

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

(19) World Intellectual Property
Organization
International Bureau



(10) International Publication Number
WO 2015/128542 A2

(43) International Publication Date
3 September 2015 (03.09.2015)

- (51) International Patent Classification: Not classified
- (21) International Application Number: PCT/FI2015/050109
- (22) International Filing Date: 25 February 2015 (25.02.2015)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data: 14/193,508 28 February 2014 (28.02.2014) US
- (71) Applicant: **NOKIA TECHNOLOGIES OY** [FI/FI]; Karaportti 3, FI-02610 Espoo (FI).
- (72) Inventors: **CHEN, Ching-Hui**; 151 Westway, Apt 102, Greenbelt, Maryland 20770 (US). **ZHOU, Hui**; 992 Belmont Ter No. 2, Sunnyvale, California 94086 (US). **AHONEN, Timo**; 1122 Whipple Ave, Apt 14, Redwood City, California 94062 (US).
- (74) Agents: **NOKIA TECHNOLOGIES OY** et al.; Ari Aarnio, IPR Department, Karakaari 7, FI-02610 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, 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, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, 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, SA, 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, ST, 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).
- Published:**
— without international search report and to be republished upon receipt of that report (Rule 48.2(g))



WO 2015/128542 A2

(54) Title: PROCESSING STEREO IMAGES

(57) Abstract: A method comprises estimating disparity values for pixels located in an occluded region of a first one of a defocused pair of stereo images using: a first set of blur map data corresponding to the first one of the defocused pair of stereo images; a first set of disparity map data corresponding to the first one of the defocused pair of stereo images; and a first set of occlusion map data corresponding to the first one of the defocused pair of stereo images.

Processing Stereo Images

Field

This specification relates to the processing of stereo images. In particular, this
5 specification relates to estimating disparity values for pixels in at least one of a pair of stereo images.

Background

Estimating disparity from a pair of stereo images is a technique that has been well studied
10 with the estimated disparity being used in the extraction of further information from the stereo images. For instance, the estimated disparity can be used for depth/distance calculation. Disparity between the location of an object in each of the pair of images and the distance of the object from the camera(s) which capture the stereo images are inversely related. As such, as the distance from the camera(s) increases, the disparity
15 decreases. Consequently, disparity estimation enables depth perception in stereo images. Specifically, using geometry and algebra, pixels in two-dimensional stereo images can be mapped as coordinates in three-dimensional space. Determining depth from two-dimensional images has many uses, not least in the field of navigation.

Summary

In a first aspect, this specification describes a method comprising estimating disparity values for pixels located in an occluded region of a first one of a defocused pair of stereo images using: a first set of blur map data corresponding to the first one of the defocused pair of stereo images; a first set of disparity map data corresponding to the first one of the
25 defocused pair of stereo images; and a first set of occlusion map data corresponding to the first one of the defocused pair of stereo images.

The method may further comprise generating a first set of blur-disparity map data corresponding to the first one of the defocused pair of stereo images, the first set of blur-disparity map data being generated based on: the first set of disparity map data; the first
30 set of occlusion map data; and the first set of blur map data; and estimating the disparity values for the pixels located in the occluded region of the first one of the defocused pair of stereo images using the first set of occlusion map data and the first set of blur-disparity map data.

35

The method may further comprise updating the first set of disparity map data with the estimated disparity values to generate a first set of updated disparity map data.

The method may further comprise applying at least one filter to the first set of updated disparity map data. Applying the at least one filter may comprise applying a first filter to the first set of updated disparity map data and applying a second filter to the first-filtered first set of updated disparity map data. The first filter may be one of a morphological erosion filter and a morphological dilation filter and the second filter may be the other of the morphological erosion filter and the morphological dilation filter.

Generating the first set of blur-disparity map data may comprise: using the first set of occlusion map data to identify pixels in the first set of disparity map data which correspond to the non-occluded pixels in the first one of the defocused pair of stereo images, each identified pixel in the set of disparity map data having a disparity value; for each different disparity value of the identified pixels: identifying a set of non-occluded pixels in the first set of blur map data, the set of non-occluded pixels in the first set of blur map data corresponding to the identified pixels in the first set of disparity map data which have the different disparity value, each identified non-occluded pixel in the first set of blur map data having a blur value; calculating the mean blur value of the set of non-occluded pixels in the first set of blur map data; and assigning the calculated mean blur value to the pixels in the first set of blur-disparity map data which correspond to the identified set of non-occluded pixels in the first set of blur map data. The method may further comprise, if none of the previously-calculated mean blur values have a contributing pixel having the blur value of an identified occluded pixel in the first set of blur map data, assigning to the occluded pixel in the first set of blur-disparity map data, which corresponds to the identified occluded pixel, the mean blur value which is closest to the blur value of the identified occluded pixel.

Each pixel in the first set of blur-disparity map data may have an associated mean blur value, and estimating the disparity values for the pixels located in the occluded region of the first one of the defocused pair of stereo images may comprise: using the first set of occlusion map data to identify at least one occluded pixel in the first set of blur-disparity map data; for each occluded pixel in the first set of blur-disparity map data: calculating a combined difference between the mean blur value of the occluded pixel and the mean blur value of each non-occluded pixel in the first set of blur-disparity map data within an $M \times N$ neighbourhood of the occluded pixel; for each mean blur value within the $M \times N$ neighbourhood, calculating the mean blur difference; and identifying the mean blur value having the lowest mean blur difference, wherein the estimated disparity value for the occluded pixel is the disparity value which corresponds to the mean blur value having the

- 3 -

lowest mean blur difference. The combined difference may be a weighted average of the absolute difference between the mean blur value of the occluded pixel and the mean blur value of the non-occluded pixel and the geometric Euclidian distance between occluded pixel and the non-occluded pixel. The mean blur difference for each mean blur value
5 within the $M \times N$ neighbourhood may be calculated by summing the combined differences of all the non-occluded pixels having the mean blur value and normalising the result by the total number of the non-occluded pixels having the mean blur value.

The method may further comprise generating the first set of disparity map data using the
10 pair of stereo images. Generating the first set of disparity map data may comprise stereo matching the pair of stereo images.

The method may further comprise generating the first set of occlusion map data using the first one of the defocused pair of stereo images and the first set of disparity map data.
15 Generating the first occlusion map data may comprise: transforming pixels of the first one of the defocused pair of stereo images in accordance with the first set of disparity map data and processing the resulting image by setting each pixel in the resulting image which has at least one pixel from the first stereo image to a first value indicating a non-occluded pixel and setting each pixel which has no pixels from the first stereo image to a second
20 value indicating an occluded pixel.

The method may further comprise generating the first set of blur map data using the first one of the defocused pair of stereo images. Generating the first set of blur map data may comprise applying a depth-from-defocus technique to the first one of the defocused the
25 pair of stereo images.

The method may further comprise: estimating disparity values for pixels located in an occluded region of a second one of a defocused pair of stereo images using: a second set of blur map data corresponding to the second one of the defocused pair of stereo images; a
30 second set of disparity map data corresponding to the second one of the defocused pair of stereo images; and a second set of occlusion map data corresponding to the second one of the defocused pair of stereo images. The method may further comprise: generating a second set of blur-disparity map data corresponding to the second one of the defocused pair of stereo images, the second set of blur-disparity map data being generated based on:
35 the second set of disparity map data; the second set of occlusion map data; and the second set of blur map data; estimating the disparity values for the pixels located in the occluded region of the second one of the defocused pair of stereo images using the second set of

- 4 -

occlusion map data and the second set of blur-disparity map data. The method may further comprise updating the second set of disparity map data with the estimated disparity values for pixels located in the occluded region of the second one of the defocused pair of stereo images, to generate a second set of updated disparity map data.

5 The method may further comprise applying at least one filter to the second set of updated disparity map data. The method may further comprise: applying a first filter to the second set of updated disparity map data; and applying a second filter to the first-filtered second set of updated disparity map data.

10 The first filter may be one of a morphological erosion filter and a morphological dilation filter and the second filter may be the other of the morphological erosion filter and the morphological dilation filter.

In a second aspect, this specification describes an apparatus comprising: at least one processor; and at least one memory including computer program code, the at least one
15 memory and the computer program code configured to, with the at least one processor, cause the apparatus at least to: estimate disparity values for pixels located in an occluded region of a first one of a defocused pair of stereo images using: a first set of blur map data corresponding to the first one of the defocused pair of stereo images; a first set of disparity map data corresponding to the first one of the defocused pair of stereo images; and a first
20 set of occlusion map data corresponding to the first one of the defocused pair of stereo images.

The at least one memory and the computer program code may be configured to, with the at least one processor, cause the apparatus to: generate a first set of blur-disparity map
25 data corresponding to the first one of the defocused pair of stereo images, the first set of blur-disparity map data being generated based on: the first set of disparity map data; the first set of occlusion map data; and the first set of blur map data; and estimate the disparity values for the pixels located in the occluded region of the first one of the defocused pair of stereo images using the first set of occlusion map data and the first set of
30 blur-disparity map data.

The at least one memory and the computer program code may be configured to, with the at least one processor, cause the apparatus to: update the first set of disparity map data with the estimated disparity values to generate a first set of updated disparity map data.

35 The at least one memory and the computer program code may be configured to, with the at least one processor, cause the apparatus to: apply at least one filter to the first set of updated disparity map data. The at least one memory and the computer program code

- 5 -

may be configured to, with the at least one processor, cause the apparatus to: apply a first filter to the first set of updated disparity map data; and apply a second filter to the first-filtered first set of updated disparity map data. The first filter may be one of a morphological erosion filter and a morphological dilation filter and the second filter may be the other of the morphological erosion filter and the morphological dilation filter.

The at least one memory and the computer program code may be configured to, with the at least one processor, cause the apparatus to generate the first set of blur-disparity map data by: using the first set of occlusion map data to identify pixels in the first set of disparity map data which correspond to the non-occluded pixels in the first one of the defocused pair of stereo images, each identified pixel in the set of disparity map data having a disparity value; for each different disparity value of the identified pixels: identifying a set of non-occluded pixels in the first set of blur map data, the set of non-occluded pixels in the first set of blur map data corresponding to the identified pixels in the first set of disparity map data which have the different disparity value, each identified non-occluded pixel in the first set of blur map data having a blur value; calculating the mean blur value of the set of non-occluded pixels in the first set of blur map data; and assigning the calculated mean blur value to the pixels in the first set of blur-disparity map data which correspond to the identified set of non-occluded pixels in the first set of blur map data. The at least one memory and the computer program code may be configured to, with the at least one processor, cause the apparatus to, if none of the previously-calculated mean blur values have a contributing pixel having the blur value of an identified occluded pixel in the first set of blur map data, assign to the occluded pixel in the first set of blur-disparity map data, which corresponds to the identified occluded pixel, the mean blur value which is closest to the blur value of the identified occluded pixel.

Each pixel in the first set of blur-disparity map data may have an associated mean blur value, and wherein the at least one memory and the computer program code may be configured to, with the at least one processor, cause the apparatus to estimate the disparity values for the pixels located in the occluded region of the first one of the defocused pair of stereo images by: using the first set of occlusion map data to identify at least one occluded pixel in the first set of blur-disparity map data; for each occluded pixel in the first set of blur-disparity map data: calculating a combined difference between the mean blur value of the occluded pixel and the mean blur value of each non-occluded pixel in the first set of blur-disparity map data within an $M \times N$ neighbourhood of the occluded pixel; for each mean blur value within the $M \times N$ neighbourhood, calculating the mean blur difference; and identifying the mean blur value having the lowest mean blur difference,

- 6 -

wherein the estimated disparity value for the occluded pixel is the disparity value which corresponds to the mean blur value having the lowest mean blur difference. The combined difference may be a weighted average of the absolute difference between the mean blur value of the occluded pixel and the mean blur value of the non-occluded pixel and the geometric Euclidian distance between occluded pixel and the non-occluded pixel.
5 The at least one memory and the computer program code may be configured to, with the at least one processor, cause the apparatus to: calculate the mean blur difference for each mean blur value within the $M \times N$ neighbourhood by summing the combined differences of all the non-occluded pixels having the mean blur value and normalising the result by the
10 total number of the non-occluded pixels having the mean blur value.

