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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
24 December 2014 (24.12.2014)

WIPO | PCT

(10) International Publication Number
WO 2014/201667 A1

(51) International Patent Classification: *H04N 13/00* (2006.01)

(21) International Application Number: PCT/CN2013/077595

(22) International Filing Date: 20 June 2013 (20.06.2013)

(25) Filing Language: English

(26) Publication Language: English

(71) Applicant *(for all designated States except US)*: **THOMSON LICENSING** [FR/FR]; 1-5 rue Jeanne d'Arc, F-92130 Issy-les-Moulineaux (FR).

(72) Inventors; and
(71) Applicants *(for US only)*: **SONG, Wenjuan** [CN/CN]; 8th Floor, Building A, Technology Fortune Center, No. 8 Xueqing Road, Haidian District, Beijing 100192 (CN). **DU, Lin** [CN/CN]; 8th Floor, Building A, Technology Fortune Center, No. 8 Xueqing Road, Haidian District, Beijing 100192 (CN). **SONG, Jianping** [CN/CN]; 8th Floor, Building A, Technology Fortune Center, No. 8 Xueqing Road, Haidian District, Beijing 100192 (CN).

(74) Agent: **LIU, SHEN & ASSOCIATES**; A0601, Huibin Building, No.8 Beichen Dong Street, Chaoyang District, Beijing 100101 (CN).

(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, BN, 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, 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, PA, 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, 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, KM, ML, MR, NE, SN, TD, TG).

[Continued on next page]

(54) Title: METHOD AND DEVICE FOR DETECTING AN OBJECT

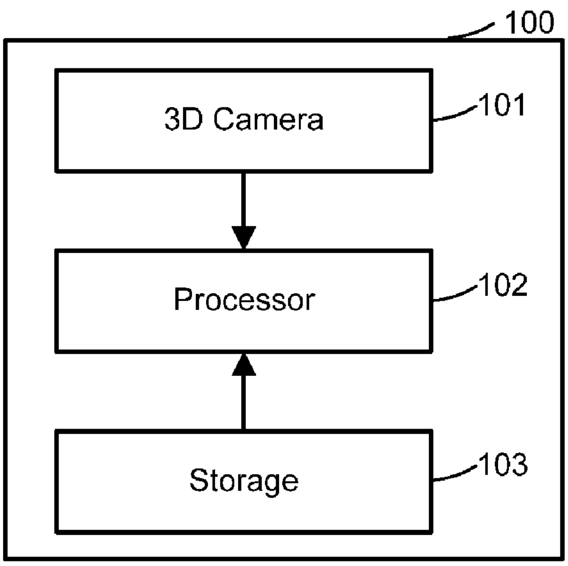

Figure 1

(57) Abstract: It is provided a method for detecting an object in a left view image and a right view image, comprising steps of receiving the left view image and the right view image; detecting a coarse region containing the object in one image of the left view image and the right view image; detecting the object within the detected coarse region in the one image; determining a coarse region in the other image of the left view image and the right view image based on the detected coarse region in the one image and offset relationship indicating position relationship of the object in a past left view image and a past right view image; and detecting the object within the determined coarse region in the other image.

WO 2014/201667 A1

Published:

— *with international search report (Art. 21(3))*

METHOD AND DEVICE FOR DETECTING AN OBJECT

TECHNICAL FIELD

The present invention relates to 3D technology, and more particularly, relates to a method and a device for detecting an object in a left view image and a right view image.

BACKGROUND

Eye tracking is the process of measuring either the point of gaze ("where one is looking") or the motion of an eye relative to the head. It has been used in diverse applications such as in visual system, in psychology, in cognitive linguistics and in product design.

