



(51) International Patent Classification:

G06K 9/00 (2006.01) **G06T 7/00** (2006.01)
G06K 9/66 (2006.01) **G06F 17/15** (2006.01)

(21) International Application Number:

PCT/FI2012/050305

(22) International Filing Date:

28 March 2012 (28.03.2012)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

1481/CHE/2011 29 April 2011 (29.04.2011) IN

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

(72) Inventors; and

(75) Inventors/Applicants (for US only): **ULIYAR, Mithun** [IN/IN]; NO. 31/169, Bandhavya, 3rd Cross, Canara Bank Colony, Chikalsandra, Karnataka, Bangalore 560061 (IN). **UKIL, Soumik** [IN/IN]; 603, 2nd Cross Road, 16th B main, Koramangala 3rd Block, Karnataka, Bangalore 560034 (IN).(74) Agents: **NOKIA CORPORATION** et al.; IPR Department, Jussi Jaatinen, Keilalahdentie 4, FI-02150 Espoo (FI).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

- with international search report (Art. 21(3))
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))

(54) Title: METHOD, APPARATUS AND COMPUTER PROGRAM PRODUCT FOR BLINK DETECTION IN MEDIA CONTENT

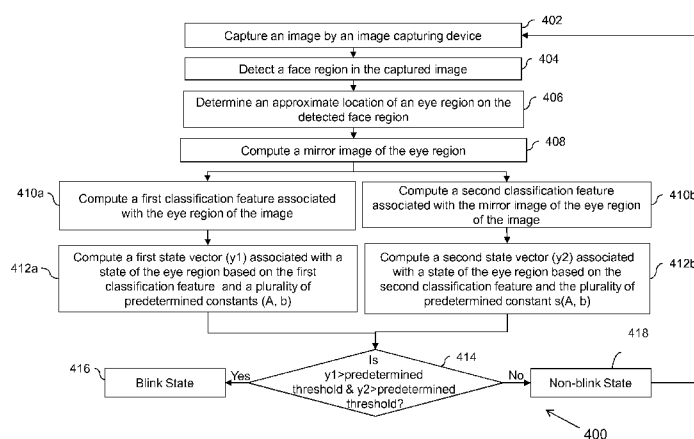


FIGURE 4

(57) Abstract: A method and an apparatus is provided for blink detection in digital images. The method comprises computing a first classification feature or set of features associated with an eye region of an image (410a). Local binary patterns (LBP) may be used in calculating the features. A first state vector associated with the eye region may be computed based on the first classification feature and a plurality of predetermined constants (412a). A state of the eye region may be determined based on a comparison of the computed first state vector with a predetermined threshold (414). A second set of features may be computed from a mirror image of the eye region (408, 410b, 412b). The state of the eye region is one of a plurality of blink states associated with the eye region (416, 418).



METHOD, APPARATUS AND COMPUTER PROGRAM PRODUCT FOR BLINK DETECTION IN MEDIA CONTENT

TECHNICAL FIELD

5

Various implementations relate generally to method, apparatus, and computer program product for blink detection in media content.

BACKGROUND

10

Blinking is a common phenomenon performed involuntarily by living beings (humans and animals) for keeping their eyes moist. Typically, while capturing media content such as images or videos of individuals, blinking produces undesirable results. The effect of blinking on the quality of media content is even more pronounced when a flash is used for capturing the media content. While capturing media content, blink detection is desirable for producing good quality results. Apart from its application in capturing media content, blink detection may be utilized in other areas as well, such as driver fatigue detection, face live detection, and the like.

15

SUMMARY OF SOME EMBODIMENTS

20

Various aspects of examples embodiments are set out in the claims.

25

In a first aspect, there is provided a method comprising: computing a first classification feature associated with an eye region of an image; computing a first state vector associated with the eye region based on the first classification feature and a plurality of predetermined constants; and determining a state of the eye region based on a comparison of the computed first state vector with a predetermined threshold, the state of the eye region being one of a plurality of blink states.

30

In a second aspect, there is provided an apparatus comprising: at least one processor; and at least one memory comprising computer program code, the at least one memory and the computer program code configured to, with the at least one processor, cause the apparatus at least to perform: computing a first classification feature associated with an eye region of an image;

computing a first state vector associated with the eye region based on the first classification feature and a plurality of predetermined constants; and determining a state of the eye region based on a comparison of the computed first state vector with a predetermined threshold, the state of the eye region being one of a plurality of blink states.

5

In a third aspect, there is provided a computer program product comprising at least one computer-readable storage medium, the computer-readable storage medium comprising a set of instructions, which, when executed by one or more processors, cause an apparatus to at least perform: computing a first classification feature associated with an eye region of an image;
10 computing a first state vector associated with the eye region based on the first classification feature and a plurality of predetermined constants; and determining a state of the eye region based on a comparison of the computed first state vector with a predetermined threshold, the state of the eye region being one of a plurality of blink states.

15 In a fourth aspect, there is provided an apparatus comprising: means for computing a first classification feature associated with an eye region of an image; means for computing a first state vector associated with the eye region based on the first classification feature and a plurality of predetermined constants; and means for determining a state of the eye region based on a comparison of the computed first state vector with a predetermined threshold, the state of the eye
20 region being one of a plurality of blink states.

In a fifth aspect, there is provided a computer program comprising program instructions which when executed by an apparatus, cause the apparatus to: compute a first classification feature associated with an eye region of an image; compute a first state vector associated with the eye
25 region based on the first classification feature and a plurality of predetermined constants; and determine a state of the eye region based on a comparison of the computed first state vector with a predetermined threshold, the state of the eye region being one of a plurality of blink states.

BRIEF DESCRIPTION OF THE FIGURES

30

Various embodiments are illustrated by way of example, and not by way of limitation, in the figures of the accompanying drawings in which:

FIGURE 1 illustrates a device in accordance with an example embodiment;

FIGURE 2 illustrates an apparatus for blink detection in media content in accordance with an example embodiment;

FIGURE 3 is a flowchart depicting an example method for blink detection in media content in accordance with an example embodiment; and

FIGURE 4 is a flowchart depicting an example method for detecting blink in the media content in accordance with another example embodiment.

DETAILED DESCRIPTION

Example embodiments and their potential effects are understood by referring to FIGURES 1 through 4 of the drawings.

FIGURE 1 illustrates a device 100 in accordance with an example embodiment. It should be understood, however, that the device 100 as illustrated and hereinafter described is merely illustrative of one type of device that may benefit from various embodiments, therefore, should not be taken to limit the scope of the embodiments. As such, it should be appreciated that at least some of the components described below in connection with the device 100 may be optional and thus in an example embodiment may include more, less or different components than those described in connection with the example embodiment of FIGURE 1. The device 100 could be any of a number of types of mobile electronic devices, for example, portable digital assistants (PDAs), pagers, mobile televisions, gaming devices, cellular phones, all types of computers (for example, laptops, mobile computers or desktops), cameras, audio/video players, radios, global positioning system (GPS) devices, media players, mobile digital assistants, or any combination of the aforementioned, and other types of communications devices.

The device 100 may include an antenna 102 (or multiple antennas) in operable communication with a transmitter 104 and a receiver 106. The device 100 may further include an apparatus, such as a controller 108 or other processing device that provides signals to and receives signals from the transmitter 104 and receiver 106, respectively. The signals may include signaling information in accordance with the air interface standard of the applicable cellular system, and/or may also include data corresponding to user speech, received data and/or user generated data. In this

regard, the device 100 may be capable of operating with one or more air interface standards, communication protocols, modulation types, and access types. By way of illustration, the device 100 may be capable of operating in accordance with any of a number of first, second, third and/or fourth-generation communication protocols or the like. For example, the device 100 may be capable of operating in accordance with second-generation (2G) wireless communication protocols IS-136 (time division multiple access (TDMA)), GSM (global system for mobile communication), and IS-95 (code division multiple access (CDMA)), or with third-generation (3G) wireless communication protocols, such as Universal Mobile Telecommunications System (UMTS), CDMA1000, wideband CDMA (WCDMA) and time division-synchronous CDMA (TD-SCDMA), with 3.9G wireless communication protocol such as evolved- universal terrestrial radio access network (E-UTRAN), with fourth-generation (4G) wireless communication protocols, or the like. As an alternative (or additionally), the device 100 may be capable of operating in accordance with non-cellular communication mechanisms. For example, computer networks such as the Internet, local area network, wide area networks, and the like; short range wireless communication networks such as include Bluetooth® networks, Zigbee® networks, Institute of Electric and Electronic Engineers (IEEE) 802.11x networks, and the like; wireline telecommunication networks such as public switched telephone network (PSTN).

