(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) **EP 3 513 380 B1**

(12) **EUROPEAN PATENT SPECIFICATION**

(45) Date of publication and mention of the grant of the patent:
27.12.2023 Bulletin 2023/52

(21) Application number: **16787824.8**

(22) Date of filing: **26.10.2016**

(51) International Patent Classification (IPC):
G06T 7/593 *(2017.01)*

(52) Cooperative Patent Classification (CPC):
G06T 7/593

(86) International application number:
PCT/EP2016/075714

(87) International publication number:
WO 2018/077394 (03.05.2018 Gazette 2018/18)

(54) **METHOD AND DEVICE FOR DEPTH DETECTION USING STEREO IMAGES**

VERFAHREN UND VORRICHTUNG FÜR TIEFENERKENNUNG UNTER VERWENDUNG VON STEREOBILDERN

PROCÉDÉ ET DISPOSITIF DE DÉTECTION DE PROFONDEUR À L'AIDE D'IMAGES STÉRÉO

(84) Designated Contracting States:
AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR

(43) Date of publication of application:
24.07.2019 Bulletin 2019/30

(73) Proprietor: **Huawei Technologies Co., Ltd.**
Longgang District
Shenzhen, Guangdong 518129 (CN)

(72) Inventors:
- **NAVARRO FRUCTUOSO, Hector**
  80992 Munich (DE)
- **MARTINEZ CORRAL, Manuel**
  46100 Burjassot
  Valencia (ES)
- **SAAVEDRA TORTOSA, Genaro**
  46100 Burjassot
  Valencia (ES)
- **SOLA PIKABEA, Jorge**
  46100 Burjassot
  Valencia (ES)
- **BARREIRO HERVAS, Juan Carlos**
  46100 Burjassot
  Valencia (ES)
- **HONG, Seokmin**
  46100 Burjassot
  Valencia (ES)

(74) Representative: **Gill Jennings & Every LLP**
The Broadgate Tower
20 Primrose Street
London EC2A 2ES (GB)

(56) References cited:
WO-A1-2011/002515
WO-A2-2015/015383
US-A- 5 522 789
US-A1- 2013 321 790

EP 3 513 380 B1

Description

TECHNICAL FIELD

[0001] The invention relates to determining depth information of scenes from stereo images of the respective scene.

BACKGROUND

[0002] Currently, depth information is obtained optically by means of two kinds of systems, namely, passive and active devices.

[0003] Active systems are based on controlling the illumination of the scene in the spatial domain, as in pattern projection techniques and/or in the temporal domains, as in time-of-flight cameras. These systems, however, present a low versatility since they can work properly only in a limited operation regime, regarding both the scenes and the illumination sources that can be used. For instance, the natural background light in a natural scene can severely compromise the signal-to-noise ratio in these devices. On the other hand, several of such devices working at the same time generate a strong crosstalk resulting in interference, leading to complex and noisy detection schemes. Finally, their need to produce and control their own adapted illumination adds an extra power consumption that limits their use in many applications as autonomous video surveillance or in portable imaging systems, such as mobile phones.

[0004] Regarding passive systems, the 3D information in the scene is extracted by use of either a single, monocular, camera or different cameras, stereo or multi-ocular. In the monocular architecture, different strategies are possible both for single shot configurations, including wavefront coding engineering and for multi-shot, time multiplexing, schemes. However, the use of complex pupil masks with a poor response with natural light in the first case and the low response time in the second one, make these approaches very limited in practical applications with conventional real scenes.

[0005] For example a first stereo image is shown in Fig. 1a, while a second stereo image is shown in Fig. 1b. In Fig 1c, an overlapping view of both stereo images is shown.

[0006] The solution that best fits the operation conditions in natural dynamic scenes is obtained when several cameras are used to capture different views of the scene. Nevertheless, some limitations in the depth estimation apply to these techniques. Since the comparison of the different images is the basis for the 3D location of the objects in the scene, the use of different sensors and lenses leads to long calibration procedures to account for the different responses of the cameras, both from a geometrical point of view - rectification, distortion compensation - and from a photo/electronic point of view - pixel response equalization.

[0007] Considering the passive system for depth estimation, stereo technology is the one that provides superior results. However, such technology has the drawback of needing to calibrate and synchronize two independent cameras. In some cases, the sensors are very expensive and/or energy consuming.

[0008] The "Catadioptric monocular stereo-camera", which corresponds to the above-mentioned single shot configuration, allows for obtaining stereo images, but due to the optical setup, the scale of the objects appearing in both images is different for each depth. This is a serious problem, since the conventional low complexity block matching algorithms, just searching for the displarity in rows, cannot be applied for real time depth estimation. Usually, stereo cameras are arranged so that the axial difference between the two cameras is negligible. The small differences can be corrected digitally. Such an arrangement is though not always possible, resulting in a different length of the optical paths of the two stereo images and thereby resulting in a different lateral scaling of the stereo images. Also, in the case of using only a single camera and switching between two different optical paths, this problem occurs, since for reasons of optical path geometry, usually different optical path lengths occur. This can readily be seen in Fig. 4 and Fig. 5.

[0009] US5522789 describes a stereo endoscope which includes a pair of objective optical systems, a pair of relay optical systems, an imagery optical system having a single optical axis, and a pair of imaging devices. The imagery optical system is a zoom optical system or a varifocal optical system. For example, an encoder is used to detect a magnitude of moving lenses in the zoom optical system. Based on the detected magnitude, a control unit allows the two imaging devices to approach or depart from each other. Thus, the stereo endoscope attains the coincidence between observation points of right and left fields of view against the displacements of display points of right and left object images resulting from variations of zooming magnifications.

[0010] US2013321790 describes a method and system for 3D surface mapping system using a plurality of image sensors, each image sensor associated with an optical flow processor, each image sensor sharing a substantially coaxial optical path from the scene to a beam splitter and having substantially non-coaxial optical paths between the beam splitter and the image sensor such that the optical magnification of each optical path varies differently with the distance between the system and the surface of interest. The ratio of detected optical flows combined with the parameters of the two optical paths and the baseline between the image sensors is used to compute the Z-distance from the optical center of the image sensors to the surface.

SUMMARY

[0011] Accordingly, an object of the present invention is to provide an apparatus and method, which allow for determining an accurate depth map with a high flexibility regarding the geometry of optical paths within the camera or cameras used for recording the stereo images.

[0012] The object is solved by the features of claim 1 for the apparatus and claim 14 for the method. The dependent claims contain further developments.

[0013] According to a first aspect of the invention, a depth detection device comprising a stereo image recording device, a scaler and a depth determiner is provided, in accordance with appended claim 1. By performing the scaling, it is possible to reduce the negative impact of the differing length of the optical paths and thereby to increase their accuracy of the determined depth map.

[0014] According to a first implementation form of the first aspect, the scaler is configured to determine the scaling of the first image or of the second image, taking stereo image parameters into account. The stereo image parameters comprise a baseline of the stereo image and/or a focal length of the stereo image recording device and/or an aperture of the stereo image recording device and/or a resolution of the stereo image recording device and/or a length of the first optical path and/or a length of the second optical path. An especially accurate determining of the optimum scaling is thereby possible.

[0015] According to a second implementation form of the first aspect or the previous implementation form, the scaler is configured to determine the scaling of the first image or of the second image, taking scene parameters into account. The scene parameters comprise a depth range, and/or a depth distribution of the objects in the scene. A further increase in determining the optimum scaling factor can thereby be achieved.

[0016] According to an implementation form of the previous two implementation forms, the scaler is configured to determine a mathematical model of the stereo image recording device and/or the scene, based upon the stereo image parameters and/or the scene parameters. The scaler is configured to determine the scaling of the first image or of the second image based on the mathematical model. A further increase of the accuracy of determining the optimum scaling factor can thereby be achieved.

[0017] The scaler is configured to determine a scaling factor M_0^{opt} as

$$M_0^{opt} = 1 + \frac{\Delta \log\frac{z_1}{z_2}}{z_2 - z_1},$$

wherein Δ is a difference in the optical length of the first optical path and the second optical path, z_1 is a lower limit of a depth range of the scene, and z_2 is an upper limit of a depth range of the scene. An especially accurate determining of an optimal scaling factor is thereby achieved.

[0018] According to a further implementation form of the first aspect or the previous implementation forms, the depth detection device comprises a calibrator, which is configured to perform a calibration of the depth detection device based upon a known reference scene and a known depth map of the known reference scene, resulting in a calibration function and perform the calibration function on every determined depth map after determining the depth map by the depth determiner, resulting in a calibrated depth map. A further increase of accuracy of determining the depth map can thereby be achieved.

[0019] According to an implementation form of the previous implementation form, the stereo image recording device is configured to record a calibration stereo image of the known reference scene. The scaler is then configured to determine a scaling of the first image or of the second image of the calibration stereo image, by minimizing differing lateral magnification of objects in the scene depicted in the first image and the second image of the calibration stereo image and perform the determined scaling resulting in a scaled calibration stereo image. The depth determiner is then configured to determine a calibration depth map of the known reference scene based upon the scaled calibration stereo image. The calibrator is then configured to determine differences of the calibration depth map and the known depth map and determine the calibration function from the differences of the calibration depth map and the known depth map. An especially accurate calibration can thereby be performed.

