

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

(10) **Patent No.:** US 6,539,100 B1
 (45) **Date of Patent:** Mar. 25, 2003

(54) **METHOD AND APPARATUS FOR ASSOCIATING PUPILS WITH SUBJECTS**

(75) **Inventors:** Arnon Amir, Cupertino, CA (US); Myron Dale Flickner, San Jose, CA (US); David Bruce Koons, San Jose, CA (US); Carlos Hitoshi Morimoto, San Jose, CA (US)

(73) **Assignee:** International Business Machines Corporation, Armonk, 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.:** 09/238,979

(22) **Filed:** Jan. 27, 1999

(51) **Int. Cl.⁷** G06K 9/00

(52) **U.S. Cl.** 382/117; 382/173

(58) **Field of Search** 382/117, 199, 382/128, 173, 286; 351/206, 221, 208, 210; 434/40, 42

(56) **References Cited**

U.S. PATENT DOCUMENTS

4,275,385	A	6/1981	White	340/312
4,625,329	A	11/1986	Ishikawa et al.	382/1
4,931,865	A	6/1990	Scarampi	358/84
5,016,282	A	* 5/1991	Tomono et al.	382/117
5,291,560	A	* 3/1994	Daugman	382/117
5,430,809	A	7/1995	Tomitaka	382/173
5,432,866	A	* 7/1995	Sakamoto	382/128
5,550,928	A	8/1996	Lu et al.	382/116
5,596,362	A	1/1997	Zhou	348/14

OTHER PUBLICATIONS

Aizawa et al., "Detection and Tracking of Facial Features," Proc. of the SPIE Com. and Image Proc. (1995), v. 2501, Taipei Taiwan, pp. 1161-1172.

Baluja et al., "Neutral Network-Based Face Detection," Proc. IEEE Conf. on Computer Vision and Pattern Recognition (1996), San Francisco, CA, pp. 203-208.

Baluja et al., "Rotation Invariant Neutral Network-Based Face Detection," Proc. IEEE Conf. on Computer Vision and Pattern Recognition (Jun. 1998), Santa Barbara, CA, pp. 38-44.

Bérard et al., "LAFTER: Lips and Face Real Time Tracker," Proc. IEEE Conf. on Computer Vision and Pattern Recognition (Jun. 1997), Puerto Rico, pp. 123-129.

(List continued on next page.)

Primary Examiner—Andrew W. Johns

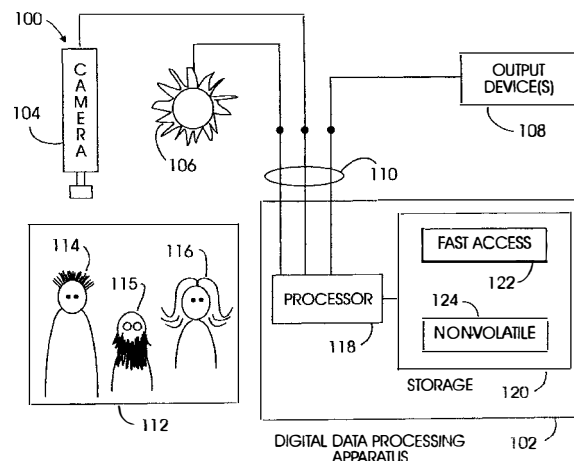
Assistant Examiner—Seyed Azarian

(74) *Attorney, Agent, or Firm*—Dan Hubert

(57) **ABSTRACT**

A method and apparatus analyzes a scene to determine which pupils correspond to which subjects. First, a machine-readable representation of the scene, such as a camera image, is generated. Although more detail may be provided, this representation minimally depicts certain visually perceivable characteristics of multiple pupil candidates corresponding to multiple subjects in the scene. A machine such as a computer then examines various features of the pupil candidates. The features under analysis include (1) visually perceivable characteristics of the pupil candidates at one given time ("spatial cues"), and (2) changes in visually perceivable characteristics of the pupil candidates over a sampling period ("temporal cues"). The spatial and temporal cues may be used to identify associated pupil pairs. Some exemplary spatial cues include interocular distance, shape, height, and color of potentially paired pupils. In addition to features of the pupils themselves, spatial cues may also include nearby facial features such as presence of a nose/mouth/eyebrows in predetermined relationship to potentially paired pupils, a similarly colored iris surrounding each of two pupils, skin of similar color nearby, etc. Some exemplary temporal cues include motion or blinking of paired pupils together, etc. With the foregoing examination, each pupil candidate can be associated with a subject in the scene.

