(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
17 February 2005 (17.02.2005)

PCT

(10) International Publication Number
WO 2005/015896 A1

(51) International Patent Classification[7]**:** **H04N 1/62**

(21) International Application Number: PCT/EP2004/008706

(22) International Filing Date: 3 August 2004 (03.08.2004)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
10/635,862 5 August 2003 (05.08.2003) US

(71) Applicant *(for all designated States except US)*: **AC-CAPELLA VISION LIMITED** [IE/IE]; Galway Business Park, Dangan, Galway (IE).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **DELUCA, Michael, J.** [US/US]; 734 Camino Gardens Lane, Boca Raton, FL 33432 (US). **PRILUTSKY, Yury** [US/US]; 1626 Shoal Drive, San Mateo, CA 94404 (US). **STEINBERG, Eran** [IL/US]; 137 Granville Way, San Francisco, CA 94127 (US). **CORCORAN, Peter** [IE/IE]; Cregg, Claregalway, Galway (IE). **POSOSIN, Alexei** [MD/IE]; 28 Arus na Tra, Cois Cuain, Dr. Colohan Road, Galway (IE).

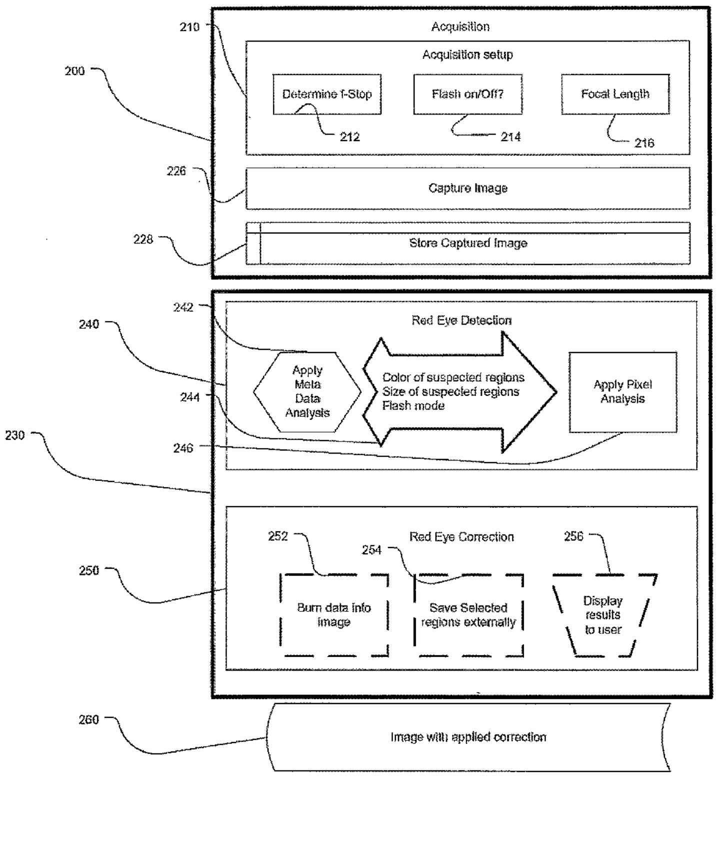
(74) Agents: BOYCE, Conor et al.; F.R. Kelly & Co., 27 Clyde Road, Ballsbridge, Dublin 4 (IE).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, 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, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),

[Continued on next page]

(54) Title: A METHOD AND SYSTEM OF FILTERING A RED-EYE PHENOMENON FROM A DIGITAL IMAGE



(57) Abstract: A method of filtering a red-eye phenomenon from an acquired digital image comprises using both anthropometric data and meta-data associated with the image to identify regions of the image potentially susceptible to red-eye artifacts. In an embodiment the image is acquired by focusing an external object onto a sensor array through a lens, and wherein the regions are identified as a function of one or more of the focal length of the lens, the size of the array, the distance to the object and the depth of field at the moment of acquisition.

WO 2005/015896 A1

European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, 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 AND SYSTEM OF FILTERING A RED-EYE PHENOMENON FROM A DIGITAL IMAGE

BACKGROUND

1. Field of the Invention

The present invention relates generally to digital photography using flash, and specifically to filtering “Red Eye” artifacts from digital images shot by digital cameras, scanned by a digital scanner, or otherwise acquired by a digital image acquisition device.

2. Description of the Related Art

i. Red Eye Phenomenon

"Red-eye" is a phenomenon in flash photography where a flash is reflected within a subject's eye and appears in a photograph as a red dot where the black pupil of the subject's eye would normally appear. The unnatural glowing red of an eye is due to internal reflections from the vascular membrane behind the retina, which is rich in blood vessels. This objectionable phenomenon is well understood to be caused in part by a small angle between the flash of the camera and the lens of the camera. This angle has decreased with the miniaturization of cameras with integral flash capabilities. Additional contributors include the relative closeness of the subject to the camera, iris color where light eyes are more susceptible to this artifact and low ambient light levels which means the pupils are dilated.

The red-eye phenomenon can be somewhat minimized by causing the iris to reduce the opening of the pupil. This is typically done with a "pre-flash", a flash or illumination of light shortly before a flash photograph is taken or a strong additional light source. This causes the iris to close. Unfortunately, these techniques typically delay the photographic exposure process by 0.5 second or more to allow for the pupil to contract. Such delay may cause the user to move, the subject to turn away, etc. Therefore, these techniques, although somewhat useful in removing the red-eye artifact, can cause new unwanted results.

ii. Digital Cameras and Red Eye Artifacts

Digital cameras are becoming more popular and smaller in size, and produce digital images each comprising a plurality of pixels indicative of colour, the pixels forming various shapes within the image. Digital cameras have several advantages over film cameras, e.g. eliminating the need for film as the image is digitally captured and stored in a memory array for display on a display screen on the camera itself. This allows photographs to be viewed and enjoyed virtually instantaneously as opposed to waiting for film processing. Furthermore, the digitally captured image may be downloaded to another display device such as a personal computer or color printer for further enhanced viewing. Digital cameras include microprocessors for image processing and compression and camera systems control. Nevertheless, without a pre-flash, both digital and film cameras can capture the red-eye phenomenon as the flash reflects within a subject's eye. Thus, what is desired is a method of eliminating red-eye phenomenon within a miniature digital camera having a flash without the distraction of a pre-flash.

An advantage of digital capture devices is that the image file contains more data than pure image (pixel) data. Such additional data is also referred to as meta-data and is usually saved in the header of the digital file containing the image. The meta-data may include (a) information about the camera or other acquisition device which is independent of the particular image being acquired at any given time (device-specific meta-data) and (b) the acquisition parameters for a particular image being acquired (acquisition-specific meta-data), i.e. information relating to those parameters that are adjustable or that may change from exposure to exposure, based on user input or otherwise.

iii. Digital Scanning and Red Eye Artifacts

In many cases images that originate from analog devices like film are scanned to create a digital image. The scanning can be either for the purpose of digitization of film-based images into digital form, or as an intermediate step as part of the printing of film based images on a digital system. Red Eye phenomenon is a well known problem even for film cameras, and in particular point and shoot cameras where the proximity of the flash and the lens is accentuated.

When an image is scanned from film, the scanner may have the option to adjust its scanning parameters in order to accommodate for exposure and color balance. In addition, for negative film, the scanner software will reverse the colors as well as remove the orange, film base mask of the negative.

The meta-data for film images is generally more limited than for digital cameras. However, most films include information about the manufacturer, the film type and even the batch number of the emulsion. Such information can be useful in evaluating the raw, uncorrected color of eyes suffering from red eye artifacts.

iv. Anthropometry

Anthropometry is defined as the study of human body measurement for use in anthropological classification and comparison. Such data, albeit extremely statistical in nature, can provide good indication as to whether an object is an eye, based on analysis of other detected human objects in the image.

v. Red-Eye Detection And Correction Algorithms

Red-eye detection algorithms typically include detecting the pupil and detecting the eye. Both of these operations may be performed in order to determine if red-eye data is red-eye or if an eye has red-eye artifact in it. The success of a red eye detection algorithm is generally dependent on the success of a correct positive detection and a minimal false detection of the two.

In the prior art red-eye detection is usually done primarily on image (pixel) data, as described, for example, in U.S. Patent 6,407,777 (DeLuca). Although DeLuca does use some meta-data, it is of an elementary nature and solely acquisition-specific (flash on/off, ambient lighting level, camera-subject distance). It is derived solely to determine whether the image acquisition conditions were conducive to red-eye and hence whether in fact to use red-eye correction at all.

US 2003/0095197 (Eastman Kodak) discloses the use of meta-data in the correction of several image defects, including red-eye. However, like DeLuca, Eastman Kodak only uses the meta-data to make a decision as to whether red-eye is likely to have occurred, but does not use such data in the correction process itself.

Meta-data has also been used in the photo-finishing industry, where a digital image may be post-processed to optimize the output from a printing system. Examples of this use of meta-data are provided in U.S. Patents 6,505,003 6,501,911 and 6,496,655. Meta-data may also be recorded onto a standard camera film and the meta-data may be subsequently recovered to assist in the post-processing of the film, US Patent 6,429,924.