[0020] According to a further implementation form of the previous two implementation forms, the calibrator is configured to determine the calibration function as a non-linear function. By use of a non-linear function, an especially accurate calibration is possible.

[0021] According to a further implementation form of the previous three implementation forms, the calibrator is configured to determine the calibration function as an image transformation matrix. This results in an especially accurate calibration.

[0022] According to a further implementation form of the first aspect or any of the previous implementation forms, the stereo image recording device is configured to record the first image of the stereo image and the second image of the stereo image displaced by a baseline. It is thereby possible to determine the depth map based upon the stereo image.

[0023] According to a further implementation form of the first aspect or the previous implementation forms, the stereo image recording device comprises a single camera and an optical splitter. The single camera is configured for recording the first image and the second image successively. The optical splitter is configured for switching between the first optical path and the second optical path successively. It is thereby possible to use only a very limited number of hardware elements, especially only a single camera. Also, it is thereby possible to remove negative influences by slightly differing cameras, in case of using two cameras.

[0024] According to an implementation form of the previous implementation form, the optical splitter comprises a beam splitter arranged in front of the single camera and a total reflection prism arranged in a beam splitting direction of the beam splitter. The first optical path leads from the scene to the total reflection prism to the beam splitter to the single camera. The second optical path leads from the scene to the beam splitter to the single camera. This allows for a very simple implementation of the depth detection device. The beam splitter may be a beam-splitting cube, for example. Preferably, the beam splitter is placed directly in front of the single camera.

[0025] According to a further implementation form of the previous two implementation forms, the optical splitter comprises a first shutter device arranged within the first optical path, but not within the second optical path and a second shutter device arranged within the second optical path, but not within the first optical path. The first shutter device is configured to shut the first optical path during recording of the second image by the single camera and open the first optical path during recording the first image by the single camera. The second shutter device is configured to shut the second optical path during recording of the first image by the single camera and open the second optical path during recording the second image by the single camera. This effectively prevents stray image information from the presently non-recorded optical path to negatively influence the presently recorded image.

[0026] According to a further implementation form of the first aspect or any of the first implementation forms of the first aspect, the stereo image recording device comprises a first camera for recording the first image and a second camera for recording the second image. The first camera and the second camera are located at different distances from the scene, resulting in the different lengths of the first optical path and the second optical path. By this configuration, the hardware effort used for beam splitting can be saved.

[0027] According to a second aspect of the invention, a depth detection method is provided in accordance with appended claim 14. By performing the scaling, it is possible to reduce the negative impact of the differing length of the optical paths and thereby to increase the accuracy of the determined depth map.

[0028] According to a first implementation form of the second aspect, the scaling is performed taking stereo image parameters into account. The stereo image parameters comprise a baseline of the stereo image and/or a focal length of the stereo image recording device and/or an aperture of the stereo image recording device and/or a resolution of the stereo image recording device and/or a length of the first optical path and/or a length of the second optical path. An especially accurate determining of the optimum scaling is thereby possible.

[0029] According to a second implementation form of the second aspect or the previous implementation form, the scaling is determined taking scene parameters into account, which comprise a depth range and/or a depth distribution of the objects in the scene. A further increase in determining the optimum scaling factor can thereby be achieved.

[0030] According to a further implementation form of the previous two implementation forms of the second aspect, a mathematical model of the stereo image recording device and/or the scene is determined based upon the stereo image parameters and/or the scene parameters. The scaling factor is determined based upon the mathematical model. A further increase of the accuracy of determining the optimum scaling factor can thereby be achieved.

[0031] The scaler is configured to determine a scaling factor M_0^{opt} as

$$M_0^{opt} = 1 + \frac{\Delta \log \frac{z_1}{z_2}}{z_2 - z_1},$$

wherein Δ is a difference in optical length of the first optical path and the second optical path, z_1 is a lower limit of a depth range of the scene and z_2 is an upper limit of a depth range of the scene. An especially accurate determining of an optimal scaling factor is thereby achieved.

[0032] According to a further implementation form of the second aspect or the previous implementation forms, a calibration of the depth detection based upon a known reference scene and a known depth map of the known reference scene is performed, resulting in a calibration function. The calibration function is performed on every determined depth map after determining the depth map, resulting in a calibrated depth map. A further increase of accuracy of determining the depth map can thereby be achieved.

[0033] According to an implementation form of the previous implementation form, a calibration stereo image of the known reference scene is recorded. A scaling of the first image and/or of the second image of the calibration stereo image is determined, by minimizing different lateral magnifications of objects in the scene depicted in the first image and the second image of the calibration stereo image. The scaling is performed on one of the images resulting in a scaled calibration stereo image. A calibration depth map is determined based upon the scaled calibration stereo image. Differences of the calibration depth map and the known depth map are then determined. Finally, a calibration function is determined from these differences. An especially accurate calibration can thereby be performed.

[0034] According to an implementation form of the previous two implementation forms, the calibration function is determined as a non-linear function. By use of a non-linear function, an especially accurate calibration is possible.

[0035] According to a further implementation form of the previous three implementation forms, the calibration function is determined as an image transformation matrix. It is thereby possible to determine the depth map based upon the stereo image.

[0036] According to a further implementation form of the second aspect or any of the previous implementation forms, the first image and the second image are recorded displaced by a baseline. It is thereby possible to determine the depth map based upon the stereo image.

[0037] According to a further implementation form of the second aspect or the previous implementation forms of the second aspect, a single camera is configured for successively recording the first image and the second image, while an optical splitter switches between the first optical path and the second optical path successively. It is thereby possible to use only a very limited number of hardware elements, especially only a single camera. Also, it is thereby possible to remove negative influences by slightly differing cameras, in case of using two cameras.

[0038] Alternatively, in a further implementation form of the second aspect or the first nine implementation forms of the second aspect, the first image is recorded by a first camera and the second image is recorded by a second camera. By this configuration, the hardware effort used for beam splitting can be saved.

[0039] Generally, it has to be noted that all arrangements, devices, elements, units and means and so forth described in the present application could be implemented by software or hardware elements or any kind of combination thereof. Furthermore, the devices may be processors or may comprise processors, wherein the functions of the elements, units and means described in the present applications may be implemented in one or more processors. All steps which are performed by the various entities described in the present application as well as the functionality described to be performed by the various entities are intended to mean that the respective entity is adapted to or configured to perform the respective steps and functionalities. Even if in the following description or specific embodiments, a specific functionality or step to be performed by a general entity is not reflected in the description of a specific detailed element of that entity which performs that specific step or functionality, it should be clear for a skilled person that these methods and functionalities can be implemented in respect of software or hardware elements, or any kind of combination thereof.

BRIEF DESCRIPTION OF DRAWINGS

[0040] The present invention is in the following explained in detail in relation to embodiments of the invention in reference to the enclosed drawings, in which

Fig. 1a shows a first image of a stereo image;

Fig. 1b shows a second image of a stereo image;

Fig. 1c shows the first image and the second image of the stereo image in an overlapping fusion;

Fig. 2 shows a first embodiment of the first aspect of the invention;

Fig. 3 shows a second embodiment of the first aspect of the invention;

Fig. 4 shows a detail of a third embodiment of the first aspect of the invention;

Fig. 5 shows a detail of a fourth embodiment of the first aspect of the invention;

Fig. 6 shows a first embodiment of the second aspect of the invention;

Fig. 7 shows a second embodiment of the second aspect of the invention, and

Fig. 8 shows results obtainable by use of the invention.

[0041] First, the concept of depth estimation by use of stereo images was described along Fig. 1a - 1c. In the following, along Fig. 2 - Fig. 5, the construction and function of different embodiments of the inventive device are shown. Along Fig. 6 - Fig. 7, the functions of different embodiments of the inventive method are described. Finally along Fig. 8, further benefits of the invention are eliminated.

[0042] Similar entities and reference numbers in different figures have been partially omitted.

DESCRIPTION OF EMBODIMENTS

[0043] In Fig. 2, a first embodiment of a depth determining device 20 is shown. The depth determining device comprises a stereo image recording device 21, a scaler 22 and a depth determiner 23.

[0044] In Fig. 3, a second embodiment of the inventive depth detection device 20 is shown. Here, the depth determining device 20 furthermore comprises a control unit 30 and a calibrator 31. The control unit 30 is connected to the units 21, 22, 23 and 31 and controls these.

[0045] The scaler 22 is moreover connected to the stereo image recording device 21 and to the depth determiner 23.

[0046] The stereo image recording device 21 is configured for recording a stereo image comprising a first image and a second image. The construction of the stereo image recording device 21 is further detailed in Fig. 4 and Fig. 5. The first image is recorded through a first optical path, while the second image is recorded through a second optical path. The optical paths are of different lengths.

[0047] After the first image and the second image have been recorded, they are handed to the scaler 22, which performs a scaling of the first image or the second image in order to compensate for differing lateral magnification of objects in the scene depicted in the first image and the second image. This is done by first determining an optimal scaling and afterwards performing the scaling on the stereo image resulting in a scaled stereo image. The scaled stereo image is handed on to the depth determiner 23, which performs a depth determination resulting in a depth map of the scene.

[0048] In the example, the calibrator 31 performs a calibration using a known reference scene and an accordingly known depth map of the known reference scene. This is done by recording a stereo image of the known reference scene using the stereo image recording device 21, performing a scaling by the scaler 22, as explained above, and by performing a depth determining by the depth determiner 23, as also explained above. The resulting calibration depth map is compared to the known depth map. The calibrator 31 determines a calibration function therefrom. The calibration function advantageously is a non-linear function. Also, the calibration function advantageously is an image transformation matrix.