65 Claims, 2 Drawing Sheets



OTHER PUBLICATIONS

- Birchfeld, "Elliptical Head Tracking Using Intensity Gradients and Color Histograms," Proc. IEEE Conf. on Computer Vision and Pattern Recognition (1998), pp. 232-237.
- Cohen et al., "Feature Extraction From Faces Using Deformable Template," International Journal of Computer Vision (1992), vol. 8, No. 2, pp. 99-111.
- Darrell et al., "Active Face Tracking and Pose Estimation in an Interactive Room," MIT Media Lab, TR-356 (1996), pp. 1-16.
- Darrell et al., "A Virtual Mirror Interface Using Real-Time Robust Face Tracking," Proc. Int'l conf. on Automatic Face and Gesture Recognition (Apr. 1998), Japan, pp. 616-621.
- Darrell et al., "Integrated Person Tracking Using Stereo, Color, and Pattern Detection," Proc. IEEE on Conf. on Computer Vision and Pattern Recognition (Jun. 1998), Santa Barbara, CA, pp. 601-608.
- Ebisawa et al., "Effectiveness of Pupil Area Detection Technique Using Two Light Sources and Image Difference Method", *Proc. of the 15th Ann. Int'l Conf. of IEEE Engineering in Medicine and Biology Society*, vol. 15, (Jan. 1993), pp. 1268-1269.
- Ebisawa, "Unconstrained Pupil Detection Technique Using Two Light Sources and the Image Difference Method", *Visualization and Intelligent Design in Engineering and Architecture* (1995), pp. 79-89.
- Fieguth et al., "Color-Based Tracking of Heads and Other Mobile Objects at Video Frame Rates", Proc. IEEE Conference on Computer Vision and Pattern Recognition (1997), pp. 21-27.
- Govindaraju et al., "A Computational Model For Face Location," Proc. of the Int'l Conf. on computer Vision (Dec. 1990), pp. 718-721, Osaka, Japan.
- Harville et al., "Tracking People With Integrated Stereo, Color, and Face Detection," Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, Jun. 1998, pp. 601-608.
- Kothari et al., "Detection of Eye Locations in Unconstrained Visual Images", Proc. Int'l Conf. on Image Processing (1996), Switzerland, pp. 519-522.
- Poggio et al., "Example-Based Learning for View-Based Human Face Detection," MIT AI Lab TR-AI-1521 (1994), pp. 1-20.
- Scassellati, "Eye Finding via Face Detection for a Foveated, Active Vision System," Proceedings of the 15th Conf. on Artificial Intelligence (AAAI-98), ISBN 0-262-51098-7, Jul. 26-30, 1998.
- Scassellati, "Real-Time Face and Eye Detection," world-wide-web, 1998.
- Stiefelhagen et al., "A Model-Based Gaze Tracking System," Proc. Joint Symposium on Intelligence and Systems (1996), pp. 304-310.
- Sirohey, "Human Face Segmentation and Identification", CAR-TR-695, CS-TR-3176, (1993), pp. 1-33.
- Waibel et al., "A Real-Time Face Tracker", Proc. of the 3rd IEEE Workshop on Applications of Computer Vision (1996), Sarasota, FL, pp. 142-147.