The at least one memory and the computer program code may be configured to, with the at least one processor, cause the apparatus to generate the first set of disparity map data using the pair of stereo images.
15

The at least one memory and the computer program code may be configured to, with the at least one processor, cause the apparatus to generate the first set of occlusion map data using the first one of the defocused pair of stereo images and the first set of disparity map data.
20

The at least one memory and the computer program code may be configured to, with the at least one processor, cause the apparatus to generate the first set of blur map data using the first one of the defocused pair of stereo images.

25 The at least one memory and the computer program code may be configured to, with the at least one processor, cause the apparatus to estimate disparity values for pixels located in an occluded region of a second one of a defocused pair of stereo images using: a second set of blur map data corresponding to the second one of the defocused pair of stereo images; a second set of disparity map data corresponding to the second one of the
30 defocused pair of stereo images; and a second set of occlusion map data corresponding to the second one of the defocused pair of stereo images. The at least one memory and the computer program code may be configured to, with the at least one processor, cause the apparatus to generate a second set of blur-disparity map data corresponding to the second one of the defocused pair of stereo images, the second set of blur-disparity map data being
35 generated based on: the second set of disparity map data; the second set of occlusion map data; and the second set of blur map data; estimate the disparity values for the pixels located in the occluded region of the second one of the defocused pair of stereo images

- 7 -

using the second set of occlusion map data and the second set of blur-disparity map data. The at least one memory and the computer program code may be configured to, with the at least one processor, cause the apparatus to: update the second set of disparity map data with the estimated disparity values for pixels located in the occluded region of the second one of the defocused pair of stereo images to generate a second set of updated disparity map data. The at least one memory and the computer program code may be configured to, with the at least one processor, cause the apparatus to: apply at least one filter to the second set of updated disparity map data. The at least one memory and the computer program code may be configured to, with the at least one processor, cause the apparatus to: apply a first filter to the second set of updated disparity map data; and apply a second filter to the first-filtered second set of updated disparity map data. The first filter may be one of a morphological erosion filter and a morphological dilation filter and the second filter may be the other of the morphological erosion filter and the morphological dilation filter.

15

In a third aspect, this specification describes a non-transitory computer-readable storage medium having stored thereon computer-readable code, which, when executed by computing apparatus, causes the computing apparatus to perform a method comprising: estimating disparity values for pixels located in an occluded region of a first one of a defocused pair of stereo images using: a first set of blur map data corresponding to the first one of the defocused pair of stereo images; a first set of disparity map data corresponding to the first one of the defocused pair of stereo images; and a first set of occlusion map data corresponding to the first one of the defocused pair of stereo images. The non-transitory computer-readable storage medium may additionally have stored thereon computer-readable code which, when executed causes the computer apparatus to perform any of the steps described with reference to the method of the first aspect.

20

25

In a fourth aspect, this specification describes computer-readable code, which when executed by computing apparatus causes the computing apparatus to perform any method as described with reference to the first aspect.

30

In a fifth aspect, this specification describes apparatus comprising: means for estimating disparity values for pixels located in an occluded region of a first one of a defocused pair of stereo images using: a first set of blur map data corresponding to the first one of the defocused pair of stereo images; a first set of disparity map data corresponding to the first one of the defocused pair of stereo images; and a first set of occlusion map data corresponding to the first one of the defocused pair of stereo images. The apparatus

35

- 8 -

according to the fifth aspect may further comprise means for performing any of the operations described with reference to the first aspect.

Brief Description of the Figures

- 5 For a more complete understanding of examples of embodiments of the invention, reference is now made to the following description taken in connection with the accompanying drawings in which:
- Figure 1 is a schematic illustration of an apparatus which may perform operations according to examples of embodiments of the invention;
- 10 Figure 2 is a schematic illustration of an image capture and processing apparatus which may perform operations according to examples of embodiments of the invention;
- Figure 3 shows schematically an image processing apparatus according to some examples of embodiments of the invention;
- Figure 4 is a flow chart illustrating operations which may be performed by the image
15 processing apparatus of Figure 3;
- Figures 5A to 5J are images depicting outputs of various ones of the operations illustrated in Figure 4; and
- Figure 6 is a graph showing the theoretical relationship between the blur of a pixel and its
20 disparity.

Detailed Description of Examples of Embodiments

The following description and the drawings relate to provision of improved methods and apparatuses for image processing and analysis. In the description and drawings, like reference numerals refer to like elements throughout.

- 25 Figure 1 is a schematic block diagram of an exemplary electronic device 1 or apparatus 1, which incorporates image processing apparatus 20 for providing improved image processing functionality.
- 30 The electronic device 1 is, in some embodiments, a portable electronic device, or more specifically, a portable image capture device. For instance the electronic device may be but is not limited to a tablet computer, a mobile phone, a media player, a PDA and a dedicated camera.
- 35 The electronic device 1 comprises image processing apparatus 20. The image processing apparatus 20 comprises processing circuitry 10 and at least one non-transitory computer readable memory medium 12. The processing circuitry 10 and the memory 12 may be

coupled/connected to one another such that the processing circuitry can read data from and write data to the memory 12.

The processing circuitry 10 may be any type of processing circuitry. For example, the
5 processing circuitry 10 may comprise one or more programmable processor 10A that interprets computer program instructions and processes data. Alternatively or additionally, the processing circuitry 10 may comprise, for example, programmable hardware with embedded firmware. The processing circuitry 10 may be termed processing means or means for processing.

10

The at least one non-transitory computer readable memory medium 12 has stored therein computer program instructions 12A which may provide the logic and routines that enables the image processing apparatus 20 to perform the functionality described below with reference to Figure 2 to 6. The computer program instructions 117, 127 may be pre-
15 programmed into the image processing apparatus 20. Alternatively, they may arrive at the image processing apparatus 20 via an electromagnetic carrier signal or be copied from a physical entity such as a non-transitory computer-readable memory medium 50. Examples of such a non-transitory computer-readable memory medium 50 (which may be referred to as a computer program product) include a non-volatile electronic memory
20 device (e.g. flash memory) and a record medium such as a CD-ROM or DVD.

The memory 12 may comprise both volatile memory and non-volatile memory. The computer program instructions 12A are typically, stored in the non-volatile memory and are executed by the processing circuitry using the volatile memory for temporary storage
25 of data 12B. Examples of volatile memory include RAM, DRAM, SDRAM etc. Examples of non-volatile memory include ROM, PROM, EEPROM, flash memory, optical storage, magnetic storage, etc.

The image processing apparatus 20 may be a single integrated circuit. It may alternatively
30 be provided as a set of integrated circuits (i.e. a chipset). The image processing apparatus 20 may alternatively be a hardwired, application-specific integrated circuit (ASIC). The image processing apparatus 20 may, in some embodiments, form part of a master controller of a device or apparatus which performs other functions which are not related to image processing.

35

The electronic device 1 may, in some embodiments, comprise first and second camera modules 14, 16 configured to capture a pair of stereo images. The first camera module 14

- 10 -

may be referred to as the left camera module and the second camera module 16 may be referred to as the right camera module. The camera modules 14, 16 may be physically configured side-by-side so as to simulate human binocular vision. Together they may be referred to as a stereo camera. The distance between the lenses of each camera module
5 (the intra-axial distance) may be, for example, the distance between an average person's eyes (known as the intra-ocular distance), which is about 6.35 cm. Alternatively, the camera modules may be configured to have a greater or lesser intra-axial distance.

Figures 5A and 5B show two images of the same scene, each of which has been captured by
10 a different one of a pair of stereoscopically-configured camera modules, such as those 14, 16 of Figure 1. Specifically, the image of Figure 5A has been captured by a left camera module 14 and the image of Figure 5B has been captured by a right camera module 16. As can be seen from Figures 5A and 5B, foreground objects in right stereo image of Figure 5B appear closer to the left-hand edge of the image than they do in the left stereo image of
15 Figure 5A. Conversely, foreground objects in the left stereo image appear closer to the right-hand edge of the image than they do in the right stereo image. This difference between the locations of objects in each of a pair of stereo images is known as disparity.

Returning now to Figure 1, each of the camera modules 14, 16 are configured to provide
20 image data representative of their respective image to the image processing apparatus 20. More specifically, the first, or left, camera module 14 is configured to provide first or left stereo image data I_L and the second, or right, camera module 16 is configured to provide second or left image data I_R .

25 The camera modules 14, 16 may be operable under the control of the image processing apparatus 20, which may control the operation of the camera module 14, 16. For instance, the image processing apparatus 20 may control the camera modules to capture their respective images. In other examples, the camera modules 14, 16 may be operable under the control of a different controller, such as a master controller of the electronic device 1
30 (not shown).

The lens position of the camera modules 14, 16 may be individually controllable such that the lens of each camera module can be set to a different position. Alternatively, the lens position of the camera modules 14, 16 may not be controllable but may be fixed at
35 respective different values. The effect of each camera module 14, 16 having a different lens position is discussed in more detail below. The lens position of each camera module 14, 16 may be un-calibrated or unknown.

In this specific example, the electronic device 1 further comprises a display 18 and a user input interface 17 (UII). The display 18 may be operable to output images to a user. Such images may include the images captured by the camera modules 14, 16 and or other image data provided by the image processing apparatus 20. The display 18 may operate under the control of the processing circuitry 10 of the image processing apparatus 20 or a different controller (not shown) such as a master controller of the device 1.

The UII 17 is configured to respond to inputs received by the user and to output signals indicative of these. These signals may be provided to the processing circuitry 10 of the image processing apparatus 20, which may perform operations in accordance with the inputs. Alternatively, the signals may be provided to a master controller 10 (not shown) which controls other components of the device 1 accordingly. The UII 17 may be of any suitable type. For instance, the UII 17, may be a touch-sensitive device which may be overlaid on the display 18 to form a touchscreen.

The electronic device 1, in this example, further comprises a transceiver (RX/TX) 19. The transceiver 19 enables communication with other electronic devices, for example via a wireless communication network. The transceiver 19 may be operable under the control of the image processing apparatus 20 or any other controller 10. The transceiver 19 may be operable to transmit data, such as image data, that is output by the image processing apparatus 20. Similarly, the transceiver 19 may be configured to provide received data to the image processing apparatus 20.

Figure 2 shows a schematic configuration view of an image processing system 2 (which may be part of the electronic device 1 of Figure 1) including the image processing apparatus 20 and the left and right camera modules 14, 16 for generating the left and right image data I_L , I_R . In other embodiments, the image processing system 2 may comprise only the image processing apparatus 20 and the left and right image data I_L , I_R may be received from an external source, for instance from a network via the transceiver 19 (not shown in Figure 2) or from an external computer-readable memory medium 50 such as that shown in Figure 1. An output of the image processing apparatus 20 may be sent to a memory medium (e.g. medium 50 shown in Figure 1) for storage or to the transceiver 19 (or simply a transmitter) for transmission to another apparatus via a network.

Figure 3 is a schematic illustration of various image processing functions/operations which may be performed by the image processing apparatus 20 of Figures 1 and 2. In

some embodiments, the various functions/operations may be performed by one or more portions of computer-readable instructions 12A stored in the memory 12 and executed by the processing circuitry 10.

5 The inputs to the image processing apparatus 20 are, in this example, left stereo image data I_L and right stereo image data I_R . The left stereo image data I_L is representative of an image captured by the leftmost of two camera modules, such as the left and right camera modules 14, 16 depicted in Figures 1 and 2, which are configured to capture stereo images. Similarly, the right stereo image data I_R is representative of an image captured by the
10 rightmost of two camera modules. An example of a left image is shown in Figure 5A, with the corresponding right image being shown in Figure 5B.