The controller 108 may include circuitry implementing, among others, audio and logic functions of the device 100. For example, the controller 108 may include, but are not limited to, one or more digital signal processor devices, one or more microprocessor devices, one or more processor(s) with accompanying digital signal processor(s), one or more processor(s) without accompanying digital signal processor(s), one or more special-purpose computer chips, one or more field-programmable gate arrays (FPGAs), one or more controllers, one or more application-specific integrated circuits (ASICs), one or more computer(s), various analog to digital converters, digital to analog converters, and/or other support circuits. Control and signal processing functions of the device 100 are allocated between these devices according to their respective capabilities. The controller 108 thus may also include the functionality to convolutionally encode and interleave message and data prior to modulation and transmission. The controller 108 may additionally include an internal voice coder, and may include an internal data modem. Further, the controller 108 may include functionality to operate one or more software programs, which may be stored in a memory. For example, the controller 108 may be

capable of operating a connectivity program, such as a conventional Web browser. The connectivity program may then allow the device 100 to transmit and receive Web content, such as location-based content and/or other web page content, according to a Wireless Application Protocol (WAP), Hypertext Transfer Protocol (HTTP) and/or the like. In an example embodiment, the controller 108 may be embodied as a multi-core processor such as a dual or quad core processor. However, any number of processors may be included in the controller 108.

The device 100 may also comprise a user interface including an output device such as a ringer 110, an earphone or speaker 112, a microphone 114, a display 116, and a user input interface, which may be coupled to the controller 108. The user input interface, which allows the device 100 to receive data, may include any of a number of devices allowing the device 100 to receive data, such as a keypad 118, a touch display, a microphone or other input device. In embodiments including the keypad 118, the keypad 118 may include numeric (0-9) and related keys (#, *), and other hard and soft keys used for operating the device 100. Alternatively or additionally, the keypad 118 may include a conventional QWERTY keypad arrangement. The keypad 118 may also include various soft keys with associated functions. In addition, or alternatively, the device 100 may include an interface device such as a joystick or other user input interface. The device 100 further includes a battery 120, such as a vibrating battery pack, for powering various circuits that are used to operate the device 100, as well as optionally providing mechanical vibration as a detectable output.

In an example embodiment, the device 100 includes a media capturing element, such as a camera, video and/or audio module, in communication with the controller 108. The media capturing element may be any means for capturing an image, video and/or audio for storage, display or transmission. In an example embodiment in which the media capturing element is a camera module 122, the camera module 122 may include a digital camera capable of forming a digital image file from a captured image. As such, the camera module 122 includes all hardware, such as a lens or other optical component(s), and software for creating a digital image file from a captured image. Alternatively, the camera module 122 may include only the hardware needed to view an image, while a memory device of the device 100 stores instructions for execution by the controller 108 in the form of software to create a digital image file from a captured image. In an example embodiment, the camera module 122 may further include a processing element such as

a co-processor, which assists the controller 108 in processing image data and an encoder and/or decoder for compressing and/or decompressing image data. The encoder and/or decoder may encode and/or decode according to a JPEG standard format or another like format. For video, the encoder and/or decoder may employ any of a plurality of standard formats such as, for example, standards associated with H.261, H.262/ MPEG-2, H.263, H.264, H.264/MPEG-4, MPEG-4, and the like. In some cases, the camera module 122 may provide live image data to the display 116. Moreover, in an example embodiment, the display 116 may be located on one side of the device 100 and the camera module 122 may include a lens positioned on the opposite side of the device 100 with respect to the display 116 to enable the camera module 122 to capture images on one side of the device 100 and present a view of such images to the user positioned on the other side of the device 100.

The device 100 may further include a user identity module (UIM) 124. The UIM 124 may be a memory device having a processor built in. The UIM 124 may include, for example, a subscriber identity module (SIM), a universal integrated circuit card (UICC), a universal subscriber identity module (USIM), a removable user identity module (R-UIM), or any other smart card. The UIM 124 typically stores information elements related to a mobile subscriber. In addition to the UIM 124, the device 100 may be equipped with memory. For example, the device 100 may include volatile memory 126, such as volatile random access memory (RAM) including a cache area for the temporary storage of data. The device 100 may also include other non-volatile memory 128, which may be embedded and/or may be removable. The non-volatile memory 128 may additionally or alternatively comprise an electrically erasable programmable read only memory (EEPROM), flash memory, hard drive, or the like. The memories may store any number of pieces of information, and data, used by the device 100 to implement the functions of the device 100.

FIGURE 2 illustrates an apparatus 200 for blink detection in the media content in accordance with an example embodiment. The apparatus 200 may be employed, for example, in the device 100 of FIGURE 1. However, it should be noted that the apparatus 200, may also be employed on a variety of other devices both mobile and fixed, and therefore, embodiments should not be limited to application on devices such as the device 100 of FIGURE 1. In an example embodiment, the apparatus 200 is a mobile phone, which may be an example of a

communication device. Alternatively or additionally, embodiments may be employed on a combination of devices including, for example, those listed above. Accordingly, various embodiments may be embodied wholly at a single device, for example, the device 100 or in a combination of devices. It should be noted that some devices or elements described below may not be mandatory and thus some may be omitted in certain embodiments.

The apparatus 200 includes or otherwise is in communication with at least one processor 202 and at least one memory 204. Examples of the at least one memory 204 include, but are not limited to, volatile and/or non-volatile memories. Some examples of the volatile memory includes, but are not limited to, random access memory, dynamic random access memory, static random access memory, and the like. Some example of the non-volatile memory includes, but are not limited to, hard disks, magnetic tapes, optical disks, programmable read only memory, erasable programmable read only memory, electrically erasable programmable read only memory, flash memory, and the like. The memory 204 may be configured to store information, data, applications, instructions or the like for enabling the apparatus 200 to carry out various functions in accordance with various example embodiments. For example, the memory 204 may be configured to buffer input data comprising media content for processing by the processor 202. Additionally or alternatively, the memory 204 may be configured to store instructions for execution by the processor 202.

An example of the processor 202 may include the controller 108. The processor 202 may be embodied in a number of different ways. The processor 202 may be embodied as a multi-core processor, a single core processor; or combination of multi-core processors and single core processors. For example, the processor 202 may be embodied as one or more of various processing means such as a coprocessor, a microprocessor, a controller, a digital signal processor (DSP), processing circuitry with or without an accompanying DSP, or various other processing devices including integrated circuits such as, for example, an application specific integrated circuit (ASIC), a field programmable gate array (FPGA), a microcontroller unit (MCU), a hardware accelerator, a special-purpose computer chip, or the like. In an example embodiment, the multi-core processor may be configured to execute instructions stored in the memory 204 or otherwise accessible to the processor 202. Alternatively or additionally, the processor 202 may be configured to execute hard coded functionality. As such, whether configured by hardware or

software methods, or by a combination thereof, the processor 202 may represent an entity, for example, physically embodied in circuitry, capable of performing operations according to various embodiments while configured accordingly. For example, if the processor 202 is embodied as two or more of an ASIC, FPGA or the like, the processor 202 may be specifically
5 configured hardware for conducting the operations described herein. Alternatively, as another example, if the processor 202 is embodied as an executor of software instructions, the instructions may specifically configure the processor 202 to perform the algorithms and/or operations described herein when the instructions are executed. However, in some cases, the processor 202 may be a processor of a specific device, for example, a mobile terminal or network
10 device adapted for employing embodiments by further configuration of the processor 202 by instructions for performing the algorithms and/or operations described herein. The processor 202 may include, among other things, a clock, an arithmetic logic unit (ALU) and logic gates configured to support operation of the processor 202.

15 A user interface 206 may be in communication with the processor 202. Examples of the user interface 206 include, but are not limited to, input interface and/or output user interface. The input interface is configured to receive an indication of a user input. The output user interface provides an audible, visual, mechanical or other output and/or feedback to the user. Examples of the input interface may include, but are not limited to, a keyboard, a mouse, a joystick, a keypad,
20 a touch screen, soft keys, and the like. Examples of the output interface may include, but are not limited to, a display such as light emitting diode display, thin-film transistor (TFT) display, liquid crystal displays, active-matrix organic light-emitting diode (AMOLED) display, a microphone, a speaker, ringers, vibrators, and the like. In an example embodiment, the user interface 206 may include, among other devices or elements, any or all of a speaker, a
25 microphone, a display, and a keyboard, touch screen, or the like. In this regard, for example, the processor 202 may comprise user interface circuitry configured to control at least some functions of one or more elements of the user interface 206, such as, for example, a speaker, ringer, microphone, display, and/or the like. The processor 202 and/or user interface circuitry comprising the processor 202 may be configured to control one or more functions of one or more
30 elements of the user interface 206 through computer program instructions, for example, software and/or firmware, stored on a memory, for example, the at least one memory 204, and/or the like, accessible to the processor 202.