SUMMARY OF THE INVENTION

The present invention provides a method of filtering a red-eye phenomenon from a digital image, the method comprising using both anthropometric data and meta-data associated with the image to identify regions of the image potentially susceptible to red-eye artifacts.

In the preferred embodiment the image is created by optically projecting an external subject onto a sensor array through a lens, and properties of said regions are identified as a function of one or more of the focal length of the lens, the aperture of the lens, the size of the array, the resolution of the array, the distance to the subject and the depth of field at the moment of acquisition.

The invention also provides a system of filtering a red-eye phenomenon from a digital image, the system comprising means using both anthropometric data and meta-data associated with the image to identify regions of the image potentially susceptible to red-eye artifacts.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows a block diagram of a digital image acquisition device operating in accordance with a preferred embodiment.

Figure 2 illustrates a high level workflow of a method of detecting red eye artifacts in digital images in accordance with a preferred embodiment.

Figures 3a-3d schematically depicts a light sensor, and the formation of a digital pixelated image on it, for use in explaining the preferred embodiment.

Figure 4 describes a high level process of using meta-data as part of red-eye detection in accordance with a preferred embodiment.

Figure 5 illustrates by means of geometric optics, a relationship between an object and an image based on a distance to the object and the focal length, where the focal length is the distance from the image principal plane of the optical system to the image focal plane, which is the plane where the image of the object situated at infinity is formed.

Figure 6 illustrates a relationship between focal length of a lens and depth of field, and an object size as it appears on an image.

Figures 7a-7c illustrate some anthropometric measurements of a human face for an adult male and female.

Figure 8 is a workflow diagram describing a statistical analysis of an image using meta-data and anthropometric data in accordance with a preferred embodiment.

Figure 9 depicts a spectral response of an acquisition system based on spectral sensitivity curves of a hypothetical three color sensor, the spectral distribution of a generic light source and the spectral characteristics of a object being photographed.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

The methods of the preferred embodiments are generally applicable to digital image acquisition devices, such as digital cameras and scanners, and to and output devices such as printers and electronic storage devices. When the terms digital camera and output device or printer are used, it is generally meant to more broadly, respectively include digital image acquisition devices and digital data output devices.

The digital camera or other acquisition device preferably has the capability of analyzing and processing images. Alternatively, the processing of the images can be done outside of the camera on a general purpose or specialized computer after downloading the images or on a device that is acting as a hosting platform for the digital camera. Such a device may be, but is not limited to, a hand held PC, a print server, a printer with built in processing capability, or cell phone equipped with a digital camera. Alternatively the acquisition process can be of an analog image, such as scanning of a film based negative or reversal film, or scanning of a photographic print.

The accuracy of a detection process may be measured by two parameters. The former is the correct detection, which relates to the percentage of objects correctly detected. The second parameter for evaluating successful detection is the amount of mis-classifications, which is also defined as false detections or beta-error. False detections relate to the objects falsely determined to have the specific characteristics, which they do not possess.

Overall, the goal of a successful detection process is to improve the accuracy of correct detections while minimizing the percentage of false detections. In many cases there is a tradeoff between the two. When the search criterion is relaxed, more images are detected but at the same time, more false detections are typically introduced, and vice versa.

In order to improve the accuracy of the red eye detection and correction, the preferred embodiment uses anthropometric information about the subject and meta-data, the latter comprising device-specific meta-data and acquisition-specific meta-data as previously described. In the case of a digital camera, device-specific meta-data may include color sensitivity, spectral response or size of the camera sensor, whether the sensor is CCD or CMOS, tone reproduction transformations of the image and color transformations from the RAW data gathered by the sensor, e.g. CCD, to a known color space such as RGB, the f-stop, or other camera-specific parameters understood by those skilled in the art, and combinations thereof. In the case of scanning device-specific meta-data may include the color sensitivity curve of the film, the color sensitivity of the scanner sensor, whether CCD or CMOS, whether linear or area sensors, the color transformations from the RAW data gathered by the scanner to a known color space such as RGB. Acquisition meta-data may include focal distance as determined by an auto focus mechanism of a digital camera, the power of the flash including whether a flash was used at all, the focal length of the lens at the moment of acquisition, the size of the CCD, the depth of field or the lens aperture, exposure duration, or other acquisition parameters understood by those skilled in the art, and combinations thereof. Anthropometric data may include first and higher order statistics, which is an average and a variability of an expected size and ratio between different parts of the human body, and particularly the facial region.

Using the aforementioned information, preferred embodiments described herein achieve a more accurate detection of eye regions potentially containing red eye artifacts. Based on this detection, a processor, whether in a camera or in a different device, can perform a correction step.

Referring to Figure 1, block 100 represents an image acquisition device which can be a digital camera in various packaging such as a digital still camera, a lens connected to a hand held computer, a cell phone with image capturing capability, a video camera with still image capturing capability, etc.

The image capture apparatus 100 comprises a light sensor 102 that can be a CCD, CMOS or other sensor array that transforms a light image into electronic form. Most cameras are equipped with a built-in flash 104, also referred to as a strobe. In many cases, the camera strobe is physically close to the lens, which tends to accentuate the occurrence and strength of the red eye artifact. In addition, the camera is equipped with a lens 106. The relevant parameters of the lens during acquisition include the aperture 114, or a f-stop, which primarily determines the depth of field, the focal length 112 which determines the enlargement of the image, and the focusing distance 116 which determines the distance to the object that the lens 106 was focused at.

Block 130 of Figure 1 represents the red eye filter that performs a process of detection and correction of red eye artifacts. The process can be done in the camera as part of the acquisition stage, in the camera at a post processing stage, during the transferring of the image from the camera to an external device such as a personal computer, or on the external device as a post processing stage, such as in image transfer software or image editing software.

The red eye filter includes two main stages. Block 132 represents a meta-data analysis module 132, where the image is evaluated based on meta-data and anthropometric data (Fig. 8). Block 138 represents a pixel-based analysis of the image data. The pixel-based analysis 138 receives information from the meta-data stage 132. Therefore, the decision on the pixel level may vary based on the conditions under which the image was captured and/or other meta-data. Block 160 represents an image storage component 160 that saves the image after the red eye correction operation.

Figure 2 is a workflow representation corresponding to the preferred camera embodiment illustrated at Figure 1. The image capture stage is represented by block 200. This operation includes the pre-acquisition setup 210, where the user and/or the camera determine preferred settings such as f-stop 212, flash on/off 214 and/or focal length 216. The image capture stage 200 also includes acquisition or picture taking 226 and temporary storage in block 228 in its final form or in a RAW form that corresponds to the image as captured by the light sensor 102 of Figure 1. As part of the capture process, the camera determines the best acquisition parameters in the pre-acquisition stage 210. Such parameters may include the right exposure, including gain, white balance and color transformation, and in particular aperture settings 212 and whether to use flash 214. In addition, the user may decide on the focal length 216 of the lens 106, which is also be referred to as the zoom position.

The image after being stored in block 228 is then processed for red eye, 230, in accordance with a preferred embodiment, among other stages of processing that may include color corrections, compression, sharpening, etc. The red-eye filter 230 is preferably performed by software in or outside the camera or other image acquisition device. The red eye filter preferably includes two main operations - red eye detection 240 and red eye correction 250.

The red eye detection 240 includes a first stage of analyzing the meta-data 242, a stage of transferring the resultant data 244, and the specific red eye detection 246 based on pixel analysis.

The red eye correction is illustrated at Figure 2 as the operation 250 where any image modifications based on the results of the detection stage 240 are applied to the image. At this stage 250, correction may be burned into the data 252, thus replacing the damaged pixels, saved as a list of the pixels that need to be changed with their new value in the header of the image or externally 254, and/or presented to the user 256, requesting the user to take an action in order to apply the corrections, or a combination of these operations. The image, with the corrections applied as described in 240, is then preferably saved in block 260.

Figures 3a-3d illustrates in detail the image as created on the light sensor 102 of Figure 1, which is located at the image plane of the optical system. Such sensor can be any electro-photosensitive device such as CCD or CMOS.

Figure 3a illustrates a grid type CCD. Each one of the smaller squares (as illustrated by block 302) is a cell, which is sensitive to light. The CCD size 304 is calculated as the diagonal of the rectangle made of width 306 and height 308.

Figure 3b illustrates how a face may be projected onto the CCD. Figure 3c illustrates how the image is pixelized, where the continuous image is transformed into a grid-based image.

Figure 3d is more specific to the image as created by a human eye. The image of the eye will include the iris 342 as well as the pupil 344, which is usually the locations where red-eye artifacts occur. The white part 346 of the eye is also a component of the human eye illustrated at Figure 3d and which can be used in red-eye detection, particularly false-detection avoidance.

Figure 4 illustrates various meta-data information that can be used as part of a preferred embodiment as input, and the potential outcome of such data analysis. For example, blocks 412, 422 and 432 illustrate an operation of red-eye detection relating to the use or non-use of flash. The information whether the flash is used or not, block 412, is forwarded at operation 422 to red-eye pre-processing 432 to determine whether there is reason to launch the red-eye filter. If a flash, as determined in 412, is not used, there is preferably no reason to apply the red-eye filter. This is a reasonable estimation for consumer lever cameras where most of the red eye is created, as described in the introduction, by the small disparity between the strobe unit and the lens.