[0049] In Fig. 4, the stereo image recording device 21a, which corresponds to the stereo image recording device 21 of Fig. 2 and Fig. 3, is shown. Here, the stereo image recording device 21a comprises a first camera 40a and a second camera 40b. Both cameras 40a, 40b are arranged side by side, but axially displaced relative to each other, resulting in a first optical path 43 and a second optical path 44 differing in length. The optical paths 43, 44 have a differing length with regard to the depicted scene. Also, the cameras 40a, 40b are laterally displaced relative to each other by a baseline.

[0050] In Fig. 5, a further alternative embodiment of the stereo image recording device 21b, which corresponds to the stereo image recording device 21 of Fig. 2 and Fig. 3, is shown. Here, the stereo image recording device 21b merely comprises a single camera 50 and an optical splitter 51. The optical splitter 51 is arranged between the camera 50 and the depicted scene. Through the optical splitter 51, the first optical path 43 and the second optical path 44 have differing lengths, as already shown in Fig. 4.

[0051] The optical splitter 51 in this embodiment comprises a beam splitter (e.g., a beam splitting cube) 52, which is arranged closely in front of the single camera 50. Moreover, it comprises a total reflection prism 53, which is arranged in a beam splitting direction of the beam splitter 52. The first optical path 43 leads from the scene to the beam splitter 52 to the single camera 50. The second optical path 43 leads from the scene to the total reflection prism 53 to the beam splitter 52 to the single camera 50.

[0052] In the example, the optical splitter 51 moreover comprises shutters 54, 55, which are arranged on the optical paths. A first shutter 54 is arranged on the first optical path 43 between the scene and the beam splitting 52. A second shutter 55 is arranged between the scene and the total reflection prism 53. The shutters 54, 55 shut the optical path which is currently not recorded. This means that while recording the first image through the first optical path 43, the first shutter 54 is open while the second shutter 55 is closed. While recording the second image through the second optical path 44, the second shutter 55 is open while the first shutter 54 is closed.

[0053] Note that the proposed system is equivalent to a virtual stereo camera with axes separated by baseline Δ, but placed at different depths, also separated by an offset Δ. The maximum light efficiency of any branch of the virtual stereo camera is of 25%, obtained as the product of the maximum efficiency of the LCL (50%) and of the CBS (50%). This implies that in terms of light efficiency, the effective f-number of any virtual camera is equal to the f-number of the objective plus 2.0 units.

[0054] In order to obtain enough luminosity, together with sufficient depth of field, we need to take into account that the depth of field, Δ_F, of a photographic camera is given by

$$\Delta_{\mathrm{F}} = k\frac{f_{\#}}{f^2},$$

where k is a proportionality factor, and $f_{\#}$ the f-number. Then the proposed camera must have an objective lens with small focal length (smaller than f = 20 mm).

[0055] An important issue to take into account is the fact that the field of view (FOV) limitation is different in the two branches of the virtual camera. In a single camera the FOV is limited by the sensor size and by the field aperture. In the proposed design the field apertures are given by the projection of reflecting elements onto the plane perpendicular to the optical axis. Such projections are represented in Fig. 1(b) by means of the virtual apertures. As it is well known, further field apertures produce stronger FOV limitation. So, FOV limitation in the left image is stronger than FOV limitation in the right image. Problems with FOV limitation can be avoided by placing the TRP and the CBS in contact or close to contact.

[0056] In the following, the functions of the different elements of the depth determining device 20 are explained in greater detail:

The invention tries to minimize the scale difference for the objects located at different depths on each image. In order to do that, we calculate the affine transformation minimizing the difference in the lateral magnification over the camera sensor for the axial interval of interest. Next this process will be described in a more detailed manner:

It must be taken into account that the distance to the objects, and the consequent magnification, is not the same for the two virtual cameras. Even the magnification offset depends on the distance to the object. In this sense we can define the function

$$M(\Delta, z) = 1 + \frac{\Delta}{z}, \qquad (3)$$

which gives a relation between the scales of left and right images. In this function Δ is the distance between virtual axes, and z is the distance between the right camera and the object. Now we define the square residual

$$\epsilon^2 = (M(\Delta, z) - M_0)^2 \qquad (4)$$

where M_0 is a value of magnification to be optimized. Next we can evaluate the sum of square residuals

$$\sigma^2 = \int_{z_1}^{z_2} \epsilon^2 \, dz, \qquad (5)$$

where z_1 and z_2 are the limits of the axial range. Finally we calculate the minimum of this function, that is

$$\frac{d\sigma^2}{dM_0} = 0, \qquad (6)$$

and therefore obtain the optimum value for M_0

$$M_0^{opt} = 1 + \frac{\Delta \log\frac{z_2}{z_1}}{z_2 - z_1}. \qquad (7)$$

[0057] As an example, we can calculate the optimum re-scaling for the following values: Δ= 20 *mm,* z_2 = 10.0 *m*, and z_1 = 1.0 m. We obtain, $M_0^{opt} = 1.0051$.

[0058] Our proposal is to scale by this factor the left image in a stereo pair and all the left images if a video-sequence is considered. This is not to be understood as the only option though. It is just as well possible to scale the other of the two images.

[0059] Once the left image has been scaled by the scaler 22, and it has been made sure that the two images are aligned along any line parallel to the baseline, a disparity map can be obtained by calculating point by point the disparity in pixel units, by the depth determiner 23. From the disparity map, expressed in pixel units, a depth map is calculated according the following formula:

$$z = \frac{d\Delta}{4g} p \ . \tag{8}$$

[0060] In the Eq. (8) d is the disparity measured in number of pixels; *g* is the gap between camera lens and the sensor (in case of object at infinity, $g = f$, being *f* the focal length), and *p* is the actual dimension of sensor pixels.

[0061] Since the use of a rescaled left image for the disparity calculation is the result of a first order approximation, a nonlinear post calibration is needed for the accurate depth recovery. For the post calibration, a table with real and calculated depth distances must be done for a dense amount of sentences within the axial range. This table permits to define a post-calibration function.

[0062] In order to explain how the post-calibration works, we have implemented a prototype in which the baseline was Δ= 30.5 *mm* and the objective of f = 50 mm operated at $f_{\#}$ = 8.0. We have tested a 3D scene composed by a series of 7$\times$7 cm 2D objects with helix symbol, placed at equidistant positions from 1 to 10 m. The images are shown in Fig. 1a - 1c.

[0063] A representation of both calculated depth and directly measured distances is shown in Fig. 8, including a parabolic fitting of both results. The value of both the slope and correlation coefficient for the curve fitting shows the good agreement of the results under test.

[0064] It can therefore be seen that there is a remaining error between the actual depth and the depth map. By determining the calibration function by use of the calibrator 31, and then applying the calibration function to each depth map, which is determined by the depth determiner 23, a calibrated depth map is achieved. This calibrated depth map has significantly less error than the non-calibrated depth map. The calibration function therein advantageously is a non-linear function, especially an image transformation matrix. For determining the calibration function, a mathematical model may be used, as explained above.

[0065] In Fig. 6, a first embodiment of the inventive depth determining method is shown. In a first step 100, a stereo image of a scene is recorded using an image recording device. A first image of the stereo image is recorded through a first optical path while a second image of the stereo image is recorded through a second optical path. The two optical paths do not have the same lengths. In a second step 101, a scaling of the first image or of the second image is determined by minimizing differing lateral magnification of objects in the scene depicted in the first image and second image. Especially, scene parameters and stereo image recording device parameters of the stereo image recording device 21 may be used therefore.

[0066] In a third step 102, an according scaling is performed on one of the first image or the second image resulting in a scaled stereo image. In a fourth step 103, a depth map of the scene is determined based upon the scaled stereo image. In an optional final step 104, a calibration of the depth map is performed using a calibration function. This results in a calibrated depth map.

[0067] An embodiment showing the calibration of step 104 of Fig. 6 is shown in Fig. 7. There, in a first step 200, optical parameters of the camera or cameras of the stereo image recording device 21 are selected. In an optional second step 201, further optical parameters of catadioptric elements, such as the total reflection prism 53 and the beam splitting 52 of Fig. 5 may be selected. In a third step 202, the elements are arranged in such a manner that the baseline of the two cameras is minimized. By minimizing the baseline, the differing optical lengths of the optical paths can be minimized therefore also minimizing the need for corrections. In a fourth step 203, an optimum scaling is determined, as explained earlier. In a fifth step 204, one or more calibration stereo images are recorded using a known reference scene. In a sixth step 205, the calibration stereo image or images are scaled according to the optimum scaling factor determined in step 203. In a seventh step 206, a calibration depth map is determined for each of the calibration stereo images. In an eighth step 207, differences between a known depth map of the known reference scene and the determined calibration depth map or maps are determined. In a final ninth step 208, a calibration function is determined based upon the differences of the calibration depth map and the known depth map. More than one calibration depth map can be taken into account here.

[0068] The invention is not limited to the examples and especially not to a specific type of construction of the stereo image recording device. The characteristics of the exemplary embodiments can be used in any advantageous combination, within the scope of the appended claims.