In an example embodiment, the apparatus 200 may include an electronic device. Some examples of the electronic device includes communication device, media capturing device with communication capabilities, computing devices, and the like. Some examples of the communication device may include a mobile phone, a personal digital assistant (PDA), and the like. Some examples of computing device may include a laptop, a personal computer, and the like. In an example embodiment, the communication device may include a user interface, for example, the UI 206, having user interface circuitry and user interface software configured to facilitate a user to control at least one function of the communication device through use of a display and further configured to respond to user inputs. In an example embodiment, the communication device may include a display circuitry configured to display at least a portion of the user interface of the communication device. The display and display circuitry may be configured to facilitate the user to control at least one function of the communication device.

In an example embodiment, the communication device may be embodied as to include a transceiver. The transceiver may be any device operating or circuitry operating in accordance with software or otherwise embodied in hardware or a combination of hardware and software. For example, the processor 202 operating under software control, or the processor 202 embodied as an ASIC or FPGA specifically configured to perform the operations described herein, or a combination thereof, thereby configures the apparatus or circuitry to perform the functions of the transceiver. The transceiver may be configured to receive media content. Examples of media content may include audio content, video content, data, and a combination thereof.

In an example embodiment, the communication device may be embodied as to include an image sensor, such as an image sensor 208. The image sensor 208 may be in communication with the processor 202 and/or other components of the apparatus 200. The image sensor 208 may be in communication with other imaging circuitries and/or software, and is configured to capture digital images or to make a video or other graphic media files. The image sensor 208 and other circuitries, in combination, may be an example of the camera module 122 of the device 100.

A media content, such as an image of a subject may be captured using an image capturing device, for example, a camera. The image may include a face region of the subject. The face region of the subject may include an eye region such that the eye region in the image may be in

one of a plurality of blink states. The term 'blink' may refer to a phenomenon of involuntarily opening and closing the eyes of the subject. The blink in the media content may refer to an image captured while the subject's eye is blinking, meaning thereby that the captured image may show a partially open or a partially closed eye. The term 'plurality of blink states' may include multiple states depending on an extent of the blink of the eye. The plurality of blink states may include a substantially open eye (may be referred to as non-blink state), a substantially closed eye (may be referred to as a blink state), and a plurality of other blink states that may vary in the extent of blink of the eye, for example, a half blink state (including 50% blink) a quarter blink state (including a 25% blink), and the like.

In an example embodiment, at least one eye region containing an eye and a neighborhood region thereof may be determined in the image. In an example embodiment, the at least one eye region may be detected by first detecting a location or region of a face in the image, and thereafter determining approximate location of the eyes, within the detected region of the face. The region of the face (hereinafter referred to as face region) may be determined by any technique known in the art. For example, the face region may be determined by using pattern recognition face detection technique. Upon determining approximate location of the face region, an approximate location of the eyes may be determined on the detected face region. In an example embodiment, the approximate location of the eyes may be determined by using pattern recognition eye locator being executed inside the detected face region. In an example embodiment, the eye region may include an intensity image of the eye region.

In an example embodiment, the processor 202 is configured to, with the content of the memory 204, and optionally with other components described herein, to cause the apparatus 200 to normalize the image containing the face region of the subject. In an example embodiment, the image may be normalized for the illumination characteristics in the image. In an example embodiment, a processing means may be configured to normalize the image. An example of the processing means may include the processor 202, which may be an example of the controller 108. In an example embodiment, a mirror image of the eye region may be normalized to generate a normalized mirror image of the eye region.

In an example embodiment, the processor 202 is configured to, with the content of the memory 204, and optionally with other components described herein, to cause the apparatus 200 to compute a first classification feature associated with the eye region of the image. In an example embodiment, the processor 202 is configured to, with the content of the memory 204, and optionally with other components described herein, to cause the apparatus 200 to compute the first classification feature by converting the normalized image into a first linear binary pattern (LBP) histogram image. The LBP features enhance edge properties of the image, for example the features of the eye region. In an example embodiment, the processor 202 is configured to, with the content of the memory 204, and optionally with other components described herein, to cause the apparatus 200 to compute a first difference image based on a difference of the first LBP histogram image and a predetermined mean LBP histogram image. In an example embodiment, the predetermined mean LBP image may be computed based on the processing of a set of eye region samples.

In an example embodiment, the processor 202 is configured to, with the content of the memory 204, and optionally with other components described herein, to cause the apparatus 200 to compute a second classification feature associated with the normalized mirror image. In an example embodiment, the second classification feature may be computed by converting the normalized mirror image into a second LBP histogram image. A second difference image may be computed based on a difference of the second LBP histogram image and the predetermined mean LBP histogram image.

For computing classification feature, such as the first classification feature and the second classification feature, a normalized image of the image and the mirror image respectively, may be divided into a plurality of non-overlapping rectangular regions. Each rectangular region (except the rectangular regions occurring at the edges of the image) may be surrounded by eight other rectangular regions. For each rectangular region, an intensity of each of the eight neighborhood regions is determined and compared with a threshold value. When the intensity value associated with the central rectangle's value is greater than that of the neighborhood rectangle, a value of '1' may be assigned. Otherwise, a value '0' may be assigned. In this way, an 8 digit binary number may be generated and a histogram of the values may be computed over the image. In an example embodiment, the histogram may be normalized for generating a feature

vector. Accordingly, an LBP feature X may be expressed as a concatenated sequence of histograms $X = (X_1, X_2, \dots, X_N)$, where N is the number of the regions. In an example embodiment, in addition to the LBP histogram image, a mean LBP histogram Image may also be generated based on the processing of a set of eye region samples.

5

In an example embodiment, the processor 202 is configured to, with the content of the memory 204, and optionally with other components described herein, to cause the apparatus 200 to compute a second classification feature associated with the mirror image of the eye region of the image. In an example embodiment, computing the second classification feature includes
10 converting the normalized mirror image of the eye region into a second LBP histogram image. A second difference image may be computed based on a difference of the second LBP histogram image and the predetermined mean LBP histogram image to generate the second classification feature. In an example embodiment, a processing means may be configured to compute a second classification feature associated with the mirror image of the eye region of the image. An
15 example of the processing means may include the processor 202, which may be an example of the controller 108.

In an example embodiment, the dimensions of the first classification feature X may be reduced using PCA (Principal Component Analysis). PCA is a multivariate data analysis procedure that
20 involves a transformation of a number of possibly correlated variables into a smaller number of uncorrelated variables known as principal components.

In an example embodiment, the processor 202 is configured to, with the content of the memory 204, and optionally with other components described herein, to cause the apparatus 200 to
25 compute a first state vector associated with a state of the eye region of the image based on the first classification feature and a plurality of predetermined constants. In an example embodiment, a processing means may be configured to compute the first state vector associated with the state of the eye region. An example of the processing means may include the processor 202, which may be an example of the controller 108.

30

In an example embodiment, the processor 202 is configured to, with the content of the memory 204, and optionally with other components described herein, to cause the apparatus 200 to

compute a second state vector associated with the eye region of the image based on the second classification feature and the plurality of the predetermined constants.

In an example embodiment, the processor 202 is configured to, with the content of the memory 204, and optionally with other components described herein, to cause the apparatus 200 to determine the plurality of predetermined constants based on a correlation between the state of the eye region and a set of image features extracted from a set of eye region samples. In an example embodiment, correlation is determined by a canonical correlation analysis (CCA).

The CCA may be utilized for measuring a linear relationship between two multidimensional variables. In an example embodiment, for performing the CCA, the input is the intensity image of the eye region. In an example embodiment, for the computation of the plurality of predetermined constants by the CCA, input data X and Y may be derived from the set of eye region samples, such that

$$X = \{x_1, x_2, \dots, x_N\} \text{ and}$$

$$Y = \{y_1, y_2, \dots, y_N\},$$

where,

X is represented by the concatenated block histograms of LBP features, computed around the detected eye regions, the dimension of X being $N \times D_x$ and that of Y being $N \times D_y$.

Y is a binary state vector e.g. {blink, non-blink}, denoting the status of the eyes, x_N is D_x dimensional, y_N is D_y dimensional, and sample number is N , and where, D_x and D_y are positive integers.

In an example embodiment, the dimensions of the input sample X may be reduced using PCA.

The result of PCA may be represented as Z . In an example embodiment, the state vector Y , may be defined as {1, 0} for closed eyes, and for open eyes the corresponding vector may be {0, 1}. In an example embodiment, the state vector Y may be created corresponding to each sample image.

In an example embodiment, CCA may be performed on the variable Z and Y as follows:

The variables C_{zz} , C_{yy} , C_{zy} , and C_{yz} may be calculated as:

$$\mathbf{C}_{zz} = \frac{1}{N} \sum_{n=1}^N \mathbf{z}_n \mathbf{z}_n^T = \frac{1}{N} \mathbf{Z} \mathbf{Z}^T$$

$$\bar{y} = \frac{1}{N} \sum_{n=1}^N y_n$$

$$\mathbf{C}_{yy} = \frac{1}{N} \sum_{n=1}^N (\mathbf{y}_n - \bar{y})(\mathbf{y}_n - \bar{y})^T$$

$$\mathbf{C}_{zy} = \frac{1}{N} \sum_{n=1}^N \mathbf{z}_n (\mathbf{y}_n - \bar{y})^T$$

$$5 \quad \mathbf{C}_{yz} = \mathbf{C}_{zy}^T$$

wherein, $\mathbf{C}_{zz}$, and $\mathbf{C}_{yy}$ are nonsingular matrices, and $\mathbf{C}_{zy} = \mathbf{C}_{yz}^T$

10 An eigen-decomposition may be performed to $\mathbf{C}_{yy}^{-1} \mathbf{C}_{yz} \mathbf{C}_{zz}^{-1} \mathbf{C}_{zy}$ to obtain eigen vector matrix $\mathbf{W}_y$ and eigen value vector.