In the past, several methods for following eye movement were developed based on light reflected on the cornea. Lately, many biometric techniques have been developed allowing identification of faces, iris, etc based on digital image processing algorithms which are called non-invasive methods. “Face and eye tracking algorithm based on digital image processing”, which was published in Systems, Man, and Cybernetics, 2001 IEEE International Conference on (Volume:2) (pages: 1178 – 1183 vol. 2, ISSN: 1062-922X, Print ISBN: 0-7803-7087-2), discloses a non-invasive interface to track eye position using digital image processing techniques. It provides a method for eye tracking including: coarse face detection, fine face detection, finding the eye region of maximum probability, map of the pupil location and pupil detection.

It is desired a method for a 3D camera (or called stereo camera) to track user’s eyes.

SUMMARY

According to an aspect of the present invention, it is provided a method for detecting an object in a left view image and a right view image, comprising steps of receiving the left view image and the right view image; detecting a coarse region containing the object in one image of the left view image and the right view image; detecting the object within the detected coarse region in the one image; determining a coarse region in the other image of the left view image and the right view image based on the detected coarse region in the one image and offset relationship indicating position relationship of the object in a past left view image and a past right view image; and detecting the object within the determined coarse region in the other image.

According to another aspect of the present invention, it is provided a device for detecting an object in a left view image and a right view image, comprising a storage for storing offset relationship indicating position relationship of the object in a past left view image and a past right view image; and a processor for receiving the left view image and the right view image; detecting a coarse region containing the object in one image of the left view image and the right view image; detecting the object within the detected coarse region in the one image; determining a coarse region in the other image of the left view image and the right view image based on the detected coarse region in the one image and offset relationship indicating position relationship of the object in a past left view image and a past right view image; and detecting the object within the determined coarse region in the other image.

It is to be understood that more aspects and advantages of the invention will be found in the following detailed description of the present invention.

BRIEF DESCRIPTION OF THE DRAWINGS

The accompanying drawings, which are included to provide a further understanding of the invention and are incorporated in and constitute a part of this application, will be used to illustrate an embodiment of the invention, as explained by the description. The invention is not limited to the embodiment.

In the drawings:

Fig. 1 is a block diagram showing a device for tracking eyes according to an embodiment of the present invention;

Fig. 2 is a diagram showing a user looking at the tablet according to the embodiment of the present invention;

Fig. 3A and 3B are diagrams showing images taken separately by the Cam0 and Cam1 according to the embodiment of the present invention;

Fig. 4 is a flow chart showing a method for tracking eyes by using a 3D camera according to the embodiment of the present invention

Fig. 5 is a diagram showing the geometric position according to the embodiment of the present invention;

Fig. 6A is a diagram showing the detected coarse region in the left image, and Fig. 6B is a diagram showing the determined coarse region in the right image according to the embodiment of the present invention;

Fig. 7A and 7B are diagrams separately showing a 3D camera mounted a display and a 3D camera mounted on a tablet according to the embodiment of present invention.

DETAILED DESCRIPTION

The embodiment of the present invention will now be described in detail in conjunction with the drawings. In the following description, some detailed descriptions of known functions and configurations may be omitted for clarity and conciseness.

Fig. 1 is a block diagram showing a device 100 for eye tracking in a 3D environment according to an embodiment of present invention. The device 100 comprises a 3D camera 101, a processor 102 and a storage 103.

The 3D camera 101 is used to capture a 3D image or a sequence of 3D images of more than one object, and output the captured data to the processor 102. Herein, one 3D image is composed of a left image corresponding to left view of the object and a right image corresponding to right view of the object, wherein the left image and the right image are taken at a same time point. So a sequence of 3D images is a sequence of pairs of left images and right images.

The storage 103 is used to store data and the processor 102 is used to process data.

Fig. 7A and 7B are diagrams showing two applicable scenario of the present invention. Specifically, Fig. 7A shows a 3D display with a 3D camera mounted, and Fig. 7B shows a tablet with a 3D camera mounted.

Taking the tablet 22 with a 3D camera 21 as example, Fig. 2 shows that a user 20 is looking at the tablet and the tablet uses the 3D camera to capture a 3D image or a sequence of 3D images of the user and uses the 3D display to display the captured 3D image or the captured sequence of 3D images.