Blocks 414, 424, 434 describe the application of meta-data including the distance to the object, the aperture, CCD size, focal length of the lens and the depth of field. This data is usually recorded on or with the image at acquisition. Based on this information, as transferred to the filter at operation 424, the filter can determine at operation 434, e.g., a range of potential sizes of eye regions in which red-eye can occur.

Blocks 416, 426, 436 relate to information that is specific to the camera. The color composition of the image is determined by a few parameters which include the CCD response curves as illustrated in Figure 9 (see below), the potential color transformations from the recorded, RAW image data such as color correction, gain adjustment and white balance to a known color space such as RGB or YCC and tone reproduction transformations of the image. Such transformations can be presented in the form of lookup tables, transformation matrices, color profiles, etc.

Based on the knowledge of the transfer from operation 426, the software can better determine a more precise range of colors at operation 436 that are good candidates for the red eye artifacts. This information can advantageously narrow down the potential red eye regions based on the variability of sensors and color correction algorithms. It may also help to eliminate colors that, without this knowledge, could be falsely identified as potential red eye region candidates, but are not such in case of a specific combination of sensor and color transformation.

Figure 5 depicts illustrative information that can be gathered to determine the relative size of the object. The ratio of the image size divided by image distance, and the object size divided by the object distance, are approximately equal, wherein the image size divided by the object size is defined as the magnification of the lens 106. If one knows three out of the four values, namely focal length 112, distance to object 116, and object size 516, one can estimate the size of the image (512):

$$\frac{\text{Object size (516)}}{\text{distance to object (116)}} = \frac{\text{image size (512)}}{\text{focal length (112)}}$$

However, the parameter values described above are usually not known precisely. Instead, distributions of values can be estimated.

Figure 6 illustrates the variability generated by the depth of field. Depth of field is defined as the range of distances from the camera to the objects where the images of the objects are captured sufficiently sharp. For a fixed length lens, the depth of field is a function of the aperture. The more open the aperture is, the shallower the depth of field is.

As can be seen in Figure 6, due to the fact that the depth of field can be rather large, the distance to the objects still in focus can vary. Therefore the parameter

$$Dis\tan ce\_to\_Subject$$

is rather a range:

$$Dis\tan ce\_to\_Subject_{Close\_range} \leq Subject \leq Dis\tan ce\_to\_Subject_{Far\_range}$$

The reason why this information is important and has to be taken into consideration is depicted in Figure 6. In this case, two objects, a tree 614 and a house 624 are located in close distance 616, and further away 626 respectively. Even though the tree, 614 and the house 634 are the same size, the sizes of the objects or the projections of the objects on the image plane are different and the tree image, 636 being closer to the camera appears much larger than the house 646.

Figure 7 includes some relevant anthropometrical values for male and female averages. Figure 7a is an average male and figure 7b is an average adult female. For example, for adult male, 700, the distance between the eyes, 714, is on average 2.36", the distance between the eyes and the nostrils, 724, is 1.5" the width of the head, 712 is 6.1" etc.

However, this is only the first order approximation. There is a second order approximation, which is the overall variability of the values. Such variability once again needs to be calculated into the formula.

Or:

$$Subject\_Size_{Small} \le Subject\_Size \le Subject\_Size_{Large}$$

The object size, in order to be considered as a candidate for being a face, and eye or any known object will be:

$$\boxed{\frac{Subject\_SizeSmall * Focal\_Length}{Distance\_To\_Object_{Far\_Range}} \le Object\_Size \le \frac{Subject\_Size_{large} * Focal\_Length}{Distance\_To\_Object_{Close\_Range}}}$$

Specifically, as seen in Figure 7c, the average size of an eyeball, 770, is roughly 1", or 24mm, and the average size of the iris, 772, is half in diameter to the full eye, or 0.5" or 12mm in diameter. The pupil, 774 can be as small as a few millimeters, and dilated to as large as the size of the iris. Fortunately, in the case of red-eye artifacts, which happen primarily in low lighting conditions that required a flash, the pupil will be on the dilated side.

The variability in this case is not only for different individuals, but also variability based on age. Luckily, in the case of eyes, the size of the eye is relatively constant as the person grows from a baby into an adult, this is the reason of the striking effect of "big eyes" that is seen in babies and young children. The average infant's eyeball measures approximately 19½ millimeters from front to back, and as described above, grows to 24 millimeters on average during the person's lifetime. Based on this data, in case of eye detection, the size of the object which is the pupil which is part of the iris, is limited, when allowing some variability to be:

$$9mm \leq Size\_Of\_Iris \leq 13mm$$

The object size as calculated above is going to be in actual physical size such as millimeters or inches. For this invention to become useful, this information needs to be presented measured in pixel sizes.

Returning to Figure 3a, the size of the sensor is depicted by 304, which is the diagonal of the sensor. Based on that, and the ratio between the width, 306 and the height, 308, the width and height can be calculated as a Pythagorean triangle.

$$Sensor\_Diagonal\_Size = \sqrt{width^2 + Height^2}$$

Knowing the sensor resolution, the size of object can now be translated into pixel size.
For example, given a ½ inch (12mm) CCD with an aspect ratio of 2:3 and a 2,000 x 3,000 CCD resolution, the width of the CCD is:

$$12mm = \sqrt{(2\alpha)^2 + (3\alpha)^2} = \sqrt{13}\alpha$$
$$\therefore$$
$$3\alpha = 3x12/\sqrt{13} \approx 3x3.3 \approx 10mm$$

and therefore, for a 3000 pixel width, a 1mm object size is equal to roughly 300 pixels.
Or

$$Image\_Size_{in\_pixels} = Image\_Size_{in\_millimeters}$$

Based on this formula, when an image of an object is detected, the image size in pixels is compared to the range allowed for an object of the kind sought for, and accordingly a decision is made as to whether the image is that of the sought object or not. Thus, for example, using meta-data and anthropometric data as well as image data, as exemplified by blocks 414, 424 and 434 of Figure 4, a human eye can be identified with a fair degree of certainty, and certainly with more certainty than a solely image-based analysis as performed by DeLuca.

An example is depicted in Figure 3d where a hypothetical eye is displayed in pixels, and in this case the iris 342 is roughly 11 pixels in diameter and the pupil 344 is roughly 6 pixels in diameter. With the added knowledge of the distance to the object and the focal length of the lens, the above calculations, performed in software, provide a decision process capable of rejecting objects that are not eyes and selecting most likely candidates to be an eye based on the sizes of the captured images of the objects.

Figure 8 is a software workflow diagram describing a statistical analysis of an image using meta-data and anthropometric data in accordance with a preferred embodiment. The input is the meta-data 414, as described in Figure 4, and human anthropometric data 800, as depicted in Figures 7a and 7b.

In step 810 the size distribution of the objects of interest, in the present case human eyes, is calculated as a function of the data 414 and 800. This calculation is described above and yields a probable range of object sizes which, for the given meta-data and anthropometric data, could be human eyes (irises and/or pupils). Note that this calculation can be done on the fly or alternatively pre-calculated values can be stored in a database to speed up the processing.

When looking for eyes in an image, given regions suspected as eyes 820, as determined by prior art methods which examine only the image (pixel) data, step 830 checks whether each region falls within the range of sizes calculated above in step 810. If the region size is too large or too small, i.e. it falls outside the probable range, step 860 decreases a probability index that the region is an eye. On the other hand, if the region size is within the range, step 850 increases the probability index. It will be understood, however, that these decisions are probabilistic results and not necessarily conclusive.

This preferred embodiment therefore applies additional anthropometric tests to refine the decision by looking for additional clues such as the existence of a second eye 832, searching for surrounding facial features 834 such as the overall shape of the face, the hair, neck etc., searching for the existence of lips in proximity to the eyes 836, searching for nostrils 838, etc.

In each step, the question asked is whether the searched feature was found, step 840. If the reply is positive, then the probability index for the region being an eye is raised, 850, and if negative, the probability index is reduced, 860. Of course, this probabilistic approach can be useful to create a better set of criteria in deciding whether the detected object is what the system is looking for. In more detail, the detection process involves two types of allowed errors also known as Type-I and Type-II errors, or also referred to as α-error, which is the acceptable probability of making a wrong decision, or a false positive and β-error, which is the acceptable probability of not detecting at all. Based on this approach, the probability index as decreased or increased in steps 850 and 860 are always compared against the two criteria α and β.

Finally, when all tests are complete for a given region, if the probability index is above some predetermined threshold, which may be empirically derived for a particular make or type of digital image acquisition device, the object is taken to be an eye.

Now, the image data is subject to pixel analysis 246, Figure 2. In this, the pixels in those regions which are taken to be eyes, and more particularly the pupil and/or iris, are examined to see whether their colour corresponds to red-eye. Techniques for testing pixels for a particular colour, or range of colors, indicative of red-eye are well known, as described for example in DeLuca. However, by using the device-specific meta-data 416, Figure 4, the range of colours indicative of red eye in an image can be more accurately determined than in the prior art.

