowledge in human reasoning framework. A Bayes net consists of four components:

1. Priors: The initial beliefs about various nodes in the Bayes net
2. Conditional Probability Matrices (CPMs): the statistical relationship between two connected nodes in the Bayes net
3. Evidences: Observations from feature detectors that are input to the Bayes net
4. Posteriors: The final computed beliefs after the evidences have been propagated through the Bayes net.

[0072] Referring to Fig. 7, a two-level Bayesian net is used in the present invention that assumes conditional independence between various feature detectors. The main subject is determined at the root node 20 and all the feature detectors are at the leaf nodes 22. There is one Bayes net active for each region (identified by the segmentation algorithm) in the image. The root node gives the posterior belief in that region being part of the main subject. It is to be understood that the present invention can be used with a Bayes net that has more than two levels without departing from the scope of the present invention.

Training Bayes nets

[0073] One advantage of Bayes nets is each link is assumed to be independent of links at the same level. Therefore, it is convenient for training the entire net by training each link separately, i.e., deriving the CPM for a given link independent of others. In general, two methods are used for obtaining CPM for each root-feature node pair:

1. Using Expert Knowledge

[0074] This is an ad-hoc method. An expert is consulted to obtain the conditional probabilities of each feature detector observing the main subject given the main subject.

2. Using Contingency Tables

[0075] This is a sampling and correlation method. Multiple observations of each feature detector are recorded along with information about the main subject. These observations are then compiled together to create contingency tables which, when normalized, can then be used as the CPM. This method is similar to neural network type of training (learning). This method is preferred in the present invention.

[0076] Consider the CPM for centrality as an example. This matrix was generated using contingency tables derived from the ground truth and the feature detector. Since the feature detector in general does not supply a binary decision (referring to Table 3), fractional frequency count is used in deriving the CPM. The entries in the CPM are determined by

$$CPM = \left[\left(\sum_{i \in I} \sum_{r \in R_i} n_i F_r^T T_r\right) P\right]^T \qquad (14)$$

$$F_r = [f_0^r f_1^r \cdots f_M^r],\ T_r = [t_0^r t_1^r \cdots t_L^r],$$

$$P = diag\{p_j\},\ p_j = \left(\sum_{i \in I} \sum_{r \in R_i} n_i t_r\right),$$

where I is the set of all training images, R_i is the set of all regions in image i, n_i is the number of observations (observers) for image i. Moreover, F_r represents an M-label feature vector for region r, T_r represents an L-level ground-truth vector, and P denotes an L x L diagonal matrix of normalization constant factors. For example, in Table 3, regions 1, 4, 5 and 7 contribute to boxes 00, 11, 10 and 01 in Table 4, respectively. Note that all the belief values have been normalized by the proper belief sensors. As an intuitive interpretation of the first column of the CPM for centrality, a "central" region is about twice as likely to be the main subject than not a main subject.

Table 3

An example of training the CPM.			
Region Number	**Ground Truth**	**Feature Detector**	**Output Contribution**
1	0	0.017	00
2	0	0.211	00
3	0	0.011	00
4	0.933	0.953	11
5	0	0.673	10
6	1	0.891	11
7	0.93	0.072	01
8	1	0.091	01

Table 4

The trained CPM.		
	Feature = 1	feature = 0
Main subject = 1	0.35 (11)	0.65 (01)
Main subject = 0	0.17 (10)	0.83 (00)

[0077] The output of the algorithm is in the form of a list of segmented regions ranked in a descending order of their likelihood as potential main subjects for a generic or specific application. Furthermore, this list can be converted into a map in which the brightness of a region is proportional to the main subject belief of the region. This "belief" map is more than a binary map that only indicates location of the determined main subject. The associated likelihood is also attached to each region so that the regions with large brightness values correspond to regions with high confidence or belief being part of the main subject. This reflects the inherent uncertainty for humans to perform such a task. However, a binary decision, when desired, can be readily obtained by applying an appropriate threshold to the belief map. Moreover, the belief information may be very useful for downstream applications. For example, different weighting factors can be assigned to different regions in determining bit allocation for image coding.

[0078] Other aspects of the invention include:

1. The method wherein the step of extracting for each of the regions at least one structural saliency feature and at least one semantic saliency feature includes using an extended neighborhood window to compute a plurality of the relative saliency features, wherein the extended neighborhood window is determined by the steps of:

   (c1) finding a minimum bounding rectangle of a region;
   (c2) stretching the minimum bounding rectangle in all four directions proportionally; and
   (c3) defining all regions intersecting the stretched minimum bounding rectangle as neighbors of the region.

2. The method as in claim 4, wherein the step of extracting for each of the regions at least one structural saliency feature and at least one semantic saliency feature includes using a centralty as the location feature, wherein the centrality feature is computed by the steps of:

   (c1) determining a probability density function of main subject locations using a collection of training data;
   (c2) computing an integral of the probability density function over an area of a region; and,
   (c3) obtaining a value of the centrality feature by normalizing the integral by the area of the region.

3. The method wherein the step of extracting for each of the regions at least one structural saliency feature and at least one semantic saliency feature includes using a hyperconvexity as the convexity feature, wherein the hyper-convexity feature is computed as a ratio of a perimeter-based convexity measure and an area-based convexity measure.

4. The method wherein the step of extracting for each of the regions at least one structural saliency feature and at least one semantic saliency feature includes computing a maximum fraction of a region perimeter shared with a neighboring region as the surroundedness feature.

5. The method wherein the step of extracting for each of the regions at least one structural saliency feature and at least one semantic saliency feature includes using an orientation-unaware borderness feature as the borderness feature, wherein the orientation-unaware borderness feature is categorized by the number and configuration of image borders a region is in contact with, and all image borders are treated equally.

6. The method wherein the step of extracting for each of the regions at least one structural saliency feature and at least one semantic saliency feature includes using an orientation-aware borderness feature as the borderness feature, wherein the orientation-aware borderness feature is categorized by the number and configuration of image borders a region is in contact with, and each image border is treated differently.

7. The method wherein the step of extracting for each of the regions at least one structural saliency feature and at least one semantic saliency feature includes using the borderness feature that is determined by what fraction of an image border is in contact with a region.

8. The method wherein the step of extracting for each of the regions at least one structural saliency feature and at least one semantic saliency feature includes using the borderness feature that is determined by what fraction of a region border is in contact with an image border.

9. The method wherein the step of integrating the structural saliency feature and the semantic feature using a probabilistic reasoning engine into an estimate of a belief that each regions is the main subject includes using a belief sensor function to convert a measurement of a feature into evidence, which is an input to a Bayes net.

10. The method wherein the step of integrating the structural saliency feature and the semantic feature using a probabilistic reasoning engine into an estimate of a belief that each region is the main subject includes outputting a belief map, which indicates a location of and a belief in the main subject.

11. The method wherein the step of extracting for each of the regions at least one structural saliency feature and at least one semantic saliency feature includes using a color distance metric defined in a color space, a spatial homogeneity constraint, and a mechanism for permitting spatial adaptivity.

12. The method wherein the step of extracting for each of the regions at least one structural saliency feature and at least one semantic saliency feature includes using either individually or in combination at least one low-level vision feature and at least one geometric feature as the structural saliency feature.

13. The method wherein the step of extracting for each of the regions at least one structural saliency feature and at least one semantic saliency feature includes using either individually or in combination a color, brightness and/or texture as a low-level vision feature; a location, size, shape, convexity, aspect ratio, symmetry, borderness, surroundedness and/or occlusion as a geometric feature; and a flesh, face, sky, grass and/or other green vegetation as the semantic saliency feature.

14. The method wherein the step of integrating the structural saliency feature and the semantic feature using a probabilistic reasoning engine into an estimate of a belief that each regions is the main subject includes using a collection of human opinions to train the reasoning engine to recognize the relative importance of the saliency features.

15. The method wherein the step of extracting for each of the regions at least one structural saliency feature and at least one semantic saliency feature includes using either individually or in combination a self-saliency feature and a relative saliency feature as the structural saliency feature.

16. The method wherein the step of extracting for each of the regions at least one structural saliency feature and at least one semantic saliency feature includes using an extended neighborhood window to compute a plurality of the relative saliency features, wherein the extended neighborhood window is determined by the steps of:

(c1) finding a minimum bounding rectangle of a region;
(c2) stretching the minimum bounding rectangle in all four directions proportionally; and,
(c3) defining all regions intersecting the stretched minimum bounding rectangle as neighbors of the region.

17. The method wherein the step of extracting for each of the regions ate least one structural saliency feature and at least one semantic saliency feature includes using a centralty as the location feature, wherein the centrality feature is computed by the steps of:

(c1) determining a probability density function of main subject locations using a collection of training data;
(c2) computing an integral of the probability density function over an area of a region; and,
(c3) obtaining a value of the centrality feature by normalizing the integral by the area of the region.

18. The method wherein the step of extracting for each of the regions at least one structural saliency feature and at least one semantic saliency feature includes using a hyperconvexity as the convexity feature, wherein the hyperconvexity feature is computed as a ratio of a perimeter-based convexity measure and an area-based convexity measure.

19. The method wherein the step of extracting for each of the regions at least one structural saliency feature and at least one semantic saliency feature includes computing a maximum fraction of a region perimeter shared with a neighboring region as the surroundedness feature.

20. The method wherein the step of extracting for each of the regions at least one structural saliency feature and at least one semantic saliency feature includes using an orientation-unaware borderness feature as the borderness feature, wherein the orientation-unaware borderness feature is categorized by the number and configuration of image borders a region is in contact with, and all image borders are treated equally.

21. The method wherein the step of extracting for each of the regions at least one structural saliency feature and at least one semantic saliency feature includes using an orientation-aware borderness feature as the borderness feature, wherein the orientation-aware borderness feature is categorized by the number and configuration of image borders a region is in contact with, and each image border is treated differently.

22. The method wherein the step of extracting for each of the regions at least one structural saliency feature and at least one semantic saliency feature includes using the borderness feature that is determined by what fraction of an image border is in contact with a region.

23. The method wherein the step of extracting for each of the regions at least one structural saliency feature and at least one semantic saliency feature includes using the borderness feature that is determined by what fraction of a region border is in contact with an image border.

24. The method wherein the step of integrating the structural saliency feature and the semantic feature using a probabilistic reasoning engine into an estimate of a belief that each region is the main subject includes using a Bayes net as the reasoning engine.

25. The method wherein the step of integrating the structural saliency feature and the semantic feature using a probabilistic reasoning engine into an estimate of a belief that each region is the main subject includes using a conditional probability matrix that is determined by using fractional frequency counting according to a collection of training data.

26. The method wherein the step of integrating the structural saliency feature and the semantic feature using a probabilistic reasoning engine into an estimate of a belief that each region is the main subject includes using a belief sensor function to convert a measurement of a feature into evidence, which is an input to a Bayes net.

27. The method wherein the step of extracting for each of the regions at least one structural saliency feature and at least one semantic saliency feature includes outputting a belief map, which indicates a location of and a belief in

the main subject.

Claims

1. A method for detecting a main subject in an image, the method comprising the steps of:

   a) receiving a digital image;
   b) extracting regions of arbitrary shape and size defined by actual objects from the digital image;
   c) extracting for each of the regions at least one structural saliency feature and at least one semantic saliency feature ; and,
   d) integrating the structural saliency feature and the semantic feature using a probabilistic reasoning engine into an estimate of a belief that each region is the main subject.

2. The method as in claim 1, wherein step (b) includes using a color distance metric defined in a color space, a spatial homogeneity constraint, and a mechanism for permitting spatial adaptivity.

3. The method as in claim 1, wherein step (c) includes using either individually or in combination at least one low-level vision feature and at least one geometric feature as the structural saliency feature.

4. The method as in claim 1 ,wherein step (c) includes using either individually or in combination a color, brightness and/or texture as a low-level vision feature; a location, size, shape, convexity, aspect ratio, symmetry, borderness, surroundedness and/or occlusion as a geometric feature; and a flesh, face, sky, grass and/or other green vegetation as the semantic saliency feature.

5. The method as in claim 1, wherein step (d) includes using a collection of human opinions to train the reasoning engine to recognize the relative importance of the saliency features.

6. The method as in claim 1, wherein step (c) includes using either individually or in combination a self-saliency feature and a relative saliency feature as the structural saliency feature.

7. The method as in claim 1, wherein step (d) includes using a Bayes net as the reasoning engine.

8. The method as in claim 1, wherein step (d) includes using a conditional probability matrix that is determined by using fractional frequency counting according to a collection of training data.

9. A method for detecting a main subject in an image, the method comprising the steps of:

   a) receiving a digital image;
   b) extracting regions of arbitrary shape and size defined by actual objects from the digital image;
   c) grouping the regions into larger segments corresponding to physically coherent objects;
   d) extracting for each of the regions at least one structural saliency feature and at least one semantic saliency feature; and,
   e) integrating the structural saliency feature and the semantic feature using a probabilistic reasoning engine into an estimate of a belief that each region is the main subject.

10. The method as in claim 9, wherein step (c) includes using either individually or in combination non-purposive grouping and purposive grouping.