In order to facilitate description, the two lenses of the 3D camera 21 are named Cam0 and Cam1 from right to left.

Fig. 3A and Fig. 3B are diagrams showing images taken separately by the Cam0 and Cam1. In the Fig. 3A and 3B, the solid lines correspond to the border lines of the images. As can be seen in the Fig. 3A and 3B, objects in the image shown by the Fig. 3A has an offset in position with the corresponding objects in the image shown by the Fig. 3B.

The present invention takes advantage of 1) the determined coarse region of user's face in one of current left image and current right image and 2) offset relationship between previous left image and previous right image to determine the coarse region in the other image. Herein, the offset relationship indicates position relationship of two corresponding points in the previous left image and the previous right image, and is used to predict the position of a point in one image after a corresponding point in the other image is detected.

Fig. 4 is a flow chart showing a method for tracking eyes by using a 3D camera according to the embodiment of present invention.

In the step 401, the processor 102 receives a pair of left image and right image outputted by the 3D camera 101.

In the step 402, the processor 102 detects a coarse region of user face in one of the left image and the right image. In this example, we use the left image. Herein, there is no order between the step 403 and the step 405, which means that they can be processes in parallel.

In the steps 403 and 404, the processor 102 detects a fine region of user face in the left image and detects eyes within the fine region for the left image. The detection method here can be used the same as disclosed in the article titled “Face and eye tracking algorithm based on digital image processing”. But it shall note that other coarse-to-fine approach can also be used here.

In the step 405, the processor 102 determines if an offset relationship between previous left image and previous right image exist. If it does not exist, it goes to step 408, and if it exists it goes to step 406.

In the step 408 the processor 102 detects a fine region of the user face in the right image.

In the step 409, the processor 102 detects eyes within the fine region for the right image.

In the step 406, the processor 102 determines coarse region of the user face in the right image based on the detected coarse region of the user face in the left image and offset relationship between previous left image and previous right image. The details about the determination will be described below after the description of the method.

In the step 407, the processor 102 determines if the user face is detected during the fine detection. If the user face is detected, it goes to the step 409, and if not, it goes to the step 408.

In the step 410, the processor 102 determines offset relationship between the current left image and the current right image based on detected eyes in the current left image and the current right image, and stores the determined the offset relationship in the storage 103.

Below explains how to determine the offset relationship and how to use the offset relationship to determine the coarse region in the other image.

It is assumed the 3D display is based on vertical parallax barriers; and in a virtual 3D coordinates system, X axis has a horizontal direction and is parallel to a plain of the display surface, Y axis has a vertical direction and is parallel to the plain of the display surface and Z axis is perpendicular to the plain of the display surface. The origin of the 3D coordinates system is set in the center of the display surface. The position of the 3D views shown in front of the screen does not depend on the position along the axis Y. Therefore, only the X and Z axes need to be considered. Fig. 5 is a diagram showing the geometric position among the 3D camera and a desired-to-detect object, i.e. eyes in the example. In the Fig. 5, f_0 and f_1 are the focal length of cam0 and cam1; d is the distance between cam0 and cam1; x_0 and x_1 are the absolute values of P in the images captured by the cam0 and cam1 and Z_p is the distance between the 3D camera and the desired-to-detect object.

In the step 410, because the processor 102 has information about the detected eyes in the left image and the detected eyes in the right image, the processor 102 can determine values of x_0 and x_1 and use the following formula (1) to calculated Z_p.

$$Z_p = \frac{d + x_0 + x_1}{\frac{x_0}{f_0} + \frac{x_1}{f_1}}$$

Formula (1)

When the processor 102 determines the coarse region in the image outputted by the Cam1 in the step 406, the following formula (2) and formula (3) are used for determining a point in the right image based on the detected corresponding point in the left image (in the formulas, X_0 and Y_0 represent a point in the left image and X_1 and Y_1 represent a corresponding point in the right image).