[0069] The invention has been described in conjunction with various embodiments herein. However, other variations to the disclosed embodiments are conceivable within the scope of the appended claims. In the claims, the word "comprising " does not exclude other elements or steps and the indefinite article "a" or "an" does not exclude a plurality. A single processor or other unit may fulfill the functions of several items recited in the claims. The mere fact that certain measures are recited in usually different dependent claims does not indicate that a combination of these measures cannot be used to advantage. A computer program may be stored/distributed on a suitable medium, such as an optical storage medium or a solid-state medium supplied together with or as part of other hardware, but may also be distributed in other forms, such as via the internet or other wired or wireless communication systems.

Claims

1. A depth detection device (20), comprising a stereo image recording device (21, 21a, 21b), a scaler (22), and a depth determiner (23),

   wherein the stereo image recording device (21, 21a, 21b) is configured to record a stereo image of a scene, wherein the stereo image recording device (21, 21a, 21b) comprises

   - a first optical path (43), configured to record a first image of the stereo image, and
   - a second optical path (44), configured to record a second image of the stereo image,

   wherein the first optical path (43) and second optical path (44) are of different length, wherein the scaler (22) is configured to

   - determine a scaling of the first image or of the second image, by minimizing differing lateral magnification of objects in the scene depicted in the first image and the second image, and
   - perform the determined scaling, resulting in a scaled stereo image,

   wherein the depth determiner (23) is configured to determine a depth map of the scene based upon the scaled stereo image, wherein the scaler (22) is configured to determine a scaling factor M_0^{opt} as

   $$M_0^{opt} = 1 + \frac{\Delta \log\frac{z_1}{z_2}}{z_2 - z_1},$$

   wherein Δ is a difference in optical length of the first optical path (43) and the second optical path (44), z_1 is a lower limit of a depth range of the scene, and z_2 is an upper limit of a depth range of the scene.

2. The depth detection device (20) according to claim 1,

   wherein the scaler (22) is configured to determine the scaling of the first image or of the second image, taking stereo image parameters into account, and
   wherein the stereo image parameters comprise a baseline of the stereo image, and/or a focal length of the stereo image recording device (21, 21a, 21b), and/or an aperture of the stereo image recording device (21, 21a, 21b), and/or a resolution of the stereo image recording device (21, 21a, 21b), and/or a length of the first optical path (43), and/or the length of the second optical path (44).

3. The depth detection device (20) according to claim 1 or 2,

   wherein the scaler (22) is configured to determine the scaling of the first image or of the second image, taking scene parameters into account, and
   wherein the scene parameters comprise a depth range, and/or a depth distribution of the objects in the scene.

4. The depth detection device (20) according to claim 2 or 3,

   wherein the scaler (22) is configured to determine a mathematical model of the stereo image recording device (21, 21a, 21b) and/or the scene, based upon the stereo image parameters and/or the scene parameters, and wherein the scaler (22) is configured to determine the scaling of the first image or of the second image based on the mathematical model.

5. The depth detection device (20) according to any of the claims 1 to 4,
   wherein the depth detection device (20) comprises a calibrator (31), configured to

   - perform a calibration of the depth detection device (20) based upon a known reference scene and a known

depth map of the known reference scene, resulting in a calibration function, and
- perform the calibration function on every determined depth map after determining the depth map by the depth determiner (23), resulting in a calibrated depth map.

6. The depth detection device (20) according to claim 5,

wherein stereo image recording device (21, 21a, 21b) is configured to record a calibration stereo image of the known reference scene,
wherein the scaler (22) is configured to

- determine a scaling of the first image or of the second image of the calibration stereo image, by minimizing differing lateral magnification of objects in the scene depicted in the first image and the second image of the calibration stereo image,
- perform the determined scaling, resulting in a scaled calibration stereo image,

wherein the depth determiner (23) is configured to determine a calibration depth map of the known reference scene based upon the scaled calibration stereo image,
wherein the calibrator (31) is configured to

- determine differences of the calibration depth map and the known depth map, and
- determine the calibration function from the differences of the calibration depth map and the known depth map.

7. The depth detection device (20) according to claim 5 or 6,
wherein the calibrator (31) is configured to determine the calibration function as a non-linear function.

8. The depth detection device (20) according to any of the claims 5 to 7,
wherein the calibrator (31) is configured to determine the calibration function as an image transformation matrix.

9. The depth detection device (20) according to any of the claims 1 to 8,
wherein the stereo image recording device (21, 21a, 21b) is configured to record the first image of the stereo image and the second image of the stereo image displaced by a baseline.

10. The depth detection device (20) according to any of the claims 1 to 9,

wherein the stereo image recording device (21, 21b) comprises a single camera (50) and an optical splitter (51),
wherein the single camera (50) is configured for recording the first image and the second image successively, and
wherein the optical splitter (51) is configured for switching between the first optical path (43) and the second optical path (44) successively.

11. The depth detection device (20) according to claim 10,

wherein the optical splitter (51) comprises

- a beam splitter (52) arranged in front of the single camera (50), and
- a total reflection prism (53) arranged in a beam splitting direction of the beam splitter (52),

wherein the first optical path (43) leads from the scene to the beam splitter (52) to the single camera (50), and
wherein the second optical path (44) leads from the scene to the total reflection prism (53) to the beam splitter (52) to the single camera (50).

12. The depth detection device (20) according to claim 10 or 11,

wherein the optical splitter (51) comprises

- a first shutter device (54) arranged within the first optical path (43) but not within the second optical path (44), and
- a second shutter device (55) arranged within the second optical path (44) but not within the first optical

path (43),

wherein the first shutter device (54) is configured to

- shut the first optical path (43) during recording the second image by the single camera, and
- open the first optical path (43) during recording the first image by the single camera, and

wherein the second shutter device (55) is configured to

- shut the second optical path (44) during recording the first image by the single camera, and
- open the second optical path (44) during recording the second image by the single camera.

13. The depth detection device (20) according to any of the claims 1 to 9, wherein the stereo image recording device (21, 21a) comprises a first camera (40a) for recording the first image, and a second camera (40b) for recording the second image, and wherein the first camera (40a) and the second camera (40b) are located at differing distances from the scene, resulting in the differing lengths of the first optical path (43) and the second optical path (44).

14. A depth detection method, comprising:

- recording (100) a stereo image of a scene, using a stereo image recording device (21, 21a, 21b), comprising a first optical path (43), and a second optical path (44), wherein a first image of the stereo image is recorded through the first optical path (43) and a second image of the stereo image is recorded through the second optical path (44), wherein the first optical path (43) and second optical path (44) are of different length,
- determining (101) a scaling of the first image or of the second image, by minimizing differing lateral magnification of objects in the scene depicted in the first image and the second image,
- performing (102) the determined scaling, resulting in a scaled stereo image, and
- determining (103) a depth map of the scene based upon the scaled stereo image, wherein determining (101) the scaling of the first image or of the second image comprises determining a scaling factor M_0^{opt} as

$$M_0^{opt} = 1 + \frac{\Delta \log\frac{z_1}{z_2}}{z_2 - z_1}$$

wherein Δ is a difference in optical length of the first optical path (43) and the second optical path (44), z_1 is a lower limit of a depth range of the scene, and z_2 is an upper limit of a depth range of the scene.

Patentansprüche

1. Tiefendetektionsvorrichtung (20), die eine Stereobildaufzeichnungsvorrichtung (21, 21a, 21b), eine Skalierungsvorrichtung (22) und eine Tiefenbestimmungsvorrichtung (23) umfasst, wobei

die Stereobildaufzeichnungsvorrichtung (21, 21a, 21b) konfiguriert ist, ein Stereobild einer Szene aufzuzeichnen, und
die Stereobildaufzeichnungsvorrichtung (21, 21a, 21b) Folgendes umfasst:

- einen ersten optischen Pfad (43), der konfiguriert ist, ein erstes Bild des Stereobilds aufzuzeichnen, und
- einen zweiten optischen Pfad (44), der konfiguriert ist, ein zweites Bild des Stereobilds aufzuzeichnen, wobei der erste optische Pfad (43) und der zweite optische Pfad (44) verschiedene Längen besitzen und die Skalierungsvorrichtung (22) konfiguriert ist zum:
- Bestimmen einer Skalierung des ersten Bilds oder des zweiten Bilds durch Minimieren einer unterschiedlichen seitlichen Vergrößerung von Objekten in der Szene, die in dem ersten Bild und dem zweiten Bild dargestellt ist, und
- Durchführen der bestimmten Skalierung, was in einem skalierten Stereobild resultiert, wobei

die Tiefenbestimmungsvorrichtung (23) konfiguriert ist, eine Tiefenkarte der Szene auf der Grundlage des skalierten Stereobilds zu bestimmen, wobei die Skalierungsvorrichtung (22) konfiguriert ist, einen Skalierungsfaktor M_o^{opt} als

$$M_o^{opt} = 1 + \frac{\Delta \log \frac{z_1}{z_2}}{z_2 - z_1}$$

zu bestimmen, wobei Δ eine Differenz der optischen Länge des ersten optischen Pfads (43) und des zweiten optischen Pfads (44) ist, z_1 eine Untergrenze eines Tiefenbereichs der Szene ist und z_2 eine Obergrenze eines Tiefenbereichs der Szene ist.