This is illustrated in Figure 9 whose graph describes the relative response 900 as a function of the visual wavelength 910, of the three sensors for blue, 932, Green 934, and Red 936, of a typical CCD type sensor. Similar graph, although with different response curve describes the response of the different layers for photographic film.

The x-axis, which is the wavelength range of the human visual system, is expanded to include infrared and ultraviolet, which may not be visible to the human eye but may record on a sensor. The y-axis is depicted in relative value as opposed to an absolute one. The three Red, Green, and Blue spectral response functions as functions of the wavelength are defined respectively as: $R(\lambda), G(\lambda), B(\lambda)$

Given a light source defined as a spectral response curve $L(\lambda)$, , the light source when reaching the three different color sensors, or color pigments on film, will generate a response for each of the colors as defined mathematically as the integral of the scalar multiplication of the curves. The range of integration is from the low wavelength region UV to the highest IR.

$$R = \int_{\lambda-UV}^{\lambda-IR} R_\lambda \times L_\lambda d\lambda \ , \ G = \int_{\lambda-UV}^{\lambda-IR} G_\lambda \times L_\lambda d\lambda \quad B = \int_{\lambda-UV}^{\lambda-IR} B_\lambda \times L_\lambda d\lambda$$

to create a tristimulus value of $\{R, G, B\}$

Those skilled in the art are familiar with the fact that different spectral responses may create the same tristimulus values due to the scalar reduction from a 2 dimensional representation to a single value. This effect is also known as Metamerizm which can be a property of the sensor's/film's metamerizm, the human visual system metamerizm, or the light source's metamerizm.

Due to the many variable parameters, it is relatively hard to find a specific color that can be a fixed-reference-point in an image. The reason is that the reflected colors are usually dependent on many factors and especially on the ambient light. However, Red Eye artifacts, as previously explained, are results of the reflection of the strobe light, which has very well defined characteristics, from the vascular membrane behind the retina, which is rich in blood vessels. In most cases, the effect of the external ambient light is relatively low, and the red-eye effect can be considered as a self-illuminating object, with more precise spectral characteristics than other objects. An example of such spectral response, which is a combination, of the flash spectral response, which is relatively broad and the blood vessels inside the eye, is depicted in block 950.

Given the spectral sensitivity of the sensor $R(\lambda), G(\lambda), B(\lambda)$ and the reflection of the flash light in the eye, as defined by 950, $E(\lambda)$, the red eye tristimulus values for this specific sensor are:

$$\boxed{\{R, G, B\}_{red-eye} = \int_{\lambda-UV}^{\lambda-IR} \{R, G, B\}_\lambda \times L_\lambda d\lambda}$$

This value of $\{R,G,B\}_{red-eye}$ is relatively constant for a given camera. However, due to the difference in the response between different sensors, these values are not constant across different cameras. However, with the knowledge of the response curves above, one can determine a much closer approximation of the range or red colors based on this information. Note that it is not only the value of the Red that may help in such determination, but also the residual response of the red eye on the Green and even less the blue sensor. One skilled in the art knows that most cameras perform additional transformations for exposure and tone reproduction for images before saving them into persistent storage. An example of such transformation will be a concatenation of color correction and tone reproduction as a function of the pixel value:

Given a Raw pixel value of:

$$\{R,G,B\}_{RAW-CCD}$$

as transformed via three lookup tables. For example for red lookup table:

$$R-LUT(Raw-Pix):\{input\_values\}\rightarrow\{output\_values\}$$

For example the Red lookup table R-Lut can be a gamma function from 10 bit raw data to 8 bits as follows:

$$R_{LUT}(Raw-Pix):\{0..1024\}\rightarrow\{0..256\}$$

$$R_{LUT}(x)=(R_{RAW-CCD}/1024)^{2.2}*256$$

and the inverse function

$$R^{-1}{}_{LUT}(x)=(R_{LUT\_RAW}/256)^{1/2.2}*1024$$

The {R,G,B} values after transformed through the lookup table will be:

$$\{R,G,B\}_{LUT\_RAW} = \{R_{LUT}(R_{RAW\text{-}CCD}), G_{LUT}(G_{RAW\text{-}CCD}), B_{LUT}(B_{RAW\text{-}CCD})\}$$

$$\{R,G,B\}_{new} = \{R,G,B)_{LUT\_RAW} \times \begin{bmatrix} RR & RG & RB \\ GR & GG & GB \\ BR & BG & BB \end{bmatrix}$$

With the internal knowledge of these transformations, one can reverse the process, to reach the RAW values as defined above.

$$\{R,G,B\}_{LUT\_RAW} = \begin{bmatrix} RR & RG & RB \\ GR & GG & GB \\ BR & BG & BB \end{bmatrix}^{-1} \times \{R,G,B\}^{T}{}_{NEW}$$

and

$$\{R,G,B\}_{RAW} = \{R^{-1}{}_{LUT}(R_{LUT\_RAW}), G^{-1}{}_{LUT}(G_{lut\_raw}), B^{-1}{}_{LUT}(B_{LUT\_RAW})\}$$

and the value of the raw tristimulus values can be then determined and used for the exact matching. Similar transformations are performed by digital scanners in order to correct for sub optimal images such as underexposure, or wrong ambient light. Reversing the process may be difficult in its pure mathematical sense e.g. the conversion function may through the transformation not be fully reversible. Such issues occur for example when the pixel values are clipped or condensed. In such cases, there is a need to define a numerical approximation to the inverse function.

The preferred embodiments described herein may involve expanded digital acquisition technology that inherently involves digital cameras, but that may be integrated with other devices such as cell-phones equipped with an acquisition component, toy cameras etc. The digital camera or other image acquisition device of the preferred embodiment has the capability to record not only image data, but also additional data referred to as meta-data. The file header of an image file, such as JPEG, TIFF, JPEG-2000, etc., may include capture information such as whether a flash was used, the distance as recorded by the auto-focus mechanism, the focal length of the lens, the sensor resolution, the shutter and the aperture. The preferred embodiments described herein serve to improve the detection of red eyes in images, while eliminating or reducing the occurrence of false positives, and to improve the correction of the detected artifacts.

While an exemplary drawing and specific embodiments of the present invention have been described and illustrated, it is to be understood that that the scope of the present invention is not to be limited to the particular embodiments discussed. Thus, the embodiments shall be regarded as illustrative rather than restrictive, and it should be understood that variations may be made in those embodiments by workers skilled in the arts without departing from the scope of the present invention, as set forth in the claims below and structural and functional equivalents thereof.

In addition, in methods that may be performed according to preferred embodiments herein and that may have been described above, the operations have been described in selected typographical sequences. However, the sequences have been selected and so ordered for typographical convenience and are not intended to imply any particular order for performing the operations, unless expressly set forth or understood by those skilled in the art being necessary.

CLAIMS

1. A method of filtering a red-eye phenomenon from a digital image, the method comprising using both anthropometric data and meta-data associated with the image to identify regions of the image potentially susceptible to red-eye artifacts.

2. The method claimed in claim 1, wherein the image is created by optically projecting an external subject onto a sensor array through a lens, and wherein properties of said regions are identified as a function of one or more of the focal length of said lens, the aperture of said lens, the size of the array, the resolution of said array, the distance to the subject and the depth of field at the moment of acquisition.

3. The method claimed in claim 2, wherein each identified region is assigned a probability of being susceptible to red-eye artifacts, and is subject to at least one further anthropometric test to increase or decrease the said probability.

4. The method claimed in claim 3, further comprising examining pixels within each identified region having a probability above a certain threshold to determine those pixels having a colour representative of red-eye artifacts.

5. The method claimed in claim 4, wherein colours representative of red-eye artifacts are defined as a function of one or more of the spectral response of the sensor array, the colour transformations of said image, the tone reproduction transformations of the image, or a combination thereof.

6. The method claimed in any preceding claim, the filtering being executed at least partially within a portable image acquisition device having no photographic film.

7. The method claimed in claim 6, wherein the filtering is executed wholly within the portable image acquisition device.

8. The method claimed in any one of claims 1 to 5, the filtering being executed at least partially as a post-processing operation on an external computation device.

9. A system of filtering a red-eye phenomenon from a digital image, the system comprising means using both anthropometric data and meta-data associated with the image to identify regions of the image potentially susceptible to red-eye artifacts.

10. The system claimed in claim 9, wherein the image is created by optically projecting an external subject onto a sensor array through a lens, and wherein properties of said regions are identified as a function of one or more of the focal length of said lens, the aperture of said lens, the size of the array, the resolution of said array, the distance to the subject and the depth of field at the moment of acquisition.

11. The system claimed in claim 10, further including means for assigning each identified region a probability of being susceptible to red-eye artifacts, and means for subjecting each identified region to at least one further anthropometric test to increase or decrease the said probability.

12. The system claimed in claim 11, further comprising means for examining pixels within each identified region having a probability above a certain threshold to determine those pixels having a colour representative of red-eye artifacts.

13. The system claimed in claim 12, wherein colours representative of red-eye artifacts are defined as a function of one or more of the spectral response of the sensor array, the colour transformations of said image, the tone reproduction transformations of the image, or a combination thereof.