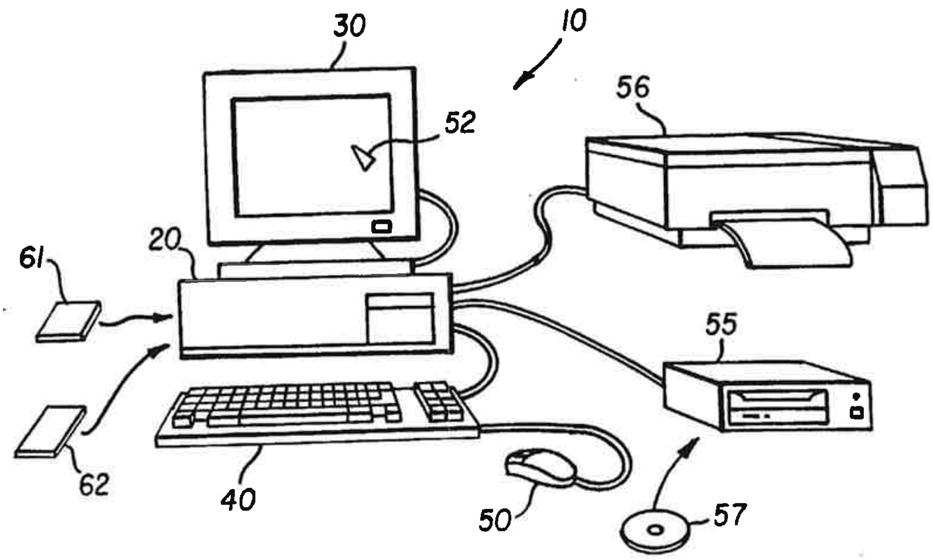


FIG. 1

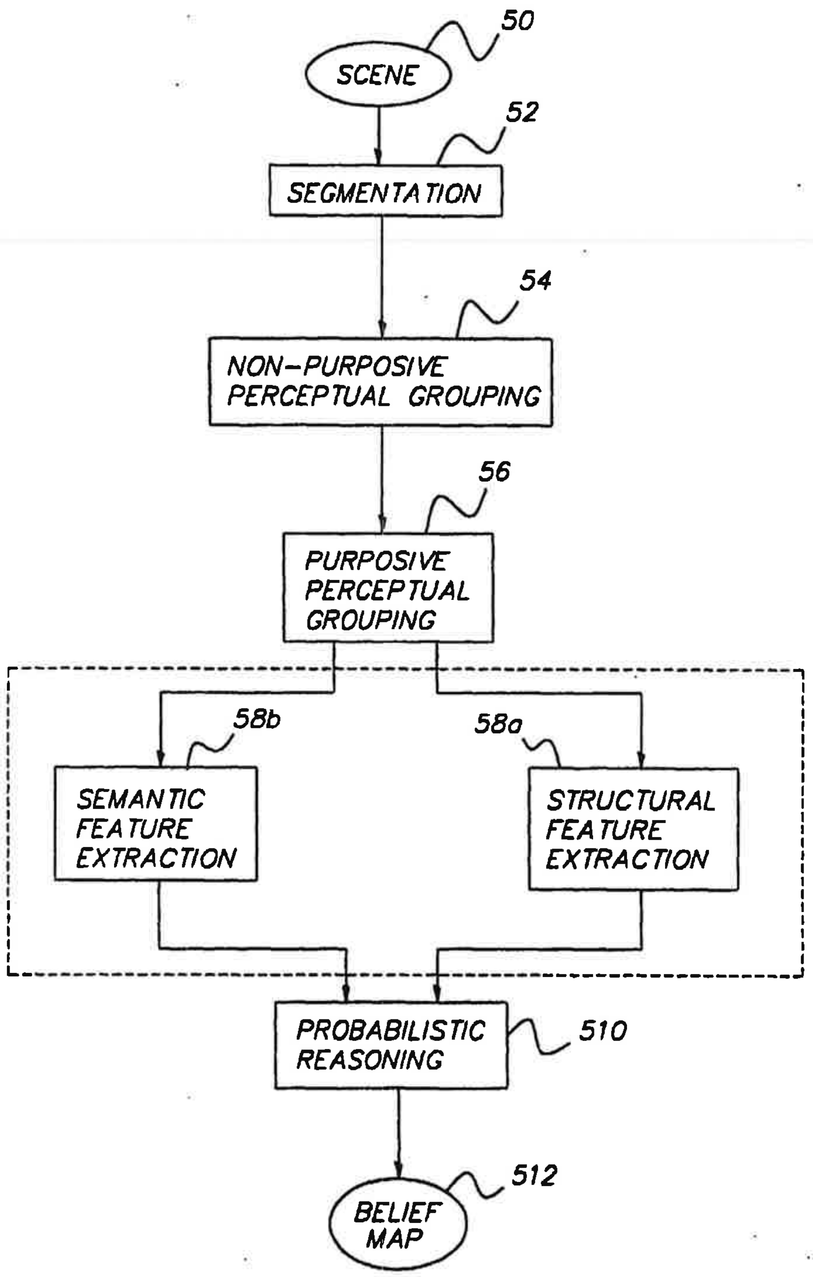

50
SCENE
52
SEGMENTATION
54
NON-PURPOSIVE PERCEPTUAL GROUPING
56
PURPOSIVE PERCEPTUAL GROUPING
58b
SEMANTIC FEATURE EXTRACTION
58a
STRUCTURAL FEATURE EXTRACTION
PROBABILISTIC REASONING
510
512
BELIEF MAP

