

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
26 July 2001 (26.07.2001)

PCT

(10) International Publication Number
WO 01/54416 A1

(51) International Patent Classification⁷: **H04N 7/34, 7/50**

(21) International Application Number: PCT/FI01/00050

(22) International Filing Date: 22 January 2001 (22.01.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
20000131 21 January 2000 (21.01.2000) FI

(71) Applicant (for all designated States except US): **NOKIA MOBILE PHONES LTD** [FI/FI]; Keilalahdentie 4, FIN-02150 Espoo (FI).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **KALEVO, Ossi** [FI/FI]; Ketunhätä 1, FIN-37800 Toijala (FI). **VAHTERI,**

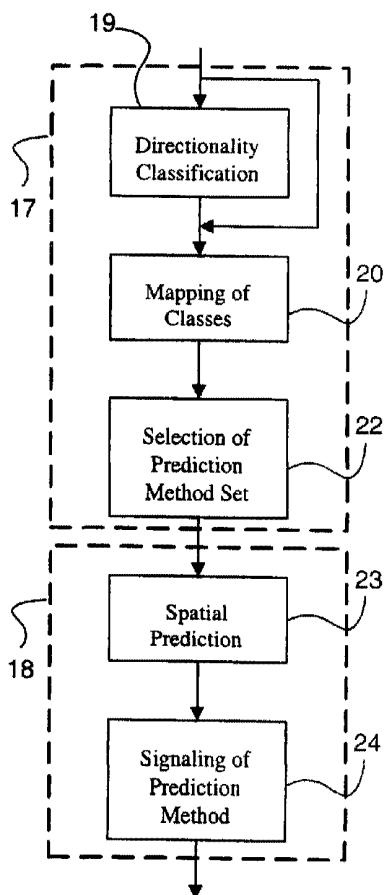
Joni [FI/FI]; Ojavainionkatu 18 B 10, FIN-33710 Tampere (FI). **DOBRIN, Bogdan-Paul** [RO/FI]; Iso-Roobertinkatu 8 B 12, FIN-00120 Helsinki (FI). **KARCZEWICZ, Marta** [PL/US]; 1224 Hidden Ridge, #3097, Irving, TX 75038 (US).

(74) Agent: **TAMPEREEN PATENTTITOIMISTO OY;** Hermiankatu 6, FIN-33720 Tampere (FI).

(81) Designated States (national): AE, AG, AL, AM, AT, AT (utility model), AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, CZ (utility model), DE, DE (utility model), DK, DK (utility model), DM, DZ, EE, EE (utility model), ES, FI, FI (utility model), GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SK (utility model), SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

[Continued on next page]

(54) Title: A METHOD FOR ENCODING IMAGES, AND AN IMAGE CODER



(57) Abstract: The invention relates to a method for encoding a digital image, in which method the digital image is divided into blocks (C, L, U, UL, UR). In the method a spatial prediction for a block (C) is performed to reduce the amount of information to be transmitted, wherein at least one prediction method (P1-P13) is defined. In the method a classification is determined for at least one neighbouring block (L, U) of said block (C) to be predicted according to the contents of said neighbouring block (L, U), and a prediction method (P1-P13) is selected for the current block (C) on the basis of at least one said classification.



WO 01/54416 A1



(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— with international search report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

A method for encoding images, and an image coder

The present invention relates to a method for encoding images according to the preamble of claim 1. The present invention also relates to a device for encoding images according to the preamble of claim 12. Furthermore, the present invention relates to an encoder according to the preamble of claim 23, to a decoder according to the preamble of claim 24, to a codec according to the preamble of claim 25, to a mobile terminal according to the preamble of claim 26, and a storage medium for storing a software program according to the preamble of claim 27.