14. The system claimed in any one of claims 9 to 13, wherein the system is embodied at least partially within a portable image acquisition device having no photographic film.

15. The system claimed in claim 14, wherein the system is embodied wholly within said portable image acquisition device.

16. The system claimed in claim 14 or 15, wherein said portable image acquisition device is a digital camera.

17. The system claimed in any one of claims 9 to 13, wherein the system is embodied at least partially as a post-processing function on an external computation device.

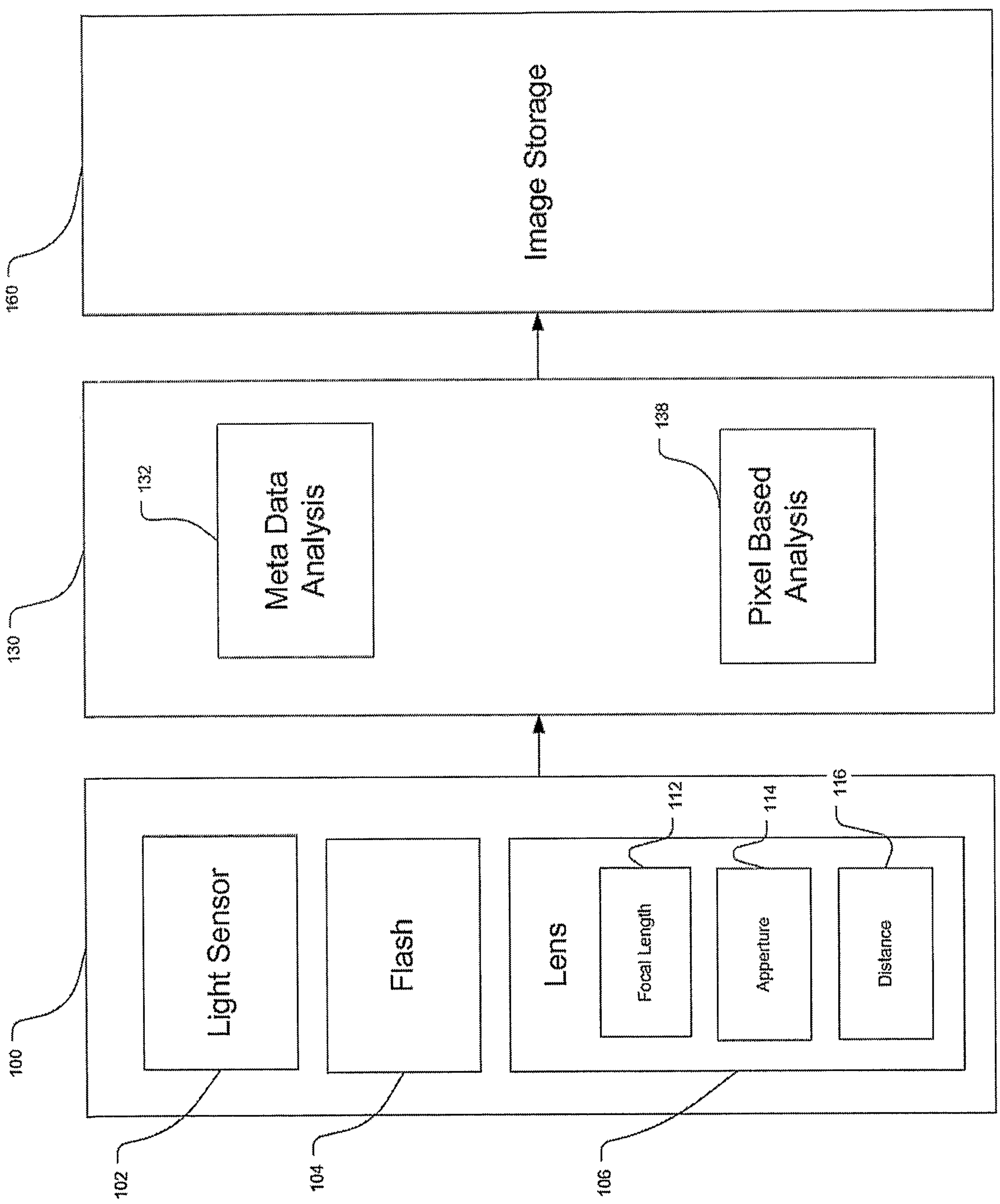


Image Storage
160
130
Meta Data
Analysis
132
Pixel Based
Analysis
138
100
Light Sensor
Flash
Lens
Focal Length
Apperture
Distance
102
104
106
112
114
116

Figure 1

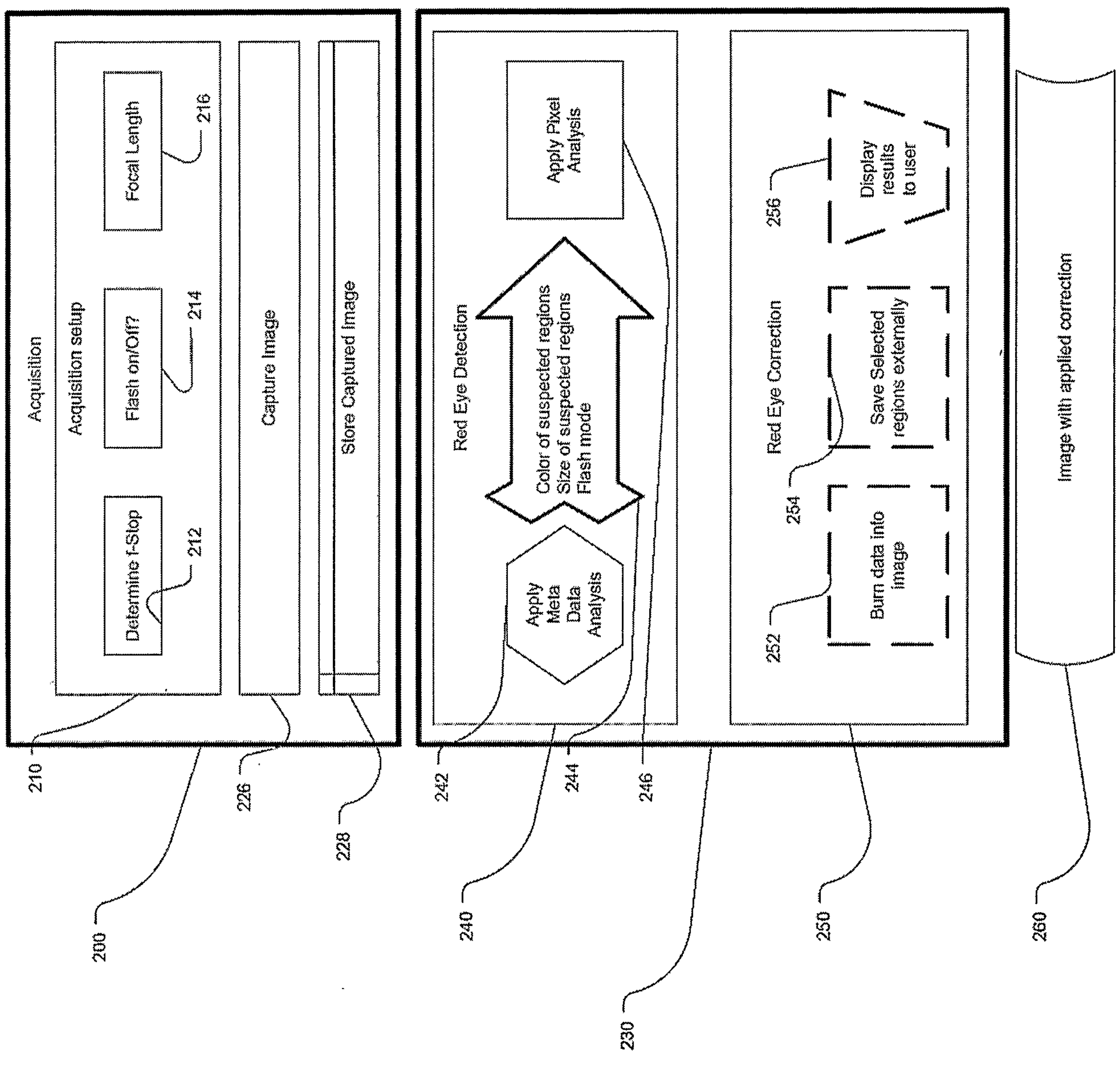

Acquisition
Acquisition setup
Determine f-Stop
Flash on/Off?
Focal Length
Capture Image
Store Captured Image
Red Eye Detection
Apply Meta Data Analysis
Color of suspected regions
Size of suspected regions
Flash mode
Apply Pixel Analysis
Red Eye Correction
Burn data into image
Save Selected regions externally
Display results to user
Image with applied correction
200
210
212
214
216
226
228
230
240
242
244
246
250
252
254
256
260