FIG. 2

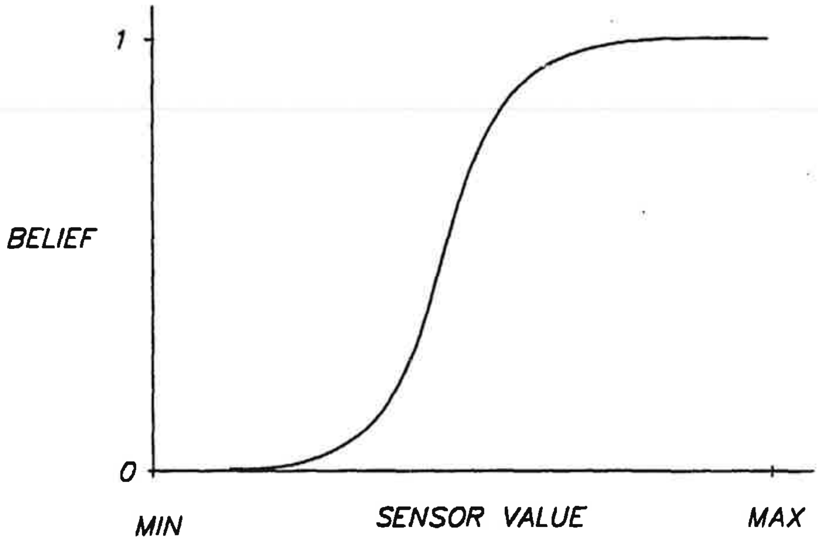

1
BELIEF
0
MIN
SENSOR VALUE
MAX

FIG. 3

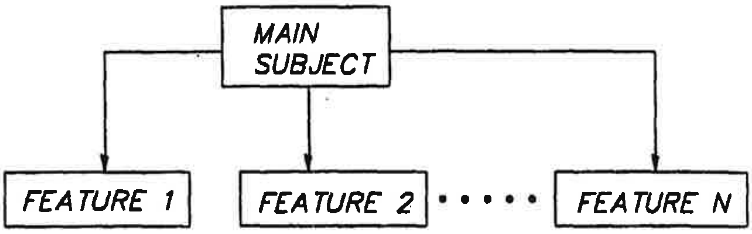

MAIN SUBJECT
FEATURE 1
FEATURE 2
FEATURE N

FIG. 7

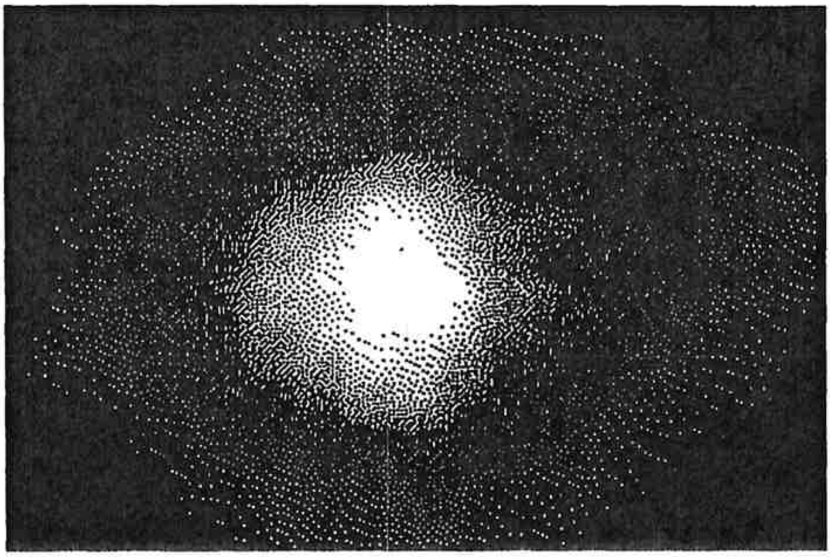


Fig. 4a

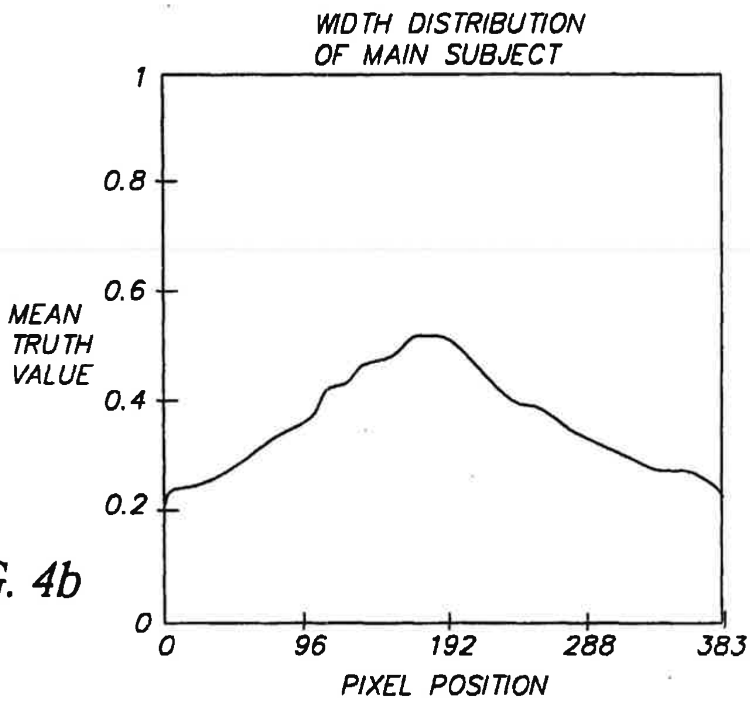

WIDTH DISTRIBUTION
OF MAIN SUBJECT
MEAN
TRUTH
VALUE
1
0.8
0.6
0.4
0.2
0
0
96
192
288
383
PIXEL POSITION

FIG. 4b

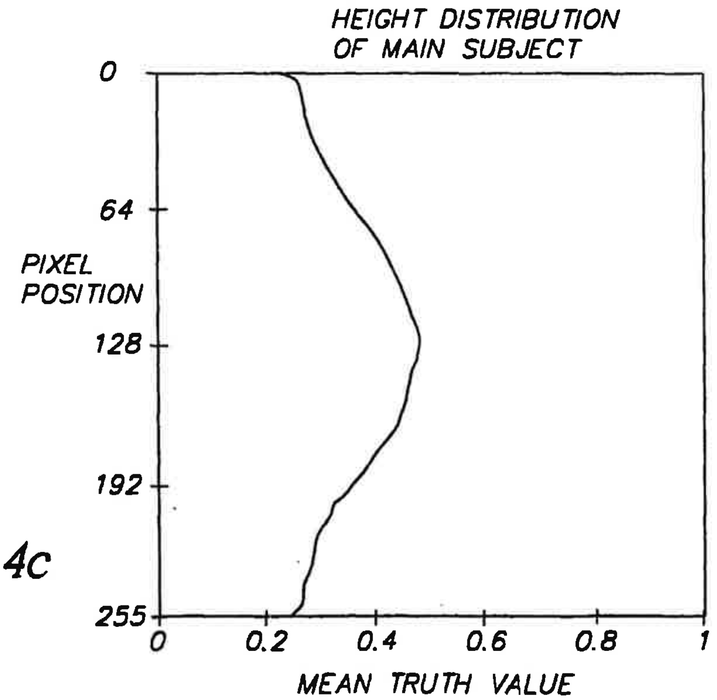

HEIGHT DISTRIBUTION
OF MAIN SUBJECT
PIXEL
POSITION
0
64
128
192
255
0
0.2
0.4
0.6
0.8
1
MEAN TRUTH VALUE

FIG. 4c

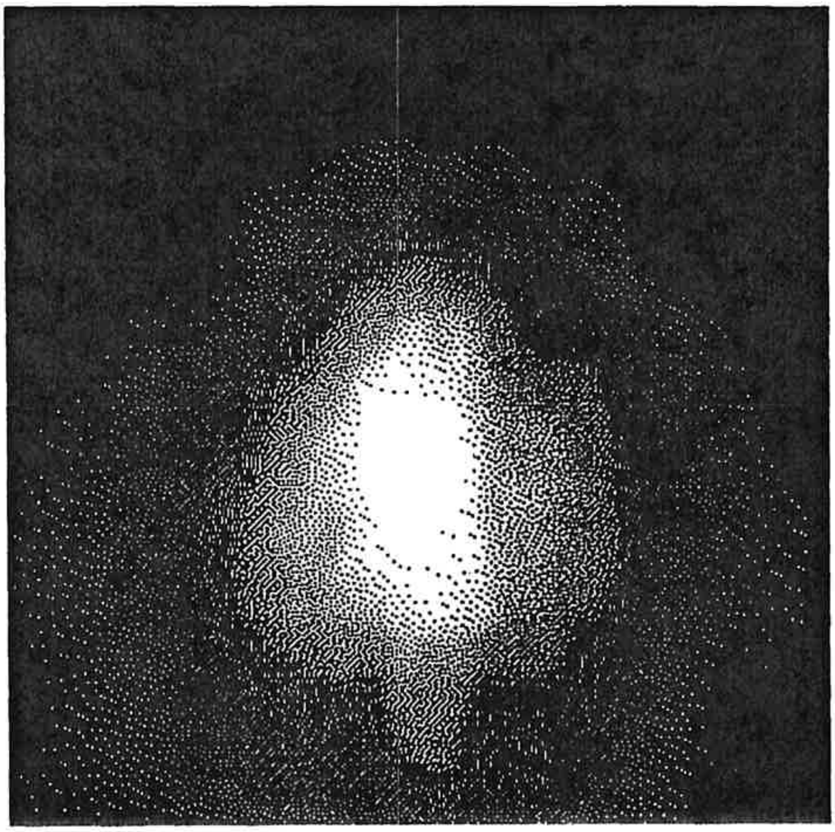


Fig. 5a

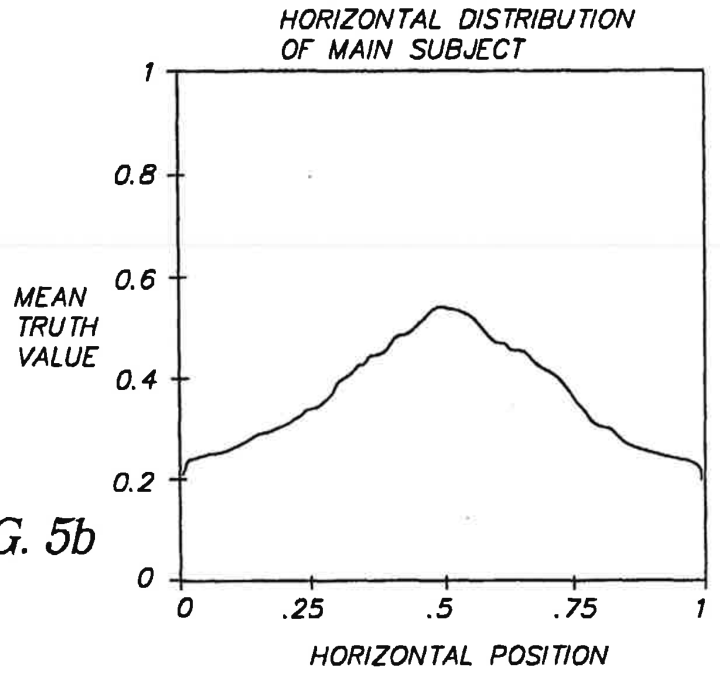

HORIZONTAL DISTRIBUTION
OF MAIN SUBJECT
MEAN
TRUTH
VALUE
1
0.8
0.6
0.4
0.2
0
0
.25
.5
.75
1
HORIZONTAL POSITION

FIG. 5b

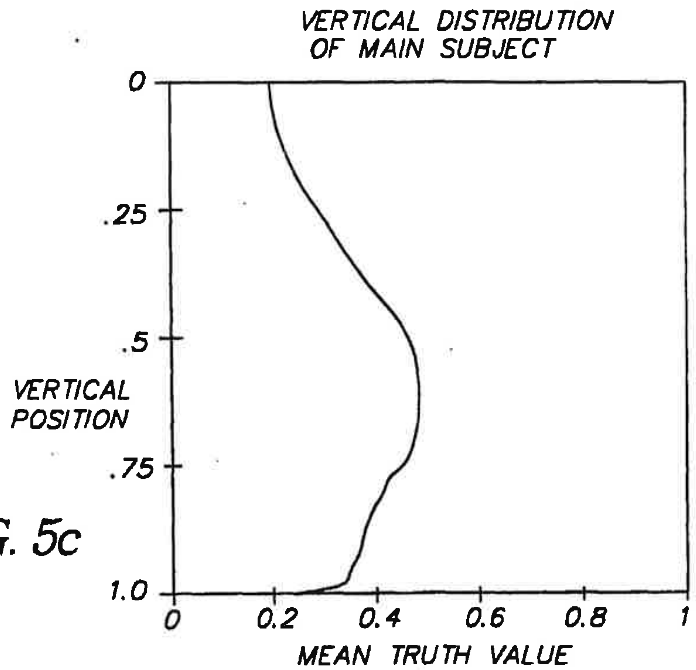

VERTICAL DISTRIBUTION
OF MAIN SUBJECT
VERTICAL
POSITION
0
.25
.5
.75
1.0
0
0.2
0.4
0.6
0.8
1
MEAN TRUTH VALUE

FIG. 5c

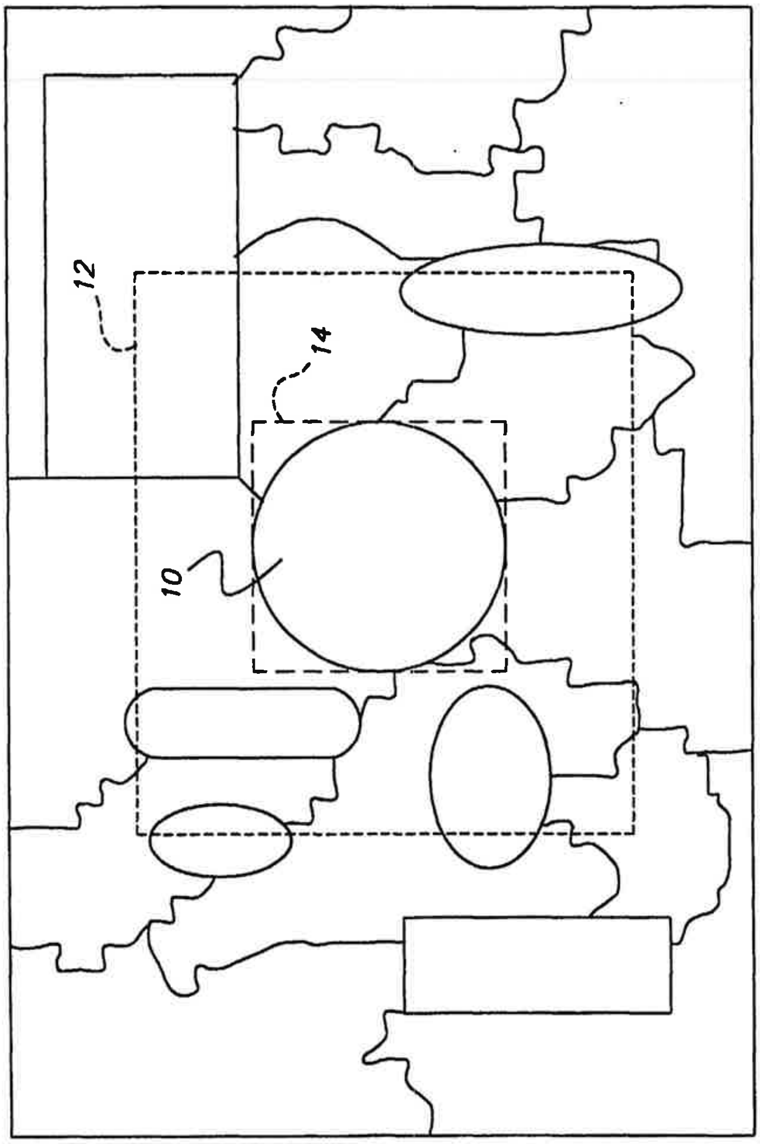


FIG. 6

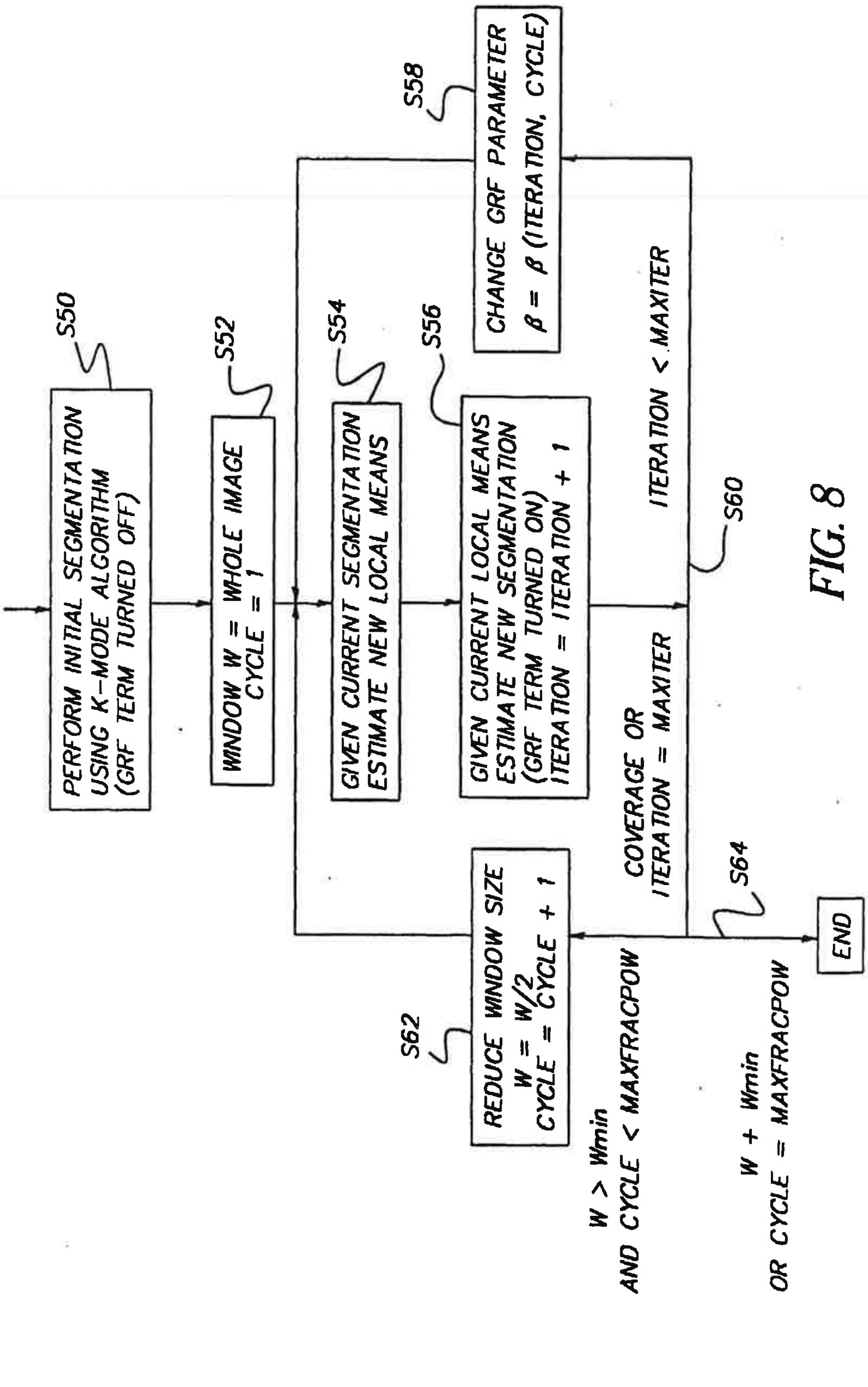

PERFORM INITIAL SEGMENTATION
USING K-MODE ALGORITHM
(GRF TERM TURNED OFF)
S50
WINDOW W = WHOLE IMAGE
CYCLE = 1
S52
GIVEN CURRENT SEGMENTATION
ESTIMATE NEW LOCAL MEANS
S54
GIVEN CURRENT LOCAL MEANS
ESTIMATE NEW SEGMENTATION
(GRF TERM TURNED ON)
ITERATION = ITERATION + 1
S56
CHANGE GRF PARAMETER
β = β (ITERATION, CYCLE)
S58
ITERATION < MAXITER
S60
COVERAGE OR
ITERATION = MAXITER
S64
END
REDUCE WINDOW SIZE
W = W/2
CYCLE = CYCLE + 1
S62
W > Wmin
AND CYCLE < MAXFRACPOW
W + Wmin
OR CYCLE = MAXFRACPOW

FIG. 8

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 1 158 464 A1

(12) **EUROPEAN PATENT APPLICATION**

(43) Date of publication:
28.11.2001 Bulletin 2001/48

(51) Int Cl.7: **G06T 7/00**, G03B 27/62, G03B 27/53

(21) Application number: **01201712.5**

(22) Date of filing: **10.05.2001**

(84) Designated Contracting States:
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE TR
Designated Extension States:
AL LT LV MK RO SI

(30) Priority: **22.05.2000 US 575784**

(71) Applicant: **EASTMAN KODAK COMPANY**
Rochester, New York 14650 (US)

(72) Inventors:
- **Luo, Jiebo, c/o Eastman Kodak Company Rochester, New York 14650-2201 (US)**
- **Gray, Robert T., c/o Eastman Kodak Company Rochester, New York 14650-2201 (US)**
- **Gindele, Edward B., c/o Eastman Kodak Company Rochester, New York 14650-2201 (US)**

(74) Representative:
Lewandowsky, Klaus, Dipl.-Ing. et al
Kodak Aktiengesellschaft,
Patentabteilung
70323 Stuttgart (DE)

(54) Method of automatically crop and zoom a digitized photographic image

(57) A method of producing an image of at least a portion of a photographic image onto a photographic receiver including receiving a digital image corresponding to the photographic image, the digital image comprising pixels, and locating the relative optical position of a photographic image, the lens, and the photographic receiver in response to pixels of the digital image and illuminating a portion of the photographic image of high subject content to produce an image of such portion onto the photographic receiver.

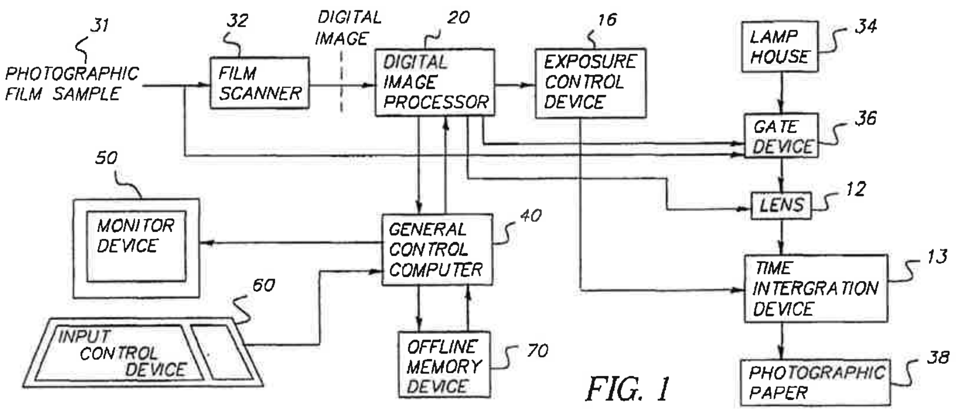


FIG. 1

EP 1 158 464 A1

Printed by Jouve, 75001 PARIS (FR)

Description

[0001] The invention relates to producing an image of a portion of a photographic image onto a receiver by using digital image processing.

[0002] For many decades, traditional commercial photofinishing systems have placed limits on the features offered to consumers to promote mass production. Among those features that are unavailable conventionally, zooming and cropping have been identified by both consumers and photofinishers as extremely useful additional features that could potentially improve the quality of the finished photographs and the subsequent picture sharing experiences. With the advent of, and rapid advances in digital imaging, many of the technical barriers that existed in traditional photography no longer stand insurmountable.

[0003] Hybrid and digital photography provide the ability to crop undesirable content from a picture, and magnify or zoom the desired content to fill the entire photographic print. In spite of the fact that some traditional cameras with zoom capability provide consumers greater control over composing the desired scene content, studies have found that photographers may still wish to perform a certain amount of cropping and zooming when viewing the finished photograph at a later time. Imprecise viewfinders of many point-and-shoot cameras, as well as simply second-guessing their initial compositions, are factors in the desirability of zoom and crop. In addition, it maybe desirable to use some other regular border templates such as ovals, heart shapes, squares, and so forth. In another scenario, some people commonly referred to as "scrapbookers" tend to perform more aggressive crop in making a scrapbook, for example, cutting along the boundary of objects.

[0004] There are significant differences in objectives and behaviors between these two types of cropping, namely album-making and scrapbook making, with the latter more difficult to understand and summarize. The invention described below performs automatic zooming and cropping for making photographic prints. One customer focus group study indicated that it would be beneficial to provide customers a double set of prints-- one regular and one zoom. Moreover, it is preferred that the cropping and zooming be done automatically. Most customers do not want to think about how the zooming and cropping is being done as long as the content and quality (for example, sharpness) of the cropped and zoomed pictures is acceptable.

[0005] There has been little research on automatic zoom and crop due to the apparent difficulty involved in performing such a task. None of the known conventional image manipulation software uses scene content in determining the automatic crop amount. For example, a program entitled "XV", a freeware package developed by John Bradley at University of Pennsylvania, USA (Department of Computer and Information Science), provides an "autocrop" function for manipulating images and operates in the following way:

1. The program examines a border line of an image, in all of the four directions, namely from the top, bottom, left and right sides;
2. The program checks the variation within the line. In grayscale images, a line has to be uniform to be cropped. In color images, both the spatial correlation and spectral correlation have to be low, except for a small percentage of pixels, for the line to be qualified for cropping. In other words, a line will not be cropped if it contains a significant amount of variation;
3. If a line along one dimension passes the criterion, the next line (row or column) inward is then examined; and
4. The final cropped image is determined when the above recursive process stops.

[0006] This program essentially tries to remove relatively homogeneous margins around the borders of an image. It does not examine the overall content of the image. In practice, the XV program is effective in cropping out the dark border generated due to imprecise alignment during the scanning process. However, disastrous results can often be produced due to the apparent lack of scene understanding. In some extreme cases, the entire image can be cropped.

[0007] Another conventional system, described by Bollman and others. in US-A-5,978,519 provides a method for cropping images based upon the different intensity levels within the image. With this system, an image to be cropped is scaled down to a grid and divided into non-overlapping blocks. The mean and variance of intensity levels are calculated for each block. Based on the distribution of variances in the blocks, a threshold is selected for the variance. All blocks with a variance higher than the threshold variance are selected as regions of interest. The regions of interest are then cropped to a bounding rectangle. However, such a system is only effective when uncropped images contain regions where intensity levels are uniform and other regions where intensity levels vary considerably. The effectiveness of such a system is expected to be comparable to that of the XV program. The difference is that the XV program examines the image in a line by line fashion to identify uniform areas, while Bollman examines the image in a block by block fashion to identify uniform areas. In summary, both techniques cannot deal with images with non-uniform background.

[0008] Some optical printing systems have the capability of changing the optical magnification of the relay lens used in the photographic copying process. In US-A-5,995,201, Sakaguchi describes a method of varying the effective mag-

nification of prints made from film originals utilizing a fixed optical lens instead of zoom lens. In US-A-5,872,619, Stephenson and others. describe a method of printing photographs from a processed photographic filmstrip having images of different widths measured longitudinally of the filmstrip and having heights measured transversely of the filmstrip. This method uses a photographic printer having a zoom lens and a printing mask to provide printed images having a selected print width and a selected print height. In US-A-4,809,064, Amos and others. describe an apparatus for printing a selected region of a photographic negative onto a photosensitive paper to form an enlarged and cropped photographic print. This apparatus includes means for projecting the photographic negative onto first and second zoom lenses, each of the zoom lenses having an adjustable magnification. In US-A-,872,643, Maeda and others. describe a film reproducing apparatus that can effectively perform zoom and crop. This apparatus includes an image pick-up device which picks up a film frame image recorded on a film to generate image data, an information reader which reads information about photographing conditions of the film frame image, and a reproducing area designator which designates a reproducing area of the film frame image.
However, the reproducing area of the film frame image is determined based on pre-recorded information about the position of the main object, as indicated by which zone of the photograph the automatic focusing (AF) operation in the camera was on - part of the recorded information about photographing conditions. In all the above mentioned optical printing systems, the position of the photographic film sample and magnification factor of the relay lens are preselected.

[0009] It is an object of the present invention to provide a method for producing a portion of an image from a photographic image by identifying the main subject of the photographic image.