Fig. 6A is a diagram showing the detected coarse region in the left image, and Fig. 6B is a diagram showing the determined coarse region in the right image. In this example, the coarse region is a rectangle and defined by top-left point and bottom-right point. Therefore, based on the formula (2) and formula (3), the top-left point (X_1, Y_1) and bottom-right point (X_1', Y_1') of the coarse region in the right image can be calculated.

$$x_1 = \frac{d - x_0\left(\frac{z_p}{f_0} - 1\right)}{\frac{z_p}{f_1} - 1}$$

Formula (2)

$$y_1 = y_0$$

Formula (3)

According to the embodiment, the offset relationship is determined and updated in response to each pair of left image and right image. According to a variant, only when the face cannot be detected in the step 407, the offset relationship is determined and updated. In other words, the offset relationship between a past left image and a past right image (but the past left image and the past right image are taken at the same time) is not updated until the processor determines the face cannot be detected within the determined coarse region.

According to a variant of present embodiment, the present invention is not limited to eye tracking. It is applicable to a system adopting the coarse-to-fine approach to track or detect an object in a sequence of 3D images, wherein the detected coarse region in one of current left image and current right image and offset relationship between previous left image and previous right image to determine coarse region in the other image. And then fine detection technique is used to 1) detect the object within the detected coarse region in one image and 2) detect the object within the determined coarse region in the other image.

According to the present embodiment, the offset relationship between the left image and the right image are determined based on the detected eyes in the left image and the right image. According to a variant, the offset relationship is determined based on determined fine regions of the user face in the left image and the right image.

It shall note the features from above embodiment and its variants can be combined and substituted.

A number of implementations have been described. Nevertheless, it will be understood that various modifications may be made. For example, elements of different implementations may be combined, supplemented, modified, or removed to produce other implementations. Additionally, one of ordinary skill will understand that other structures and processes may be substituted for those disclosed and the resulting implementations will perform at least substantially the same function(s), in at least substantially the same way(s), to achieve at least substantially the same result(s) as the implementations disclosed. Accordingly, these and other implementations are contemplated by this application and are within the scope of the invention as defined by the appended claims.

CLAIMS

1. A method for detecting an object in a left view image and a right view image, comprising steps of:

receiving the left view image and the right view image;

detecting a coarse region containing the object in one image of the left view image and the right view image;

detecting the object within the detected coarse region in the one image;

determining a coarse region in the other image of the left view image and the right view image based on the detected coarse region in the one image and offset relationship indicating position relationship of the object in a past left view image and a past right view image; and

detecting the object within the determined coarse region in the other image.

2. The method of the claim 1, wherein if failing to detect the object within the determined coarse region in the other image, the method further comprising

detecting a coarse region containing the object in the other image; and

detecting the object within the detected coarse region in the other image.

3. The method of the claim 2, further comprising

determining offset relationship of the object in the left view image and right view image based on the detected object in the left view image and the detected object in the right view image.

4. The method of the claim 1, wherein the past left view image is a previous left view image and the past right view image is a previous right view image, the method further comprising

determining offset relationship of the object in the left view image and right view image based on the detected object in the left view image and the detected object in the right view image.

5. A device for detecting an object in a left view image and a right view image, comprising:

a storage for storing offset relationship indicating position relationship of the object in a past left view image and a past right view image; and

a processor for receiving the left view image and the right view image; detecting a coarse region containing the object in one image of the left view image and the right view image; detecting the object within the detected coarse region in the one image; determining a coarse region in the other image of the left view image and the right view image based on the detected coarse region in the one image and offset relationship indicating position relationship of the object in a past left view image and a past right view image; and detecting the object within the determined coarse region in the other image.

6. The device of the claim 5, wherein further comprising

a 3D camera for capturing and outputting the left view image and the right view image.

7. The device of the claim 5, wherein the processor is further used for detect the object within the determined coarse region in the other image; and if determining failure to detect the object within the determined coarse region in the other image, detecting a coarse region containing the object in the other image and detecting the object within the detected coarse region in the other image.