Figure 2

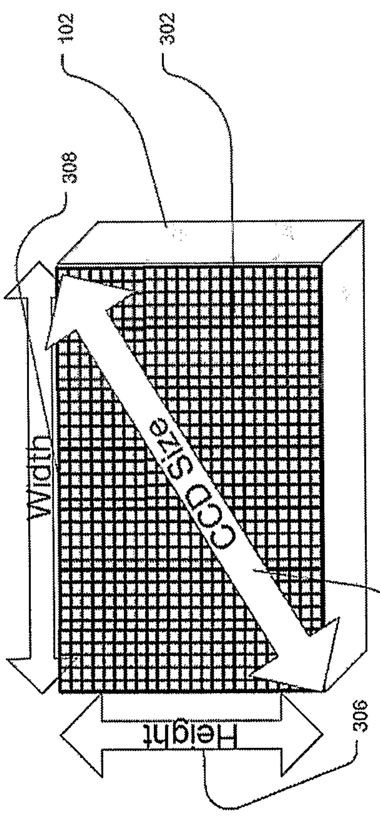


Figure 3a

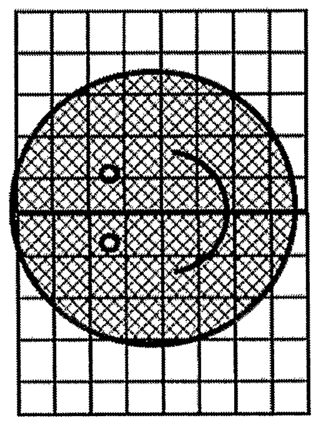


Figure 3b

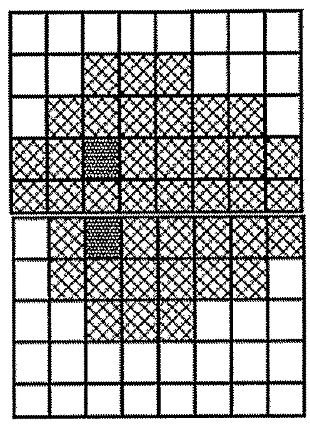


Figure 3c

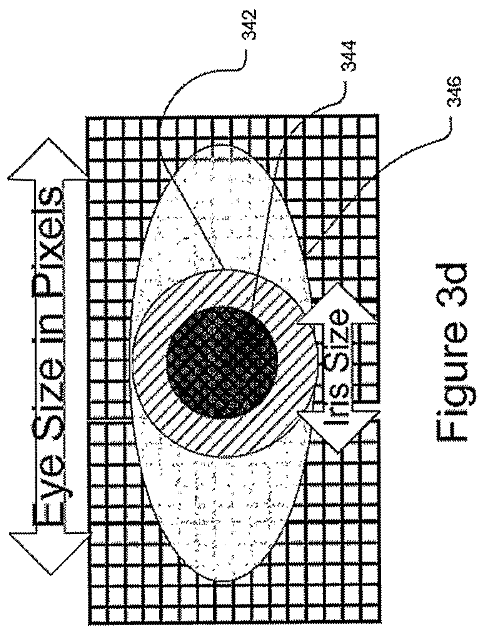


Figure 3d

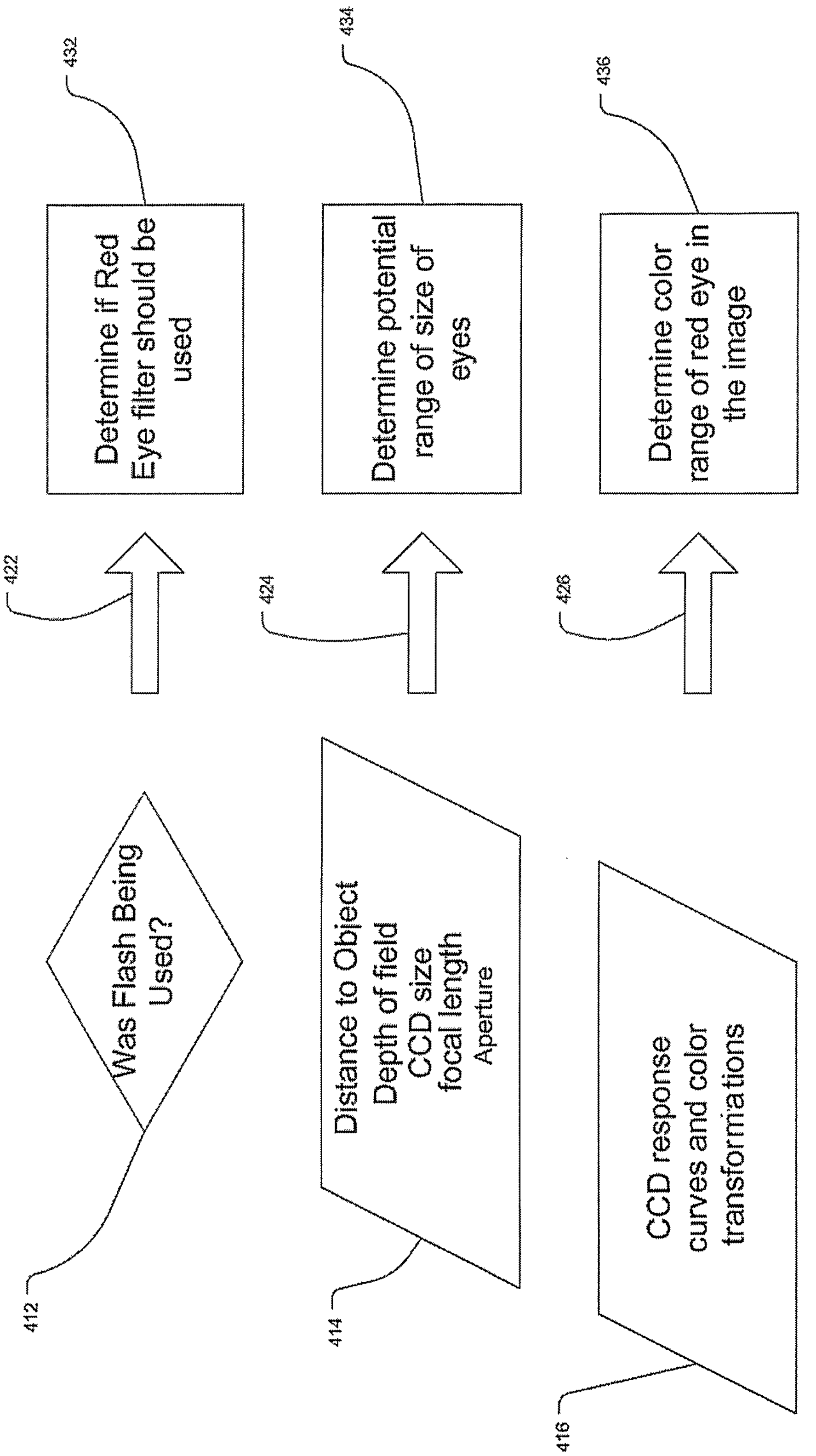

412
Was Flash Being Used?
414
Distance to Object
Depth of field
CCD size
focal length
Aperture
416
CCD response curves and color transformations
422
424
426
432
Determine if Red Eye filter should be used
434
Determine potential range of size of eyes
436
Determine color range of red eye in the image