$\mathbf{W}_z$ may be computed as:

$$\mathbf{W}_z = \mathbf{C}_{zz}^{-1} \mathbf{C}_{zy} \mathbf{W}_y$$

The columns of $\mathbf{W}_z$ may be normalized.

$$15 \quad \mathbf{y}'_n = \mathbf{W}_y^T \mathbf{y}_n, \mathbf{z}'_n = \mathbf{W}_z^T \mathbf{z}_n, \\ \mathbf{Y}' = \{\mathbf{y}'_1, \mathbf{y}'_2, \dots, \mathbf{y}'_N\}, \mathbf{Z}' = \{\mathbf{z}'_1, \mathbf{z}'_2, \dots, \mathbf{z}'_N\}$$

wherein, dimensions of

$$\mathbf{C}_{zz} = D_z \times D_z, \mathbf{C}_{yy} = D_y \times D_y, \mathbf{C}_{zy} = D_z \times D_y \text{ \& } \mathbf{C}_{yz} = D_y \times D_z \text{ \& }$$

$$\mathbf{W}_y = D_y \times D_y \text{ \& } \mathbf{r} = D_y \times 1 \text{ \& } \mathbf{W}_z = D_z \times D_y \text{ \& }$$

$$20 \quad \mathbf{y}'_n = D_y \times 1 \text{ \& } \mathbf{z}'_n = D_z \times 1 \text{ \& }$$

$$\mathbf{Y}' = D_y \times N \text{ \& } \mathbf{Z}' = D_z \times N$$

25 If a scatter plot shows a linear relationship between two quantitative variables, least-squares regression is a method for finding a line that summarizes the relationship between the two variables, at least within the domain of the explanatory variable, x. The least-squares regression line (LSRL) is a mathematical model for the data. A least square regression between $\mathbf{Z}'$ and $\mathbf{Y}'$ may be performed as follows:

Let dth row of $\mathbf{Y}'$ and $\mathbf{Z}'$ be

$$\mathbf{y}'_d = [\mathbf{y}'_{d,1}, \mathbf{y}'_{d,2}, \dots, \mathbf{y}'_{d,N}]^T \text{ \& }$$

$$\mathbf{z}'_d = [z'_{d,1}, z'_{d,2}, \dots, z'_{d,N}]^T \text{ respectively}$$

Also, let

$$y'_{n,d} = a_d z'_{n,d} + b_d, \text{ then}$$

$$\mathbf{y}'_d = a_d \mathbf{z}'_d + b_d = \mathbf{H}_d [a_d, b_d]^T$$

5 where $\mathbf{H}_d = [\mathbf{z}'_d, \mathbf{1}]$

$$\text{then } [a_d, b_d]^T = (\mathbf{H}_d^T \mathbf{H}_d)^{-1} \mathbf{H}_d^T \mathbf{y}'_d$$

$$\text{Let } \mathbf{a}' = [a_1, a_2, \dots, a_{D_y}]^T \ \& \ \mathbf{b}' = [b_1, b_2, \dots, b_{D_y}]^T \ \& \ \mathbf{A}' = \text{diag}(\mathbf{a}'), \text{ then}$$

$$\mathbf{y}'_n = \mathbf{A}' \mathbf{z}'_n + \mathbf{b}'$$

10 If two variables are correlated, then knowing the score on one variable allows prediction of the score on the other variable. The stronger the correlation, the closer the scores may fall to a regression line, thereby providing a more accurate prediction. In multiple regression, one variable is predicted on the basis of several other variables.

15 The multi-regression of X and Y may be derived as follows:

$$\mathbf{W}_y^T \mathbf{y}_n = \mathbf{y}'_n = \mathbf{A}' \mathbf{z}'_n + \mathbf{b}' = \mathbf{A}' \mathbf{W}_z^T \mathbf{z}_n + \mathbf{b}' = \mathbf{A}' \mathbf{W}_z^T \mathbf{P}^T \mathbf{x}_n + \mathbf{b}'$$

$$\mathbf{y}_n = \mathbf{W}_y^{-T} (\mathbf{A}' \mathbf{W}_z^T \mathbf{P}^T \mathbf{x}_n + \mathbf{b}') = \mathbf{W}_y^{-T} \mathbf{A}' \mathbf{W}_z^T \mathbf{P}^T \mathbf{x}_n + \mathbf{W}_y^{-T} \mathbf{b}'$$

$$= \mathbf{W}_y^{-T} \mathbf{A}' \mathbf{W}_z^T \mathbf{P}^T (\mathbf{x}_n + \bar{\mathbf{x}}) + \mathbf{W}_y^{-T} \mathbf{b}' - \mathbf{W}_y^{-T} \mathbf{A}' \mathbf{W}_z^T \mathbf{P}^T \bar{\mathbf{x}} = \mathbf{A} (\mathbf{x}_n + \bar{\mathbf{x}}) + \mathbf{b}$$

$$\mathbf{A} = \mathbf{W}_y^{-T} \mathbf{A}' \mathbf{W}_z^T \mathbf{P}^T$$

$$\mathbf{b} = \mathbf{W}_y^{-T} \mathbf{b}' - \mathbf{W}_y^{-T} \mathbf{A}' \mathbf{W}_z^T \mathbf{P}^T \bar{\mathbf{x}}$$

20

Accordingly, the values of the plurality of predetermined constants, as computed based on the plurality of sample images are:

$$\mathbf{A} = \mathbf{W}_y^{-T} \mathbf{A}' \mathbf{W}_z^T \mathbf{P}^T \text{ and } \mathbf{b} = \mathbf{W}_y^{-T} \mathbf{b}' - \mathbf{W}_y^{-T} \mathbf{A}' \mathbf{W}_z^T \mathbf{P}^T \bar{\mathbf{x}}$$

For an input data $\mathbf{x}$ corresponding to a sample of an eye region in an image, these values of the predetermined constants terms A and b may be utilized for computing the value of the state vector $\mathbf{y}$ by first normalizing $\mathbf{x}$ and inputting the values of the predetermined constants A, b and $\bar{\mathbf{x}}$ in the following equation:

25

$$y = Ax + b, \text{ wherein}$$

A is a PCA_CCA_Vector, and

b is a linear vector

5 In an example embodiment, the output samples of the state vector may occur in the range 0-1. In an example embodiment, the processor 202 is configured to, with the content of the memory 204, and optionally with other components described herein, to cause the apparatus 200 to determine the state of the eye region of the image based on a comparison of the computed first state vector with a predetermined threshold. In an example embodiment, the state of the eye region may be
10 determined to be a blink state when the first state vector is determined to be greater than the predetermined threshold. In an example embodiment, the value of the predetermined threshold may be 0.55. In the present embodiment, for the state of eye to be a blink state, the predetermined threshold may be (>0.55 , <0.55), and for the state of the eye to be the non-blink state, the predetermined threshold may be (<0.55 , >0.55). In an example embodiment, a
15 processing means may be configured to determine the one of the plurality of blink states of the eye region of the image based on the comparison of the computed first state vector with the predetermined threshold. An example of the processing means may include the processor 202, which may be an example of the controller 108.

20 In the present embodiment, the processor 202 is configured to, with the content of the memory 204, and optionally with other components described herein, to cause the apparatus 200 to determine a state of the eye region of the image based on a comparison of the computed first state vector and the second state vector with the predetermined threshold. In an example embodiment, the state of the eye portion is determined to be a blink state when the values of the
25 first state vector and the second state vector are determined to be greater than the predetermined threshold. A method for detecting blink in the media content is explained in FIGURE 3.

FIGURE 3 is a flowchart depicting an example method 300 for detecting blink in the media content in accordance with an example embodiment. The method 300 depicted in the flow chart
30 may be executed by, for example, the apparatus 200 of FIGURE 2. Examples of the apparatus 200 include, but are not limited to, digital cameras, camcorders, mobile phones, personal digital assistants (PDAs), laptops, and any equivalent devices.

The method 300 describes steps for detecting blink in the media content. Examples of the media content may include, but are not limited to digital images, video content, and the like. Detecting the blink in the media content may include determining a state of an eye region in the media content. The state of the eye region may be one of a plurality of blink states. In an example embodiment, the media content, such as the image having a subject's face may be normalized.