8. The device of the claim 5, wherein the processor is further used for determining offset relationship of the object in the left view image and right view image based on the detected object in the left view image and the detected object in the right view image.

9. The device of the claim 5, wherein the past left view image is a previous left view image and the past right view image is a previous right view image, the processor is further used for determining offset relationship of the object in the left view image and right view image based on the detected object in the left view image and the detected object in the right view image.

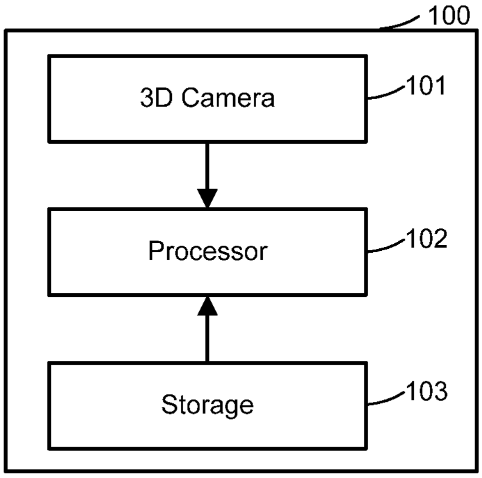

Figure 1

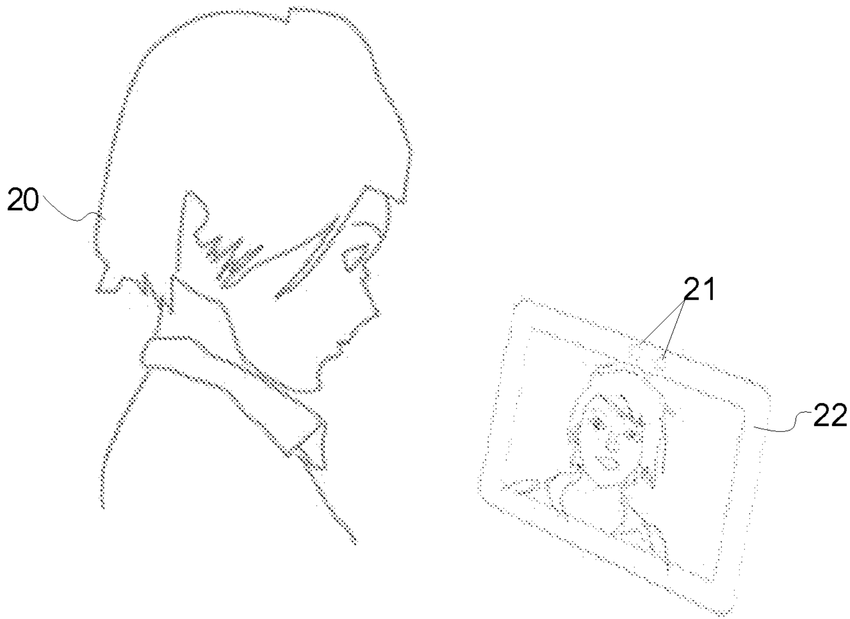

Figure 2

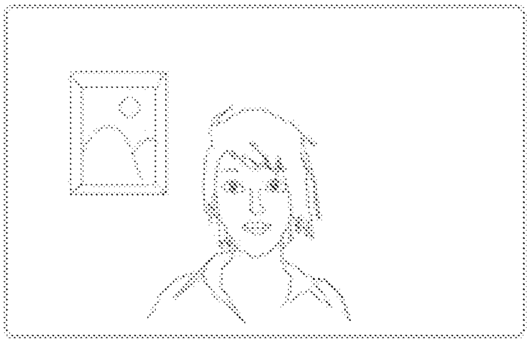

Figure 3A

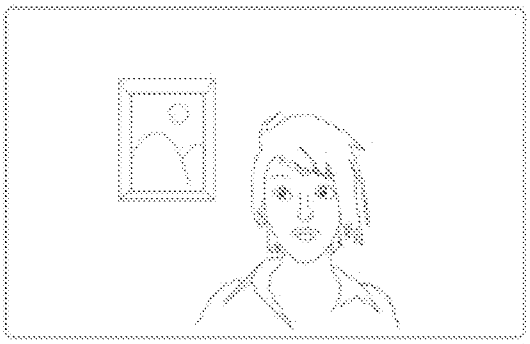

Figure 3B

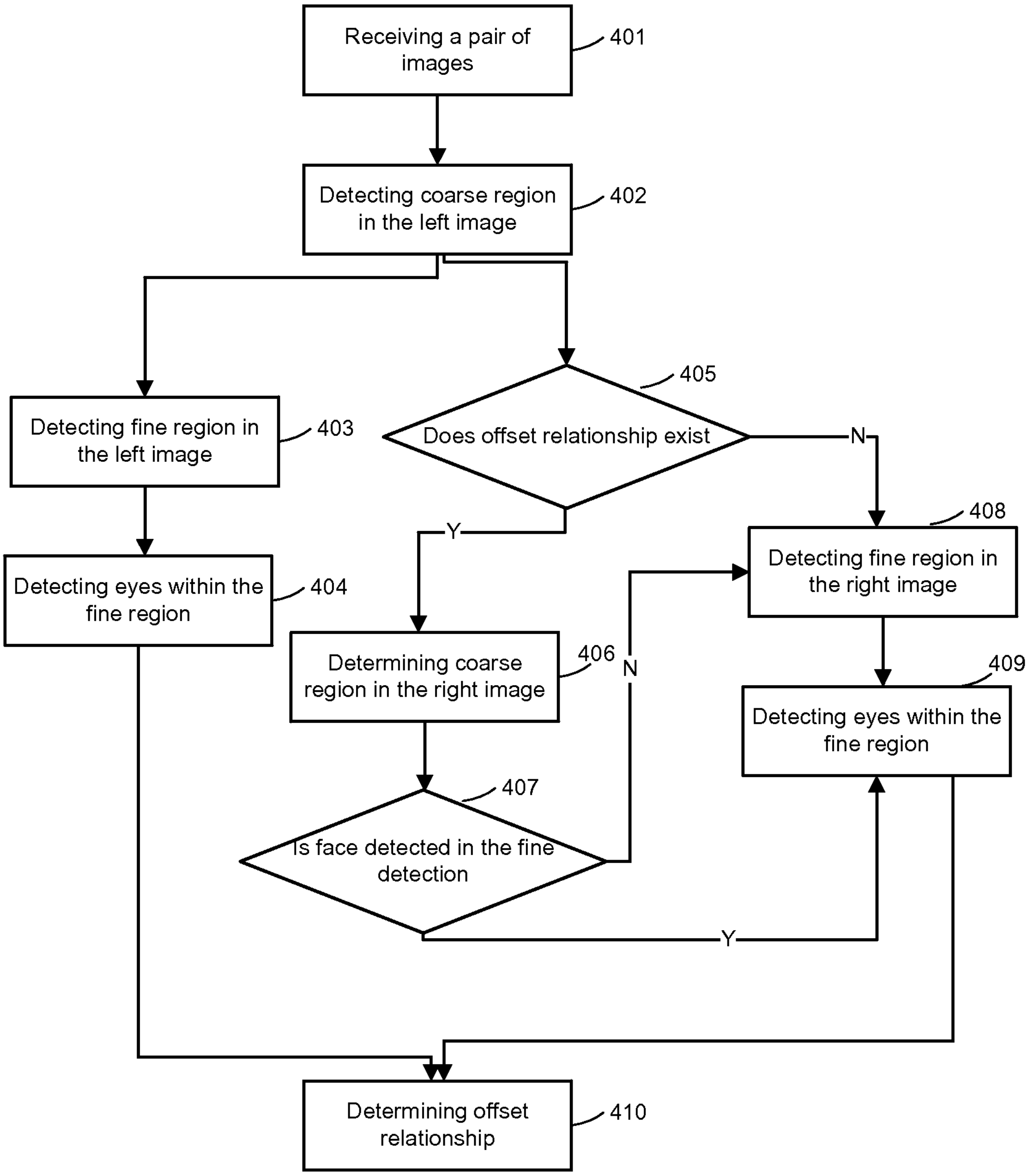
Receiving a pair of images
401
Detecting coarse region in the left image
402
Detecting fine region in the left image
403
Detecting eyes within the fine region
404
Does offset relationship exist
405
N
Y
Determining coarse region in the right image
406
N
Is face detected in the fine detection
407
Y
Detecting fine region in the right image
408
Detecting eyes within the fine region
409
Determining offset relationship
410