The left and right stereo images are asymmetrically defocused. Put another way, the cameras 14, 16 used to capture the left and right stereo image data I_L , I_R are configured to
15 have different focuses because of different lens positions of the two camera modules. As such, one of the stereo images (in the example of Figures 5A and 5B, the left stereo image) is focused on the foreground and the other image (in the example of Figures 5A and 5B, the right stereo image) is focused on the background. This can be seen in Figures 5A and 5B in which the marker pen in the foreground is in focus in the left stereo image of Figure
20 5A and the chess board in the background is in focus in right stereo image of Figure 5B. Such asymmetrically defocused stereo images may be referred to as a “defocused stereo image pair”.

The focal length and/or focus setting (lens position) associated with each of the stereo
25 images may be un-calibrated (or unknown). This is because the embodiments described herein associate the blur level associated with an object in an image with its stereo disparity. As mentioned above, the disparity of an object in an image is related to its distance from the camera and as such, embodiments described herein enable a relationship between the blur level and the distance from the camera to be determined.

30 Although it has been described that the left and right image data I_L , I_R is captured using two different camera modules 14, 16, this may not be the case. In other embodiments, the first and second image data I_L , I_R may be captured by a single camera module whose location and orientation are adjusted between capture operations. The rotation and
35 translation of the camera may then be recovered, and the images rectified, thereby to simulate a stereo camera configuration. For instance, the image capture could be carried

- 13 -

out using existing smart phones or tablet computers which have only a single camera module facing in a particular direction.

5 The images originally captured by the one or more camera modules 14, 16 may have undergone some pre-processing in order to produce the left and right stereo images. For instance, rectification may have been applied to the captured images in order to produce the stereo image pair which is acquired by the image processing apparatus 20. This may be particularly useful when the images are captured using a single camera module as it is more likely that the captured images may not be entirely horizontally aligned.

10

As can be seen from the left and right images of Figures 5A and 5B, there are certain regions of each image which are not visible in the other image. This is because foreground objects obscure different regions of the background in each image. These regions, which only appear in a single one of the stereo images are known as “occlusion” or “occluded” regions (or “single-occluded” regions), which are made up of one or more occluded pixels. For illustrative purposes, one occluded region in each image has been highlighted using white rectangles.

15

As will be understood from the below explanation, existing techniques for estimating the disparity of stereo images do not produce good results in occlusion regions. This is because, in the occlusion regions, there is no pixel correspondence between the stereo images.

20

Returning now to Figure 3, in this example, the processing apparatus 20 comprises a disparity map generator 201 (which may also be referred to as a disparity map generation function or means for generating a disparity map). The disparity map generator 201 is configured to receive and process the left and right stereo image data I_L , I_R and to output data sets representative of first and second disparity maps D_{L-R} , D_{R-L} . For reasons which will be understood from the below discussion, the first disparity map may be referred to as a “left-to-right” (or simply “left”) disparity map and the second disparity map may be referred to as a “right-to-left” (or simply “right”) disparity map.

25

30

The disparity map generator 201 may be configured to generate the disparity map data D_{L-R} , D_{R-L} by applying a technique known as “stereo matching”. Any known stereo matching technique may be used. However, when such known techniques are used to generate the disparity map data D_{L-R} , D_{R-L} , the disparity map generator 201 may be configured to treat the left and right images as being “all-in-focus”, even though this is not actually the case.

35

The stereo matching performed by the disparity map generator 201 may comprise, for each pixel in one of the images, identifying all possible corresponding pixels in the other image and then determining which one of the possible corresponding pixels in the other image is the best match for the pixel under consideration. The displacement between the location of the pixel under consideration and the location of the best matching pixel in the other image is then calculated. The disparity map data D_{L-R} , D_{R-L} comprises, for each pixel in each image, a value indicative of the displacement between the pixel and the best matching pixel in the other image.

10

The left-to-right disparity map data D_{L-R} comprises, for each pixel in the left stereo image, a value of the displacement (or disparity) between the pixel in the left stereo image and the best matching pixel in the right stereo image. The right-to-left disparity map data D_{R-L} comprises, for each pixel in the right stereo image, a value of the displacement (or disparity) between the pixel in the right stereo image and the best matching pixel in the left stereo image.

15

Figure 5C shows an example of a left-to-right disparity map calculated using the left and right stereo images of Figures 5A and 5B and Figure 5D shows an example of the corresponding right-to-left disparity map. In these disparity maps, the lighter shades of grey represent larger disparity (or displacements).

20

Ideally, in the non-occlusion regions, each pixel from one image has one, and only one, corresponding pixel in the other image. In contrast, occlusion regions of an image are constituted by pixels which do not have a corresponding pixel in the other image (i.e. occluded pixels). However, because the disparity map generator 201 does not know the location of the occlusion regions when it performs the stereo matching, it nonetheless identifies the best matching pixels for those occluded pixels in the occlusion region. This can result in a wide range of, mostly incorrect, disparities being determined for the pixels in the regions of the disparity maps which correspond to the occlusion regions of the stereo images. This can be seen in Figures 5C and 5D in which there are pixels of a number of different shades (i.e. different disparities) within the regions of the disparity maps which correspond to the occlusion regions in the left and right images. For simplicity, the regions in the disparity maps (and indeed in the below-discussed blur maps, occlusion maps, and blur-disparity maps) which correspond to occlusion regions in the left and right stereo images may also be referred to as occlusion regions, with the

30

35

- 15 -

pixels therein also being referred to as occluded pixels. Examples of occlusion regions are denoted in the disparity maps of Figures 5C and 5D using dashed rectangles.

Although the disparity map generator 201 may be configured, as described above, to
5 assume the stereo images are “all-in-focus”, it may in other examples be configured to use a disparity estimation approach which takes into account the fact that the stereo images are asymmetrically focused. In such examples, the disparity map generator 201 may be configured to estimate disparity using an approach such as that described in “Dual-focus
stereo imaging” by F. Li, J. Sun, J. Wang, and J. Yu (J. Electron. Imaging. 19(4), 043009,
10 February 09, 2010, July 26, 2010, August 30, 2010, December 06, 2010).

The image processing apparatus 20 in the example of Figure 3 comprises first and second occlusion map generators 202, 203. The first occlusion map generator 202 is configured to receive the first (or left) stereo image data I_L and the first (or left-to-right) disparity map
15 data D_{L-R} and to output data representative of the first (or left-to-right) occlusion map O_{L-R} . The first occlusion map generator 202 may be referred to as the “left-to-right occlusion map generator”.

The second occlusion map generator 203 is configured to receive the second (or right)
20 stereo image data I_R and the second (or right-to-left) disparity map data D_{R-L} and to output data representative of the second (or right-to-left) occlusion map O_{R-L} . The second occlusion map generator 203 may be referred to as the “right-to-left occlusion map generator”.

25 The occlusion map data corresponding to a one of the stereo images comprises, for each pixel in that image, a value which indicates whether or not the pixel has a corresponding pixel in the other stereo image. More specifically, each occlusion map is a binary image in which the first value is assigned to pixels in non-occlusion regions and in which the second value is assigned to pixels in the occlusion regions. Put another way, the value of each
30 pixel in an occlusion map indicates whether or not that pixel is an occluded pixel. An example of an occlusion map is illustrated in Figure 5E, with the white regions representing the occlusion regions and the black regions representing the non-occlusion regions. The occlusion map of Figure 5E is the right-to-left (or simply right) occlusion map and so is generated by the right-to-left occlusion map generator 202 on the basis of
35 the data I_R representative of the right stereo image and the data D_{R-L} representative of the right-to-left disparity map. In a right-to-left occlusion map, such as that of Figure 5E, pixels of the second value (in this example, the white pixels) indicate pixels in the right

stereo image, which do not have a corresponding pixel in the left stereo image. Similarly, in a left-to-right occlusion map, pixels of the second value would indicate pixels in the original left stereo image, which do not have a corresponding pixel in the right stereo image.

5

Each occlusion map generator 202, 203 is configured to generate the occlusion map data O_{L-R} , O_{R-L} by transforming each pixel in the received stereo image in accordance with the disparity value of the corresponding pixel in the received disparity map. In the resulting image (i.e. the image that is produced following transformation of each pixel from the received stereo image), those pixels at locations onto which one or more pixels from the received stereo image have been transformed are set to the first value (i.e. to indicate that they are non-occluded pixels) and those pixels at locations onto which none of the pixels of the received stereo image are transformed are set to the second value (i.e. to indicate that they are occluded pixels). The post-transformation processing that is applied may, for example, be expressed as follows:

15

$O(x,y) = 0$ (zero) if pixel (x,y) has at least one source pixel following the transformation; and

$O(x,y) = 1$ (one) if pixel (x,y) has no source pixels following the transformation,

20

where $O(x,y)$ is the pixel at location x, y in the occlusion map.

As will be appreciated from the above, the left-to-right occlusion map generator 202 is configured to transform each pixel from the left image in accordance with the disparity value of the corresponding pixel of the left-to-right disparity map and then to perform the above-described post-transformation processing. Similarly, the right-to-left occlusion map generator 203 is configured to perform the above-described post-transformation processing after transforming each pixel from the right stereo image in accordance with the disparity value of the corresponding pixel of the right-to-left disparity map.

30

The image processing apparatus may further comprise first and second blur map generators 204, 205. The blur map generators 204, 205 are together configured to receive the data I_L , I_R representative of the defocused stereo image pair and to output a pair of blur maps, one corresponding to each of the stereo images. Specifically, the first blur map generator 204 is configured to receive the first (or left) stereo image data I_L and to output data representative of the first (or left) blur map B_L . The second blur map generator 205 is

35

- 17 -

configured to receive the second (or right) stereo image data I_R and to output data representative of the second (or right) blur map B_R .

The blur map generators 204, 205 may be configured to generate the blur maps using any
5 known “depth-from-defocus” technique. For example, one of the techniques described in
“Learning to Estimate and Remove Non-uniform Image Blur” (2013 IEEE Conference on
Computer Vision and Pattern Recognition) by F. Couzinie-Devy, J. Sun; K. Alahari and J.
Ponce, or “Absolute Depth Estimation from a Single Defocused Image” (IEEE
Transactions on Image Processing, Volume22 , Issue 11) by J. Lin, X. Ji, W. Xu, and Q.
10 Dai, which are hereby incorporated by reference, may be used

In the resultant sets of blur map data B_L , B_R , the value for each pixel is dependent on the
measured blur associated with that pixel. Put another way, the blur map data
corresponding to a one of the stereo images comprises, for each pixel in that image, a
15 value which indicates the level of blur (or the blur value) associated with that pixel. Figure
5F shows an example of the blur map for the right stereo image of Figure 5B. In this
example, the darker the pixel, the higher the blur value associated with that pixel.

The image processing apparatus 20 comprises first and second blur-disparity map
20 generators 206, 207. The first and second blur-disparity map generators 206, 207 are
configured to output first and second sets of blur-disparity map data BD_L , BD_R , each set
corresponding to a different one of the stereo images. In general terms, the blur disparity
map data for a one of the stereo images comprises, for each pixel of the stereo image, a
mean blur value which is determined based the disparity value of the pixel

25
The first blur-disparity map generator 206 is configured to receive the first (or left)blur
map data B_L , the first (or left-to-right) disparity map data D_{L-R} and the first (or left-to-
right) occlusion map data O_{L-R} and to generate, based on the received data sets, data BD_L
representative of a first (or left) blur-disparity map. The first blur-disparity map
30 generator 206 may be referred to as the “left blur-disparity map generator”.

The second blur-disparity map generator 207 is configured to receive the second (or right)
blur map data B_R , the second (or right-to-left) disparity map data D_{R-L} and the second (or
right-to-left) occlusion map data O_{R-L} and to generate, based on the received data sets,
35 data BD_R representative of a second (or right) blur-disparity map. The second blur-
disparity map generator 207 may be referred to as the “right blur-disparity map
generator”.