At block 302, a first classification feature associated with an eye region of an image may be computed. In an example embodiment, the first classification feature may be computed by converting the normalized image into a first LBP histogram image. A first difference image may be computed based on a difference of the first LBP histogram image and a predetermined mean LBP histogram image. The predetermined mean LBP image may be computed based on the processing of a set of eye region samples.

At block 304, a first state vector associated with a state of the eye region may be computed based on the first classification feature and a plurality of predetermined constants. In an example embodiment, the values of the plurality of predetermined constants are determined based on a correlation between the state of the eye region and a set of image features extracted from a set of eye region samples. In an example embodiment, the correlation is determined by the CCA.

At block 306, a state of the eye region may be determined based on a comparison of the computed first state vector with a predetermined threshold. In an example embodiment, the state of the eye region is the blink state when the computed first state vector is determined to be greater than the predetermined threshold. In an example embodiment, the value of the predetermined threshold may be 0.55. In the present embodiment, for the state of eye region to be a blink state, the predetermined threshold may be (>0.55 , <0.55), and for the state of the eye to be the non-blink state, the predetermined threshold may be (<0.55 , >0.55).

In an example embodiment, a processing means may be configured to perform some or all of: computing a first classification feature associated with an eye region of an image; computing a first state vector associated with the eye region based on the first classification feature and a plurality of predetermined constants; and determining a state of the eye region based on a comparison of the computed first state vector with a predetermined threshold. An example of the

processing means may include the processor 202, which may be an example of the controller 108.

FIGURE 4 is a flowchart depicting an example method 400 for blink detection in media content in accordance with another example embodiment. The method 400 depicted in flow chart may be executed by, for example, the apparatus 200 of FIGURE 2.

Operations of the flowchart, and combinations of operation in the flowchart, may be implemented by various means, such as hardware, firmware, processor, circuitry and/or other device associated with execution of software including one or more computer program instructions. For example, one or more of the procedures described in various embodiments may be embodied by computer program instructions. In an example embodiment, the computer program instructions, which embody the procedures, described in various embodiments may be stored by at least one memory device of an apparatus and executed by at least one processor in the apparatus. Any such computer program instructions may be loaded onto a computer or other programmable apparatus (for example, hardware) to produce a machine, such that the resulting computer or other programmable apparatus embody means for implementing the operations specified in the flowchart. These computer program instructions may also be stored in a computer-readable storage memory (as opposed to a transmission medium such as a carrier wave or electromagnetic signal) that may direct a computer or other programmable apparatus to function in a particular manner, such that the instructions stored in the computer-readable memory produce an article of manufacture the execution of which implements the operations specified in the flowchart. The computer program instructions may also be loaded onto a computer or other programmable apparatus to cause a series of operations to be performed on the computer or other programmable apparatus to produce a computer-implemented process such that the instructions, which execute on the computer or other programmable apparatus provide operations for implementing the operations in the flowchart. The operations of the method 400 are described with help of apparatus 200. However, the operations of the method 400 can be described and/or practiced by using any other apparatus.

It may be understood that for describing the method 200, references herein may be made to FIG. 1. The method may be explained with reference to a digital image, however, the method may be equally applicable for the determination of blink in other digital media content, such as a video.

- 5 At block 402, an image of a subject, such as a human face may be captured using an image capturing device, for example, a camera. An eye region in the image may be in one of a plurality of blink states.

At block 404, a face region may be detected in the captured image. The region of the face may
10 be determined by any technique known in the art. For example, the face region may be determined by using pattern recognition face detection technique. Upon determining approximate location of the face region, an approximate location of an eye on the detected face region may be determined at block 406. In an example embodiment, the approximate location of the eyes may be determined by using pattern recognition eye locator being executed inside the
15 detected face region. In an example embodiment, the detected eye region may be the intensity image of the eye region. Upon determining approximate eye region, a neighborhood region of the eye may be selected. In an example embodiment, the neighborhood region of the eye may be selected by a processor, such as the processor 202. The eye along with a neighborhood region thereof may hereinafter be referred to as the eye region.

20

At block 408, a mirror image of the eye region may be computed. In an example embodiment, the mirror image of the eye region may be computed by a processing means. Example of a processing means may include a processor, such as the processor 202.

- 25 At block 410a, a first classification feature associated with the eye region is computed. In an example embodiment, the first classification feature may be computed by converting the intensity image of the eye region into a first LBP histogram image. In an example embodiment, the first classification feature may be represented by concatenated block histograms of LBP features, computed around the detected eye regions. The first classification feature may be
30 represented by $X1$. Also, a first difference image may be computed based on a difference of the first LBP histogram image and a predetermined mean LBP histogram image. In an example

embodiment, the predetermined mean LBP image may be computed based on the processing of a set of eye region samples.

Also, at block 410b, a second classification feature associated with the mirror image of the eye region is computed. In an example embodiment, the second classification feature may be computed by first converting the intensity image of the mirror image of the eye region into a second LBP histogram image. In an example embodiment, the second classification feature $X2$ may be computed from the second classification feature. Also, a second difference image may be computed based on a difference of the second LBP histogram image and the predetermined mean LBP histogram image.

In an example embodiment, the dimensions of the $X1$ and $X2$ may be reduced using PCA (Principal Component Analysis). PCA is a multivariate data analysis procedure that involves a transformation of a number of possibly correlated variables into a smaller number of uncorrelated variables known as principal components.

At block 412a, a first state vector associated with a state of the eye region may be computed based on the first classification feature and the plurality of predetermined constants. Also, at block 412b, a second state vector associated with the eye region may be computed based on the second classification feature and the plurality of predetermined constants. For an input x corresponding to a sample of an eye region in an image, the values of the terms A and b may be utilized for calculating the value of the state vector y by first normalizing x and inputting the values of A , b and x in the following equation:

$$y = Ax + b, \text{ wherein}$$

A is PCA_CCA_Vector and
 b is the linear vector

In an example embodiment, the values of the plurality of predetermined constants may be computed based on a correlation between the state of the eye region and a set of image features extracted from a set of eye region samples. In an example embodiment, the correlation CCA. In an example embodiment, for the computation of the plurality of predetermined constants by the CCA, input data X and Y may be derived from the set of sample, such that

$$X = \{x_1, x_2, \dots, x_N\} \text{ and}$$

$$Y = \{y_1, y_2, \dots, y_N\},$$

where,

X is represented by the concatenated block histograms of LBP features, computed around the
5 detected eye regions,

Y is a binary state vector e.g. {blink, non-blink}, denoting the status of the eyes, and

x_n is D_x dimensional, y_n is D_y dimensional, and sample number is N .

In an example embodiment, the values of the plurality of the predetermined constants (A, b) may
10 be computed as per the following expressions:

$$A = W_y^{-T} A' W_z^T P^T,$$

$$b = W_y^{-T} b' - W_y^{-T} A' W_z^T P^T \bar{x}$$

At block 414, it may be determined whether the computed values of the first state vector $Y1$ and
15 the second state vector $Y2$ are greater than a predetermined threshold. In an example
embodiment, the predetermined threshold may assume different values for determining one of the
plurality of blink state of the eye. For example, for the state of eye to be a blink state, the
predetermined threshold may be ($>0.55, <0.55$), and for the state of the eye to be the non-blink
state, the predetermined threshold may be ($<0.55, >0.55$). In an example embodiment, the
20 predetermined threshold may assume different values for determining a degree or extent of blink
of the eye. If it is determined at block 414 that the values of the first and the second state vectors
is greater than the predetermined threshold, a blink state of the detected eye region may be
determined at block 416. If, however, the values of the first and second state vectors are
determined to be less than the predetermined threshold, a non-blink state of the detected eye
25 region may be determined, at block 418.

It will be understood that although the method 400 of FIGURE 4 shows a particular order, the
order need not be limited to the order shown, and more or fewer blocks may be executed,
without providing substantial change to the scope of the present disclosure.

Without in any way limiting the scope, interpretation, or application of the claims appearing below, a technical effect of one or more of the example embodiments disclosed herein is to detect blink in the media content. A blink detector based on the disclosed blink detection algorithm is proposed that utilizes CCA along with LBP histograms for the purpose of blink
5 detection, thereby providing an efficient blink detection with very good detection accuracy. The procedure of training the blink detector is time efficient since only a series of multiplications needs to be performed without any need for comparisons. Moreover, the blink detector facilitates detection of closed eyes during flash conditions on the standalone images, thereby improving the quality of images captured during flash conditions. Also, the memory requirement of the blink
10 detector may be significantly less as only two matrices needs to be stored. In an example embodiment, the eyes may be classified based in the degree of blink.