[0010] This object is achieved by a method of producing an image of at least a portion of a photographic image onto a photographic receiver, comprising the steps of:

a) receiving a digital image corresponding to the photographic image, the digital image comprising pixels; and
b) locating the relative optical position of a photographic image, the lens, and the photographic receiver in response to pixels of the digital image and illuminating a portion of the photographic image of high subject content to produce an image of such portion onto the photographic receiver.

ADVANTAGES

[0011] One advantage of the invention lies in the ability to automatically crop and zoom photographic images based upon the scene contents. The digital image processing steps employed by the present invention includes a step of identifying the main subject within the digital image. The present invention uses the identified main subject of the digital image to automatically zoom and crop the image. Therefore, the present invention produces high-quality zoomed or cropped images automatically, regardless whether the background is uniform or not.

[0012] The foregoing and other objects, aspects and advantages will be better understood from the following detailed description of a preferred embodiment of the invention with reference to the drawings, in which:

FIG. 1 is a schematic diagram of a system embodiment of the invention;
FIG. 2 is a schematic architectural diagram of an embodiment of the invention;
FIG. 3 is a schematic architectural diagram of an embodiment of the invention;
FIG. 4 is a schematic architectural diagram of an embodiment of the invention;
FIG. 5 is a schematic architectural diagram of an embodiment of the invention;
FIG. 6 illustrates the application of the invention to a simulated photograph;
FIG. 7 illustrates the application of the invention to a simulated photograph;
FIG. 8 illustrates the application of the invention to a simulated photograph;
FIG. 9 illustrates the application of the invention to a simulated photograph;
FIG. 10 illustrates the application of the invention to a simulated photograph;
FIG. 11 illustrates the application of the invention to a simulated photograph;
FIG. 12 illustrates the application of the invention to a simulated photograph;
FIG. 13 is an exemplary uncropped photograph;
FIG. 14 is a belief map of the image shown in FIG. 13;
FIG. 15 is a cropped version of the image shown in FIG. 13;
FIG. 17 is a belief map of the image shown in FIG. 16; and
FIG. 18 is a cropped version of the image shown in FIG. 16.

[0013] The invention automatically zooms and crops digital images according to an analysis of the main subject in the scene. Previously, a system for detecting main subjects (for example, main subject detection or "MSD") in a consumer-type photographic image from the perspective of a third-party observer has been developed. Main subject detection provides a measure of saliency or relative importance for different regions that are associated with different subjects in an image. Main subject detection enables a discriminative treatment of the scene content for a number of applications related to consumer photographic images, including automatic crop and zoom.

[0014] Conventional wisdom in the field of computer vision, which reflects how a human observer would perform such tasks as main subject detection and cropping, calls for a problem-solving path via object recognition and scene content determination according to the semantic meaning of recognized objects. However, generic object recognition remains a largely unsolved problem despite decades of effort from academia and industry.

[0015] The MSD system is built upon mostly low-level vision features with semantic information integrated whenever available. This MSD system has a number of sub-tasks, including region segmentation, perceptual grouping, feature extraction, and probabilistic and semantic reasoning. In particular, a large number of features are extracted for each segmented region in the image to represent a wide variety of visual saliency properties, which are then input into a tunable, extensible probability network to generate a belief map containing a continuum of values.

[0016] Using MSD, regions that belong to the main subject are generally differentiated from the background clutter in the image. Thus, automatic zoom and crop becomes possible. Automatic zoom and crop is a nontrivial operation that was considered impossible for unconstrained images, which do not necessarily contain uniform background, without a certain amount of scene understanding. In the absence of content-driven cropping, conventional systems have concentrated on simply using a centered crop at a fixed zoom (magnification) factor, or removing the uniform background touching the image borders. The centered crop has been found unappealing to customers.

[0017] The output of MSD used by the invention is a list of segmented regions ranked in descending order of their likelihood (or belief) as potential main subjects for a generic or specific application. This list can be readily converted into a map in which the brightness of a region is proportional to the main subject belief of the region. Therefore, this map can be called a main subject "belief map. This "belief" map is more than a binary map that only indicates location of the determined main subject. The associated likelihood is also attached to each region so that regions with large values correspond to regions with high confidence or belief of being part of the main subject.

[0018] To some extent, this belief map reflects the inherent uncertainty for humans to perform such a task as MSD because different observers may disagree on certain subject matter while agreeing on other subject matter in terms of main subjects. However, a binary decision, when desired, can be readily obtained by using an appropriate threshold on the belief map. Moreover, the belief information may be very useful for downstream applications. For example, different weighting factors can be assigned to different regions (subject matters) in determining the amount of crop.

[0019] For determination of crop, the invention uses the main subject belief map instead of a binarized version of the map to avoid making a bad cropping decision that is irreversible. Furthermore, using the continuous values of the main subject beliefs helps trade-off different regions under the constraints encountered in cropping. A binary decision on what to include and what not to include, once made, leaves little room for trade-off. For example, if the main subject region is smaller than the crop window, the only reasonable choice, given a binary main subject map, is to leave equal amounts of margin around the main subject region. On the other hand, secondary main subjects are indicated by lower belief values in the main subject belief map, and can be included according to a descending order of belief values once the main subject of highest belief values are included. Moreover, if an undesirable binary decision on what to include/exclude is made, there is no recourse to correct the mistake. Consequently, the cropping result becomes sensitive to the threshold used to obtain the binary decision. With a continuous-valued main subject belief map, every region or object is associated with a likelihood of being included or a belief value in its being included.

[0020] To reduce the degrees of freedom in determining the amount of crop, in particular for making photographic prints, in one embodiment, the invention restricts the set of allowable zoom factors to {1.5x, 2x, 3x, 4x}. This is based on the findings in the customer focus studies. Those skilled in the art would recognize that the present invention could be used with any the zoom factor. In one example, the default zoom factor is set at 1.5X.

[0021] "Digital image understanding technology" as used herein means technology which is specifically designed to exclusively perform on a digital image representing at least a portion of an original scene captured by an image capture device other than a copier or scanner, the functions of:

(i) digitally processing such digital image (other than by color and/or luminance, that is, non-spatial analysis of color) to recognize and thereby assign useful meaning to:

a human understandable object other than (a) a human-created graphic object, including without limitation, a diagram, chart, alphanumeric character, mark, symbol, pattern and any combination of the foregoing, (b) the entire original scene, or (c) any part of such original scene completely designated by a human in such original scene;

an object attribute (including without limitation, characteristics, movement or direction, of such human understandable object) in such original scene (other than for using such object attribute to control the functions, including image signal processing functions, of the image capture device); or

a scene condition (other than for using such scene condition to control the functions, including image signal processing functions, of the image capture device) of such original scene; and

(ii) based on the results of step (i) and in accordance with such useful meaning, modifying, transforming, compressing, expanding, enhancing, editing, classifying, formatting, transmitting, or storing the digital image representing at least a portion of the original scene, or causing an operation other than image processing to be performed.

General Description of an Optical Printer System Application

[0022] Referring to FIG. 1, the following description relates to an optical printing system. A photographic film sample 31 is received by a film scanner 32 which produces a source digital image 10 relating to the spatial density distribution of the photographic film sample. This source digital image is received by a digital image processor 20. The digital image processor 20 may be connected to a general control computer 40 under operator control from an input control device 60. The monitor device 50 displays diagnostic information about the optical printing system. The general control computer 40 keeps track of the lens magnification setting.

[0023] Referring to FIG. 2, a zoom factor 11, which corresponds to the lens magnification setting may also received by the image processor 20 from the general control computer 40 under operator control. The image processor 20 receives the source digital image 10 and uses the zoom factor 11 and the source digital image 10 to calculate the proper position for the photographic film sample in the form of a film sample position 9. The photographic film sample is positioned in a gate device 36 which holds the film negative in place during the exposure. The gate device 36 receives the film sample position 9 to position the photographic film sample to adjust which portion of the imaging area of the photograph will be printed.

[0024] Referring to FIG. 1, a lamp house 34 provides the illumination source which is transmitted through the photographic film sample 31 and focused by a lens 12 onto photographic paper 38. The time integration device 13 opens and closes a shutter for a variable length of time allowing the focused light from the lamp house 34 to expose the photographic paper 38. The exposure control device 16 receives a brightness balance value from the digital image processor 20. The exposure control device 16 uses the brightness balance value to regulate the length of time the shutter of the time integration device stays open.

[0025] A block diagram of the inventive cropping process (for example, the digital image understanding technology) is shown in FIG. 3, which is discussed in relation to FIGS. 6-12. FIGS. 6-12 illustrate the inventive process being applied to an original image shown in FIG. 6.

[0026] In item 200, the belief map is created using MSD. The invention selects a zoom factor (for example 1.5X) and a crop window 80 (as shown in FIG. 7), as referred to in item 201 of FIG. 3. This zoom factor may be selected by an operator, or by an automatic method based directly on the main subject belief map (for example, an estimate of the size of the main subject). The crop window is typically a rectangular window with a certain aspect ratio. If the zoom factor is selected by the operator, the value of the zoom factor is sent from the general control computer 40 to the digital image processor 20 shown in FIG. 1. For both the operator selected case or the main subject belief map case, the zoom factor is used to communicate with the lens 12 to adjust the lens magnification setting. This adjustment allows the lens 12 to image the appropriate size of the photographic film sample 31 onto the photographic paper 38.

[0027] In item 202, regions of the belief map are clustered and the lowest belief cluster (for example, the background belief) is set to zero using a predefined threshold. As discussed in greater detail below, sections of the image having a belief value below a certain threshold are considered background sections. In item 202 such sections are given a belief of zero for purposes of this embodiment of the invention.

[0028] Then, in item 203 the centroid, or center-of-mass, of nonzero beliefs are computed and, as shown in item 204, the crop window position is centered at the centroid, as shown in FIG. 7. More specifically, in FIG. 6 the subject having the highest belief in the belief map is the woman and the stroller. FIG. 7 illustrates that the centroid of this subject is approximately the top of the baby's head.

[0029] The centroid $(\hat{x},\hat{y})$ of a belief map is calculated using the following procedure:

$$\hat{x} = \sum_i x_i bel(x_i, y_i), \ \hat{y} = \sum_i y_i bel(x_i, y_i),$$

where x_i and y_i denote that coordinates of a pixel in the belief map and $bel(x_i,y_i)$ represents the belief value at this pixel location.

[0030] The crop window is then moved so that the entire crop window is within the original image (for example item 205) as shown in FIG. 8. In item 207, the crop window 80 is moved again so that all the regions of the highest belief values ("main subject") are included within the crop window and to create a margin 81, as shown in FIG. 9. This process (for example, 207) captures the entire subject of interest. Therefore, as shown in FIG. 9, the top of the woman's head is included in the crop window. Compare this to FIG. 8 where the top of the woman's head was outside the crop window.

[0031] Decision box 208 determines whether an acceptable solution has been found, that is, whether it is possible to include at least the regions of the highest belief values in the crop window.

[0032] If an acceptable solution exists, the window is again moved, as shown in item 210, to optimize a subject content index for the crop window. The preferred embodiment of the present invention defines the subject content index as the sum of belief values within the crop window. It should be noted that the present invention specifies higher numerical belief values corresponding to higher main subject probability. Therefore, finding a numerical maximum of the sum of the belief values is equivalent to finding an optimum of the subject content index. This is shown in FIG. 10 where the secondary objects (for example flowers) are included within the crop window 80 to increase the sum of beliefs. The sum of beliefs for a crop window is computed as follows.

$$sum(w) = \sum_{(x,y)\in w} bel(x,y),$$

where *bel*(*x*,*y*) represents the belief value at a given pixel location (*x*,*y*) within the crop window *w*.

[0033] Provided that the primary subjects are included, moving the crop window so that more of the secondary subjects are included would increase the sum of belief values within the crop window. Recall that the primary subjects are indicated by the highest belief values and the secondary subjects are indicated by belief values lower than those of the primary subjects but higher than those of the background subjects. The goal is to find the crop window that has the highest sum of belief values while ensuring that the primary subjects are completely included in the crop window, that is,

$$\tilde{w} = \max_{w\in W} sum(w),$$

where *W* denotes the set of all possible crop windows that satisfy all the aforementioned constraints (for example, those that are completely within the uncropped image and those that encompass the entire primary subjects).

[0034] Then, in item 212, the position of the center of the crop window is used to calculate the translational component of the film sample position 9. The gate device 36, shown in FIG. 1, receives the film sample position 9 and uses this information to control the position of the photographic film sample 31 relative to the lens 12. Those skilled in the art will recognize that either or both of the lens 12 and the photographic film sample 31 may be moved to achieve the centering of the effective cropped image region on the photographic paper 38.

[0035] Referring to FIG. 3, if decision box 208 does not produce an acceptable solution, the digital image is rotated by 90 degrees, as shown in item 206, and the processing returns to item 205. At the second and greater pass through decision box 209, if the crop window 80 has already been rotated by 90 degrees and there is still not an acceptable solution (from decision box 208), then the sum of the beliefs for the centroid-based crop window 80 are computed for both orientations (for example, original and 90 degrees), as shown in item 211. Then, the position of the center of the crop window is used to calculate the translational component of the film sample position 9 and the rotational component of the film sample position 9 is set to 90 degrees. The gate device 36, shown in FIG. 1, receives the film sample position 9 and uses this information to control the position of the photographic film sample 31 relative to the lens 12. The gate device 36 then rotates the photographic film sample 31 by 90 degrees.

[0036] As would be known by one ordinarily skilled in the art given this disclosure, the photographic film sample may be rotated at an arbitrary angles (for example 45°,180°, and so forth.) in item 206. Further, the decision box 209 can allow multiple passes of the process described in items 205, 207 at different rotation angles in an attempt to find the orientation having the highest sum of beliefs.

[0037] The simulated image example shown in FIGS. 6-12 illustrates the progress the invention makes as it moves through the process shown in FIG. 3. One could formulate the problem as a global exhaustive search for the best solution. The procedure used in the invention is considered a "greedy" searching approach and is certainly more efficient than conventional processes.

[0038] FIGS. 4 and 5 illustrate variations of the process shown in FIG. 3 of determining the film sample position 9. FIG. 4 compares a landscape crop with a portrait crop, regardless of the initial (not necessarily the correct) picture orientation. The second variation in FIG. 5 creates a series of "ring-around" cropped versions with a zoom factor series (for example, {1x, 1.5x, 2x, 3x, and 4x}) for presentation to a human operator.

[0039] More specifically, in FIG. 4 items 300-303 are substantially identical to items 200-203 in FIG. 3. In item 304, the crop window 80 is positioned centered at the centroid without any rotation. Simultaneously, in a parallel processing step, the position of the crop window 80 is centered at the centroid with a rotation (for example, 90 degree rotation). In items 306 and 307 the optimal crop window 80 is determined, as discussed above with respect to FIG. 3. This produces cropped images 308 and 309. In item 310 the better (for example, higher sum of beliefs) of the two cropped images is selected and the corresponding translational and rotational components of the film sample position 9 are recorded.

[0040] Referring to FIG. 5, items 400-407 are substantially identical to items 300-307 discussed above. However, in the process shown in FIG. 5, decision box 410 observes the cropped images 408, 409 produced by the parallel processing in items 404-407 and determines whether an acceptable level has been attained. If the processing in item 410 determines that a different zoom factor or a different crop window 80 should be evaluated, processing returns to item 401. This process is repeated until an acceptable result is produced, ending in item 411. In similar fashion, the corresponding translational and rotational components of the film sample position 9 are recorded.