A theoretical relationship exists between the blur and disparity of pixels of a defocused stereo image pair. This theoretical relationship can be derived from the well-known thin lens model and the disparity model and is as set out in Equation, 1 below. “Dual-focus stereo imaging” by F. Li , J. Sun, J. Wang, and J. Yu (J. Electron. Imaging. 19(4)) and “Depth estimation and image restoration using defocused stereo pairs,” by A. N. Rajagopalan, S. Chaudhuri, and U. Mudenagudi, (IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 26, no. 11, pp. 1521–1525, 2004) are both relevant to the derivation of this equation.

$$b = D |\gamma_p - \gamma| / (B - \gamma) \quad (\text{Equation 1})$$

where b = blur radius for a specific pixel;

γ_p = disparity for the corresponding pixel in the defocused stereo image (e.g. I_L);

γ = disparity for a corresponding pixel in a theoretical “all-in-focus” image;

D is a function of the camera module focal length and aperture; and

B is a function of the stereo baseline and intrinsic parameters.

Figure 6 is a graph showing this theoretical, piecewise linear relationship between the disparity γ_p of a pixel and the blur radius b of that pixel, when B and D are assumed to be constant.

In reality, because of the error in the disparity estimation and noise in the stereo images and the single image blur estimation, the relationship between the blur radius and stereo disparity may not be linear. Consequently, the above theoretical relationship is not used. Instead, the blur-disparity map generators 206, 207 are configured to estimate the blur-disparity mapping in the following way:

Firstly, each blur-disparity map generator 206, 207 is configured to use the received occlusion map data O_{L-R} , O_{R-L} to identify the pixels in the occlusion regions (the “occluded pixels”) and the pixels outside the occlusion regions (the “non-occluded pixels”). Next, for all the non-occluded pixels in the received disparity map, the blur-disparity map generator 206, 207 is configured to identify all pixels that have the same disparity value.

Subsequently, for all pixels having a particular disparity value, the blur-disparity map generator 206, 207 sums the blur values of the corresponding pixels in the received blur map and then divides the total by the number of pixels having that particular disparity value. Put another way, the blur-disparity map generator 206, 207 is configured to

- 19 -

calculate the “mean blur” for all non-occluded pixels having the particular disparity value. All non-occluded pixels which contribute towards a particular mean blur value may be referred to as the contributing set of pixels for that mean blur value.

5 Calculation of the mean blur is repeated for each different disparity value, such that each different disparity value has a corresponding mean blur value. The blur-disparity map generator 206, 207 then generates the non-occlusion regions of the blur disparity map by assigning the corresponding mean blur value to all the pixels which correspond to pixels in the disparity map which have a particular disparity value. It will thus be appreciated that,
10 if two non-occluded pixels in the disparity map have the same disparity, the corresponding pixels in the blur-disparity map will also have the same value, that same value being the mean-blur value corresponding to the disparity.

A partial blur-disparity map for only the non-occluded regions for the right image is
15 shown in Figure 5G. In this example, the occluded regions are shown in black.

Next, the blur-disparity map generator 206, 207 generates the blur-disparity map pixel values for the occluded pixels.

20 In some embodiments, this may be performed by identifying from the blur map data B_L , B_R the blur value for a particular occluded pixel and then identifying, out of all mean blur values which were previously-calculated for the non-occluded pixels, the mean blur value which is closest or most-similar to the identified blur value. The pixel value for the particular occluded pixel in the blur-disparity map is then set as the identified closest
25 mean blur value. This process is then repeated for all occluded pixels.

However, this approach for generating the blur-disparity map pixel values for the occluded pixels may, in some instances, be suboptimal. This is because each mean blur value is calculated on the basis of many different original blur values, which may be spread over a
30 relatively large range. Consequently, generating the blur-disparity map pixel values for the occluded pixels in this way may produce a noisy result. As such, in some alternative embodiments, the blur-disparity map generator 206, 207 may be configured to generate the blur-disparity map pixel values for the occluded pixels in a different way.

35 More specifically, the blur-disparity map generator 206, 207 may be configured, when calculating the mean blur values for the non-occluded pixels, to count and store, for each different disparity value, the numbers of pixels which have each different blur value. Put

- 20 -

another way, the blur-disparity map generator 206, 207 may determine the numbers of pixels which have each different blur value that contribute towards a particular mean blur. Put yet another way, the blur-disparity map generator 206, 207 may determine the number of pixels in the contributing set of pixels for a mean blur value which have each
 5 different blur value.

The blur-disparity map generator 206, 207 then sets the pixel value for a particular occluded pixel in the blur disparity-map as the mean blur value whose contributing set of pixels had the largest number of pixels having the same blur value as the occluded pixel
 10 under consideration. This process is then repeated for each occluded pixel until each pixel in the blur-disparity map is assigned a mean blur value.

By way of example, let us consider two different disparities \mathbf{d}_1 and \mathbf{d}_2 and let us assume that there are 80 non-occluded pixels having each disparity. For the first disparity
 15 \mathbf{d}_1 , there are 10 pixels having a first blur value \mathbf{b}_1 , 50 pixels having a second blur value \mathbf{b}_2 and 20 pixels having a third blur value \mathbf{b}_3 and, consequently, the mean blur value for the pixels having first disparity \mathbf{d}_1 is $\mathbf{M}_1 = (10\mathbf{b}_1 + 50\mathbf{b}_2 + 20\mathbf{b}_3)/80$. For the second disparity \mathbf{d}_2 , there are 40 pixels having the first blur value \mathbf{b}_1 , 30 pixels having the second blur value \mathbf{b}_2 and 10 pixels having a third blur value \mathbf{b}_3 and, consequently, the mean blur for the
 20 second disparity \mathbf{d}_2 is $\mathbf{M}_2 = (40\mathbf{b}_1 + 30\mathbf{b}_2 + 10\mathbf{b}_3)/80$.

When the blur-disparity map generator 206, 207 estimates a pixel value for an occluded pixel in the blur-disparity map, the blur value of which is the first blur value \mathbf{b}_1 , it determines which of the first and second mean blur values \mathbf{M}_1 , \mathbf{M}_2 has the largest number
 25 of contributing pixels having the first blur value \mathbf{b}_1 . It then sets the pixel value of the occluded pixel in the blur-disparity map as the determined mean blur value. In this case, there were 10 pixels having the first blur value \mathbf{b}_1 which contributed to the first mean blur value \mathbf{M}_1 and 40 pixels having the first blur value \mathbf{b}_1 which contributed to the second mean blur value \mathbf{M}_2 . As a higher number of pixels having the first blur value \mathbf{b}_1
 30 contributed to the second mean blur value \mathbf{M}_2 , the value of the occluded pixel in the blur-disparity map is set to the second mean blur value \mathbf{M}_2 .

If none of the mean blur values have, in their contributing set of pixels, a pixel having the same blur value as an occluded pixel, the blur-disparity map generator 206, 207 is
 35 configured to assign to the corresponding occluded pixel in the blur-disparity map the mean blur value which is closest to the blur value of the occluded pixel. In terms of the above-described example, let us assume that an occluded pixel in the blur map has a

- 21 -

fourth blur value \mathbf{b}_4 . In this case, as neither of the first and second mean blur values \mathbf{M}_1 and \mathbf{M}_2 have in their contributing sets a pixel having the fourth blur value \mathbf{b}_4 , the blur-disparity map generator 206, 207 determines which of the mean blur values \mathbf{M}_1 and \mathbf{M}_2 is closest, or most similar, to the fourth mean blur value \mathbf{b}_4 . The pixel of the blur-disparity map which corresponds to the pixel in the blur map having the fourth pixel value \mathbf{b}_4 is then set to the determined closest mean blur value.

Figure 5H shows the complete blur-disparity map (including the occluded regions) for the right stereo image, the pixel values for the occluded regions of which were estimated using the second-described approach.

The image processing apparatus 20 may further comprise first and second disparity map updaters 208, 209.

The first disparity map updater 208 is configured to receive the first (or left) blur-disparity map data BD_L , the first (or left-to-right) disparity map data D_{L-R} and the first (or left-to-right) occlusion map data O_{L-R} and to generate, based on the received data sets, data UD_{L-R} representative of a first (or left-to-right) updated disparity map. The first disparity map updater 208 may be referred to as the “left-to-right disparity map updater”.

The second disparity map updater 209 is configured to receive the second (or right) blur-disparity map data BD_R , the second (or right-to-left) disparity map data D_{R-L} and the second (or right-to-left) occlusion map data O_{R-L} and to generate, based on the received data sets, data UD_{R-L} representative of a second (or right-to-left) updated disparity map. The second disparity map updater 209 may be referred to as the “right-to-left disparity map updater”.

Each disparity map updater 208, 209 is configured to update its respective disparity map as follows:

Firstly, the disparity map updater 208, 209 calculates, using the received blur-disparity map data BD_L , BD_R , for each occluded pixel op in the blur-disparity map data, a combined difference, CD , between the occluded pixel op and any other non-occluded pixel q within an $M \times N$ neighborhood. The combined difference between the occluded pixel op and a non-occluded pixel q is the weighted average of the absolute difference between the mean blur of the occluded pixel op and the mean blur of the non-occluded pixel q and the geometric Euclidian distance of the two pixels. More specifically, the combined difference

- 22 -

between an occluded pixel in the blur-disparity map and a non-occluded pixel q within the $M \times N$ neighbourhood may be determined using the following equation:

$$CD = \frac{1}{(a + b)} \left(a \times |MeanBlur(op) - MeanBlur(q)| + b \times \sqrt{[x(op) - x(q)]^2 + [y(op) - y(q)]^2} \right) ($$

Equation 2)

5

where: a and b are pre-defined constants for weighting;
 $x(op)$, $y(op)$ are coordinates of the occluded pixel op ; and
 $x(q)$, $y(q)$ are coordinates of the non-occluded pixel q .

10 Next, for each possible mean blur value within the $M \times N$ neighbourhood of the occluded pixel op , the disparity map updater 208, 209 calculates a corresponding mean blur difference, MBD (or mean combined difference). The mean blur difference for a particular mean blur value is calculated by summing all combined differences CD of all non-occluded pixels within the $M \times N$ neighbourhood of the occluded pixel op which have the particular
 15 mean blur value and by normalizing the result using the total number of such non-occluded pixels.

The disparity map updater 208, 209 then assigns the mean blur value which has the lowest mean blur difference as the mean blur value for the occluded pixel op . The
 20 disparity map updater 208, 209 then updates the value of the occluded pixel op in the disparity map to the disparity value which corresponds to the assigned mean blur value. The corresponding disparity value is determined using the blur-disparity map.

This process is repeated for each occluded pixel in the disparity map and the result is the
 25 updated disparity map. An example of the right-to-left updated disparity map can be seen in Figure 5I.

It should be noted that, theoretically, there may be two different disparity values for a given blur, as illustrated in Figure 6. However, the image processing apparatus 20 is
 30 configured to assume that one image of the defocused stereo pair is focused on the foreground and that the other image is focused on the background. Consequently, it is clear, for each image, which of the two possible disparity values is correct.

In some examples, the image processing apparatus 20 further comprises a noise reduction
 35 module 210. The noise reduction module 210 is configured to receive the first and second

- 23 -

updated disparity map data UD_{L-R} , UD_{R-L} and to process the data to reduce the noise therein.

5 The noise reduction module 210 may be configured to apply at least one filter to each of the sets of data thereby to reduce the noise. More specifically, the noise reduction module 210 may apply a first filter to the first and second updated disparity map data UD_{L-R} , UD_{R-L} . In some examples it may further apply a second filter to the first-filtered updated disparity map data. Each of the first and second filters may be applied one or more times. The first filter may, in some examples, be one of a morphological erosion filter and a
10 morphological dilation filter. The second filter may be the other one of the morphological erosion filter and a morphological dilation filter. In one specific example, the morphological erosion filter may be applied one or more times to each of the sets of updated disparity map data UD_{L-R} , UD_{R-L} and, subsequently, the morphological dilation filter may be applied the same number of times to the morphologically eroded updated
15 disparity data. In other examples, different de-noising filters, for instance a median filter, may be used.