Figure 4

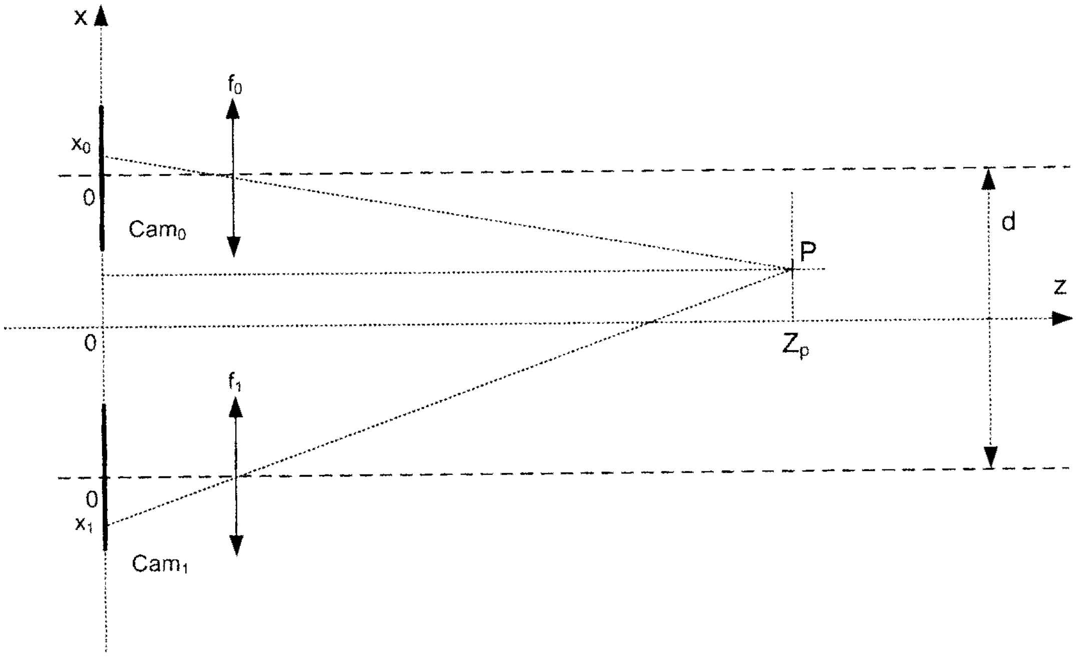

Figure 5

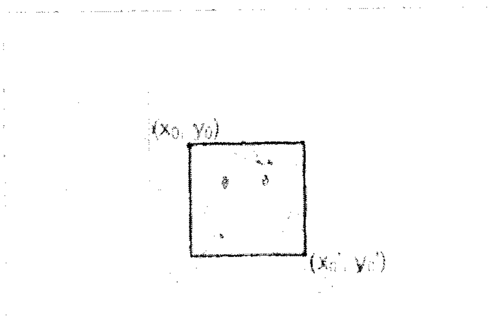

Figure 6A

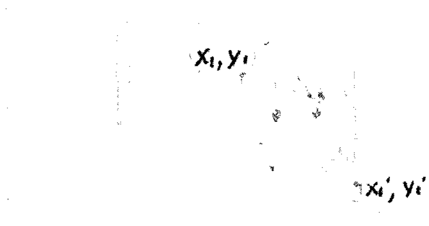

Figure 6B

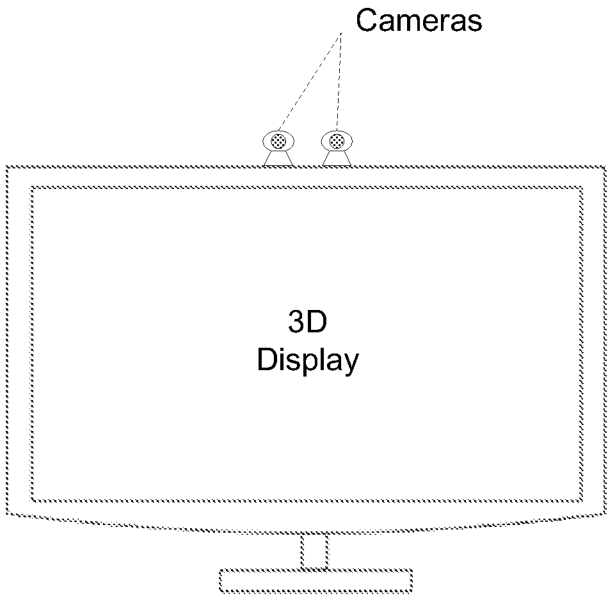

Figure 7A

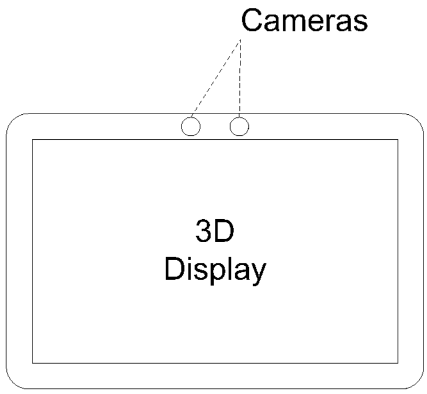

Figure 7B

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2013/077595

A. CLASSIFICATION OF SUBJECT MATTER

H04N 13/00 (2006.01) i

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: H04N, G06T

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 practicable, search terms used)

CNPAT, WPI, EPODOC, CNKI: 3D, three dimension+, stereo, left, right, face, coarse region, object, eye, offset, derivation, variation, difference

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	CN 101959042 A (FUJI FILM CORP.) 26 January 2011 (26.01.2011) description, paragraphs [0036]-[0093], and figures 1-9	1-9
Y	CN 101321269 A (UNIV. TONGJI) 10 December 2008 (10.12.2008) description, page 4, line 19 to page 6, line 15 and figure 2	1-9
A	CN 102435172 A (UNIV. BEIJING POSTS&TELECOM) 02 May 2012 (02.05.2012) the whole document	1-9
A	US 2008152214 A1 (FUJI FILM CORP.) 26 June 2008 (26.06.2008) the whole document	1-9

☐ Further documents are listed in the continuation of Box C. ☒ See patent family 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 application or patent 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
10 February 2014 (10.02.2014)	**27 Mar. 2014 (27.03.2014)**
Name and mailing address of the ISA/CN The State Intellectual Property Office, the P.R.China 6 Xitucheng Rd., Jimen Bridge, Haidian District, Beijing, China 100088 Facsimile No. 86-10-62019451	Authorized officer XIA,Kan Telephone No. (86-10)61648255

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2013/077595

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 101959042 A	26.01.2011	US 2011012991 A1	20.01.2011
		EP 2278819 A2	26.01.2011
		JP 2011024003 A	03.02.2011
CN 101321269 A	10.12.2008	CN 101321269 B	14.09.2011
CN 102435172 A	02.05.2012	None	
US 2008152214 A1	26.06.2008	JP 2008160381 A	10.07.2008
		JP 4403172 B2	20.01.2010