[0041] The invention utilizes a built-in "k-means" clustering process to determine proper thresholds of MSD beliefs for each application. The invention also uses clustering, as discussed below to enhance the cropping process. In one preferred embodiment, it is sufficient to use three levels to quantize MSD beliefs, namely "high", "medium", and "low." As would be known by one ordinarily skilled in the art, the invention is not limited to simply three levels of classification, but instead can utilize a reasonable number of classification levels to reduce the (unnecessary) variation in the belief map. These three levels allow for the main subject (high), the background (low), and an intermediate level (medium) to capture secondary subjects, or uncertainty, or salient regions of background. Therefore, the invention can perform a k-means clustering with k = 3 on the MSD belief map to "quantize" the beliefs. Consequently, the belief for each region is replaced by the mean belief of the cluster in that region. Note that a k-means clustering with k = 2 essentially produces a binary map with two clusters, "high" and "low," which is undesirable for cropping based on earlier discussion.

[0042] There are two major advantages in performing such clustering or quantization. First, clustering helps background separation by grouping low-belief background regions together to form a uniformly low-belief (for example, zero belief) background region. Second, clustering helps remove noise in belief ordering by grouping similar belief levels together. The centroiding operation does not need such quantization (nor should it be affected by the quantization). The main purpose of the quantization used here is to provide a threshold for the background.

[0043] The k-means clustering effectively performs a multi-level thresholding operation to the belief map. After clustering, two thresholds can be determined as follows:

$$threshold_{low} = (C_{low} + C_{medium})/2, \quad threshold_{high} = (C_{medium} + C_{high})/2$$

where $\{C_{low}, C_{medium}, C_{high}\}$ is the set of centroids (average belief values) for the three clusters, and $threshold_{low}$ and $threshold_{high}$ are the low and high thresholds, respectively.

[0044] Regions with belief values below the lower threshold are considered "background" and their belief values are set to zero in items 202, 302 and 402 discussed above. Regions with belief values above the higher threshold are considered part of the main subject and need to be included in their entirety, whenever possible. Regions with intermediate belief values (for example, less than or equal to the higher threshold and greater than or equal to the lower threshold) are considered part of the "secondary subject" and will be included as a whole or partially, if possible, to maximize the sum of main subject belief values retained by the crop window. Note that the variance statistics on the three clusters can be used to set the thresholds more accurately to reflect cluster dispersions.

[0045] The invention initializes the k-means process by finding the maximum value $bel_{maximum}$ and minimum values $bel_{minimum}$ of the belief map, computing the average value $bel_{average}$ of the maximum and minimum values for item in the belief map, and setting the initial centroids (denoted by a superscript of 0) at these three values, that is,

$$C^0_{low} = bel_{minimum}, \quad C^0_{medium} = bel_{medium}, \quad C^0_{high} = bel_{maximum}$$

Other ways of initialization may apply. For more about the k-means process, see *Sonka, Hlavac, and Boyle, Image Procesing Analysis, and Machine Vision, PWS Publishing, 1999 pagse 307-308.* For typical MSD belief maps, the k-means process usually converges in fewer than 10 iterations.

[0046] In applications where a zoom version of the cropped area is desired, there are two scenarios to consider. First, the zoom version effectively requires higher spatial resolution than the highest resolution of the original data. However, a visible loss of image sharpness is likely of concern in the situation. Second, the zoom version effectively requires lower spatial resolution than the highest resolution of the original data. In both cases, the invention uses an interpolation process to resample the data in order to retain a maximum amount of image detail. In general, edge or detail-preserving image interpolation processes such as cubic-spline interpolation are preferred because they tend to preserve the detail and sharpness of the original image better.

[0047] Example consumer photographs and their various cropped versions are shown in pictures "house" (for example, FIGS. 13-15) and "volleyball" (FIGS. 16-18). More specifically, FIGS. 13 and 16 illustrate uncropped original photographic images. FIGS. 14 and 17 illustrate belief maps, with lighter regions indicating higher belief values. As would be known by one ordinarily skilled in the art given this disclosure, the light intensity variations shown in FIGS. 14 and 17 are readily converted into numerical values for calculating the sum of the belief values discussed above. Finally, FIGS. 15 and 18 illustrate images cropped according to the invention.

[0048] For the "house" picture, both Bradley and Bollman (US-A-5,978,519) would keep the entire image and not be able to produce a cropped image because of the shadows at the bottom and the tree extending to the top border of the uncropped image (FIG. 13). There are no continuous flat background regions extending from the image borders in this picture, as required by US-A-5,978,519. Similarly, the top of the tree in FIG. 16 would not be cropped in the system disclosed in US-A-5,978,519.

[0049] Secondary subjects can lead to a more balanced cropped picture. For the "volleyball" picture (FIG. 16), the inclusion of some parts of the tree by the algorithm leads to more interesting cropped pictures than simply placing the main subjects (players) in the center of the cropped image (FIG. 18). The invention was able to do so because the trees are indicated to be of secondary importance based on the belief map FIG. 17. It is obvious that the art taught by Bradley and Bollman in US-A-5,978,519 would not be able to produce such a nicely cropped image. In fact, both Bradley and Bollman (US-A-5,978,519) would at best remove the entire lower lawn portion of the picture and keep the tree branches in the upper-left of the uncropped image.

[0050] A computer program product may include one or more storage medium, for example; magnetic storage media such as magnetic disk (such as a floppy disk) or magnetic tape; optical storage media such as optical disk, optical tape, or machine readable bar code; solid-state electronic storage devices such as random access memory (RAM), or read-only memory (ROM); or any other physical device or media employed to store a computer program having instructions for practicing a method according to the present invention.

[0051] Other features of the invention are included below.

[0052] The method further comprising repeating said positioning procedure with a second zoom factor and a second crop window and determining if said second zoom factor and said second crop window produces a higher subject content index.

[0053] The method wherein said zoom factor is selected according to a size of said main subject cluster.

[0054] The method further comprising positioning said crop window such that the subject content index of said crop window is at an optimum.

[0055] The method further comprising positioning said crop window such that said crop window includes all of said main subject cluster.

[0056] The method further comprising positioning said crop window to include a buffer around said main subject cluster.

[0057] A computer storage product having at least one computer storage medium having instructions stored therein.

PARTS LIST

[0058]

9	film sample position
10	source digital image
11	zoom factor
12	lens
13	time integration device
20	digital image processor
31	photographic film sample
32	film scanner
34	lamp house
36	gate device
38	photographic paper
40	general control computer
50	monitor device
60	input control device
80	crop window
81	margin
200	image and belief map
201	decision box for zoom factor selection

202 decision box for clustering regions
203 decision box for computing centroids
204 decision box for positioning the crop window
205 decision box moving a window
206 decision box for rotating the image
207 decision box for moving a window with the highest belief
208 decision box for determining if a solution exists
209 decision box for determining if the image is rotated
210 decision box for computing the sum of beliefs
211 decision box for computing the sum of beliefs
300 image and belief map
301 decision box for selecting a zoom factor and crop window
302 decision box for setting the background belief to 0
303 decision box for computing the centroid of nonzero beliefs
304 decision box for positioning the crop window
305 decision box for positioning the crop window
306 decision box for determining the optimal crop
307 decision box for determining the optimal crop
308 cropped image # 1
309 cropped image # 2
310 decision box for selecting the better crop
311 cropped image
400 image and belief map
401 decision box for selecting a zoom factor and crop window
402 decision box for setting the background belief to 0
403 decision box for computing the centroid of nonzero beliefs
404 decision box for positioning the crop window
405 decision box for positioning the crop window
406 decision box for determining the optimal crop
407 decision box for determining the optimal crop
408 cropped image # n
409 cropped image # n+1
410 decision box for selecting a different zoom factor or different crop window
411 decision box for ending the process

Claims

1. A method of producing an image of at least a portion of a photographic image onto a photographic receiver, comprising the steps of:

    a) receiving a digital image corresponding to the photographic image, the digital image comprising pixels; and
    b) locating the relative optical position of a photographic image, the lens, and the photographic receiver in response to pixels of the digital image and illuminating a portion of the photographic image of high subject content to produce an image of such portion onto the photographic receiver.

2. A method of producing an image of at least a portion of a photographic image onto a photographic receiver, comprising the steps of:

    a) receiving a digital image corresponding to the photographic image, the digital image comprising pixels;
    b) developing a belief map of the digital image, by using pixels of the digital image to determine a series of features, and using such features to assign the probability of the location of a main subject of the digital image in the belief map; and
    c) locating the relative optical position of a photographic image, the lens, and the photographic receiver in response to the belief map and illuminating a portion of the photographic image of high subject content to produce an image of such portion onto the photographic receiver.

3. A method of producing an image of a portion of at least a portion of a photographic image onto a photographic

receiver, comprising the steps of:

a) receiving a digital image corresponding to the photographic image, the digital image comprising pixels;
b) developing a belief map of the digital image, by using pixels of the digital image to determine a series of features, and using such features to assign the probability of the location of a main subject of the digital image in the belief map;
c) selecting a crop window of a size and shape;
d) selecting an initial position of the crop window;
e) using the belief values corresponding to the crop window to select the position of the crop window; and
f) locating the relative optical position of a photographic image, the lens, and the photographic receiver in response to the position of the crop window within the digital image and illuminating a portion of the photographic image to produce an image onto the photographic receiver.

4. A method as set forth in claim 3 wherein step c) further comprises the steps of:

i) receiving a magnification factor; and
ii) using the magnification factor to select the size of the crop window.

5. A method as set forth in claim 3 wherein step e) further comprising the steps of:

i) calculating a subject content index for the crop window derived from the belief values;
ii) following a positioning procedure of repeating step i) for at least two positions of the crop window; and
iii) use the subject content index values to select the crop window position.

6. A method as set forth in claim 5 wherein step d) further comprises the steps of:

i) using the belief map to determine the center of the mass of the belief map; and
ii) using the center of mass of the belief map to select the initial position of the crop window.

7. The method of claim 6 wherein the crop window is completely within the digital image.

8. The method of claim 3 wherein step b) further comprises the step of performing a clustering of the belief map to identify at least a cluster of highest belief values corresponding to main subject, a cluster of intermediate belief values corresponding to secondary subjects, and a cluster of lowest belief values corresponding to the background.

9. The method of claim 8 wherein said clustering includes setting said background portions to a zero belief value.

10. The method of claim 5 further comprising repeating said positioning procedure with a rotated image and determining if said rotated image produces a higher subject content index.

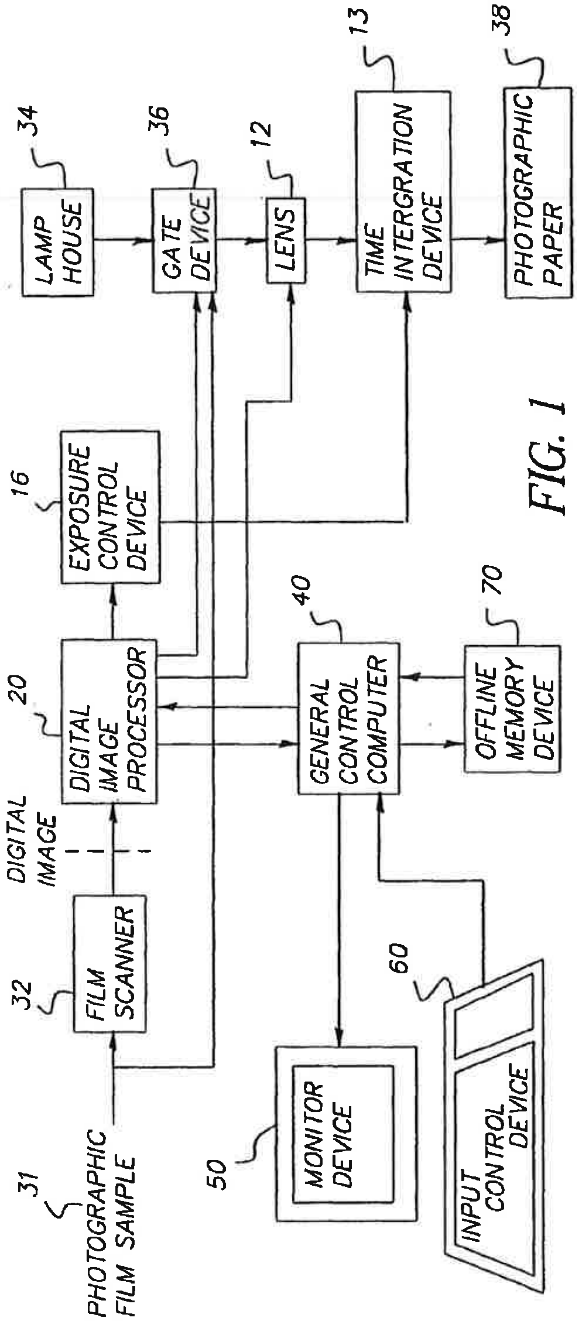

LAMP HOUSE
34
GATE DEVICE
36
LENS
12
TIME INTERGRATION DEVICE
13
PHOTOGRAPHIC PAPER
38
EXPOSURE CONTROL DEVICE
16
DIGITAL IMAGE PROCESSOR
20
GENERAL CONTROL COMPUTER
40
OFFLINE MEMORY DEVICE
70
DIGITAL IMAGE
FILM SCANNER
32
PHOTOGRAPHIC FILM SAMPLE
31
MONITOR DEVICE
50
INPUT CONTROL DEVICE
60