The output from the noise reduction module 210 is first and second sets of improved disparity map data ID_{L-R} , ID_{R-L} . The sets of improved disparity map data ID_{L-R} , ID_{R-L} may
20 be used for, for example, determining the distance of one or more objects in the stereo images from the camera modules. Other uses of the improved disparity map include, for example, 3D rendering, region segmentation and scene analysis.

Although the schematic illustration of Figure 3 shows the image processing apparatus 20
25 receiving the original image data I_L , I_R , and outputting the improved disparity map data ID_{L-R} , ID_{R-L} , in some other examples, image processing apparatus 20A according to examples of embodiments may instead receive pre-generated occlusion map O_{L-R} , O_{R-L} and blur data B_L , B_R in addition to the disparity map data D_{L-R} , D_{R-L} and may output the updated disparity map data UD_{L-R} , UD_{R-L} . Put another way, image processing apparatus
30 20A according to examples of the invention may comprise only one or both of the first and second blur-disparity map generators 206, 207 and one or both of the first and second disparity map updaters 208, 209. In other examples, the image processing apparatus may further comprise the noise reduction module 210 in addition to the blur-disparity map generators 206, 207 and the disparity map updaters 208, 209. Alternatively or
35 additionally, the image processing apparatus may additionally comprise any combination of the disparity map generator 201, the occlusion map generators 202, 203 and the blur map generators 204, 205.

As will, of course, be appreciated each of the functional blocks illustrated in Figure 4 may be referred to as “means for” performing the function of that functional block. As such, the image processing apparatus 20A may be said to comprise, for example, means 206, 207 for generating first and second blur-disparity map data and means 208, 209 for updating first and second disparity map data.

It will also be understood that the schematic of Figure 3 is purely for illustrative purposes. As such, although in Figure 3 a function (e.g. disparity map generation by the disparity map generator 201) is shown to be performed by a single functional block, it will be understood that the function may instead be performed by two or more functional blocks. Similarly, although particular functions are shown to be performed by separate functional blocks, they may instead be performed by a single functional block. For instance, the left and right blur map generation may be performed by a single functional block instead of the two blocks 204, 205 illustrated in Figure 3. Likewise, although two occlusion map generators 202, 203 have been illustrated in Figure 3, these may be replaced by a single occlusion map generator which outputs both sets of occlusion map data O_{L-R} , O_{R-L} . In addition or alternatively, functions of different types (e.g. blur map generation and occlusion map generation) may be performed by a single functional block.

Figure 4 is a flow chart illustrating various operations which may be performed by the image processing apparatus 20.

In operation S5-1, the left and right (or first and second) stereo image data I_L , I_R is acquired. The left and right images are a “defocused pair” of stereo images. Acquiring the image data may occur in any suitable way. For instance, the image processing apparatus 20 may control first and second camera modules 14, 16 to capture the image data I_L , I_R and to pass it to the image processing apparatus 20. Alternatively, the stereo image data I_L , I_R may be received at the image processing apparatus 20 via a wired or wireless connection with a network or via a removable memory medium 50. In another alternative, the stereo image data I_L , I_R may be captured using a single camera module whose position and orientation is altered between capturing each of the pair of images. The first and second sets of image data may have been subjected to some processing between capture and provision to the image processing apparatus 20. For instance, the captured images could be subjected to rectification to generate the stereo image data received at the image processing apparatus 20.

- 25 -

In operation S5-2, the image processing apparatus 20 generates the left-to-right (or left or first) disparity map data D_{L-R} and the right-to-left (or right or first) disparity map data D_{R-L} based on the acquired stereo image data I_L, I_R . This may be performed as described with reference to the disparity map generator 201 of Figure 3.

5

In operation S5-3, the image processing apparatus 20 processes the left stereo image data I_L based on the left-to-right disparity map data D_{L-R} to generate the left-to-right (or left or first) occlusion map data O_{L-R} . The image processing apparatus 20 also processes the right image data I_R based on the right-to-left disparity map data D_{R-L} to generate the right-to-
10 left (or right or second) occlusion map data O_{R-L} . The occlusion map generation may be performed as described with reference to the occlusion map generators 202, 203 of Figure 3.

In operation S5-4, the image processing apparatus 20 processes the left stereo image data
15 I_L to produce the left (or first) blur map data. The image processing apparatus also processes the right stereo image data I_R to produce the right (or left) blur map data. Generation of the blur map data may be performed using any depth-from-defocus technique.

20 In operation S5-5, the image processing apparatus 20, 20A generates the blur-disparity map data for the non-occluded regions of the respective stereo images I_L, I_R .

In operation S5-6, the image processing apparatus 20, 20A generates the blur-disparity map data for the occluded regions of the respective stereo images I_L, I_R . This is combined
25 with the blur-disparity map data for the non-occluded regions to form the complete blur-disparity map data BD_L, BD_R .

The left (or first or left-to-right) blur-disparity map data BD_L is generated, in operations S5-5 and S5-6, based on the left-to-right disparity map data D_{L-R} produced in operation
30 S5-2, the left-to-right occlusion map data O_{L-R} generated in operation S5-3, and the left blur map data B_L generated in operation S5-4. The right (or second or right-to-left) blur-disparity map data BD_R is generated, in operations S5-5 and S5-6, based on the right-to-left disparity map data D_{R-L} produced in operation S5-2, the right-to-left occlusion map data O_{R-L} generated in operation S5-3, and the right blur map data B_L generated in
35 operation S5-4. The blur-disparity map data BD_L, BD_R for the non-occluded and occluded regions is generated as described with reference to blur-disparity map generators 206, 207 of Figure 3.

In operation S5-7, the image processing apparatus 20, 20A updates the disparity map data D_L , D_R generated in operation S5-2, to produce the updated disparity map data UD_L , UD_R . More specifically, the updated left-to-right disparity map UD_{L-R} data is calculated using
5 the left blur-disparity map data BD_L , the left-to-right occlusion map data O_{L-R} , and the left-to-right disparity map data D_{L-R} . The updated right-to-left disparity map UD_{R-L} data is calculated using the right blur-disparity map data BD_R , the right-to-left occlusion map data O_{R-L} , and the right-to-left disparity map data D_{R-L} . The calculation of the updated disparity map data UD_{L-R} , UD_{R-L} is performed as described with reference to the disparity
10 map updaters 208, 209 of Figure 3.

In operation S5-8, the image processing apparatus 20 applies a first filter to the sets of updated disparity map data UD_{L-R} , UD_{R-L} . One or multiple iterations of the filtering of operation S5-8 may be performed. The first filter may be one of a morphological erosion
15 filter and a morphological dilation filter.

In operation S5-9, the image processing apparatus 20 applies a second filter to the first-filtered sets of updated disparity map data UD_{L-R} , UD_{R-L} . One or multiple iterations of the filtering of operation S5-9 may be performed. The same number of iterations of the
20 second filtering may be performed as were performed for the first filtering. The second filter may be the other one of a morphological erosion filter and a morphological dilation filter. In some embodiments, the morphological erosion filter is applied in operation S5-8 and the morphological dilation filter is subsequently applied in operation S5-9.

25 The result of operation S5-9 is the first and second sets of improved disparity map data ID_{L-R} , ID_{R-L} .

It will of course be appreciated that the various operations of the flow chart of Figure 4 may be performed in a different order and/or concurrently with each other. Similarly, one
30 or more of the operations may be omitted or combined. For instance, in some embodiments, at least one of operations S5-1, S5-2, S5-3, S5-4 may be performed by a different image processing apparatus to that which performs operations S5-5 to S5-7. Alternatively or additionally, operations S5-8 and S5-9 may be performed by a different processing apparatus.

35

In some alternative examples, the disparity values for occluded pixels may be determined using only the mean blur values from the blur disparity map data (in other words without

- 27 -

calculating the combined difference and mean blur difference etc). However, these examples may provide less accurate results than when the above described method is used. This is because the above-described method not only takes into account the observation that the disparity of an occluded pixel op will be influenced by the disparity of non-occluded pixel p with similar blur (no matter how far away p is from op) but also takes into account the fact that the further pixel p is from occluded pixel op , the higher the possible error that is associated with using the disparity of p to update op (even though the pixels have similar blur).

10 It should be understood that the terms “stereo image pair” and “pair of stereo images” as used herein are not limited to pairs of images in which objects are only horizontally offset from one another. Indeed the terms are also intended to cover pairs of images in which objects are only vertically offset and pairs of images which are both vertically and horizontally offset from one another.

15

Although, the image processing apparatus 20, 20A is shown in Figure 1 as part of, for instance, a portable electronic device it will be appreciated that image processing apparatus may 20, 20A may instead be part of, for example, an image processing server or a personal computer.

20

Reference to “computer-readable memory medium”, “computer program product”, “tangibly embodied computer program” etc, or a “processor” or “processing circuitry” etc. should be understood to encompass not only computers having differing architectures such as single/multi processor architectures and sequencers/parallel architectures, but also specialised circuits such as field programmable gate arrays FPGA, application specify circuits ASIC, signal processing devices and other devices. References to computer program, instructions, code etc. should be understood to express software for a programmable processor firmware such as the programmable content of a hardware device as instructions for a processor or configured or configuration settings for a fixed function device, gate array, programmable logic device, etc.

30

As used in this application, the term ‘circuitry’ refers to all of the following: (a) hardware-only circuit implementations (such as implementations in only analogue and/or digital circuitry) and (b) to combinations of circuits and software (and/or firmware), such as (as applicable): (i) to a combination of processor(s) or (ii) to portions of processor(s)/software (including digital signal processor(s)), software, and memory(ies) that work together to cause an apparatus, such as a mobile phone or server, to perform various functions) and

35

- 28 -

(c) to circuits, such as a microprocessor(s) or a portion of a microprocessor(s), that require software or firmware for operation, even if the software or firmware is not physically present.

5 This definition of ‘circuitry’ applies to all uses of this term in this application, including in any claims. As a further example, as used in this application, the term “circuitry” would also cover an implementation of merely a processor (or multiple processors) or portion of a processor and its (or their) accompanying software and/or firmware. The term
10 “circuitry” would also cover, for example and if applicable to the particular claim element, a baseband integrated circuit or applications processor integrated circuit for a mobile phone or a similar integrated circuit in server, a cellular network device, or other network device.”

15 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.

20 Although various aspects of the invention are set out in the independent claims, other aspects of the invention 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.

25 It is also noted herein that while the above describes example embodiments of the invention, 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 invention as defined in the appended claims.

Claims

1. A method comprising:
estimating disparity values for pixels located in an occluded region of a first one of
5 a defocused pair of stereo images using:
a first set of blur map data corresponding to the first one of the defocused
pair of stereo images;
a first set of disparity map data corresponding to the first one of the
defocused pair of stereo images; and
10 a first set of occlusion map data corresponding to the first one of the
defocused pair of stereo images.
2. The method of claim 1 comprising:
generating a first set of blur-disparity map data corresponding to the first one of
15 the defocused pair of stereo images, the first set of blur-disparity map data being
generated based on:
the first set of disparity map data;
the first set of occlusion map data; and
the first set of blur map data;
20 estimating the disparity values for the pixels located in the occluded region of the
first one of the defocused pair of stereo images using the first set of occlusion map data
and the first set of blur-disparity map data.
3. The method of claim 1 or claim 2 comprising updating the first set of disparity map
25 data with the estimated disparity values to generate a first set of updated disparity map
data.
4. The method of claim 3, comprising:
applying at least one filter to the first set of updated disparity map data.
30
5. The method of claim 4, comprising:
applying a first filter to the first set of updated disparity map data; and
applying a second filter to the first-filtered first set of updated disparity map data.
- 35 6. The method of claim 5, wherein the first filter is one of a morphological erosion
filter and a morphological dilation filter and wherein the second filter is the other of the
morphological erosion filter and the morphological dilation filter.