Figure 4

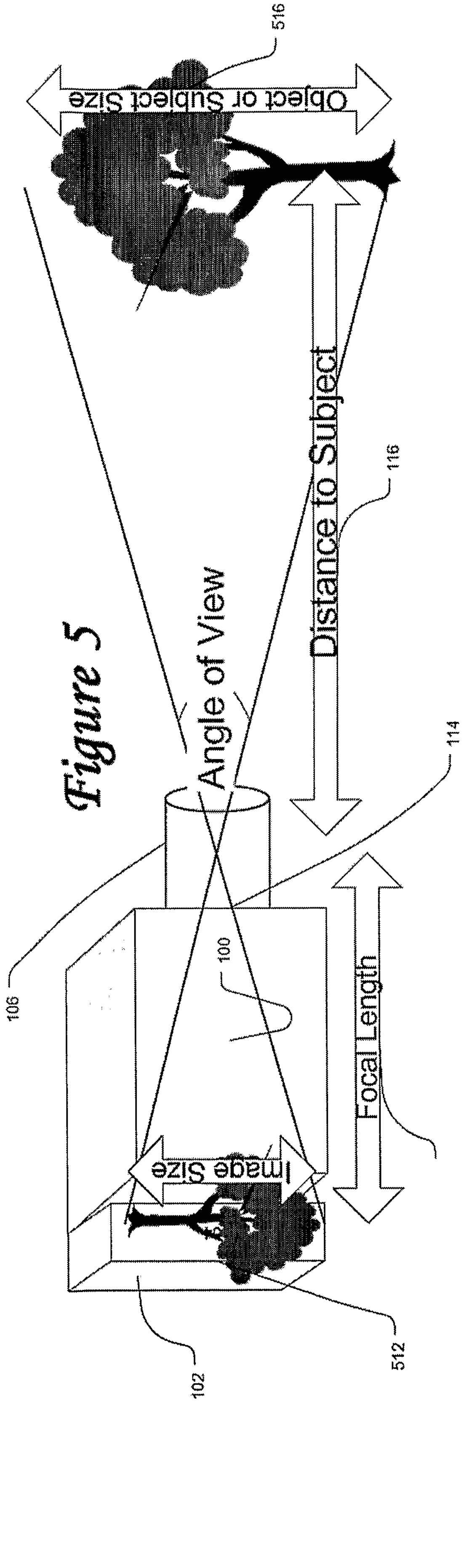


Figure 5

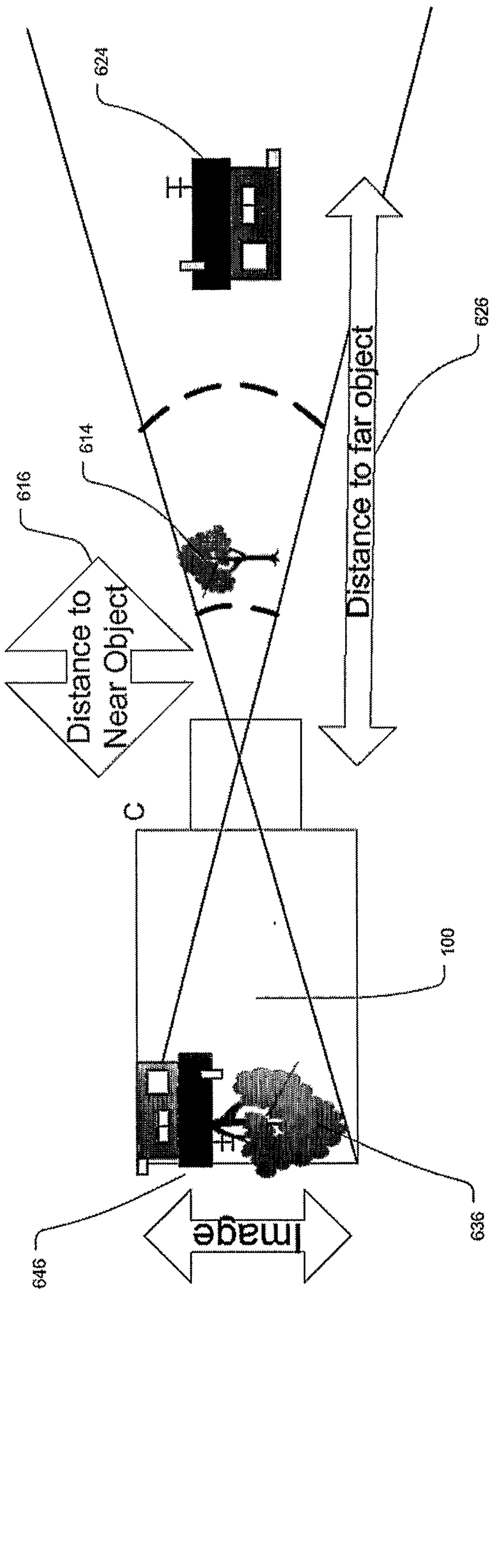


Figure 6

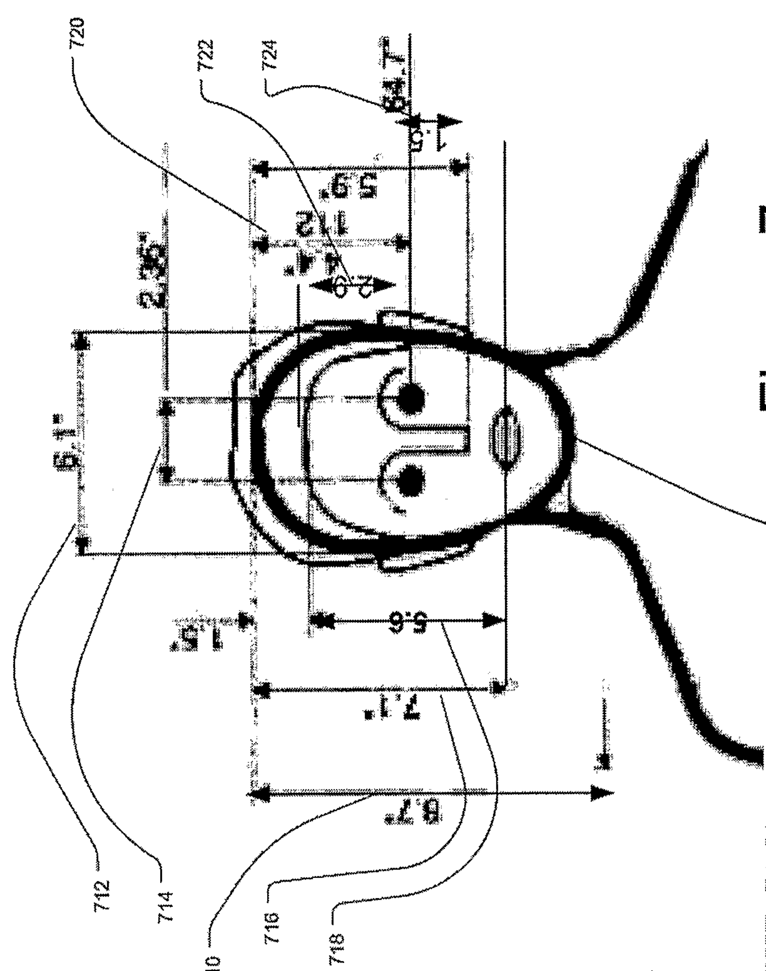


Figure 7a

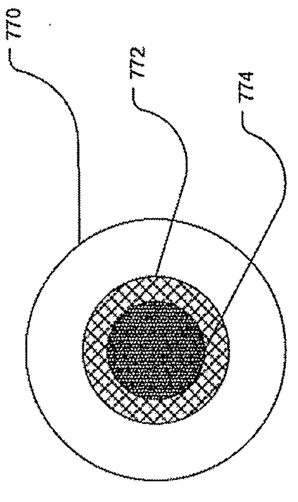


Figure 7c

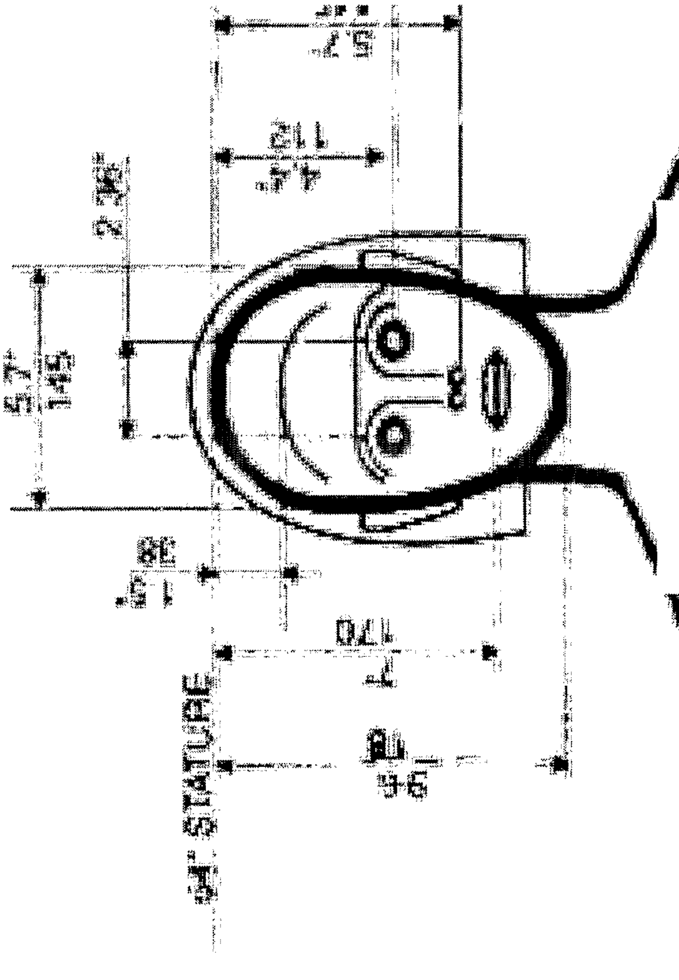


Figure 7b

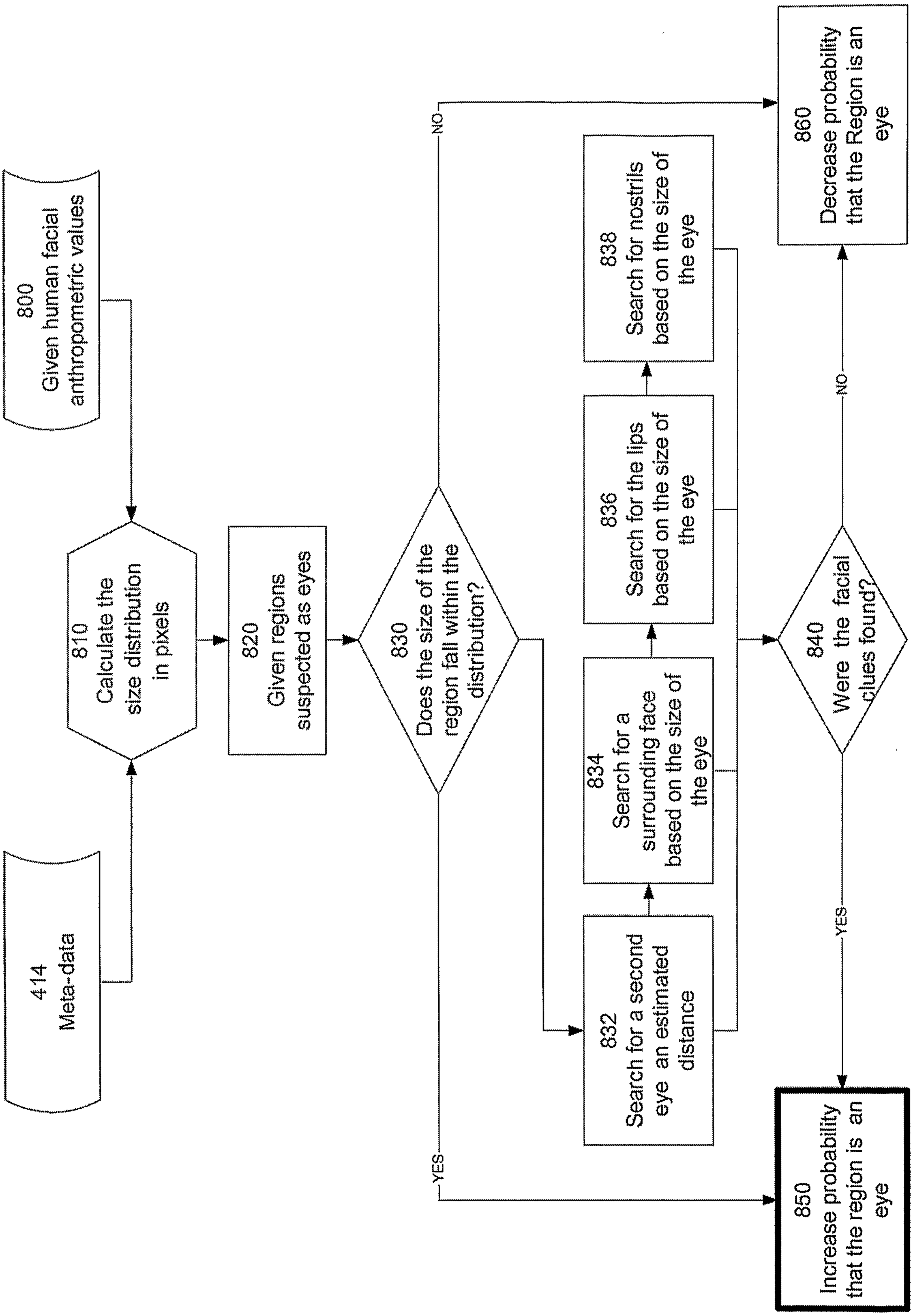

800
Given human facial anthropometric values
414
Meta-data
810
Calculate the size distribution in pixels
820
Given regions suspected as eyes
830
Does the size of the region fall within the distribution?
NO
YES
832
Search for a second eye an estimated distance
834
Search for a surrounding face based on the size of the eye
836
Search for the lips based on the size of the eye
838
Search for nostrils based on the size of the eye
840
Were the facial clues found?
NO
YES
850
Increase probability that the region is an eye
860
Decrease probability that the Region is an eye

Figure 8

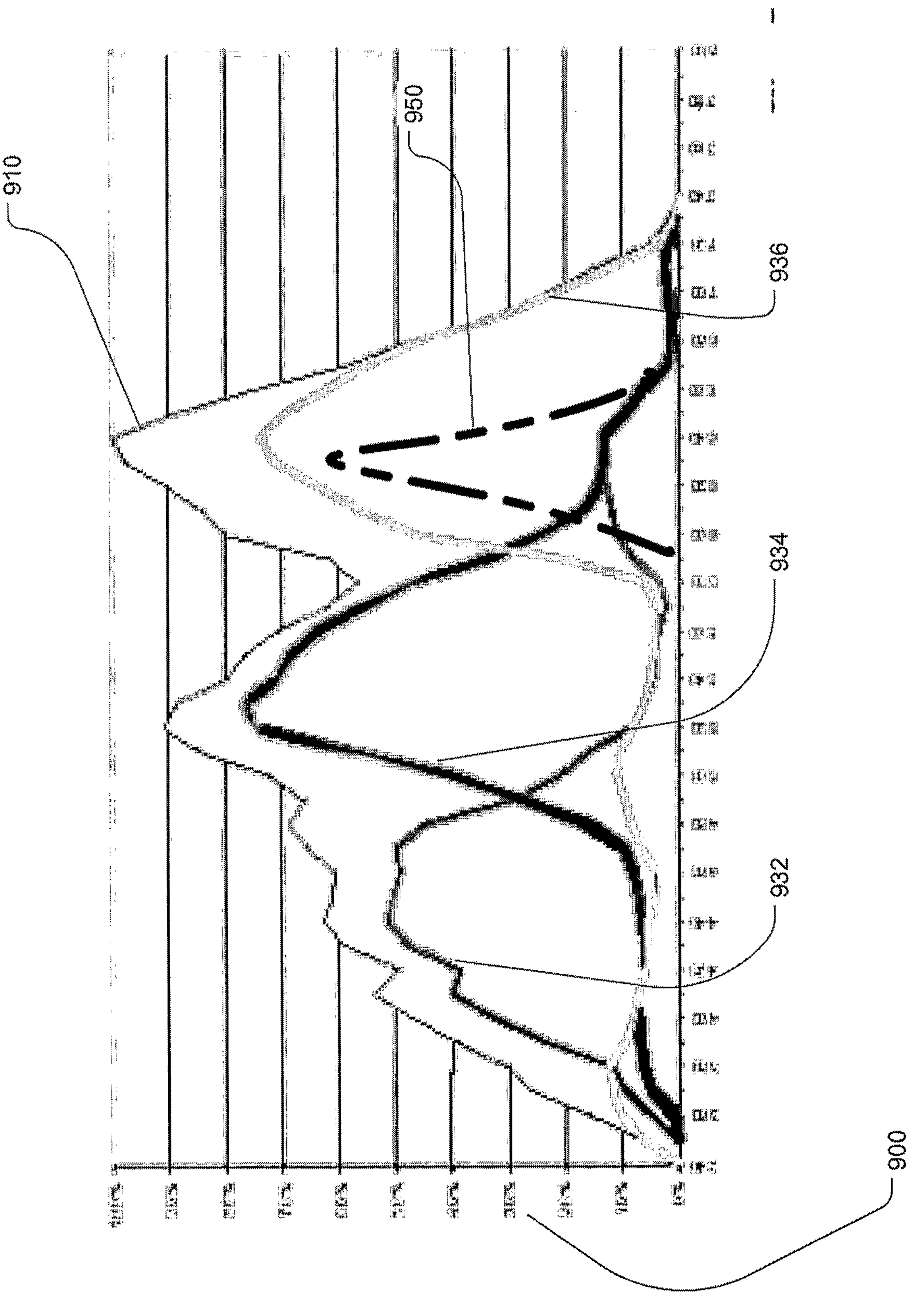


Figure 9

INTERNATIONAL SEARCH REPORT

International Application No
PCT/EP2004/008706

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 H04N1/62

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 7 H04N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

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

EPO-Internal

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X Y	US 2002/136450 A1 (YAN JIE ET AL) 26 September 2002 (2002-09-26) paragraph ‘0005! paragraph ‘0033! - paragraph ‘0034! paragraph ‘0046!	1,8,9, 16,17 2-7, 10-15