The image can be any digital image, a video image, a TV image, an image generated by a video recorder, a computer animation, a still image, etc. In general, a digital image consists of pixels, arranged in horizontal and vertical lines, the number of which in a single image is typically tens of thousands. In addition, the information generated for each pixel contains, for instance, luminance information relating to the pixel, typically with a resolution of eight bits, and in colour applications also chrominance information, e.g. a chrominance signal. This chrominance signal generally consists of two components, Cb and Cr, which are both typically transmitted with a resolution of eight bits. On the basis of these luminance and chrominance values, it is possible to form information corresponding to the original pixel on the display device of a receiving video terminal. In this example, the quantity of data to be transmitted for each pixel is 24 bits uncompressed. Thus, the total amount of information for one image amounts to several megabits. In the transmission of a moving image, several images are transmitted per second. For instance in a TV image, 25 images are transmitted per second. Without compression, the quantity of information to be transmitted would amount to tens of megabits per second. However, for example in the Internet data network, the data transmission rate can be in the order of 64 kbits per second, which makes uncompressed real time image transmission via this network practically impossible.

To reduce the amount of information to be transmitted, a number of different compression methods have been developed, such as the

CONFIRMATION COPY

JPEG, MPEG and H.263 standards. In the transmission of video, image compression can be performed either as inter-frame compression, intra-frame compression, or a combination of these. In inter-frame compression, the aim is to eliminate redundant information in successive image frames. Typically, images contain a large amount of non-varying information, for example a motionless background, or slowly changing information, for example when the subject moves slowly. In inter-frame compression, it is also possible to utilize motion compensated prediction, wherein the aim is to detect elements in the image which are moving, wherein motion vector and prediction error information are transmitted instead of transmitting the pixel values.

To enable the use of image compression techniques in real time, the transmitting and receiving video terminal should have a sufficiently high processing speed that it is possible to perform compression and decompression in real time.

In several image compression techniques, an image signal in digital format is subjected to a discrete cosine transform (DCT) before the image signal is transmitted to a transmission path or stored in a storage means. Using a DCT, it is possible to calculate the frequency spectrum of a periodic signal, i.e. to perform a transformation from the time domain to the frequency domain. In this context, the word discrete indicates that separate pixels instead of continuous functions are processed in the transformation. In a digital image signal, neighbouring pixels typically have a substantial spatial correlation. One feature of the DCT is that the coefficients established as a result of the DCT are practically uncorrelated; hence, the DCT conducts the transformation of the image signal from the time domain to the (spatial) frequency domain in an efficient manner, reducing the redundancy of the image data. As such, use of transform coding is an effective way of reducing redundancy in both inter-frame and intra-frame coding.

Current block-based coding methods used in still image coding and video coding for independently coded key frames (intra-frames) use a block-based approach. In general, an image is divided into NxM blocks

that are coded independently using some kind of transform coding. Pure block-based coding only reduces the inter-pixel correlation within a particular block, without considering the inter-block correlation of pixels. Therefore, pure block-based coding produces rather high bit rates even when using transform-based coding, such as DCT coding, which has very efficient energy packing properties for highly correlated data. Therefore, current digital image coding standards exploit certain methods that also reduce the correlation of pixel values between blocks.

Current digital image coding methods perform prediction in the transform domain, i.e. they try to predict the DCT coefficients of a block currently being coded using the previous coded blocks and are thus coupled with the compression method. Typically a DCT coefficient that corresponds to the average pixel value within an image block is predicted using the same DCT coefficient from the previous coded block. The difference between the actual and predicted coefficient is sent to decoder. However, this scheme can predict only the average pixel value, and it is not very efficient.

Prediction of DCT coefficients can also be performed using spatially neighbouring blocks. For example, a DCT coefficient that corresponds to the average pixel value within a block is predicted using the DCT coefficient(s) from a block to the left or above the current block being coded. DCT coefficients that correspond to horizontal frequencies (i.e. vertical edges) can be predicted from the block above the current block and coefficients that correspond to vertical frequencies (i.e. horizontal edges) can be predicted from the block situated to the left. Similar to the previous method, differences between the actual and predicted coefficients are coded and sent to the decoder. This approach allows prediction of horizontal and vertical edges that run through several blocks.

In MPEG-2 compression, the DCT is performed in blocks using a block size of 8×8 pixels. The luminance level is transformed using full spatial resolution, while both chrominance signals are subsampled. For

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.