7. The method of any preceding claim, wherein generating the first set of blur-disparity map data comprises:

using the first set of occlusion map data to identify pixels in the first set of disparity map data which correspond to the non-occluded pixels in the first one of the defocused pair of stereo images, each identified pixel in the set of disparity map data having a disparity value;

for each different disparity value of the identified pixels:

identifying a set of non-occluded pixels in the first set of blur map data, the set of non-occluded pixels in the first set of blur map data corresponding to the identified pixels in the first set of disparity map data which have the different disparity value, each identified non-occluded pixel in the first set of blur map data having a blur value;

calculating the mean blur value of the set of non-occluded pixels in the first set of blur map data; and

assigning the calculated mean blur value to the pixels in the first set of blur-disparity map data which correspond to the identified set of non-occluded pixels in the first set of blur map data.

8. The method of claim 7, comprising, if none of the previously-calculated mean blur values have a contributing pixel having the blur value of an identified occluded pixel in the first set of blur map data, assigning to the occluded pixel in the first set of blur-disparity map data, which corresponds to the identified occluded pixel, the mean blur value which is closest to the blur value of the identified occluded pixel.

9. The method of any preceding claim, wherein each pixel in first set of blur-disparity map data has an associated mean blur value, and wherein estimating the disparity values for the pixels located in the occluded region of the first one of the defocused pair of stereo images comprises:

using the first set of occlusion map data to identify at least one occluded pixel in the first set of blur-disparity map data;

for each occluded pixel in the first set of blur-disparity map data:

calculating a combined difference between the mean blur value of the occluded pixel and the mean blur value of each non-occluded pixel in the first set of blur-disparity map data within an $M \times N$ neighbourhood of the occluded pixel;

for each mean blur value within the $M \times N$ neighbourhood, calculating the mean blur difference; and

- 31 -

identifying the mean blur value having the lowest mean blur difference, wherein the estimated disparity value for the occluded pixel is the disparity value which corresponds to the mean blur value having the lowest mean blur difference.

- 5 10. The method of claim 9, wherein combined difference is a weighted average of the absolute difference between the mean blur value of the occluded pixel and the mean blur value of the non-occluded pixel and the geometric Euclidian distance between occluded pixel and the non-occluded pixel.
- 10 11. The method of claim 9 or claim 10, wherein the mean blur difference for each mean blur value within the $M \times N$ neighbourhood is calculated by summing the combined differences of all the non-occluded pixels having the mean blur value and normalising the result by the total number of the non-occluded pixels having the mean blur value.
- 15 12. The method of any preceding claim, comprising:
generating the first set of disparity map data using the pair of stereo images.
13. The method of any preceding claim comprising:
generating the first set of occlusion map data using the first one of the defocused
20 pair of stereo images and the first set of disparity map data.
14. The method of any preceding claim, comprising:
generating the first set of blur map data using the first one of the defocused pair of
stereo images.
- 25 15. The method of any preceding claim comprising:
estimating disparity values for pixels located in an occluded region of a second one
of a defocused pair of stereo images using:
a second set of blur map data corresponding to the second one of the
30 defocused pair of stereo images;
a second set of disparity map data corresponding to the second one of the
defocused pair of stereo images; and
a second set of occlusion map data corresponding to the second one of the
defocused pair of stereo images.
- 35 16. The method of claim 15 comprising:

- 32 -

generating a second set of blur-disparity map data corresponding to the second one of the defocused pair of stereo images, the second set of blur-disparity map data being generated based on:

the second set of disparity map data;

5 the second set of occlusion map data; and

the second set of blur map data;

estimating the disparity values for the pixels located in the occluded region of the second one of the defocused pair of stereo images using the second set of occlusion map data and the second set of blur-disparity map data.

10

17. The method of claim 15 or claim 16 comprising updating the second set of disparity map data with the estimated disparity values for pixels located in the occluded region of the second one of the defocused pair of stereo images, to generate a second set of updated disparity map data.

15

18. The method of claim 17, comprising:
applying at least one filter to the second set of updated disparity map data.

19. The method of claim 18, comprising:

20 applying a first filter to the second set of updated disparity map data; and
applying a second filter to the first-filtered second set of updated disparity map data.

20. The method of claim 19, wherein the first filter is one of a morphological erosion filter and a morphological dilation filter and wherein the second filter is the other of the morphological erosion filter and the morphological dilation filter.

25

21. An apparatus comprising:

at least one processor; and

30 at least one memory including 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:

estimate disparity values for pixels located in an occluded region of a first one of a defocused pair of stereo images using:

35

a first set of blur map data corresponding to the first one of the defocused pair of stereo images;

- 33 -

a first set of disparity map data corresponding to the first one of the defocused pair of stereo images; and

a first set of occlusion map data corresponding to the first one of the defocused pair of stereo images.

5

22. The apparatus of claim 21, the at least one memory and the computer program code being configured to, with the at least one processor, cause the apparatus to:

generate a first set of blur-disparity map data corresponding to the first one of the defocused pair of stereo images, the first set of blur-disparity map data being generated

10 based on:

the first set of disparity map data;

the first set of occlusion map data; and

the first set of blur map data; and

15 estimate the disparity values for the pixels located in the occluded region of the first one of the defocused pair of stereo images using the first set of occlusion map data and the first set of blur-disparity map data.

23. The apparatus of claim 21 or claim 22, the at least one memory and the computer program code being configured to, with the at least one processor, cause the apparatus to:

20 update the first set of disparity map data with the estimated disparity values to generate a first set of updated disparity map data.

24. The apparatus of claim 23, the at least one memory and the computer program code being configured to, with the at least one processor, cause the apparatus to:

25 apply at least one filter to the first set of updated disparity map data.

25. The apparatus of claim 24, the at least one memory and the computer program code being configured to, with the at least one processor, cause the apparatus to:

apply a first filter to the first set of updated disparity map data; and

30 apply a second filter to the first-filtered first set of updated disparity map data.

26. The apparatus of claim 25, wherein the first filter is one of a morphological erosion filter and a morphological dilation filter and wherein the second filter is the other of the morphological erosion filter and the morphological dilation filter.

35

- 34 -

27. The apparatus of any of claims 21 to 26, the at least one memory and the computer program code being configured to, with the at least one processor, cause the apparatus to generate the first set of blur-disparity map data by:

using the first set of occlusion map data to identify pixels in the first set of disparity map data which correspond to the non-occluded pixels in the first one of the defocused pair of stereo images, each identified pixel in the set of disparity map data having a disparity value;

for each different disparity value of the identified pixels:

identifying a set of non-occluded pixels in the first set of blur map data, the set of non-occluded pixels in the first set of blur map data corresponding to the identified pixels in the first set of disparity map data which have the different disparity value, each identified non-occluded pixel in the first set of blur map data having a blur value;

calculating the mean blur value of the set of non-occluded pixels in the first set of blur map data; and

assigning the calculated mean blur value to the pixels in the first set of blur-disparity map data which correspond to the identified set of non-occluded pixels in the first set of blur map data.

28. The apparatus of claim 27, the at least one memory and the computer program code being configured to, with the at least one processor, cause the apparatus to:

if none of the previously-calculated mean blur values have a contributing pixel having the blur value of an identified occluded pixel in the first set of blur map data, assign to the occluded pixel in the first set of blur-disparity map data, which corresponds to the identified occluded pixel, the mean blur value which is closest to the blur value of the identified occluded pixel.

29. The apparatus of any of claims 21 to 28, wherein each pixel in first set of blur-disparity map data has an associated mean blur value, and wherein the at least one memory and the computer program code are configured to, with the at least one processor, cause the apparatus to estimate the disparity values for the pixels located in the occluded region of the first one of the defocused pair of stereo images by:

using the first set of occlusion map data to identify at least one occluded pixel in the first set of blur-disparity map data;

for each occluded pixel in the first set of blur-disparity map data:

- 35 -

calculating a combined difference between the mean blur value of the occluded pixel and the mean blur value of each non-occluded pixel in the first set of blur-disparity map data within an $M \times N$ neighbourhood of the occluded pixel;

for each mean blur value within the $M \times N$ neighbourhood, calculating the mean blur difference; and

identifying the mean blur value having the lowest mean blur difference, wherein the estimated disparity value for the occluded pixel is the disparity value which corresponds to the mean blur value having the lowest mean blur difference.

10 30. The apparatus of claim 29, wherein the combined difference is a weighted average of the absolute difference between the mean blur value of the occluded pixel and the mean blur value of the non-occluded pixel and the geometric Euclidian distance between occluded pixel and the non-occluded pixel.

15 31. The apparatus of claim 29 or claim 30, wherein the at least one memory and the computer program code are configured to, with the at least one processor, cause the apparatus to calculate the mean blur difference for each mean blur value within the $M \times N$ neighbourhood by summing the combined differences of all the non-occluded pixels having the mean blur value and normalising the result by the total number of the non-occluded pixels having the mean blur value.

20 32. The apparatus of any of claims 21 to 31, the at least one memory and the computer program code being configured to, with the at least one processor, cause the apparatus to: generate the first set of disparity map data using the pair of stereo images.

25 33. The apparatus of any of claims 21 to 32, the at least one memory and the computer program code being configured to, with the at least one processor, cause the apparatus to: generate the first set of occlusion map data using the first one of the defocused pair of stereo images and the first set of disparity map data.

30 34. The apparatus of any of claims 21 to 33, the at least one memory and the computer program code being configured to, with the at least one processor, cause the apparatus to: generate the first set of blur map data using the first one of the defocused pair of stereo images.

35 35. The apparatus of any of claims 21 to 34, the at least one memory and the computer program code being configured to, with the at least one processor, cause the apparatus to:

- 36 -

estimate disparity values for pixels located in an occluded region of a second one of a defocused pair of stereo images using:

a second set of blur map data corresponding to the second one of the defocused pair of stereo images;

5 a second set of disparity map data corresponding to the second one of the defocused pair of stereo images; and

a second set of occlusion map data corresponding to the second one of the defocused pair of stereo images.

10 36. The apparatus of claim 35, the at least one memory and the computer program code being configured to, with the at least one processor, cause the apparatus to:

generate a second set of blur-disparity map data corresponding to the second one of the defocused pair of stereo images, the second set of blur-disparity map data being generated based on:

15 the second set of disparity map data;

the second set of occlusion map data; and

the second set of blur map data;

estimate the disparity values for the pixels located in the occluded region of the second one of the defocused pair of stereo images using the second set of occlusion map data and the second set of blur-disparity map data.

20

37. The apparatus of claim 35 or claim 36, the at least one memory and the computer program code being configured to, with the at least one processor, cause the apparatus to: update the second set of disparity map data with the estimated disparity values for pixels located in the occluded region of the second one of the defocused pair of stereo images to generate a second set of updated disparity map data.

25

38. The apparatus of claim 37, the at least one memory and the computer program code being configured to, with the at least one processor, cause the apparatus to:

30 apply at least one filter to the second set of updated disparity map data.

39. The apparatus of claim 38, the at least one memory and the computer program code being configured to, with the at least one processor, cause the apparatus to:

apply a first filter to the second set of updated disparity map data; and

35 apply a second filter to the first-filtered second set of updated disparity map data.

40. The apparatus of claim 39, wherein the first filter is one of a morphological erosion filter and a morphological dilation filter and wherein the second filter is the other of the morphological erosion filter and the morphological dilation filter.

5 41. The apparatus of any of claims 21 to 40, wherein the apparatus is portable electronic device.