FIG. 1

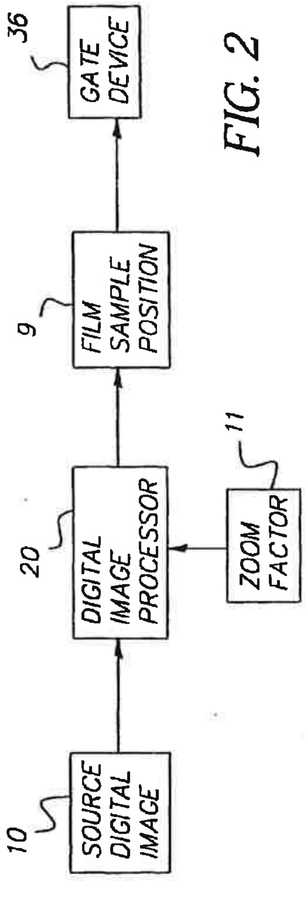

SOURCE DIGITAL IMAGE
10
DIGITAL IMAGE PROCESSOR
20
ZOOM FACTOR
11
FILM SAMPLE POSITION
9
GATE DEVICE
36

FIG. 2

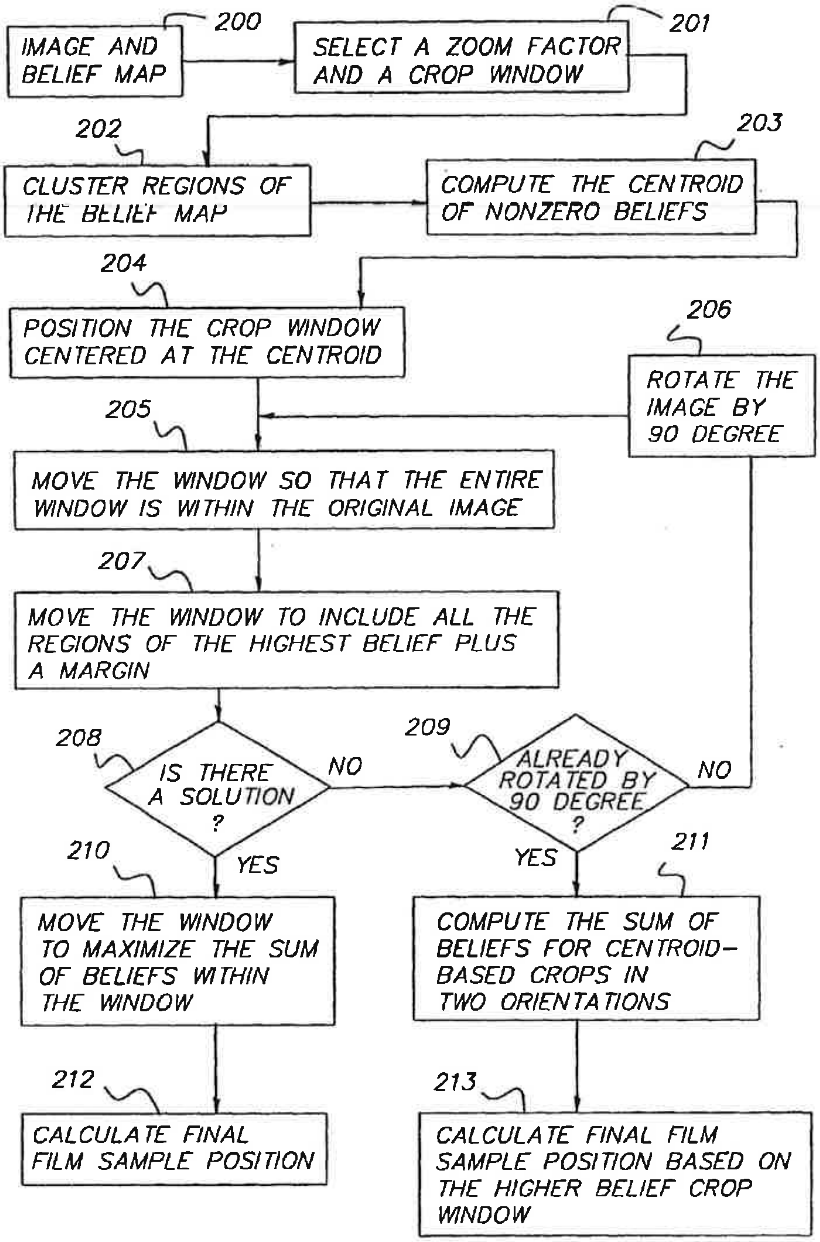

200
IMAGE AND BELIEF MAP
201
SELECT A ZOOM FACTOR AND A CROP WINDOW
202
CLUSTER REGIONS OF THE BELIEF MAP
203
COMPUTE THE CENTROID OF NONZERO BELIEFS
204
POSITION THE CROP WINDOW CENTERED AT THE CENTROID
206
ROTATE THE IMAGE BY 90 DEGREE
205
MOVE THE WINDOW SO THAT THE ENTIRE WINDOW IS WITHIN THE ORIGINAL IMAGE
207
MOVE THE WINDOW TO INCLUDE ALL THE REGIONS OF THE HIGHEST BELIEF PLUS A MARGIN
208
IS THERE A SOLUTION ?
NO
209
ALREADY ROTATED BY 90 DEGREE ?
NO
210
YES
YES
211
MOVE THE WINDOW TO MAXIMIZE THE SUM OF BELIEFS WITHIN THE WINDOW
COMPUTE THE SUM OF BELIEFS FOR CENTROID-BASED CROPS IN TWO ORIENTATIONS
212
213
CALCULATE FINAL FILM SAMPLE POSITION
CALCULATE FINAL FILM SAMPLE POSITION BASED ON THE HIGHER BELIEF CROP WINDOW

FIG. 3

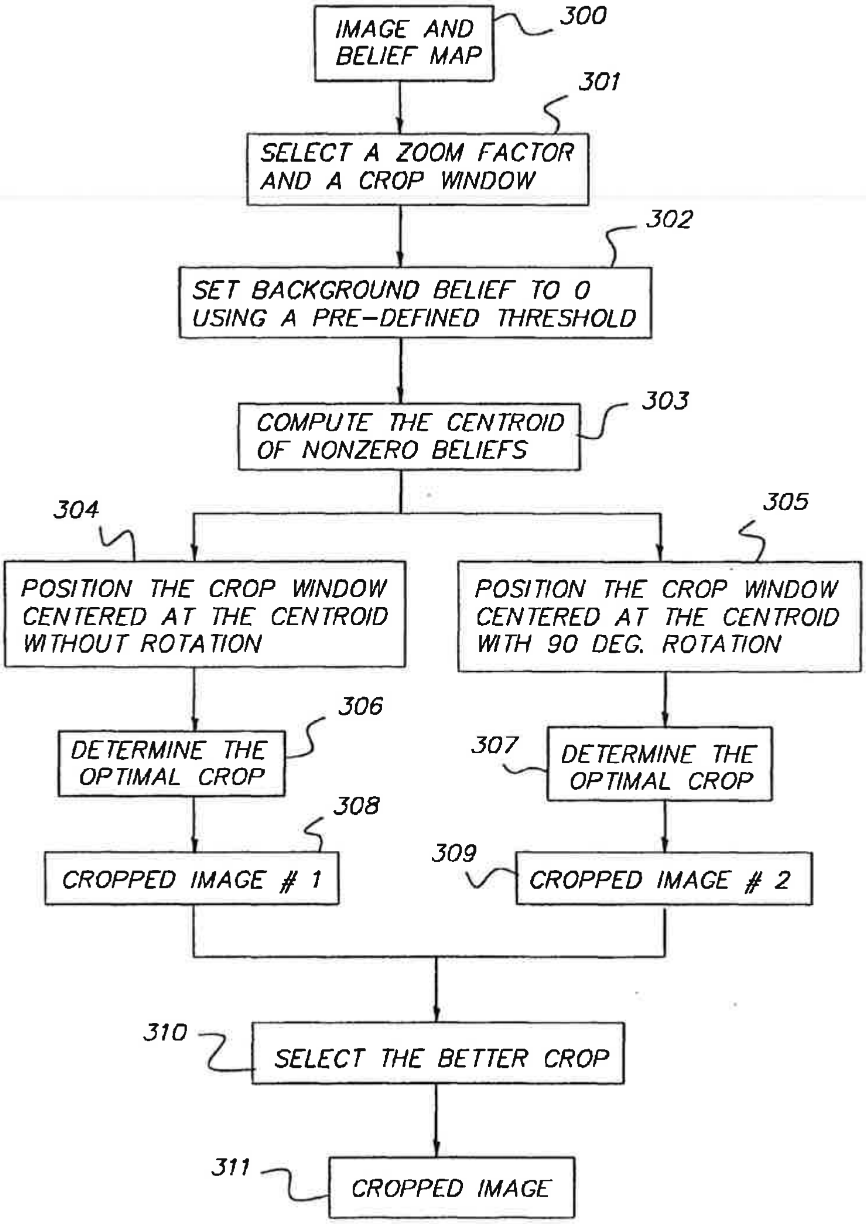


IMAGE AND BELIEF MAP
300
301
SELECT A ZOOM FACTOR AND A CROP WINDOW
302
SET BACKGROUND BELIEF TO 0 USING A PRE-DEFINED THRESHOLD
303
COMPUTE THE CENTROID OF NONZERO BELIEFS
304
305
POSITION THE CROP WINDOW CENTERED AT THE CENTROID WITHOUT ROTATION
POSITION THE CROP WINDOW CENTERED AT THE CENTROID WITH 90 DEG. ROTATION
306
DETERMINE THE OPTIMAL CROP
307
DETERMINE THE OPTIMAL CROP
308
CROPPED IMAGE # 1
309
CROPPED IMAGE # 2
310
SELECT THE BETTER CROP
311
CROPPED IMAGE

FIG. 4

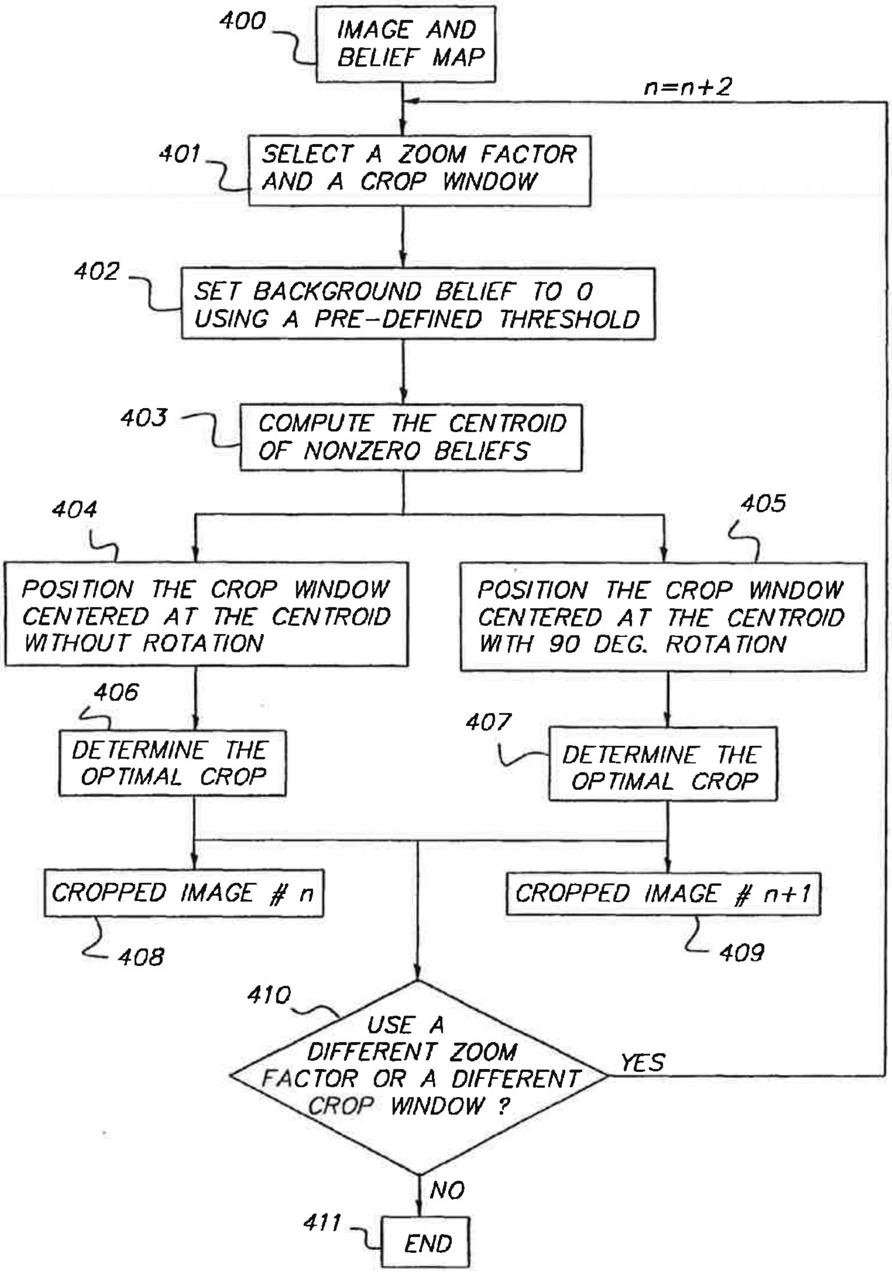

400
IMAGE AND BELIEF MAP
n=n+2
401
SELECT A ZOOM FACTOR AND A CROP WINDOW
402
SET BACKGROUND BELIEF TO 0 USING A PRE-DEFINED THRESHOLD
403
COMPUTE THE CENTROID OF NONZERO BELIEFS
404
405
POSITION THE CROP WINDOW CENTERED AT THE CENTROID WITHOUT ROTATION
POSITION THE CROP WINDOW CENTERED AT THE CENTROID WITH 90 DEG. ROTATION
406
DETERMINE THE OPTIMAL CROP
407
DETERMINE THE OPTIMAL CROP
CROPPED IMAGE # n
CROPPED IMAGE # n+1
408
409
410
USE A DIFFERENT ZOOM FACTOR OR A DIFFERENT CROP WINDOW ?
YES
NO
411
END