2. Tiefendetektionsvorrichtung (20) nach Anspruch 1, wobei

die Skalierungsvorrichtung (22) konfiguriert ist, die Skalierung des ersten Bilds oder des zweiten Bilds unter Berücksichtigung von Stereobildparametern zu bestimmen, und
die Stereobildparameter eine Grundlinie des Stereobilds und/oder eine Brennweite der Stereobildaufzeichnungsvorrichtung (21, 21a, 21b) und/oder eine Blende der Stereobildaufzeichnungsvorrichtung (21, 21a, 21b) und/oder eine Auflösung der Stereobildaufzeichnungsvorrichtung (21, 21a, 21b) und/oder eine Länge des ersten optischen Pfads (43) und/oder die Länge des zweiten optischen Pfads (44) umfassen.

3. Tiefendetektionsvorrichtung (20) nach Anspruch 1 oder 2, wobei
die Skalierungsvorrichtung (22) konfiguriert ist, die Skalierung des ersten Bilds oder des zweiten Bilds unter Berücksichtigung von Szenenparametern zu bestimmen und die Szenenparameter einen Tiefenbereich und/oder eine Tiefenverteilung der Objekte in der Szene umfassen.

4. Tiefendetektionsvorrichtung (20) nach Anspruch 2 oder 3, wobei
die Skalierungsvorrichtung (22) konfiguriert ist, ein mathematisches Modell der Stereobildaufzeichnungsvorrichtung (21, 21a, 21b) und/oder der Szene auf der Grundlage der Stereobildparameter und/oder der Szenenparameter zu bestimmen, und die Skalierungsvorrichtung (22) konfiguriert ist, die Skalierung des ersten Bilds oder des zweiten Bilds auf der Grundlage des mathematischen Modells zu bestimmen.

5. Tiefendetektionsvorrichtung (20) nach einem der Ansprüche 1 bis 4, wobei
die Tiefendetektionsvorrichtung (20) einen Kalibrator (31) umfasst, der konfiguriert ist zum:

- Durchführen einer Kalibrierung der Tiefendetektionsvorrichtung (20) auf der Grundlage einer bekannten Referenzszene und einer bekannten Tiefenkarte der bekannten Referenzszene, was in einer Kalibrierungsfunktion resultiert, und
- Durchführen der Kalibrierungsfunktion an jeder bestimmten Tiefenkarte nach einem Bestimmen der Tiefenkarte durch die Tiefenbestimmungsvorrichtung (23), was in einer kalibrierten Tiefenkarte resultiert.

6. Tiefendetektionsvorrichtung (20) nach Anspruch 5, wobei

die Stereobildaufzeichnungsvorrichtung (21, 21a, 21b) konfiguriert ist, ein Kalibrierungsstereobild der bekannten Referenzszene aufzuzeichnen, und
die Skalierungsvorrichtung (22) konfiguriert ist zum:

- Bestimmen einer Skalierung des ersten Bilds oder des zweiten Bilds des Kalibrierungsstereobilds durch Minimieren einer unterschiedlichen seitlichen Vergrößerung von Objekten in der Szene, die in dem ersten Bild und dem zweiten Bild des Kalibrierungsstereobilds dargestellt ist, und
- Durchführen der bestimmten Skalierung, was in einem skalierten Kalibrierungsstereobild resultiert, wobei

die Tiefenbestimmungsvorrichtung (23) konfiguriert ist, eine Kalibrierungstiefenkarte der bekannten Referenzszene auf der Grundlage des skalierten Kalibrierungsstereobilds zu bestimmen, und
der Kalibrator (31) konfiguriert ist zum:

- Bestimmen von Differenzen der Kalibrierungstiefenkarte und der bekannten Tiefenkarte und

- Bestimmen der Kalibrierungsfunktion aus den Differenzen der Kalibrierungstiefenkarte und der bekannten Tiefenkarte.

7. Tiefendetektionsvorrichtung (20) nach Anspruch 5 oder 6, wobei der Kalibrator (31) konfiguriert ist, die Kalibrierungsfunktion als eine nichtlineare Funktion zu bestimmen.

8. Tiefendetektionsvorrichtung (20) nach einem der Ansprüche 5 bis 7, wobei der Kalibrator (31) konfiguriert ist, die Kalibrierungsfunktion als eine Bildtransformationsmatrix zu bestimmen.

9. Tiefendetektionsvorrichtung (20) nach einem der Ansprüche 1 bis 8, wobei die Stereobildaufzeichnungsvorrichtung (21, 21a, 21b) konfiguriert ist, das erste Bild des Stereobilds und das zweite Bild des Stereobilds um eine Grundlinie verlagert aufzuzeichnen.

10. Tiefendetektionsvorrichtung (20) nach einem der Ansprüche 1 bis 9, wobei

die Stereobildaufzeichnungsvorrichtung (21, 21b) eine einzelne Kamera (50) und einen optischen Teiler (51) umfasst,
die einzelne Kamera (50) zum aufeinanderfolgenden Aufzeichnen des ersten Bilds und des zweiten Bilds konfiguriert ist und
der optische Teiler (51) zum aufeinanderfolgenden Schalten zwischen dem ersten optischen Pfad (43) und dem zweiten optischen Pfad (44) konfiguriert ist.

11. Tiefendetektionsvorrichtung (20) nach Anspruch 10, wobei

der optische Teiler (51) Folgendes umfasst:

- einen Strahlteiler (52), der vor der einzelnen Kamera (50) angeordnet ist, und
- ein Totalreflexionsprisma (53), das in einer Strahlteilungsrichtung des Strahlteilers (52) angeordnet ist, wobei

der erste optische Pfad (43) von der Szene zum Strahlteiler (52) zu der einzelnen Kamera (50) führt und
der zweite optische Pfad (44) von der Szene zum Totalreflexionsprisma (53) zum Strahlteiler (52) zu der einzelnen Kamera (50) führt.

12. Tiefendetektionsvorrichtung (20) nach Anspruch 10 oder 11, wobei

der optische Teiler (51) Folgendes umfasst:

- eine erste Verschlussvorrichtung (54), die im ersten optischen Pfad (43), jedoch nicht im zweiten optischen Pfad (44) angeordnet ist, und
- eine zweite Verschlussvorrichtung (55), die im zweiten optischen Pfad (44), jedoch nicht im ersten optischen Pfad (43) angeordnet ist, wobei die erste Verschlussvorrichtung (54) konfiguriert ist zum:
- Schließen des ersten optischen Pfads (43) während des Aufzeichnens des zweiten Bilds durch die einzelne Kamera und
- Öffnen des ersten optischen Pfads (43) während des Aufzeichnens des ersten Bilds durch die einzelne Kamera und

die zweite Verschlussvorrichtung (55) konfiguriert ist zum:

- Schließen des zweiten optischen Pfads (44) während des Aufzeichnens des ersten Bilds durch die einzelne Kamera und
- Öffnen des zweiten optischen Pfads (44) während des Aufzeichnens des zweiten Bilds durch die einzelne Kamera.

13. Tiefendetektionsvorrichtung (20) nach einem der Ansprüche 1 bis 9, wobei

die Stereobildaufzeichnungsvorrichtung (21, 21a) eine erste Kamera (40a) zum Aufzeichnen des ersten Bilds und eine zweite Kamera (40b) zum Aufzeichnen des zweiten Bilds umfasst und die erste Kamera (40a) und die zweite Kamera (40b) bei unterschiedlichen Entfernungen von der Szene angeordnet sind, was in den unterschiedlichen Längen des ersten optischen Pfads (43) und des zweiten optischen Pfads (44) resultiert.

14. Tiefendetektionsverfahren, das Folgendes umfasst:

- Aufzeichnen (100) eines Stereobilds einer Szene unter Verwendung einer Stereobildaufzeichnungsvorrichtung (21, 21a, 21b), die einen ersten optischen Pfad (43) und einen zweiten optischen Pfad (44) umfasst, wobei ein erstes Bild des Stereobilds über den ersten optischen Pfad (43) aufgezeichnet wird und ein zweites Bild des Stereobilds über den zweiten optischen Pfad (44) aufgezeichnet wird und der erste optische Pfad (43) und der zweite optische Pfad (44) verschiedene Längen besitzen,
- Bestimmen (101) einer Skalierung des ersten Bilds oder des zweiten Bilds durch Minimieren einer unterschiedlichen seitlichen Vergrößerung von Objekten in der Szene, die in dem ersten Bild und dem zweiten Bild dargestellt ist,
- Durchführen (102) der bestimmten Skalierung, was in einem skalierten Stereobild resultiert, und
- Bestimmen (103) einer Tiefenkarte der Szene auf der Grundlage des skalierten Stereobilds, wobei das Bestimmen (101) der Skalierung des ersten Bilds oder des zweiten Bilds ein Bestimmen eines Skalierungsfaktors M_o^{opt} als

$$M_o^{opt} = 1 + \frac{\Delta \log\frac{z_1}{z_2}}{z_2 - z_1}$$

umfasst, wobei Δ eine Differenz der optischen Länge des ersten optischen Pfads (43) und des zweiten optischen Pfads (44) ist, z_1 eine Untergrenze eines Tiefenbereichs der Szene ist und z_2 eine Obergrenze eines Tiefenbereichs der Szene ist.