Y	US 6 407 777 B1 (DELUCA MICHAEL JOSEPH) 18 June 2002 (2002-06-18) cited in the application column 3, line 46 - line 53 column 4, line 57 - line 67 column 5, line 40 - line 51	2-4, 10-12
Y	EP 1 296 510 A (EASTMAN KODAK CO) 26 March 2003 (2003-03-26) paragraph ‘0047! paragraph ‘0121! - paragraph ‘0123!	2-4, 10-12

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"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)

"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

"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

"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

"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.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
12 November 2004	19/11/2004

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31–70) 340–2040, Tx. 31 651 epo nl, Fax: (+31–70) 340–3016	Winne, D

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 2002/131770 A1 (MEIER ROLAND ET AL) 19 September 2002 (2002-09-19) paragraph '0074! paragraph '0080! paragraph '0041! - paragraph '0043! paragraph '0087! -----	5,13
Y	US 2003/142285 A1 (ENOMOTO JUN) 31 July 2003 (2003-07-31) figure 3 abstract -----	6,7,14, 15

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/EP2004/008706

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2002136450 A1	26-09-2002	NONE	
US 6407777 B1	18-06-2002	NONE	
EP 1296510 A	26-03-2003	US 2003095197 A1	22-05-2003
		EP 1296510 A2	26-03-2003
		JP 2003163861 A	06-06-2003
US 2002131770 A1	19-09-2002	EP 1227659 A1	31-07-2002
		CA 2367928 A1	19-07-2002
		JP 2002305666 A	18-10-2002
US 2003142285 A1	31-07-2003	CN 1437160 A	20-08-2003
		JP 2003283849 A	03-10-2003