Various embodiments described above may be implemented in software, hardware, application logic or a combination of software, hardware and application logic. The software, application
15 logic and/or hardware may reside on at least one memory, at least one processor, an apparatus or, a computer program product. In an example embodiment, the application logic, software or an instruction set is maintained on any one of various conventional computer-readable media. In the context of this document, a “computer-readable medium” may be any media or means that can contain, store, communicate, propagate or transport the instructions for use by or in connection
20 with an instruction execution system, apparatus, or device, such as a computer, with one example of an apparatus described and depicted in FIGURES 1 and/or 2. A computer-readable medium may comprise a computer-readable storage medium that may be any media or means that can contain or store the instructions for use by or in connection with an instruction execution system, apparatus, or device, such as a computer.

25 If desired, the different functions discussed herein may be performed in a different order and/or concurrently with each other. Furthermore, if desired, one or more of the above-described functions may be optional or may be combined.

30 Although various aspects of the embodiments are set out in the independent claims, other aspects comprise other combinations of features from the described embodiments and/or the dependent

claims with the features of the independent claims, and not solely the combinations explicitly set out in the claims.

5 It is also noted herein that while the above describes example embodiments, these descriptions should not be viewed in a limiting sense. Rather, there are several variations and modifications which may be made without departing from the scope of the present disclosure as defined in the appended claims.

CLAIMS:

1. A method comprising:

computing a first classification feature associated with an eye region of an image;

5 computing a first state vector associated with the eye region based on the first classification feature and a plurality of predetermined constants; and

determining a state of the eye region based on a comparison of the computed first state vector with a predetermined threshold, the state of the eye region being one of a plurality of blink states associated with the eye region.

10 2. The method as claimed in claim 1 further comprising normalizing the image prior to computing the first classification feature.

15 3. The method as claimed in claim 2, wherein computing the first classification feature comprises:

converting the normalized image into a first linear binary pattern (LBP) histogram image; and

20 computing a first difference image based on a difference of the first LBP histogram image and a predetermined mean LBP histogram image, the predetermined mean LBP image being computed based on the processing of a set of eye region samples.

25 4. The method as claimed in claim 1, wherein the plurality of predetermined constants are determined based on a correlation between the state of the eye region and a set of image features extracted from a set of eye region samples.

5. The method as claimed in claim 4, wherein the correlation is determined by a canonical correlation analysis (CCA).

6. The method as claimed in claim 1, further comprising:

30 normalizing a mirror image of the eye region to generate a normalized mirror image of the eye region;

computing a second classification feature associated with the normalized mirror image;

computing a second state vector associated with the state of the eye region based on the second classification feature and the plurality of predetermined constants; and

determining a state of the eye region based on a comparison of the computed first state vector and the second state vector with the predetermined threshold.

5

7. The method as claimed in claim 6, wherein computing the second classification feature comprises:

converting the normalized mirror image into a second LBP histogram image; and

10 computing a second difference image based on a difference of the second LBP histogram image and a predetermined mean LBP histogram image, the predetermined mean LBP image being computed based on the processing of a set eye region samples.

8. The method as claimed in claim 7, wherein the state of the eye region is the determined based on the comparison of the computed first state vector and the second state
15 vector with the predetermined threshold.

9. An apparatus comprising:

at least one processor; and

20 at least one memory comprising computer program code, the at least one memory and the computer program code configured to, with the at least one processor, cause the apparatus at least to perform:

computing a first classification feature associated with an eye region of an image;

computing a first state vector associated with the eye region based on the first classification feature and a plurality of predetermined constants ; and

25 determining a state of the eye region based on a comparison of the computed first state vector with a predetermined threshold, the state of the eye region being one of a plurality of blink states associated with the eye region.

10. The apparatus as claimed in claim 9, wherein the apparatus is further caused, at
30 least in part, to perform: normalizing of the image prior to computing the first classification feature.

11. The apparatus as claimed in claim 10, wherein the apparatus is further caused, at least in part, to perform computing the first classification feature by:

converting the normalized image into a first linear binary pattern histogram image; and

5 computing a first difference image based on a difference of the first LBP histogram image and a predetermined mean LBP histogram image, the predetermined mean LBP image being computed based on the processing of a set of eye region samples.

12. The apparatus as claimed in claim 9, wherein the apparatus is further caused, at least in part, to perform: determining the plurality of predetermined constants based on a
10 correlation between the state of the eye region and a set of image features extracted from a set of eye region samples.

13. The apparatus as claimed in claim 12, wherein the the apparatus is further caused, at least in part, to perform: determining the correlation by CCA.

14. The apparatus as claimed in claim 9, wherein the apparatus is further caused, at least in part, to further perform:

normalizing a mirror image of the eye region of the image to generate a normalized mirror image of the eye region;

20 computing a second classification feature associated with the normalized mirror image;

computing a second state vector associated with the state of the eye region based the second classification feature and a plurality of predetermined constants; and

determining a state of the eye region based on a comparison of the computed first state vector and the second state vector with the predetermined threshold.

15. The apparatus as claimed in claim 14, wherein the apparatus is further caused, at least in part, to further perform computing the second classification feature by:

converting the normalized mirror image into a second LBP histogram image; and

30 computing a second difference image based on a difference of the second LBP histogram image and a predetermined mean LBP histogram image, the predetermined mean LBP image being computed based on the processing of a set eye region samples.

16. The apparatus as claimed in claim 15, wherein the state of the eye region is determined based on the comparison of the computed first state vector and the second state vector with the predetermined threshold.

5 17. The apparatus as claimed in claim 9, wherein the apparatus comprises a communication device comprising:

a user interface circuitry and user interface software configured to facilitate a user to control at least one function of the communication device through use of a display and further configured to respond to user inputs; and

10 a display circuitry configured to display at least a portion of a user interface of the communication device, the display and display circuitry configured to facilitate the user to control at least one function of the communication device.

15 18. The apparatus as claimed in claim 17, wherein the communication device comprises an image sensor configured to capture images and videos.

19. The apparatus as claimed in claim 17, wherein the communication device comprises a mobile phone.

20 20. A computer program comprising a set of instructions, which, when executed by one or more processors, cause an apparatus at least to perform:

computing a first classification feature associated with an eye region of an image;

computing a first state vector associated with the eye region based on the first classification feature and a plurality of predetermined constants; and

25 determining a state of the eye region based on a comparison of the computed first state vector with a predetermined threshold, the state of the eye region being one of a plurality of blink states associated with the eye region.

30 21. The computer program as claimed in claim 20, wherein the apparatus is further caused, at least in part, to perform: normalizing the image prior to computing the first classification feature.

22. The computer program as claimed in claim 21, wherein the apparatus is further caused, at least in part, to perform: computing the first classification feature by:

converting the normalized image into a first LBP histogram image; and

5 computing a first difference image based on a difference of the first LBP histogram image and a predetermined mean LBP histogram image, the predetermined mean LBP image being computed based on the processing of a set of eye region samples.

23. The computer program as claimed in claim 20, wherein the apparatus is further caused, at least in part, to perform: determining the plurality of predetermined constants based on
10 a correlation between the state of the eye region and a set of image features extracted from a set of eye region samples.

24. The computer program as claimed in claim 23, wherein the the apparatus is further caused, at least in part, to perform: determining the correlation by CCA.

15

25. The computer program as claimed in claim 20, wherein the apparatus is further caused, at least in part, to perform:

normalizing a mirror image of the eye region of the image to generate a normalized mirror image of the eye region;

20 computing a second classification feature associated with the normalized mirror image;

computing a second state vector associated with a state of the eye region based on the second classification feature and a plurality of predetermined constants; and

determining a state of the eye region based on a comparison of the computed first state vector and the second state vector with the predetermined threshold.

25

26. The computer program as claimed in claim 25, wherein the apparatus is further caused, at least in part, to further perform computing the second classification feature by:

converting the normalized mirror image into a second LBP histogram image; and

30 computing a second difference image based on a difference of the second LBP histogram image and a predetermined mean LBP histogram image, the predetermined mean LBP image being computed based on the processing of a set eye region samples.

27. The computer program as claimed in claim 26, wherein the apparatus is further caused, at least in part, to perform: determining the state of the eye region based on the comparison of the computed first state vector and the second state vector with the predetermined threshold.

5

28. The computer program as claimed in any of the previous claims, wherein the computer program is comprised in a computer program product comprising a computer readable medium.

10

29. An apparatus comprising:

means for computing a first classification feature associated with an eye region of an image;

means for computing a first state vector associated with the eye region based on the first classification feature and a plurality of predetermined constants; and

15

means for determining a state of the eye region based on a comparison of the computed first state vector with a predetermined threshold, the state of the eye region being one of a plurality of blink states associated with the eye region.

20

30. The apparatus as claimed in claim 29, wherein the apparatus comprises means for normalizing the image prior to computing the first classification feature.

31. The apparatus as claimed in claim 30, wherein the apparatus comprises means for computing the first classification feature by:

converting the normalized image into a first LBP histogram image; and

25

computing a first difference image based on a difference of the first LBP histogram image and a predetermined mean LBP histogram image, the predetermined mean LBP image being computed based on the processing of a set of eye region samples.