Revendications

1. Dispositif de détection de profondeur (20), comprenant un dispositif d'enregistrement d'image stéréo (21, 21a, 21b), un dispositif de mise à l'échelle (22), et un dispositif de détermination de profondeur (23), le dispositif d'enregistrement d'image stéréo (21, 21a, 21b) étant configuré pour enregistrer une image stéréo d'une scène,

le dispositif d'enregistrement d'image stéréo (21, 21a, 21b) comprenant :

- un premier chemin optique (43), configuré pour enregistrer une première image de l'image stéréo, et
- un second chemin optique (44), configuré pour enregistrer une seconde image de l'image stéréo,

le premier chemin optique (43) et le second chemin optique (44) étant de longueur différente, le dispositif de mise à l'échelle (22) étant configuré pour

- déterminer une mise à l'échelle de la première image ou de la seconde image, en minimisant les différences d'agrandissement latéral des objets de la scène représentés dans la première image et la seconde image, et
- réaliser la mise à l'échelle déterminée, résultant en une image stéréo mise à l'échelle, le dispositif de détermination de profondeur (23) étant configuré pour déterminer une carte de profondeur de la scène sur la base de l'image stéréo mise à l'échelle, le dispositif de mise à l'échelle (22) étant configuré pour déterminer un facteur de mise à l'échelle M_0^{opt} comme suit :

$$M_0^{opt} = 1 + \frac{\Delta \log\frac{z_1}{z_2}}{z_2 - z_1},$$

Δ étant une différence de longueur optique entre le premier chemin optique (43) et le second chemin optique (44), z_1 étant une limite inférieure d'une plage de profondeur de la scène et z_2 étant une limite supérieure d'une plage de profondeur de la scène.

2. Dispositif de détection de profondeur (20) selon la revendication 1,

le dispositif de mise à l'échelle (22) étant configuré pour déterminer la mise à l'échelle de la première image ou de la seconde image en tenant compte des paramètres de l'image stéréo, et
les paramètres de l'image stéréo comprenant une ligne de base de l'image stéréo, et/ou une longueur focale du dispositif d'enregistrement d'image stéréo (21, 21a, 21b), et/ou une ouverture du dispositif d'enregistrement d'image stéréo (21, 21a, 21b), et/ou une résolution du dispositif d'enregistrement d'image stéréo (21, 21a, 21b), et/ou une longueur du premier chemin optique (43), et/ou la longueur du second chemin optique (44).

3. Dispositif de détection de profondeur (20) selon la revendication 1 ou 2,

le dispositif de mise à l'échelle (22) étant configuré pour déterminer la mise à l'échelle de la première image ou de la seconde image, en tenant compte des paramètres de la scène, et
les paramètres de la scène comprenant une plage de profondeur et/ou une distribution de profondeur des objets dans la scène.

4. Dispositif de détection de profondeur (20) selon la revendication 2 ou 3,
le dispositif de mise à l'échelle (22) étant configuré pour déterminer un modèle mathématique du dispositif d'enregistrement d'image stéréo (21, 21a, 21b) et/ou de la scène, sur la base des paramètres de l'image stéréo et/ou des paramètres de la scène, et le dispositif de mise à l'échelle (22) étant configuré pour déterminer la mise à l'échelle de la première image ou de la seconde image sur la base du modèle mathématique.

5. Dispositif de détection de profondeur (20) selon l'une quelconque des revendications 1 à 4,
le dispositif de détection de profondeur (20) comprenant un dispositif d'étalonnage (31) configuré pour

- réaliser un étalonnage du dispositif de détection de profondeur (20) sur la base d'une scène de référence connue et d'une carte de profondeur connue de la scène de référence connue, résultant en une fonction d'étalonnage, et
- réaliser la fonction d'étalonnage sur chaque carte de profondeur déterminée après avoir déterminé la carte de profondeur par le dispositif de détermination de profondeur (23), résultant en une carte de profondeur étalonnée.

6. Dispositif de détection de profondeur (20) selon la revendication 5,

le dispositif d'enregistrement d'image stéréo (21, 21a, 21b) étant configuré pour enregistrer une image stéréo d'étalonnage de la scène de référence connue,
le dispositif de mise à l'échelle (22) étant configuré pour

- déterminer une mise à l'échelle de la première image ou de la seconde image de l'image stéréo d'étalonnage, en minimisant les différences de grossissement latéral des objets de la scène représentés dans la première image et la seconde image de l'image stéréo d'étalonnage,
- réaliser la mise à l'échelle déterminée, résultant en une image stéréo d'étalonnage mise à l'échelle,

le dispositif de détermination de profondeur (23) étant configuré pour déterminer une carte de profondeur d'étalonnage de la scène de référence connue sur la base de l'image stéréo d'étalonnage mise à l'échelle,
le dispositif d'étalonnage (31) étant configuré pour

- déterminer les différences entre la carte de profondeur d'étalonnage et la carte de profondeur connue, et
- déterminer la fonction d'étalonnage à partir des différences entre la carte de profondeur d'étalonnage et la carte de profondeur connue.

7. Dispositif de détection de profondeur (20) selon la revendication 5 ou 6,
le dispositif d'étalonnage (31) étant configuré pour déterminer la fonction d'étalonnage comme une fonction non linéaire.

8. Dispositif de détection de profondeur (20) selon l'une quelconque des revendications 5 à 7,
le dispositif d'étalonnage (31) étant configuré pour déterminer la fonction d'étalonnage en tant que matrice de transformation d'image.

9. Dispositif de détection de profondeur (20) selon l'une quelconque des revendications 1 à 8, le dispositif d'enregistrement d'image stéréo (21, 21a, 21b) étant configuré pour enregistrer la première image de l'image stéréo et la seconde image de l'image stéréo décalée d'une ligne de base.

10. Dispositif de détection de profondeur (20) selon l'une quelconque des revendications 1 à 9,

le dispositif d'enregistrement d'image stéréo (21, 21b) comprenant une caméra unique (50) et un séparateur optique (51),
la caméra unique (50) étant configurée pour enregistrer successivement la première image et la seconde image, et
le séparateur optique (51) étant configuré pour commuter entre le premier chemin optique (43) et le second chemin optique (44) successivement.

11. Dispositif de détection de profondeur (20) selon la revendication 10,

le séparateur optique (51) comprenant :

- un séparateur de faisceau (52) agencé devant la caméra unique (50), et
- un prisme à réflexion totale (53) agencé dans la direction de séparation du faisceau du séparateur de faisceau (52),

le premier chemin optique (43) conduisant de la scène au séparateur de faisceau (52) jusqu'à la caméra unique (50), et
le second chemin optique (44) conduisant de la scène au prisme à réflexion totale (53), au séparateur de faisceau (52) jusqu'à la caméra unique (50).

12. Dispositif de détection de profondeur (20) selon la revendication 10 ou 11,

le séparateur optique (51) comprenant :

- un premier dispositif d'obturation (54) agencé à l'intérieur du premier chemin optique (43) mais pas à l'intérieur du second chemin optique (44), et
- un second dispositif d'obturation (55) agencé à l'intérieur du second chemin optique (44) mais pas à l'intérieur du premier chemin optique (43),

le premier dispositif d'obturation (54) étant configuré pour

- fermer le premier chemin optique (43) pendant l'enregistrement de la seconde image par la caméra unique, et
- ouvrir le premier chemin optique (43) pendant l'enregistrement de la première image par la caméra unique, et

le second dispositif d'obturation (55) étant configuré pour

- fermer le second chemin optique (44) pendant l'enregistrement de la première image par la caméra unique, et
- ouvrir le second chemin optique (44) pendant l'enregistrement de la seconde image par la caméra unique.

13. Dispositif de détection de profondeur (20) selon l'une quelconque des revendications 1 à 9,

le dispositif d'enregistrement d'image stéréo (21, 21a) comprenant une première caméra (40a) pour l'enregistrement de la première image, et une seconde caméra (40b) pour l'enregistrement de la seconde image, et
la première caméra (40a) et la seconde caméra (40b) étant situées à des distances différentes de la scène, résultant en des longueurs différentes du premier chemin optique (43) et du second chemin optique (44).

14. Procédé de détection de profondeur, comprenant :

- l'enregistrement (100) d'une image stéréo d'une scène, à l'aide d'un dispositif d'enregistrement d'image stéréo

(21, 21a, 21b), comprenant un premier chemin optique (43), et un second chemin optique (44), une première image de l'image stéréo étant enregistrée au moyen du premier chemin optique (43) et une seconde image de l'image stéréo étant enregistrée au moyen du second chemin optique (44),
le premier chemin optique (43) et le second chemin optique (44) étant de longueur différente,
- la détermination (101) d'une mise à l'échelle de la première image ou de la seconde image, en minimisant les différences d'agrandissement latéral des objets de la scène représentés dans la première image et la seconde image,
- la réalisation (102) de la mise à l'échelle déterminée, résultant en une image stéréo mise à l'échelle, et
- la détermination (103) d'une carte de profondeur de la scène sur la base de l'image stéréo mise à l'échelle,

la détermination (101) de la mise à l'échelle de la première image ou de la seconde image comprenant la détermination d'un facteur de mise à l'échelle M_0^{opt} comme suit :

$$M_0^{opt} = 1 + \frac{\Delta \log\frac{z_1}{z_2}}{z_2 - z_1},$$

Δ étant une différence de longueur optique entre le premier chemin optique (43) et le second chemin optique (44), z_1 étant une limite inférieure d'une plage de profondeur de la scène et z_2 étant une limite supérieure d'une plage de profondeur de la scène.