* cited by examiner

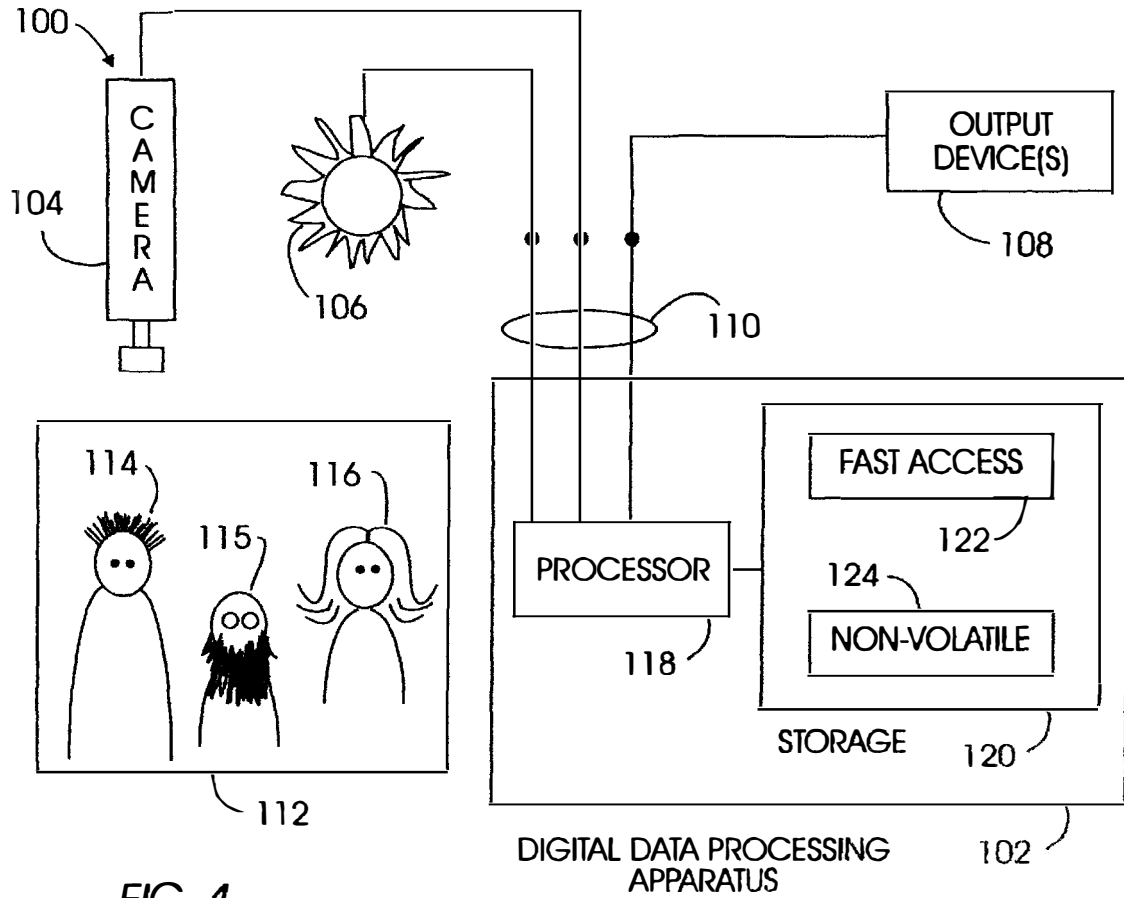


FIG. 1

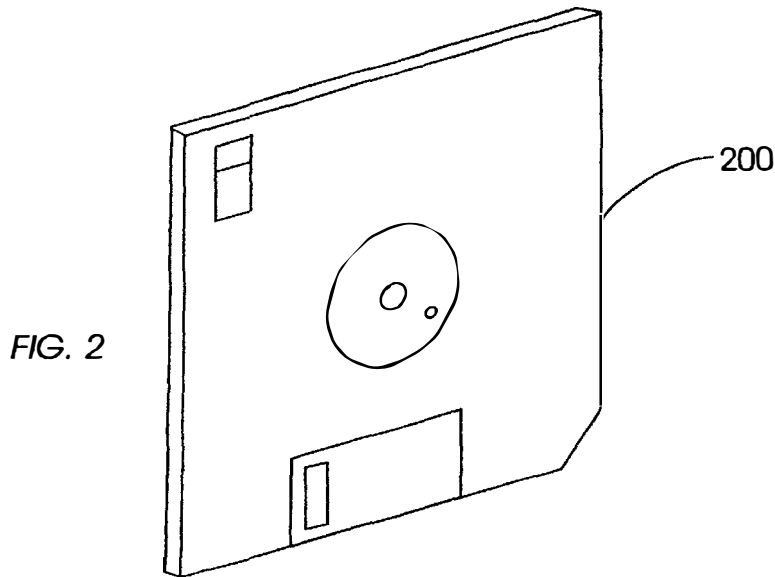


FIG. 2

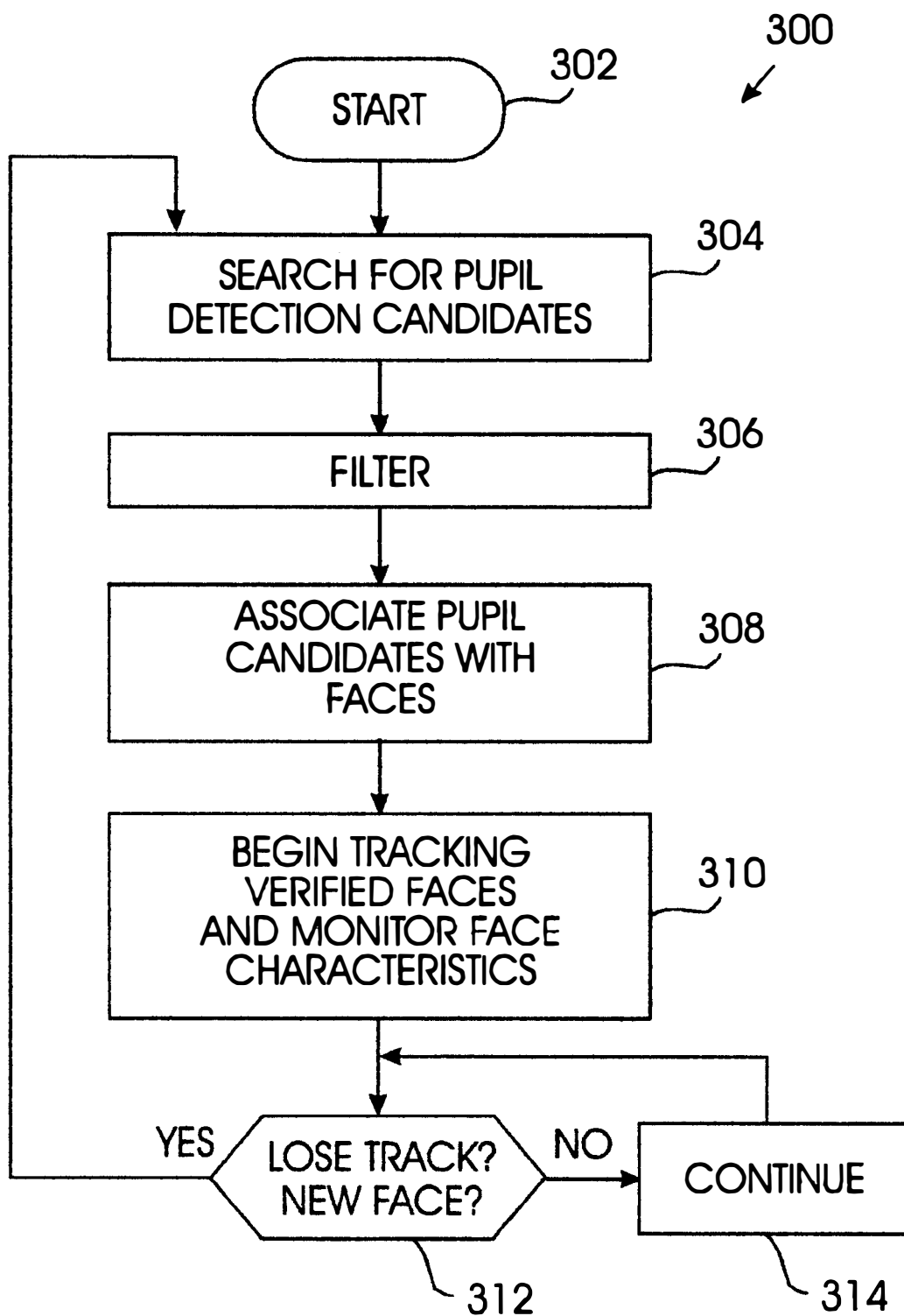


FIG. 3

METHOD AND APPARATUS FOR ASSOCIATING PUPILS WITH SUBJECTS

BACKGROUND OF THE INVENTION

1. Field of the Invention

The present invention relates to sophisticated interfaces between humans and machines. More particularly, the invention concerns a method and apparatus for analyzing a scene containing multiple subjects to determine which pupils correspond to which subjects.

2. Description of the Related Art

As more powerful human-machine interfaces are being developed, many such interfaces include the capability to perform user detection. By detecting the presence of a human user, a machine can manage its own functions more efficiently, and more reliably respond to human input. For example, a computer may employ user detection to selectively activate a screen saver when no users are present, or to display advertising banners only when a user is present. As another application, in home-based television viewing monitors for assessing "Nielsen" ratings, it may be useful to determine how many people are watching a television. User detection techniques such as face detection may also be used as a valuable precursor to eye gaze detection. In addition, face detection will likely be an important component of future human-machine interfaces that consider head and facial gestures to supplement mouse, voice, keyboard, and other user input. Such head and facial gestures may include nodding, leaning forward, head shaking, and the like. Thus, user detection is an important tool that enables a more natural human-machine interface.