30

32. The apparatus as claimed in claim 29, wherein the apparatus comprises means for determining the values of the plurality of predetermined constants based on a correlation between the state of the eye region and a set of image features extracted from a set of eye region samples.

33. The apparatus as claimed in claim 32, wherein the apparatus comprises means for determining the correlation by CCA.

34. The apparatus as claimed in claim 29, wherein the apparatus comprises:

5 means for normalizing a mirror image of the eye region of the image to generate a normalized mirror image of the eye region;

means for computing a second classification feature associated with the normalized mirror image;

10 means for computing a second state vector associated with the eye region based on a the second classification feature and a plurality of predetermined constants; and

means for determining a state of the eye region based on a comparison of the computed first state vector and the second state vector with the predetermined threshold.

35. The apparatus as claimed in claim 34, wherein the apparatus comprises:

15 means for converting the normalized mirror image into a second LBP histogram image; and

means for computing a second difference image based on a difference of the second LBP histogram image and a predetermined mean LBP histogram image, the predetermined mean LBP image being computed based on the processing of a set eye region samples.

20

36. The apparatus as claimed in claim 35, wherein the apparatus comprises means for determining the state of the eye region as the blink state based on the comparison of the computed first state vector and the second state with the predetermined threshold.

25

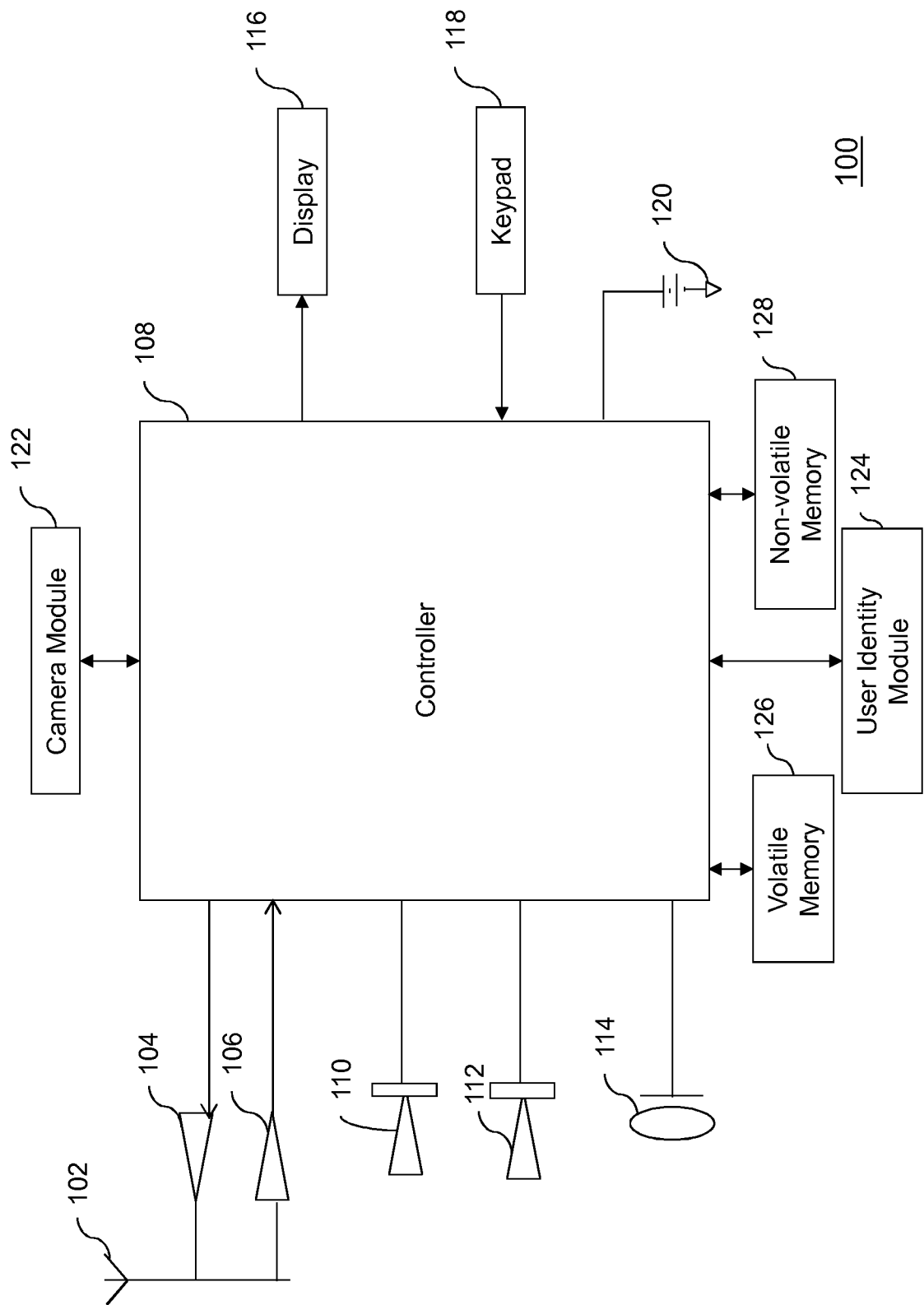


FIGURE 1

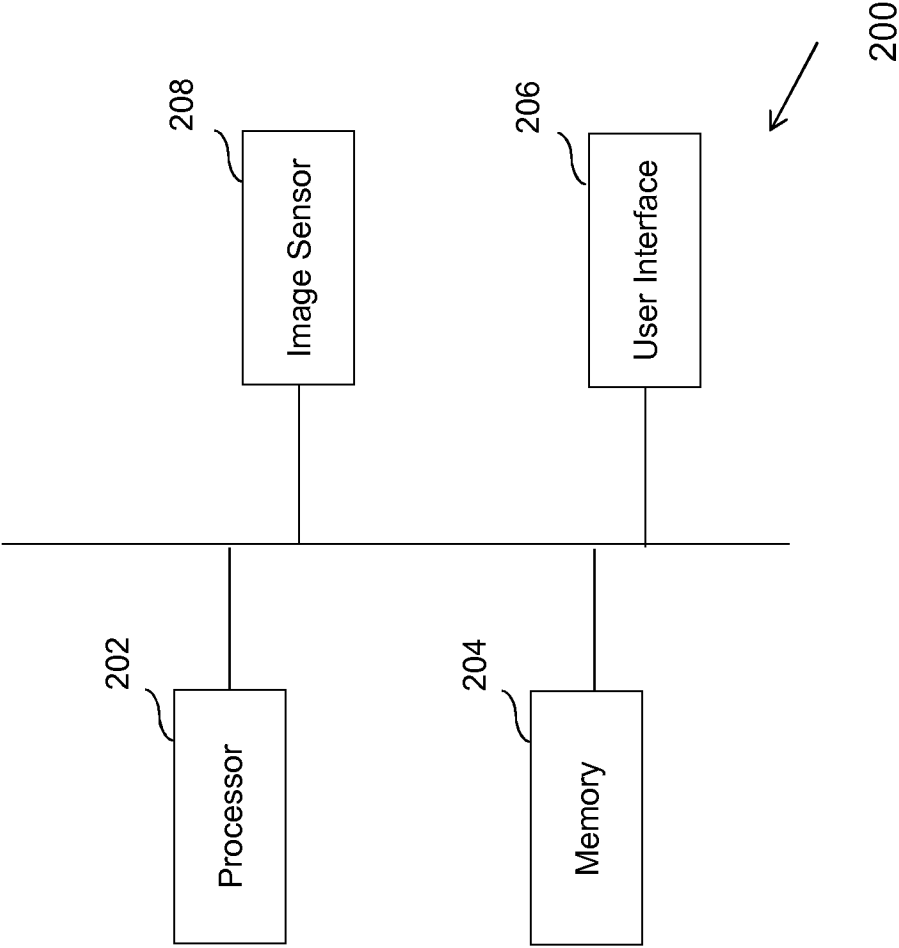


FIGURE 2

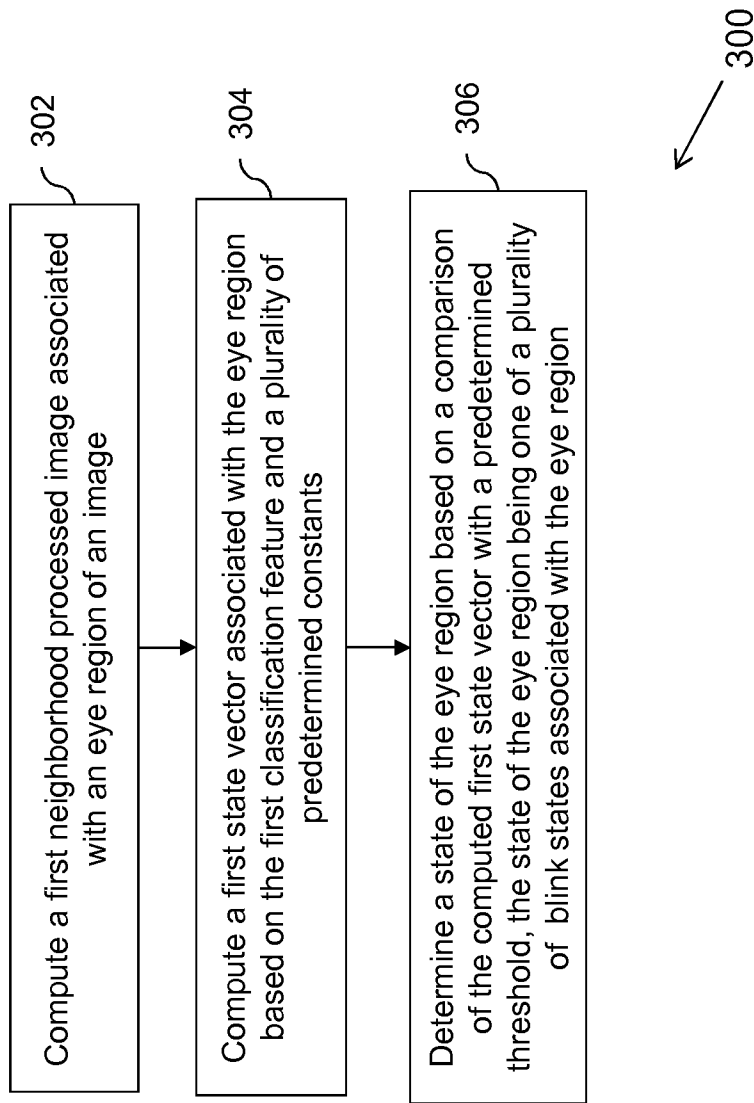


FIGURE 3

4/4

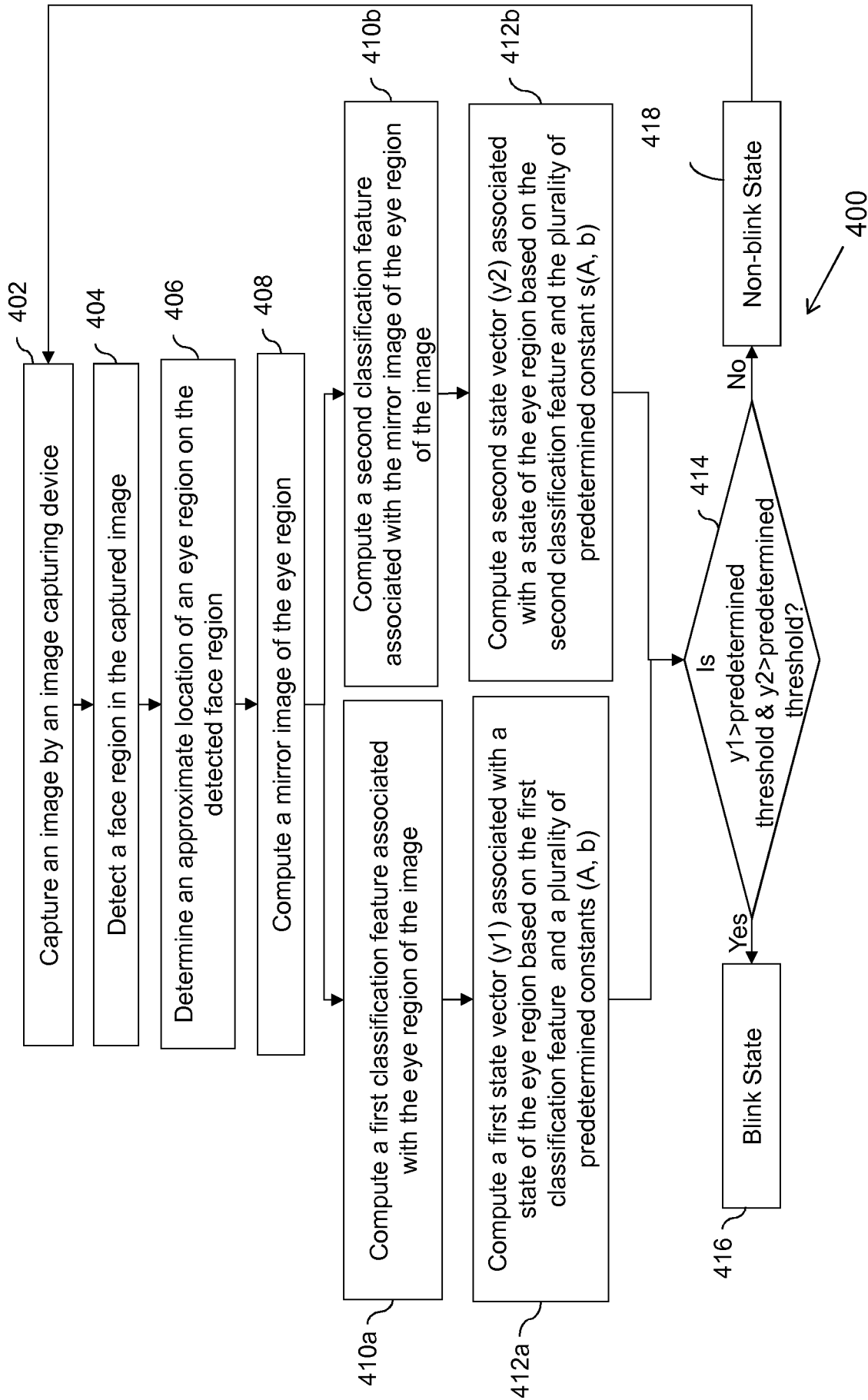


FIGURE 4

INTERNATIONAL SEARCH REPORT

International application No.

PCT/FI2012/050305

A. CLASSIFICATION OF SUBJECT MATTER

See extra sheet

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC: G06K, G06T, G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
FI, SE, NO, DK

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal, WPI, XPI3E, INSPEC, XPESP

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	XU, C. et al. 'Eye States Detection by Boosting Local Binary Pattern Histogram Features', 15th IEEE International Conference on Image Processing (ICIP 2008). October 12, 2008, pp. 1480-1483. The whole document.	1-36