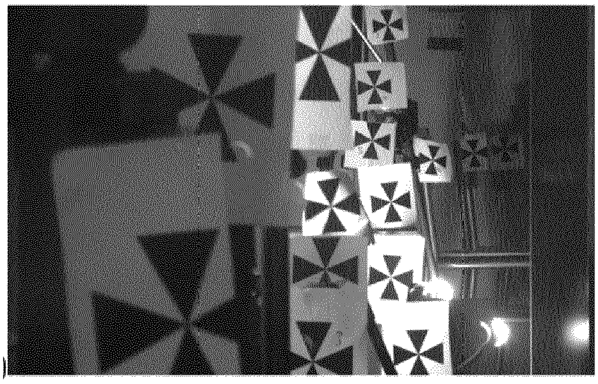


Fig. 1a

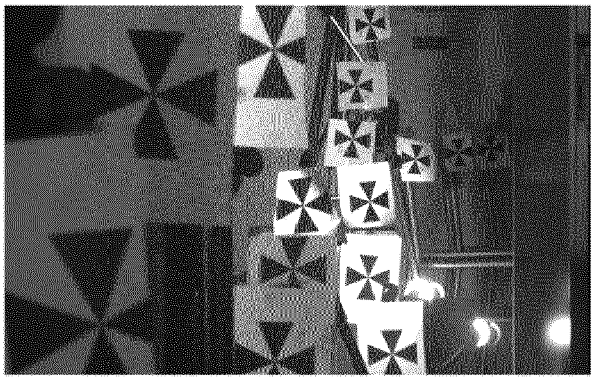


Fig. 1b

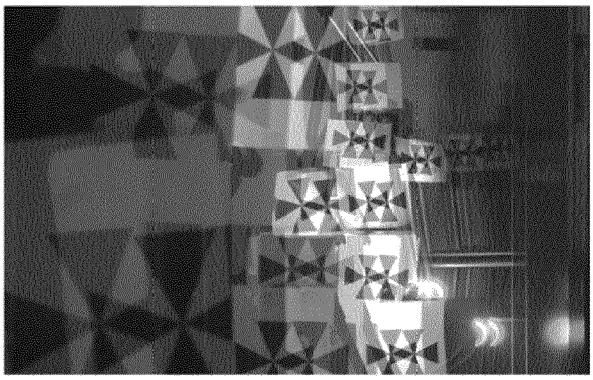


Fig. 1c

Fig. 2

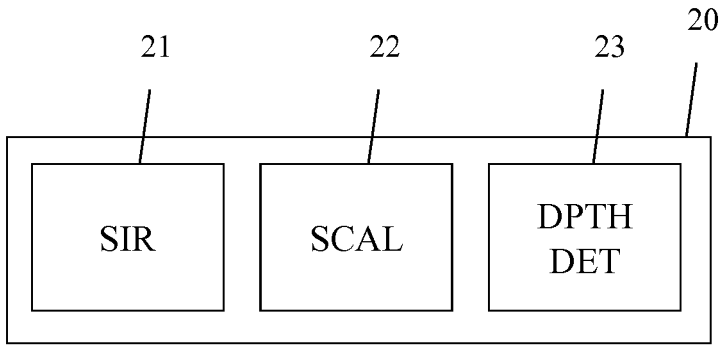
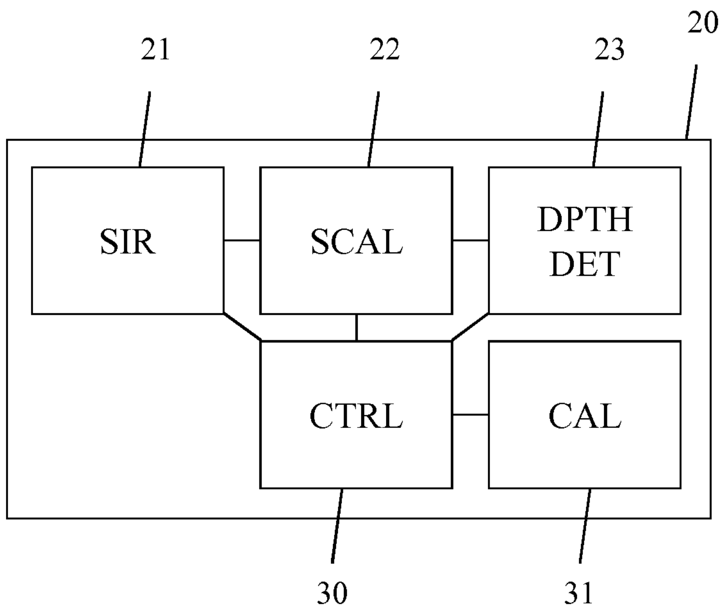