42. A non-transitory computer-readable storage medium having stored thereon computer-readable code, which, when executed by computing apparatus, causes the
10 computing apparatus to perform a method comprising:
estimating disparity values for pixels located in an occluded region of a first one of a defocused pair of stereo images using:
a first set of blur map data corresponding to the first one of the defocused
pair of stereo images;
15 a first set of disparity map data corresponding to the first one of the
defocused pair of stereo images; and
a first set of occlusion map data corresponding to the first one of the
defocused pair of stereo images.

20 43. Computer-readable code, which when executed by computing apparatus causes the computing apparatus to perform a method according to any of claims 1 to 20.

44. Apparatus comprising:
means for estimating disparity values for pixels located in an occluded region of a
25 first one of a defocused pair of stereo images using:
a first set of blur map data corresponding to the first one of the defocused
pair of stereo images;
a first set of disparity map data corresponding to the first one of the
defocused pair of stereo images; and
30 a first set of occlusion map data corresponding to the first one of the
defocused pair of stereo images.

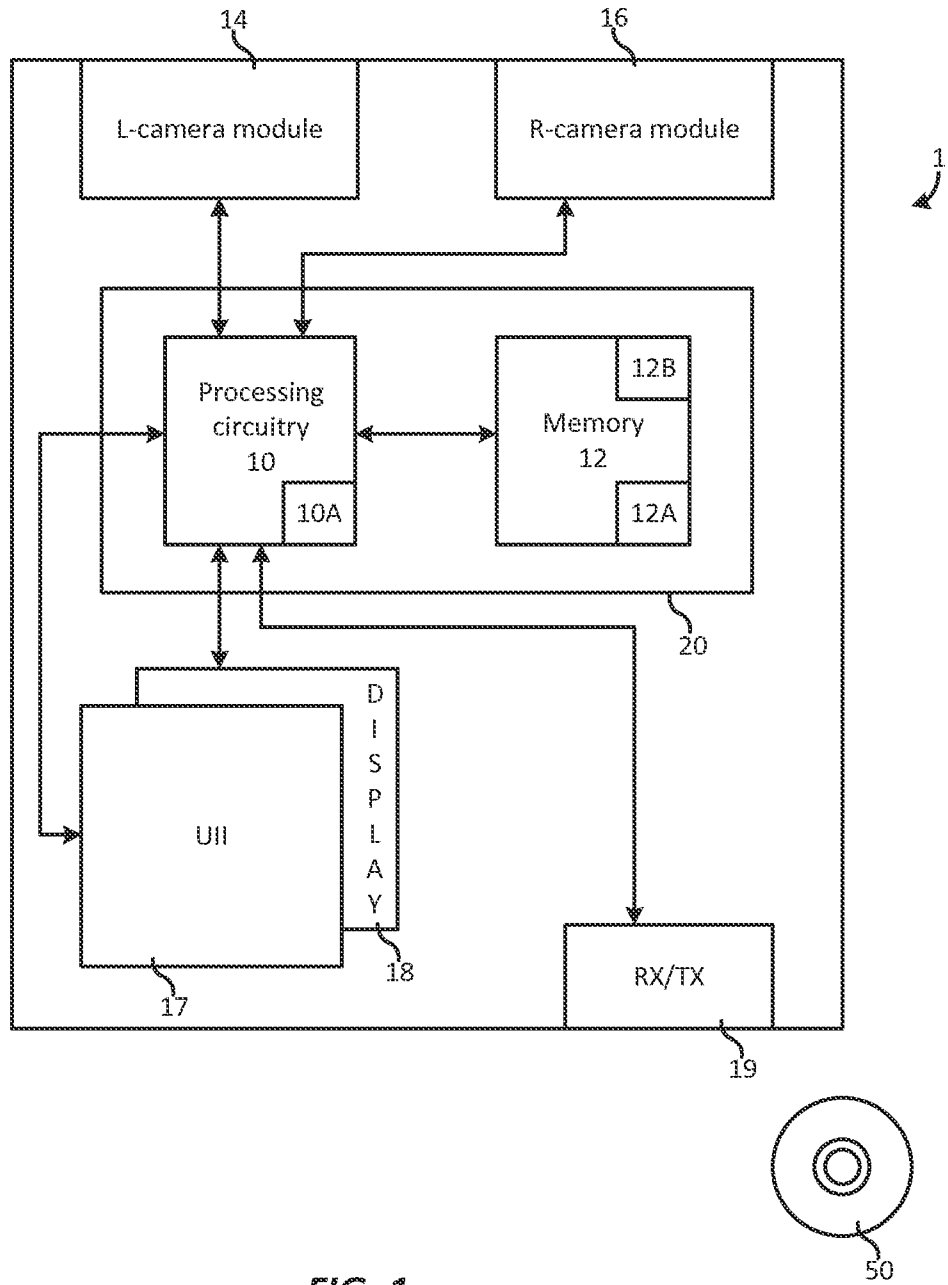


FIG. 1

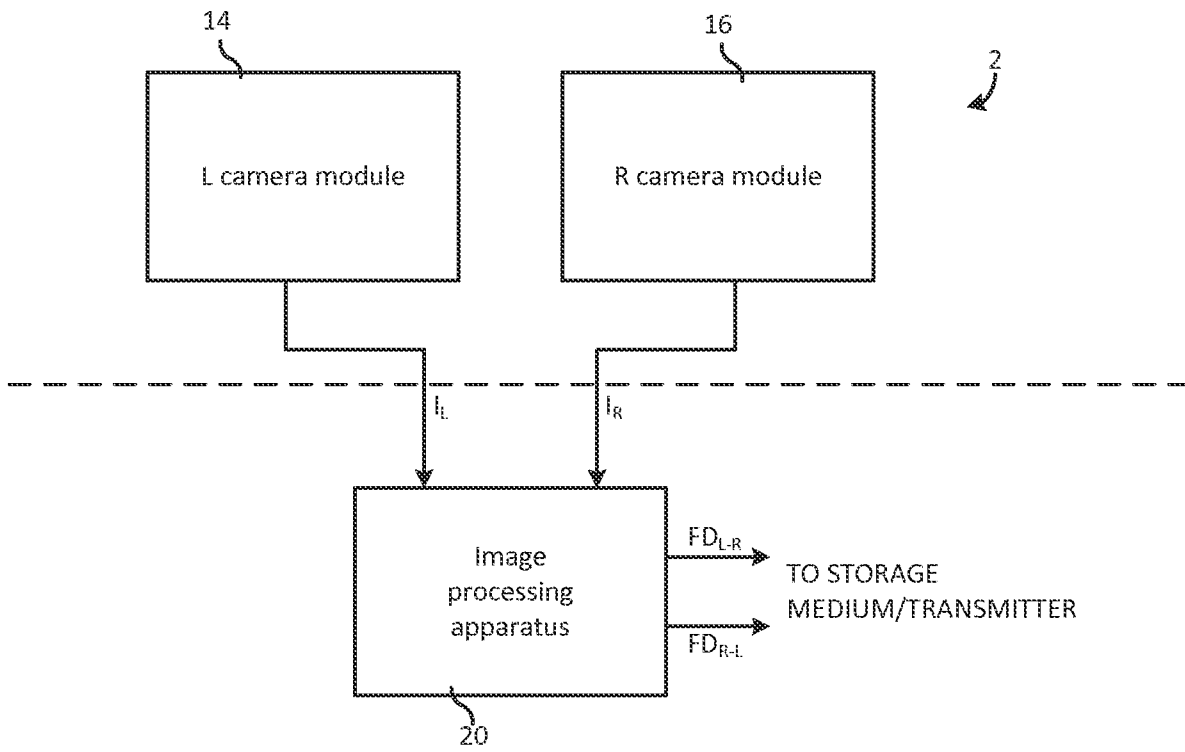


FIG. 2

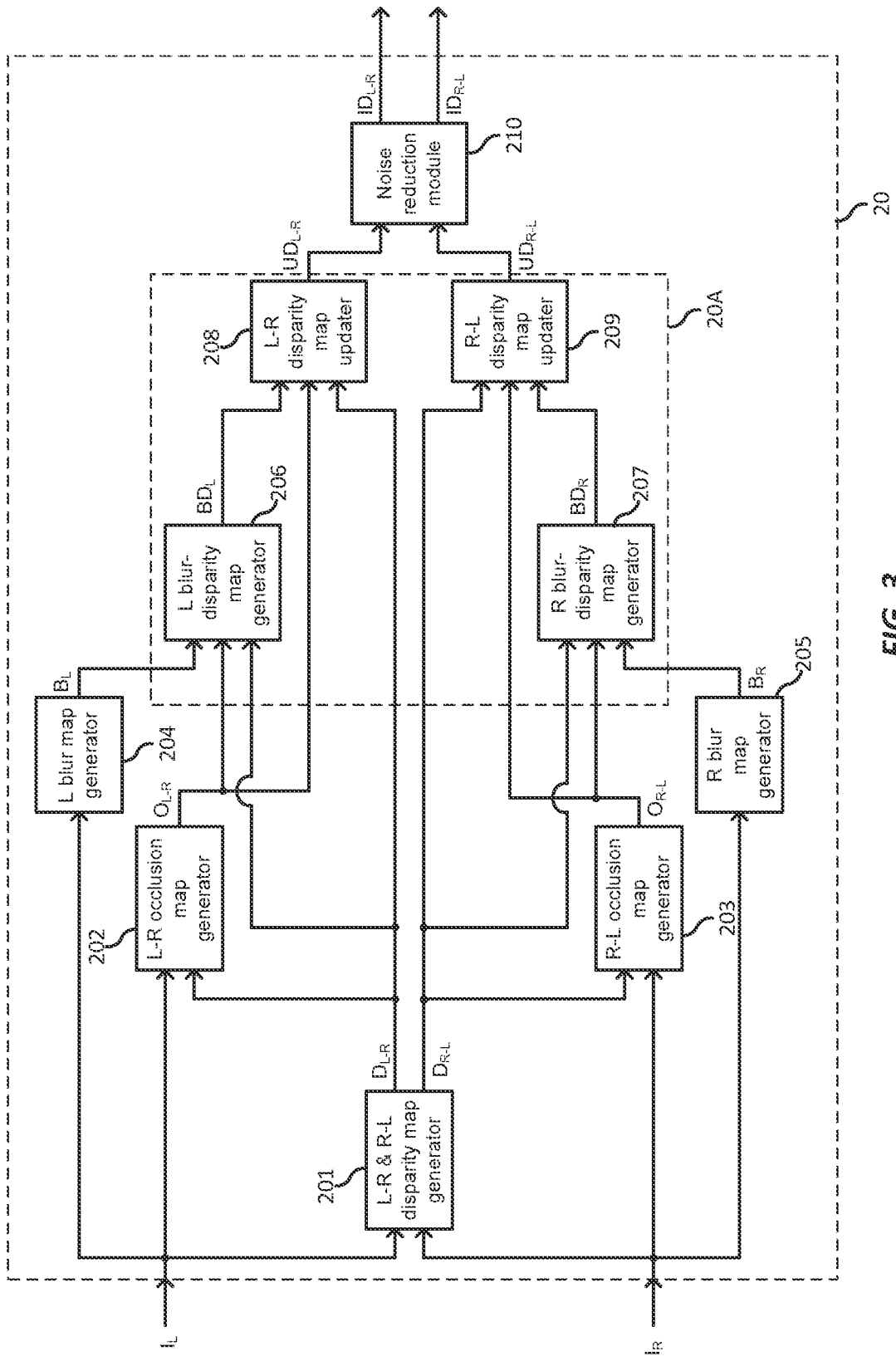


FIG. 3

4 of 8

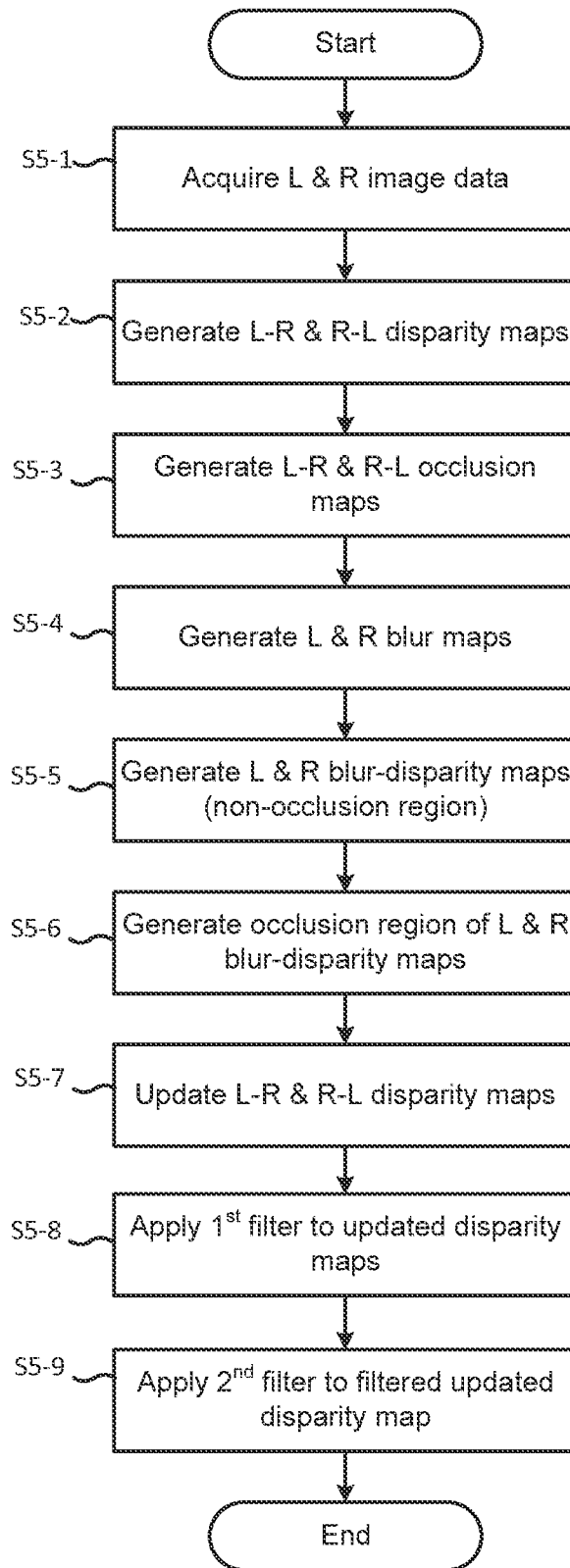


FIG. 4

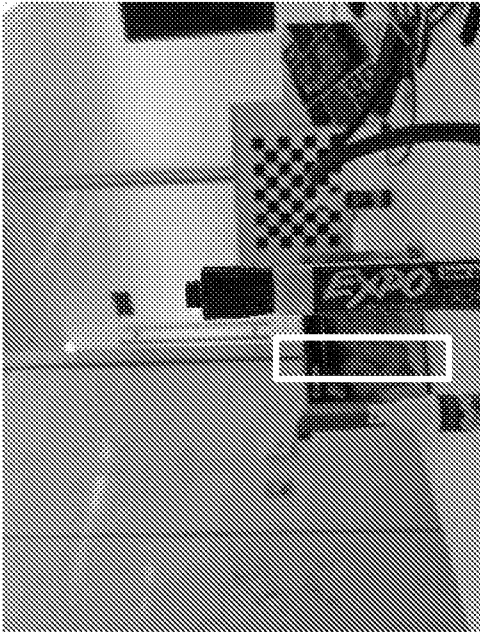


FIG. 5A
Left stereo image

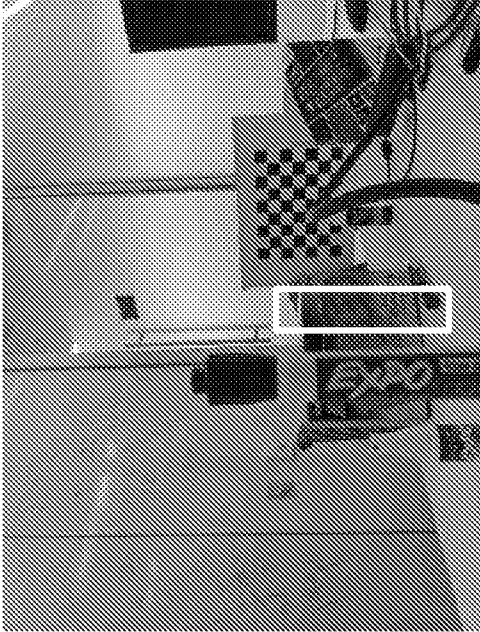


FIG. 5B
Right stereo image

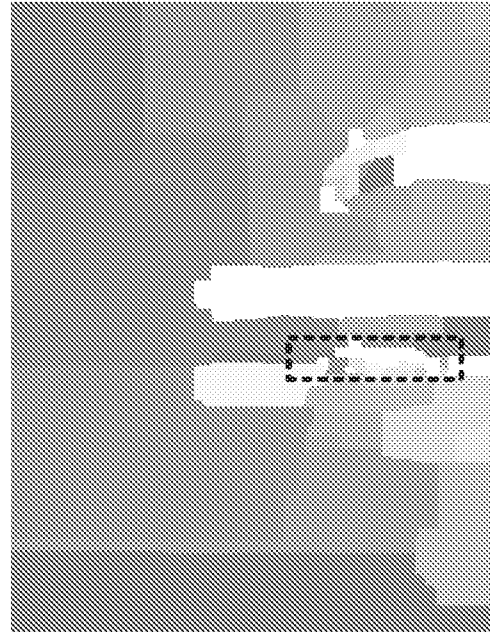


FIG. 5C
Left-to-right disparity map

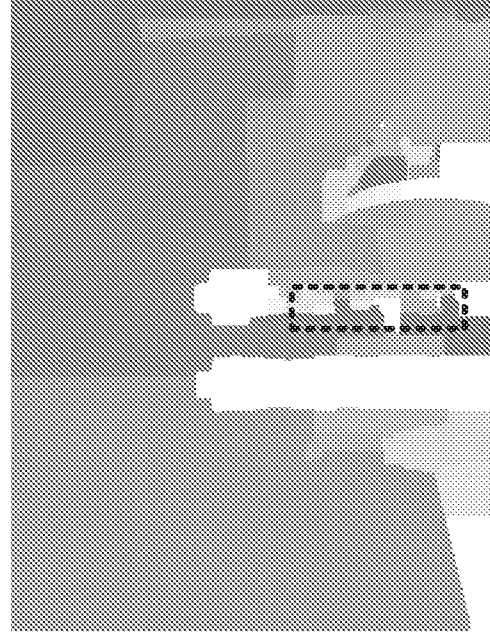


FIG. 5D
Right-to-left disparity map

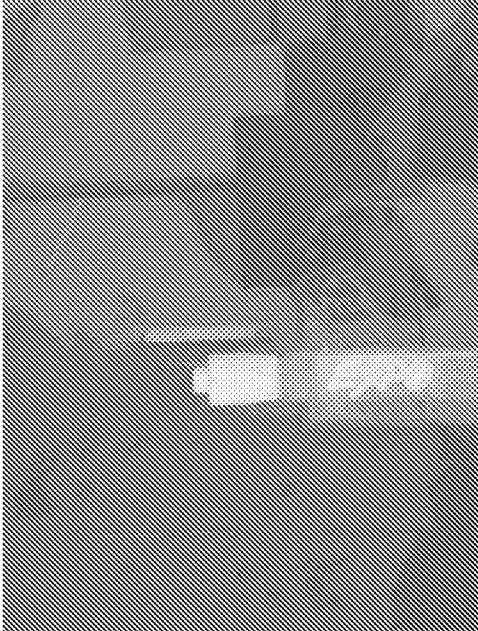


FIG. 5F

Right blur map



FIG. 5H

Right blur-disparity map inc. occlusion region



FIG. 5E

Right occlusion map



FIG. 5G

Right blur-disparity map ex. occlusion region



FIG. 5I

Final right-to-left disparity map



FIG. 5J

Updated right-to-left disparity map

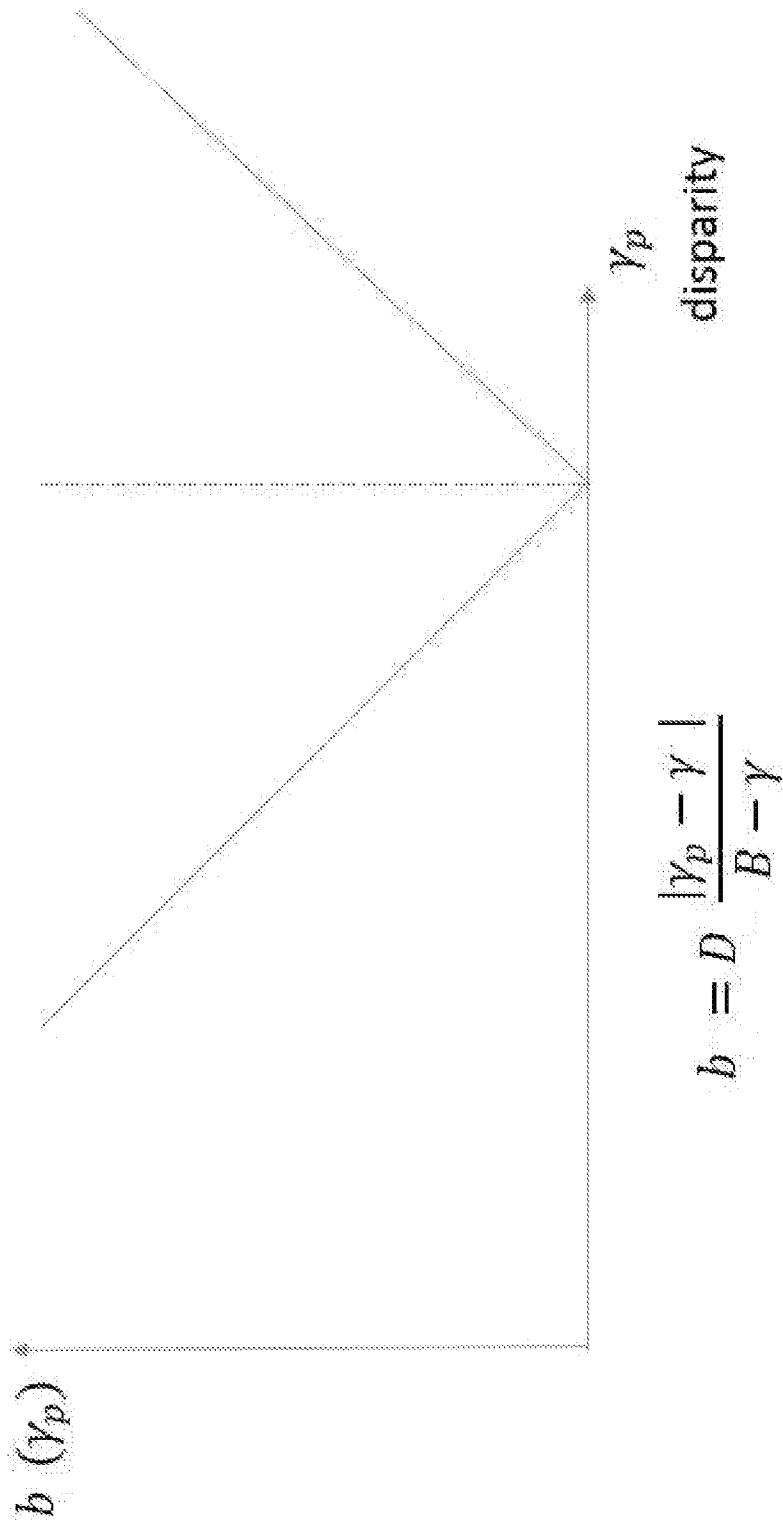


FIG. 6