X	XU, C. et al. 'Efficient Eye States Detection in Real-time for Drowsy Driving Monitoring System', IEEE International Conference on Information and Automation (ICIA 2008). June 20, 2008, pp. 170-174. The whole document.	1-36
X	US 2006210121 A1 (NAKANO YUUSUKE et al.) 21 September 2006 (21.09.2006) abstract; Figures 1-15; paragraphs [0038]-[0039], [0041], [0051]-[0053], [0078]-[0084].	1-2, 4-6, 9-10, 12-14, 17-21, 23-25, 28-30, 32-34

☒ Further documents are listed in the continuation of Box C.
 ☒ See patent family annex.

* Special categories of cited documents:	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be of particular relevance	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E" earlier application or patent but published on or after the international filing date	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&" document member of the same patent family
"O" document referring to an oral disclosure, use, exhibition or other means	
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search 17 August 2012 (17.08.2012)	Date of mailing of the international search report 22 August 2012 (22.08.2012)
Name and mailing address of the ISA/FI National Board of Patents and Registration of Finland P.O. Box 1160, FI-00101 HELSINKI, Finland Facsimile No. +358 9 6939 5328	Authorized officer Mika Inki Telephone No. +358 9 6939 500

INTERNATIONAL SEARCH REPORT

International application No.

PCT/FI2012/050305

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	KR 100880256B B1 (DAOU XILICON TECHNOLOGY CO LTD et al.) 28 January 2009 (28.01.2009) Figures 3a-3b, 5a-5b, 6 & abstract [online] EPOQUENET EPODOC & WPI & machine translation into English by Thomson Reuters [online] [retrieved 17.8.2012] paragraphs [0045], [0070]-[0078], [0101]-[0102].	
P, X	ULIYAR, M. et al. 'A Fast Blink Detector Using Canonical Correlation Analysis', IEEE International Conference on Consumer Electronics (ICCE 2012). January 13, 2012, pp. 33-34. The whole document.	1-36
P, X	RAHMAN, M. et al. 'A Novel Haptic Jacket Based Alerting Scheme in a Driver Fatigue Monitoring System', IEEE International Workshop on Haptic Audio Visual Environments and Games (HAVE 2011). October 14, 2011, pp. 112-117. The whole document.	1-36

INTERNATIONAL SEARCH REPORT

International application No.
PCT/FI2012/050305

CLASSIFICATION OF SUBJECT MATTER

Int.Cl.

G06K 9/00 (2006.01)

G06K 9/66 (2006.01)

G06T 7/00 (2006.01)

G06F 17/15 (2006.01)

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/FI2012/050305

Patent document cited in search report	Publication date	Patent family members(s)	Publication date
US 2006210121 A1	21/09/2006	JP 2006260397 A	28/09/2006
.....			
KR 100880256B B1	28/01/2009	None	
.....			