Some user detection techniques are already known. For instance, a number of techniques focus on face detection using a combination of attributes such as color, shape, motion, and depth. Some of these approaches, for example, include template matching as described in U.S. Pat. No. 5,550,928 to Lu et al., and skin color analysis as described in U.S. Pat. No. 5,430,809 to Tomitaka. Another approach is the "Interval" system. The Interval system obtains range information using a sophisticated stereo camera system, gathers color information to evaluate as flesh tones, and analyzes face candidate inputs with a neural network trained to find faces. One drawback of the Interval system is the substantial computation expense. An example of the Interval system is described in Darrell et al., "Tracking People With Integrated Stereo, Color, and Face Detection," Perceptual User Interface Workshop, 1997. Although the Interval system may be satisfactory for some applications, certain users with less powerful or highly utilized computers may be frustrated with the interval system's computation requirements. The following references discuss some other user detection schemes: (1) T. Darrell et al., "Integrated person Tracking Using Stereo, Color, and Pattern Detection," 1998, and (2) T. Darrell et al., "Active Face Tracking and Pose Estimation in an Interactive Room," 1996.

As a different approach, some techniques perform user detection based on pupil detection. Pupil characteristics may be further analyzed to track eye position and movement, as described in U.S. Pat. No. 5,016,282 to Ptomain et al. Although the '282 patent and other pupil detection schemes may be satisfactory for some applications, such approaches are unable to process multiple faces and multiple pupils in an input image. Some difficulties include determining which pupils belong to the same face, and accounting for a partially off-screen person with only one pupil showing.

Thus, when multiple people and multiple pupils are present in an image, there may be considerable difficulty in associating pupils with people in order to detect how many people are present. In this respect, known approaches are not completely adequate for some applications due to certain unsolved problems.

SUMMARY OF THE INVENTION

Broadly, the present invention concerns a method and apparatus for analyzing a scene containing multiple subjects to determine which pupils correspond to which subjects. First, a machine-readable representation of the scene, such as a camera image, is generated. Although more detail may be provided, this representation minimally depicts certain visually perceptible characteristics (such as relative locations, shape, size, etc.) of multiple pupil candidates corresponding to multiple subjects in the scene. A computer analyzes various characteristics of the pupil candidates, such as: (1) visually perceivable characteristics of the pupil candidates at one given time ("spatial cues"), and (2) changes in visually perceivable characteristics of the pupil candidates over a sampling period ("temporal cues"). The spatial and temporal cues may be used to identify associated pupil pairs, i.e., two pupils belonging to the same subject/face. Some exemplary spatial cues include interocular distance between potentially paired pupils, horizontal alignment of pupils, same shape/size of pupils, etc. In addition to features of the pupils themselves, spatial cues may also include nearby facial features such as presence of a nose/mouth/eyebrows in predetermined relationship to potentially paired pupils, similarly colored irises surrounding the pupils, nearby skin of similar color, etc. Some exemplary temporal cues include motion or blinking of paired pupils together. With the foregoing analysis, each pupil candidate can be associated with a subject in the scene.

In one embodiment, the invention may be implemented to provide a method for analyzing a scene containing multiple subjects to determine which pupils correspond to which subjects. In another embodiment, the invention may be implemented to provide a computer-driven apparatus programmed to analyze a scene containing multiple subjects to determine which pupils correspond to which subjects. In still another embodiment, the invention may be implemented to provide a signal-bearing medium tangibly embodying a program of machine-readable instructions executable by a digital data processing apparatus to perform operations for analyzing a scene containing multiple subjects to determine which pupils correspond to which subjects. Still another embodiment involves a logic circuit configured to analyze a scene containing multiple subjects to determine which pupils correspond to which subjects.

The invention affords its users with a number of distinct advantages. First, unlike prior techniques, the invention is capable of determining which pupils belong to which faces/subjects in a scene with multiple subjects. In a scene with multiple subjects, understanding the pupil-subject relationship is an important prerequisite for tracking facial expressions, tracking movement, tracking user presence/absence, etc. As another advantage, the invention is inexpensive to implement when compared to other detection and tracking systems. For example, no dense range sensing is required. Also, an inexpensive camera may be used when a suitable lighting scheme is employed to cancel noise. The analysis provided by the invention is particularly robust because it is based on the grouping of multiple cues, both spatial and temporal. The invention also provides a number of other advantages and benefits, which should be apparent from the following description of the invention.

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts



Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research



With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips



Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS

Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS

Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS

Sync your system to PACER to automate legal marketing.