FIG. 5

Fig.6

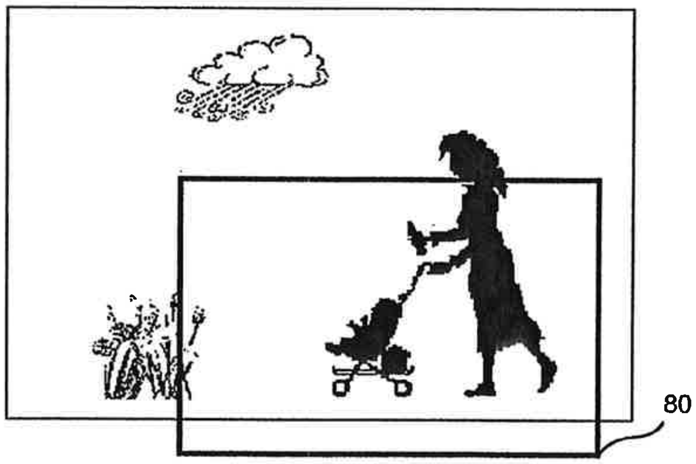


Fig.7

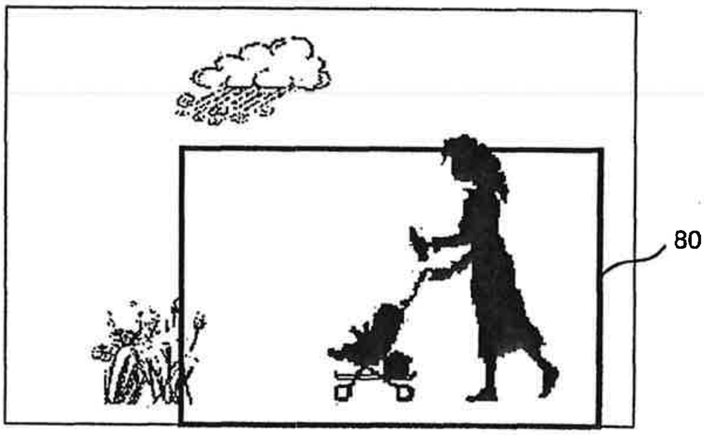


Fig. 8

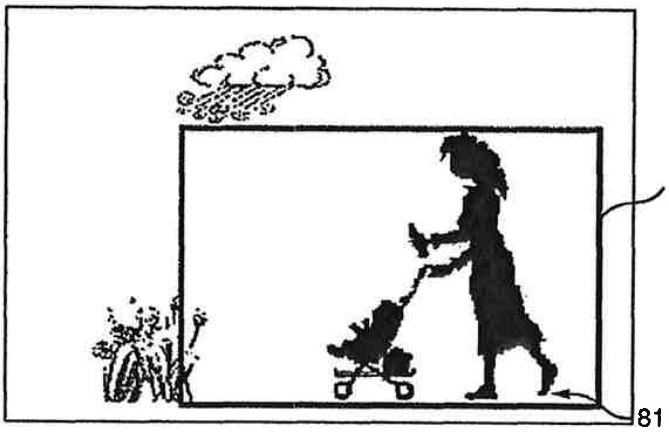


Fig. 9

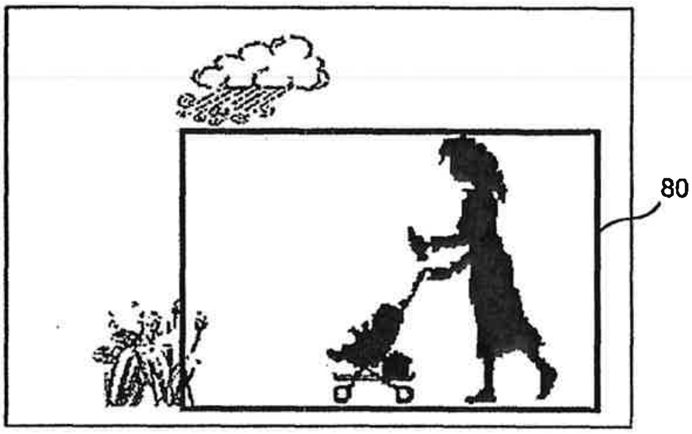


Fig. 10

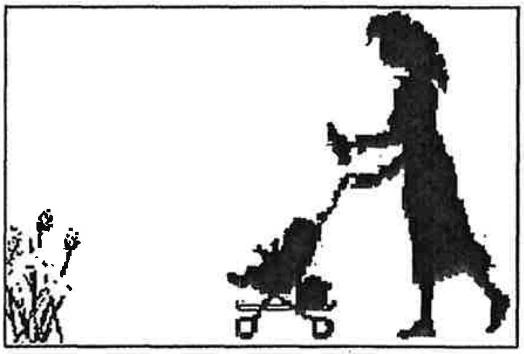


Fig. 11

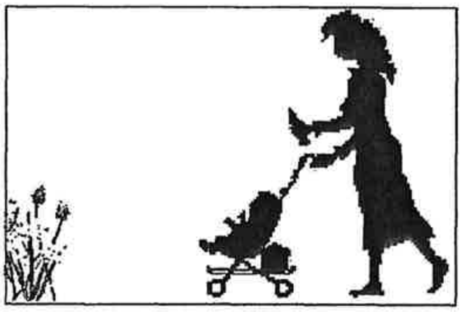


Fig. 12

Fig. 13

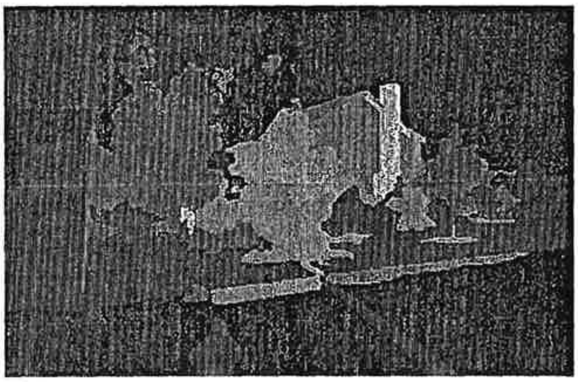


Fig. 14

Fig 15

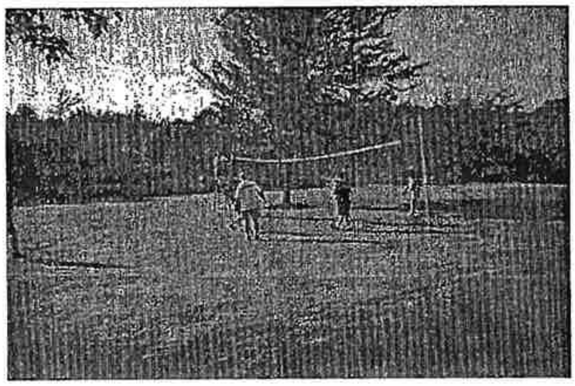


Fig. 16

Fig. 17

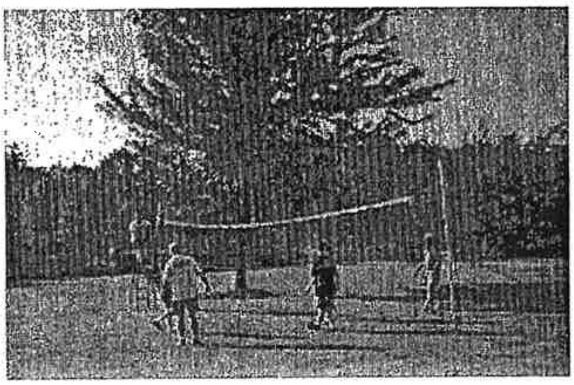


Fig 18

European Patent Office

EUROPEAN SEARCH REPORT

Application Number
EP 01 20 1712

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.7)
E	EP 1 120 742 A (EASTMAN KODAK CO) 1 August 2001 (2001-08-01) * the whole document *	1-10	G06T7/00 G03B27/62 G03B27/53
X,D	US 5 978 519 A (BOLLMAN JAMES E ET AL) 2 November 1999 (1999-11-02) * the whole document *	1	

TECHNICAL FIELDS SEARCHED (Int.Cl.7)
G06T
G03B

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
MUNICH	11 September 2001	Aratari, R

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)

ANNEX TO THE EUROPEAN SEARCH REPORT ON EUROPEAN PATENT APPLICATION NO. EP 01 20 1712

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report. The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

11-09-2001

Patent document cited in search report		Publication date	Patent family member(s)			Publication date
EP 1120742	A	01-08-2001	NONE			
US 5978519	A	02-11-1999	EP	0824246	A	18-02-1998
			JP	10105694	A	24-04-1998

EPO FORM P0459

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82

L Number	Hits	Search Text	DB	Time stamp
-	288	382/291.ccls.	USPAT	2002/12/30 11:42
-	350	382/173.ccls.	USPAT	2002/12/30 11:42
-	111	382/180.ccls.	USPAT	2002/12/30 11:42
-	110	382/225.ccls.	USPAT	2002/12/30 11:42
-	301	345/620-628.ccls.	USPAT	2002/12/30 11:43
-	208	382/282.ccls.	USPAT	2002/12/30 11:43
-	6	382/282.ccls. and ((auto or automatic$5) near10 (crop$4 or zoom$4))	USPAT	2002/12/30 11:58
-	15	((382/291.ccls. 382/173.ccls. 382/180.ccls. 382/225.ccls. 358/453.ccls. 345/620-628.ccls. 382/282.ccls.) and ((auto or automatic$5) near10 (crop$4 or zoom$4))) not (382/282.ccls. and ((auto or automatic$5) near10 (crop$4 or zoom$4)))	USPAT	2002/12/30 11:59
-	85	(382/291.ccls. 382/173.ccls. 382/180.ccls. 382/225.ccls. 358/453.ccls. 345/620-628.ccls. 382/282.ccls.) and ((auto or automatic$5) near10 (segment$6))	USPAT	2002/12/30 16:29
-	21	(382/291.ccls. 382/173.ccls. 382/180.ccls. 382/225.ccls. 358/453.ccls. 345/620-628.ccls. 382/282.ccls.) and ((auto or automatic$5) near10 (crop$4 or zoom$4))	USPAT	2002/12/30 16:30
-	4	((382/291.ccls. 382/173.ccls. 382/180.ccls. 382/225.ccls. 358/453.ccls. 345/620-628.ccls. 382/282.ccls.) and ((auto or automatic$5) near10 (segment$6))) and ((382/291.ccls. 382/173.ccls. 382/180.ccls. 382/225.ccls. 358/453.ccls. 345/620-628.ccls. 382/282.ccls.) and ((auto or automatic$5) near10 (crop$4 or zoom$4)))	USPAT	2002/12/31 10:39
-	1591	382/291.ccls. 382/173.ccls. 382/180.ccls. 382/225.ccls. 358/453.ccls. 345/620-628.ccls. 382/282.ccls.	USPAT	2002/12/30 16:35
-	60	(382/291.ccls. 382/173.ccls. 382/180.ccls. 382/225.ccls. 358/453.ccls. 345/620-628.ccls. 382/282.ccls.) and kodak	USPAT	2002/12/30 16:36
-	0	(382/291.ccls. 382/173.ccls. 382/180.ccls. 382/225.ccls. 358/453.ccls. 345/620-628.ccls. 382/282.ccls.) and (belief$4 near10 map$4)	USPAT	2002/12/30 16:35
-	9	((382/291.ccls. 382/173.ccls. 382/180.ccls. 382/225.ccls. 358/453.ccls. 345/620-628.ccls. 382/282.ccls.) and kodak) and (((382/291.ccls. 382/173.ccls. 382/180.ccls. 382/225.ccls. 358/453.ccls. 345/620-628.ccls. 382/282.ccls.) and ((auto or automatic$5) near10 (segment$6))) ((382/291.ccls. 382/173.ccls. 382/180.ccls. 382/225.ccls. 358/453.ccls. 345/620-628.ccls. 382/282.ccls.) and ((auto or automatic$5) near10 (crop$4 or zoom$4))))	USPAT	2002/12/31 13:07
-	89	707/505.ccls.	USPAT	2002/12/31 10:40
-	251	382/296.ccls.	USPAT	2002/12/31 10:40
-	93	382/297.ccls.	USPAT	2002/12/31 10:40
-	305	382/190.ccls.	USPAT	2002/12/31 10:40
-	29	("3618032" | "4047811" | "4593989" | "4907033" | "5020115" | "5053885" | "5191429" | "5212568" | "5220649" | "5260805" | "5280367" | "5438430" | "5483606" | "5546474" | "5557728" | "5596655" | "5623681" | "5713070" | "5784487" | "5838836" | "5880858" | "5943679" | "5960448" | "5974199" | "5978519" | "6011635" | "6043823" | "6240204" | "6389434").PN.	USPAT	2002/12/31 12:35

-	8	("4809064" | "5115271" | "5363209" | "5450502" | "5485568" | "5608544" | "5640468" | "5666503").PN.	USPAT	2002/12/31 12:41
-	5211	(photo or photograph$4 or image) and (subject near7 (find$3 or locat$4))	USPAT	2002/12/31 12:56
-	65	((photo or photograph$4 or image) and (subject near7 (find$3 or locat$4))) and crop$4 and scal$4	USPAT	2002/12/31 12:56
-	1166	((auto or automatic$5) near10 (crop$4 or zoom$4)) and (scal$4 or enlarg$4)	USPAT	2002/12/31 13:11
-	65	(((auto or automatic$5) near10 (crop$4 or zoom$4)) and (scal$4 or enlarg$4)) and 382/$.ccls.	USPAT	2002/12/31 13:11
-	1166	((auto or automatic$5) near10 ((crop$4 or zoom$4)) and (scal$4 or enlarg$4))	USPAT	2003/01/06 11:58
-	65	(((auto or automatic$5) near10 ((crop$4 or zoom$4)) and (scal$4 or enlarg$4))) and 382/$.ccls.	USPAT	2002/12/31 13:11
-	0	((auto or automatic$5) near10 ((crop$4 and zoom$4)) near10 (photo or photograph$4 or imag$4))	USPAT	2003/01/06 12:00
-	19	((auto or automatic$5) near10 crop$4) and ((auto or automatic$5) near10 zoom$4) and (photo or photograph$4 or imag$4)	USPAT	2003/01/06 17:23
-	1	6282317.pn.	USPAT	2003/01/07 10:52
-	7	("4975975" | "5724456" | "5809322" | "5850352" | "5862260" | "5983218" | "6014461").PN.	USPAT	2003/01/06 12:16
-	0	"Jiebo Luo".in.	USPAT	2003/01/06 17:24
-	0	Jiebo-Luo.in.	USPAT	2003/01/06 17:24
-	424	Luo.in.	USPAT	2003/01/06 17:24
-	11	Luo.in. and Kodak.as.	USPAT	2003/01/06 17:27
-	3	(Luo.in. and Kodak.as.) and (crop$6 or zoom$4)	USPAT	2003/01/06 17:28
-	362	358/453.ccls.	USPAT	2003/01/07 10:44
-	7	("4975975" | "5724456" | "5809322" | "5850352" | "5862260" | "5983218" | "6014461").PN.	USPAT	2003/01/07 10:48
-	358	382/203.ccls.	USPAT	2003/01/07 10:52
-	24	("4644582" | "4821029" | "4851616" | "5048103" | "5067015" | "5105186" | "5280610" | "5376947" | "5440744" | "5459636" | "5487115" | "5551027" | "5581758" | "5630017" | "5640468" | "5640558" | "5652880" | "5652882" | "5657426" | "5748804" | "5764283" | "5845007" | "5947169" | "6335985").PN.	USPAT	2003/01/07 11:06

L Number	Hits	Search Text	DB	Time stamp
1	1678	382/291.ccls. 382/173.ccls. 382/180.ccls. 382/225.ccls. 358/453.ccls. 345/620-628.ccls. 382/282.ccls.	USPAT	2003/06/12 17:03
2	88	(382/291.ccls. 382/173.ccls. 382/180.ccls. 382/225.ccls. 358/453.ccls. 345/620-628.ccls. 382/282.ccls.) and @pd>20021220	USPAT	2003/06/12 17:04
3	1	((382/291.ccls. 382/173.ccls. 382/180.ccls. 382/225.ccls. 358/453.ccls. 345/620-628.ccls. 382/282.ccls.) and @pd>20021220) and (photo or photograph$4 or image) and (subject near7 (find$3 or locat$4))	USPAT	2003/06/12 17:05
-	288	382/291.ccls.	USPAT	2002/12/30 11:42
-	350	382/173.ccls.	USPAT	2002/12/30 11:42
-	111	382/180.ccls.	USPAT	2002/12/30 11:42
-	110	382/225.ccls.	USPAT	2002/12/30 11:42
-	301	345/620-628.ccls.	USPAT	2002/12/30 11:43
-	200	382/282.ccls.	USPAT	2002/12/30 11:43
-	6	382/282.ccls. and ((auto or automatic$5) near10 (crop$4 or zoom$4))	USPAT	2002/12/30 11:58
-	15	((382/291.ccls. 382/173.ccls. 382/180.ccls. 382/225.ccls. 358/453.ccls. 345/620-628.ccls. 382/282.ccls.) and ((auto or automatic$5) near10 (crop$4 or zoom$4))) not (382/282.ccls. and ((auto or automatic$5) near10 (crop$4 or zoom$4)))	USPAT	2002/12/30 11:59
-	85	(382/291.ccls. 382/173.ccls. 382/180.ccls. 382/225.ccls. 358/453.ccls. 345/620-628.ccls. 382/282.ccls.) and ((auto or automatic$5) near10 (segment$6))	USPAT	2002/12/30 16:29
-	21	(382/291.ccls. 382/173.ccls. 382/180.ccls. 382/225.ccls. 358/453.ccls. 345/620-628.ccls. 382/282.ccls.) and ((auto or automatic$5) near10 (crop$4 or zoom$4))	USPAT	2002/12/30 16:30
-	4	((382/291.ccls. 382/173.ccls. 382/180.ccls. 382/225.ccls. 358/453.ccls. 345/620-628.ccls. 382/282.ccls.) and ((auto or automatic$5) near10 (segment$6))) and ((382/291.ccls. 382/173.ccls. 382/180.ccls. 382/225.ccls. 358/453.ccls. 345/620-628.ccls. 382/282.ccls.) and ((auto or automatic$5) near10 (crop$4 or zoom$4)))	USPAT	2002/12/31 10:39
-	60	(382/291.ccls. 382/173.ccls. 382/180.ccls. 382/225.ccls. 358/453.ccls. 345/620-628.ccls. 382/282.ccls.) and kodak	USPAT	2002/12/30 16:36
-	0	(382/291.ccls. 382/173.ccls. 382/180.ccls. 382/225.ccls. 358/453.ccls. 345/620-628.ccls. 382/282.ccls.) and (belief$4 near10 map$4)	USPAT	2002/12/30 16:35
-	9	((382/291.ccls. 382/173.ccls. 382/180.ccls. 382/225.ccls. 358/453.ccls. 345/620-628.ccls. 382/282.ccls.) and kodak) and (((382/291.ccls. 382/173.ccls. 382/180.ccls. 382/225.ccls. 358/453.ccls. 345/620-628.ccls. 382/282.ccls.) and ((auto or automatic$5) near10 (segment$6))) ((382/291.ccls. 382/173.ccls. 382/180.ccls. 382/225.ccls. 358/453.ccls. 345/620-628.ccls. 382/282.ccls.) and ((auto or automatic$5) near10 (crop$4 or zoom$4))))	USPAT	2002/12/31 13:07
-	89	707/505.ccls.	USPAT	2002/12/31 10:40
-	251	382/296.ccls.	USPAT	2002/12/31 10:40
-	93	382/297.ccls.	USPAT	2002/12/31 10:40
-	305	382/190.ccls.	USPAT	2002/12/31 10:40

<table>
<tr><td>-</td><td></td><td>29</td><td>("3618032" | "4047811" | "4593989" | "4907033" | "5020115" | "5053885" | "5191429" | "5212568" | "5220649" | "5260805" | "5280367" | "5438430" | "5483606" | "5546474" | "5557728" | "5596655" | "5623681" | "5713070" | "5784487" | "5838836" | "5880858" | "5943679" | "5960448" | "5974199" | "5978519" | "6011635" | "6043823" | "6240204" | "6389434").PN.</td><td>USPAT</td><td>2002/12/31 12:35</td></tr>
<tr><td>-</td><td></td><td>8</td><td>("4809064" | "5115271" | "5363209" | "5450502" | "5485568" | "5608544" | "5640468" | "5666503").PN.</td><td>USPAT</td><td>2002/12/31 12:41</td></tr>
<tr><td>-</td><td></td><td>5211</td><td>(photo or photograph$4 or image) and (subject near7 (find$3 or locat$4))</td><td>USPAT</td><td>2002/12/31 12:56</td></tr>
<tr><td>-</td><td></td><td>65</td><td>((photo or photograph$4 or image) and (subject near7 (find$3 or locat$4))) and crop$4 and scal$4</td><td>USPAT</td><td>2002/12/31 12:56</td></tr>
<tr><td>-</td><td></td><td>1166</td><td>((auto or automatic$5) near10 (crop$4 or zoom$4)) and (scal$4 or enlarg$4)</td><td>USPAT</td><td>2002/12/31 13:11</td></tr>
<tr><td>-</td><td></td><td>65</td><td>(((auto or automatic$5) near10 (crop$4 or zoom$4)) and (scal$4 or enlarg$4)) and 382/$.ccls.</td><td>USPAT</td><td>2002/12/31 13:11</td></tr>
<tr><td>-</td><td></td><td>1166</td><td>((auto or automatic$5) near10 ((crop$4 or zoom$4)) and (scal$4 or enlarg$4))</td><td>USPAT</td><td>2003/01/06 11:58</td></tr>
<tr><td>-</td><td></td><td>65</td><td>(((auto or automatic$5) near10 ((crop$4 or zoom$4)) and (scal$4 or enlarg$4))) and 382/$.ccls.</td><td>USPAT</td><td>2002/12/31 13:11</td></tr>
<tr><td>-</td><td></td><td>0</td><td>((auto or automatic$5) near10 ((crop$4 and zoom$4)) near10 (photo or photograph$4 or imag$4))</td><td>USPAT</td><td>2003/01/06 12:00</td></tr>
<tr><td>-</td><td></td><td>19</td><td>((auto or automatic$5) near10 crop$4) and ((auto or automatic$5) near10 zoom$4) and (photo or photograph$4 or imag$4)</td><td>USPAT</td><td>2003/01/06 17:23</td></tr>
<tr><td>-</td><td></td><td>1</td><td>6282317.pn.</td><td>USPAT</td><td>2003/01/07 10:52</td></tr>
<tr><td>-</td><td></td><td>7</td><td>("4975975" | "5724456" | "5809322" | "5850352" | "5862260" | "5983218" | "6014461").PN.</td><td>USPAT</td><td>2003/01/06 12:16</td></tr>
<tr><td>-</td><td></td><td>0</td><td>"Jiebo Luo".in.</td><td>USPAT</td><td>2003/01/06 17:24</td></tr>
<tr><td>-</td><td></td><td>0</td><td>Jiebo-Luo.in.</td><td>USPAT</td><td>2003/01/06 17:24</td></tr>
<tr><td>-</td><td></td><td>424</td><td>Luo.in.</td><td>USPAT</td><td>2003/01/06 17:24</td></tr>
<tr><td>-</td><td></td><td>11</td><td>Luo.in. and Kodak.as.</td><td>USPAT</td><td>2003/01/06 17:27</td></tr>
<tr><td>-</td><td></td><td>3</td><td>(Luo.in. and Kodak.as.) and (crop$6 or zoom$4)</td><td>USPAT</td><td>2003/01/06 17:28</td></tr>
<tr><td>-</td><td></td><td>362</td><td>358/453.ccls.</td><td>USPAT</td><td>2003/01/07 10:44</td></tr>
<tr><td>-</td><td></td><td>7</td><td>("4975975" | "5724456" | "5809322" | "5850352" | "5862260" | "5983218" | "6014461").PN.</td><td>USPAT</td><td>2003/01/07 10:48</td></tr>
<tr><td>-</td><td></td><td>358</td><td>382/203.ccls.</td><td>USPAT</td><td>2003/01/07 10:52</td></tr>
<tr><td>-</td><td></td><td>24</td><td>("4644582" | "4821029" | "4851616" | "5048103" | "5067015" | "5105186" | "5280610" | "5376947" | "5440744" | "5459636" | "5487115" | "5551027" | "5581758" | "5630017" | "5640468" | "5640558" | "5652880" | "5652882" | "5657426" | "5748804" | "5764283" | "5845007" | "5947169" | "6335985").PN.</td><td>USPAT</td><td>2003/01/07 11:06</td></tr>
<tr><td>-</td><td></td><td>1</td><td>5892837.pn.</td><td>USPAT</td><td>2003/05/20 14:46</td></tr>
<tr><td>-</td><td></td><td>1</td><td>6151403.pn.</td><td>USPAT</td><td>2003/05/20 14:44</td></tr>
<tr><td>-</td><td></td><td>1</td><td>6072893.pn.</td><td>USPAT</td><td>2003/05/20 14:47</td></tr>
<tr><td>-</td><td></td><td>1</td><td>6282317.pn.</td><td>USPAT</td><td>2003/05/20 17:00</td></tr>
<tr><td>-</td><td></td><td>7</td><td>("4975975" | "5724456" | "5809322" | "5850352" | "5862260" | "5983218" | "6014461").PN.</td><td>USPAT</td><td>2003/05/20 14:58</td></tr>
<tr><td>-</td><td></td><td>4</td><td>6282317.URPN.</td><td>USPAT</td><td>2003/05/20 16:44</td></tr>
<tr><td>-</td><td></td><td>4</td><td>6282317.URPN.</td><td>USPAT</td><td>2003/05/20 16:59</td></tr>
<tr><td>-</td><td></td><td>1</td><td>6545743.pn.</td><td>USPAT</td><td>2003/05/20 17:00</td></tr>
<tr><td>-</td><td></td><td>8</td><td>("4809064" | "5675400" | "5872619" | "5872643" | "5978519" | "5995201" | "6014461" | "6282317").PN.</td><td>USPAT</td><td>2003/05/20 17:10</td></tr>
</table>

CLAIMS ONLY

SERIAL NO. | FILING DATE

APPLICANT(S)

CLAIMS

	AS FILED		AFTER 1st AMENDMENT		AFTER 2nd AMENDMENT			*		*		*	
	IND.	DEP.	IND.	DEP.	IND.	DEP.		IND.	DEP.	IND.	DEP.	IND.	DEP.
1	|						51						
2		|					52						
3		|					53						
4		|					54						
5		|					55						
6		|					56						
7		|					57						
8		|					58						
9		|					59						
10		|					60						
11		|					61						
12		|					62						
13		|					63						
14		|					64						
15	|	[illegible]					65						
16		|					66						
17		|					67						
18		|					68						
19		|					69						
20		|					70						
21		|					71						
22		|					72						
23		|					73						
24		|					74						
25		|					75						
26		|					76						
27		|					77						
28		|					78						
29							79						
30							80						
31							81						
32							82						
33							83						
34							84						
35							85						
36							86						
37							87						
38							88						
39							89						
40							90						
41							91						
42							92						
43							93						
44							94						
45							95						
46							96						
47							97						
48							98						
49							99						
50							100						
TOTAL IND.	2						TOTAL IND.						
TOTAL DEP.	26						TOTAL DEP.						
TOTAL CLAIMS	28						TOTAL CLAIMS						

MAY BE USED FOR ADDITIONAL CLAIMS OR ADMENDMENTS

FORM PTO-2022 (1-98)

U.S.DEPARTMENT OF COMMERCE
Patent and Trademark Office

*U.S. GPO: 1998-443-593/89152

PATENT APPLICATION FEE DETERMINATION RECORD
Effective October 1, 2000

Application or Docket Number

CLAIMS AS FILED - PART I

	(Column 1)	(Column 2)
TOTAL CLAIMS	28	
FOR	NUMBER FILED	NUMBER EXTRA
TOTAL CHARGEABLE CLAIMS	28 minus 20=	* 8
INDEPENDENT CLAIMS	2 minus 3 =	* 0
MULTIPLE DEPENDENT CLAIM PRESENT		☐

* If the difference in column 1 is less than zero, enter "0" in column 2

SMALL ENTITY TYPE ☐		OR	OTHER THAN SMALL ENTITY	
RATE	FEE		RATE	FEE
BASIC FEE	355.00	OR	BASIC FEE	710.00
X$ 9=		OR	X$18=	144
X40=		OR	X80=	
+135=		OR	+270=	
TOTAL		OR	TOTAL	854

CLAIMS AS AMENDED - PART II

AMENDMENT A

	(Column 1) CLAIMS REMAINING AFTER AMENDMENT		(Column 2) HIGHEST NUMBER PREVIOUSLY PAID FOR	(Column 3) PRESENT EXTRA
Total	*	Minus	**	=
Independent	*	Minus	***	=
FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM ☐				

SMALL ENTITY RATE	ADDITIONAL FEE	OR	OTHER THAN SMALL ENTITY RATE	ADDITIONAL FEE
X$ 9=		OR	X$18=	
X40=		OR	X80=	
+135=		OR	+270=	
TOTAL ADDIT. FEE		OR	TOTAL ADDIT. FEE	

AMENDMENT B

	(Column 1) CLAIMS REMAINING AFTER AMENDMENT		(Column 2) HIGHEST NUMBER PREVIOUSLY PAID FOR	(Column 3) PRESENT EXTRA
Total	*	Minus	**	=
Independent	*	Minus	***	=
FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM ☐				

RATE	ADDITIONAL FEE	OR	RATE	ADDITIONAL FEE
X$ 9=		OR	X$18=	
X40=		OR	X80=	
+135=		OR	+270=	
TOTAL ADDIT. FEE		OR	TOTAL ADDIT. FEE	

AMENDMENT C

	(Column 1) CLAIMS REMAINING AFTER AMENDMENT		(Column 2) HIGHEST NUMBER PREVIOUSLY PAID FOR	(Column 3) PRESENT EXTRA
Total	*	Minus	**	=
Independent	*	Minus	***	=
FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM ☐				

RATE	ADDITIONAL FEE	OR	RATE	ADDITIONAL FEE
X$ 9=		OR	X$18=	
X40=		OR	X80=	
+135=		OR	+270=	
TOTAL ADDIT. FEE		OR	TOTAL ADDIT. FEE	

* If the entry in column 1 is less than the entry in column 2, write "0" in column 3.
** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 20, enter "20."
***If the "Highest Number Previously Paid For" IN THIS SPACE is less than 3, enter "3."
The "Highest Number Previously Paid For" (Total or Independent) is the highest number found in the appropriate box in column 1.

FORM **PTO-875**
(Rev. 8/00)

Patent and Trademark Office, U.S. DEPARTMENT OF COMMERCE
*U.S. GPO: 2000-460-706/30103

3/24 mm

ATTACH
DISK/FICHE/CD-ROM
ENVELOPE
HERE

(RIGHT INSIDE)

[redacted]

SEARCHED

Class	Sub.	Date	Exmr.
382	241 173 180 225 282	12/30/02 - 1/07/03	AWC
345	620 - 628		
358	453		
UPDATED AS ABOVE		5/20/03 - 6/12/05	AWC

SEARCH NOTES
(INCLUDING SEARCH STRATEGY)

	Date	Exmr.
EAST CLASS SEARCH	12/30/02 - 1/07/03	AWC
EAST TEXT SEARCH	12/30/02 - 1/07/03	AWC
IEEE SEARCH	5/20/03	AWC
INVENTOR NAME SEARCH	5/20/03	AWC
UPDATED EAST CLASS & TEXT SEARCH	5/20/03 - 6/12/03	AWC

INTERFERENCE SEARCHED

Class	Sub.	Date	Exmr.
SEARCHED AS ABOVE		6/12/03	AWC

(RIGHT OUTSIDE)

ISSUE SLIP STAPLE AREA (for additional cross references)

POSITION	INITIALS	ID NO.	DATE
FEE DETERMINATION			
O.I.P.E. CLASSIFIER		CS	1/16/01
FORMALITY REVIEW	AM	917	02-07-01
RESPONSE FORMALITY REVIEW	A.M	JC 580	03-23-01

INDEX OF CLAIMS

✓	Rejected	N	Non-elected
=	Allowed	I	Interference
— (Through numeral)	Canceled	A	Appeal
÷	Restricted	O	Objected

Claim Final	Claim Original	Date 1/19/03	Date 6/12/03								
1	1	✓	=								
2	2	O	↓								
3	3	✓									
4	4	O									
5	5	O									
6	6	O									
7	7	O									
8	8	✓									
9	9	O									
10	10	O									
11	11	O									
12	12	✓									
13	13	✓									
14	14	✓									
15	15	✓									
16	16	O									
17	17	✓									
18	18	O									
19	19	O									
20	20	O									
21	21	O									
22	22	✓									
23	23	O									
24	24	O									
25	25	O									
26	26	✓									
27	27	✓									
28	28	✓	↓								
	29										
	30										
	31										
	32										
	33										
	34										
	35										
	36										
	37										
	38										
	39										
	40										
	41										
	42										
	43										
	44										
	45										
	46										
	47										
	48										
	49										
	50										

Claim Final	Claim Original	Date
	51	
	52	
	53	
	54	
	55	
	56	
	57	
	58	
	59	
	60	
	61	—
	62	
	63	
	64	
	65	
	66	
	67	
	68	
	69	
	70	
	71	
	72	
	73	
	74	
	75	
	76	
	77	
	78	
	79	
	80	
	81	
	82	
	83	
	84	
	85	
	86	
	87	
	88	
	89	
	90	
	91	
	92	
	93	
	94	
	95	
	96	
	97	
	98	
	99	
	100	

Claim Final	Claim Original	Date
	101	
	102	
	103	
	104	
	105	
	106	
	107	
	108	
	109	
	110	
	111	
	112	
	113	
	114	
	115	
	116	
	117	
	118	
	119	
	120	
	121	
	122	
	123	
	124	
	125	
	126	
	127	
	128	
	129	
	130	
	131	
	132	
	133	
	134	
	135	
	136	
	137	
	138	
	139	
	140	
	141	
	142	
	143	
	144	
	145	
	146	
	147	
	148	
	149	
	150	

If more than 150 claims or 10 actions
staple additional sheet here

(LEFT INSIDE)

PATENT APPLICATION

09736825

jc784 U.S. PTO
09/736825
12/14/00

INITIALS JAN 16 01

CONTENTS

	Date Received (Incl. C. of M.) or Date Mailed
1. Application 10 papers.	
2. Ltr re CompAP	02-07-01
3. Drawings	03-16-01
4. IDS	12-14-00
5. Non-Final Reject (3 mos)	1-14-03
6. Asso. Power of Attorney	3-5-03
7. Amdt. A	4-2-03
8. Terminal Disclaimer (NE)	6/2/03
9. Terminal Disclaimer (NE)	6/2/03
10. Interview Summary	6.18.03
11. Allowance	6.18.03
12. Prior Art	9/29/03
13. Letter	[illegible]
14.	
15.	
16.	
17.	
18.	
19.	
20.	
21.	
22.	
23.	
24.	
25.	
26.	
27.	
28.	
29.	
30.	
31.	
32.	
33.	
34.	
35.	
36.	
37.	
38.	
39.	
40.	
41.	
42.	
43.	
44.	
45.	
46.	
47.	
48.	
49.	
50.	
51.	
52.	
53.	
54.	
55.	
56.	
57.	
58.	
59.	
60.	
61.	
62.	
63.	
64.	
65.	
66.	
67.	
68.	
69.	
70.	
71.	
72.	
73.	
74.	
75.	
76.	

(LEFT O