20
21
22
23
SIR
SCAL
DPTH
DET

Fig. 3


20
21
22
23
SIR
SCAL
DPTH
DET
CTRL
CAL
30
31

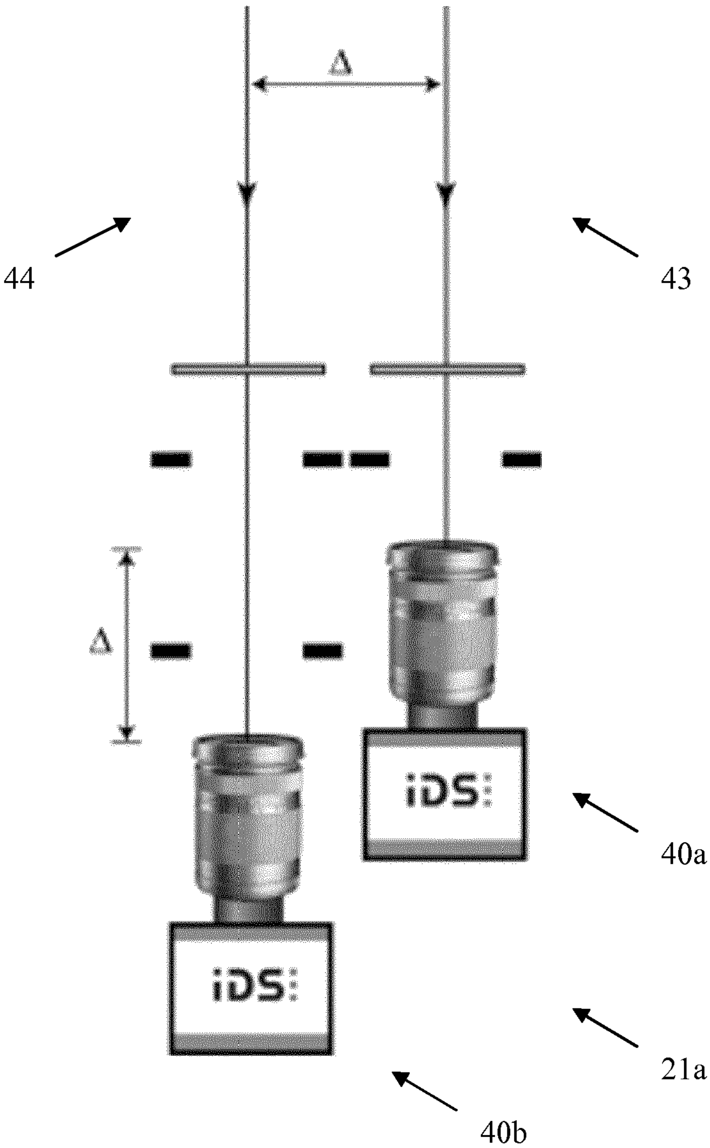


Fig. 4

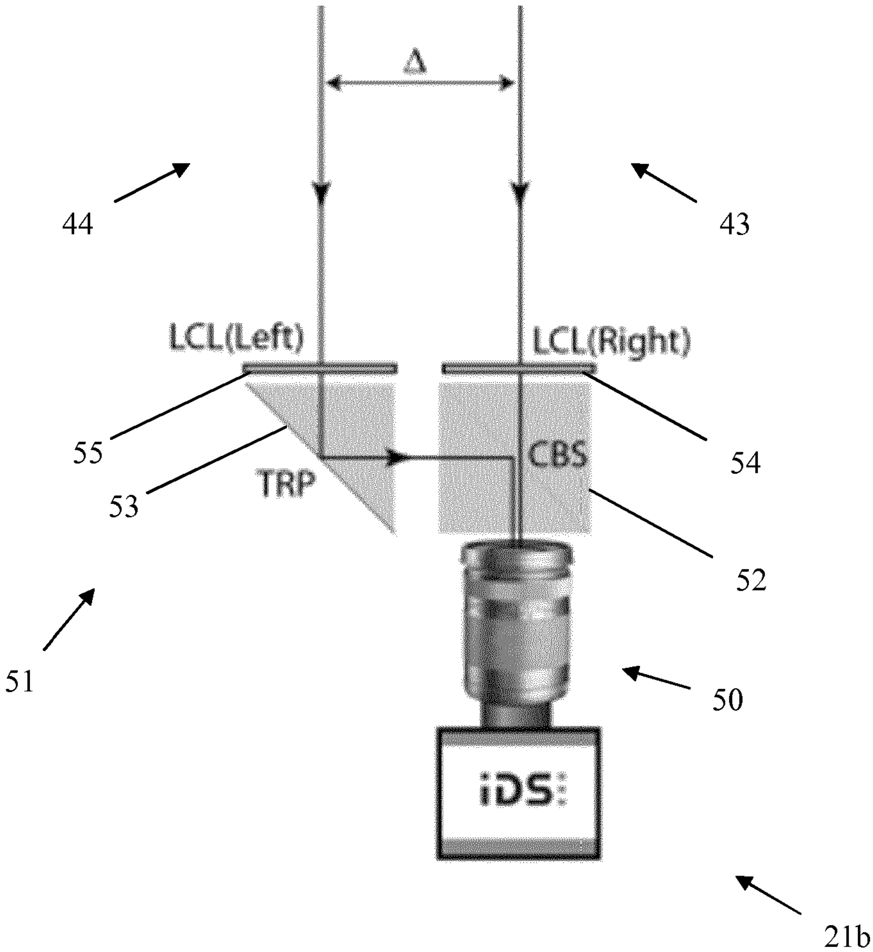


Fig. 5

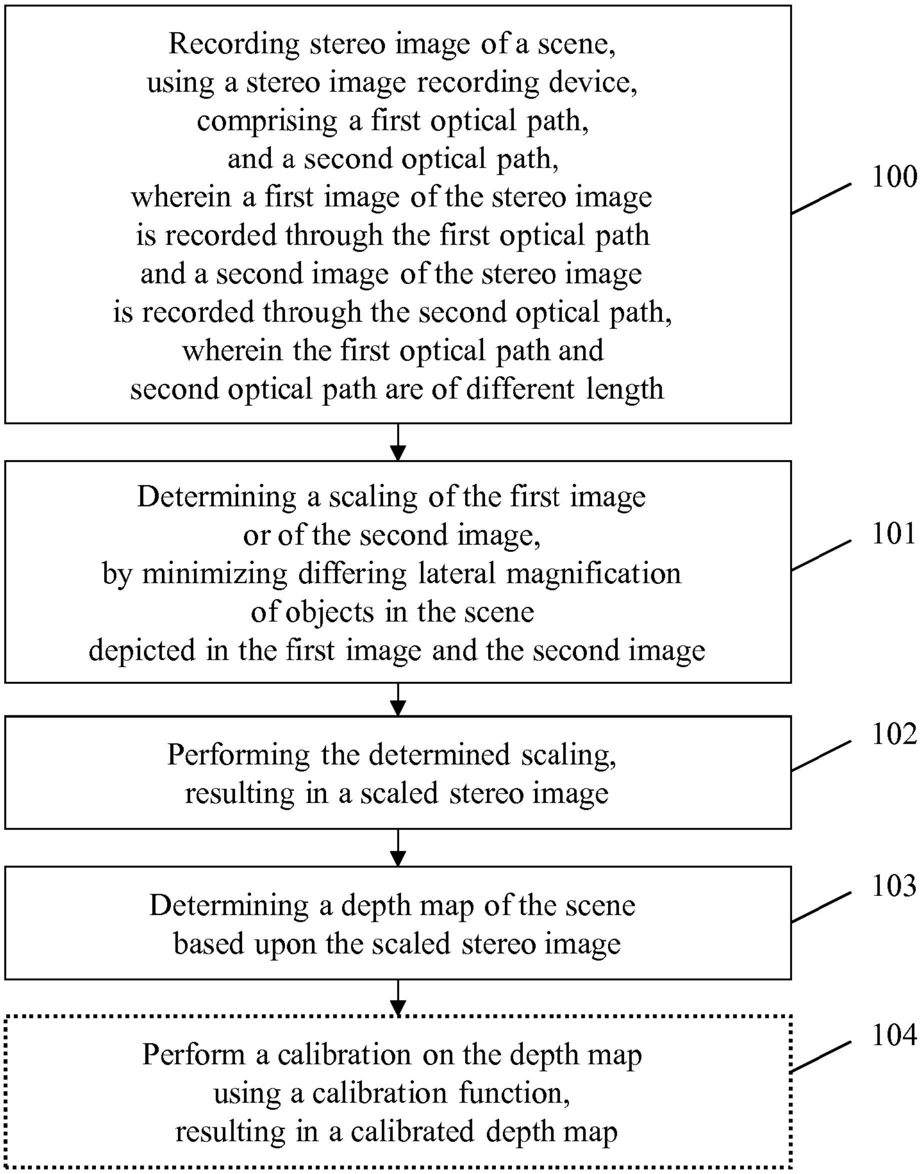

Recording stereo image of a scene, using a stereo image recording device, comprising a first optical path, and a second optical path, wherein a first image of the stereo image is recorded through the first optical path and a second image of the stereo image is recorded through the second optical path, wherein the first optical path and second optical path are of different length
100
Determining a scaling of the first image or of the second image, by minimizing differing lateral magnification of objects in the scene depicted in the first image and the second image
101
Performing the determined scaling, resulting in a scaled stereo image
102
Determining a depth map of the scene based upon the scaled stereo image
103
Perform a calibration on the depth map using a calibration function, resulting in a calibrated depth map
104

Fig. 6

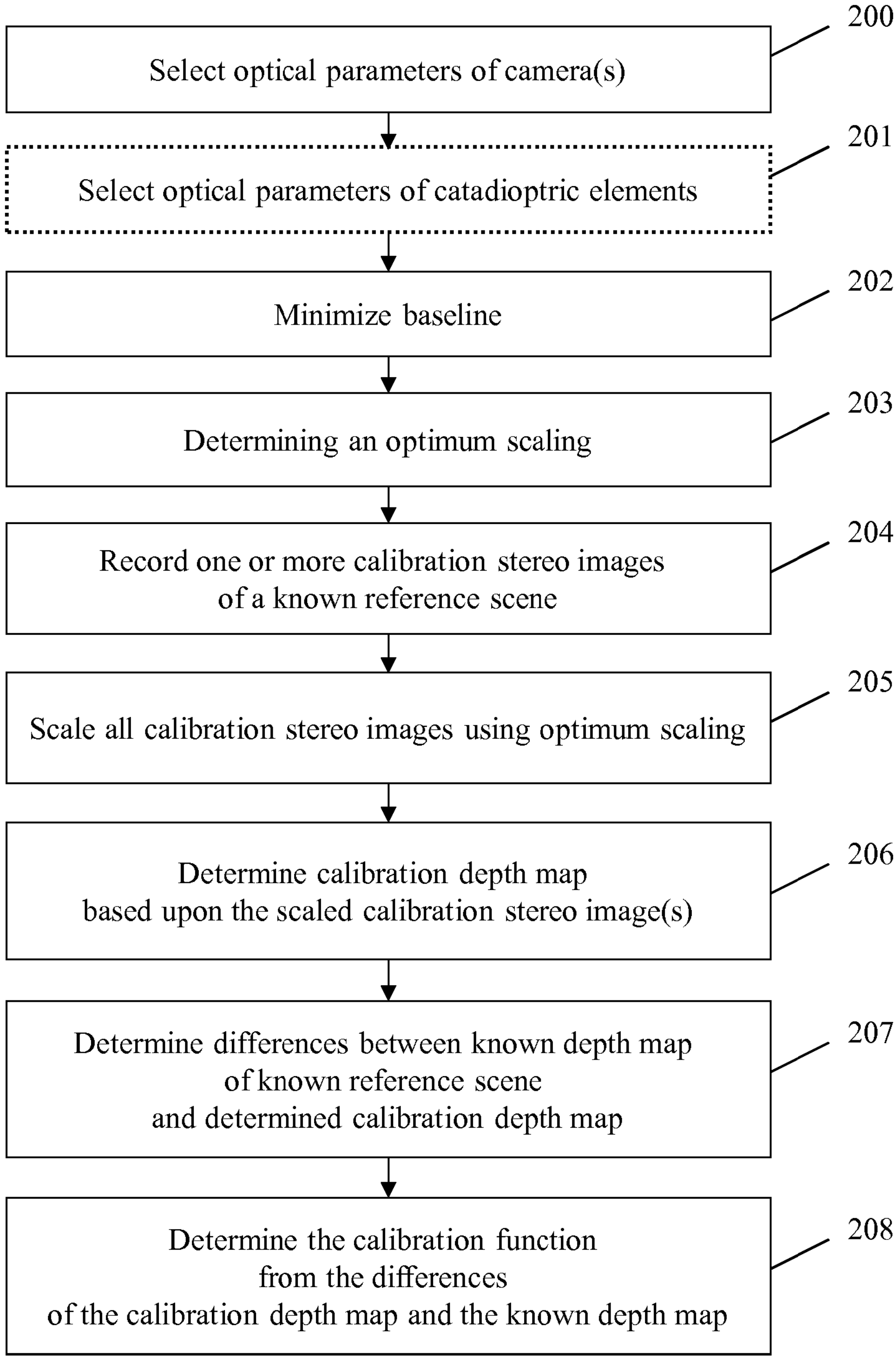

Select optical parameters of camera(s)
200
Select optical parameters of catadioptric elements
201
Minimize baseline
202
Determining an optimum scaling
203
Record one or more calibration stereo images of a known reference scene
204
Scale all calibration stereo images using optimum scaling
205
Determine calibration depth map based upon the scaled calibration stereo image(s)
206
Determine differences between known depth map of known reference scene and determined calibration depth map
207
Determine the calibration function from the differences of the calibration depth map and the known depth map
208

Fig. 7

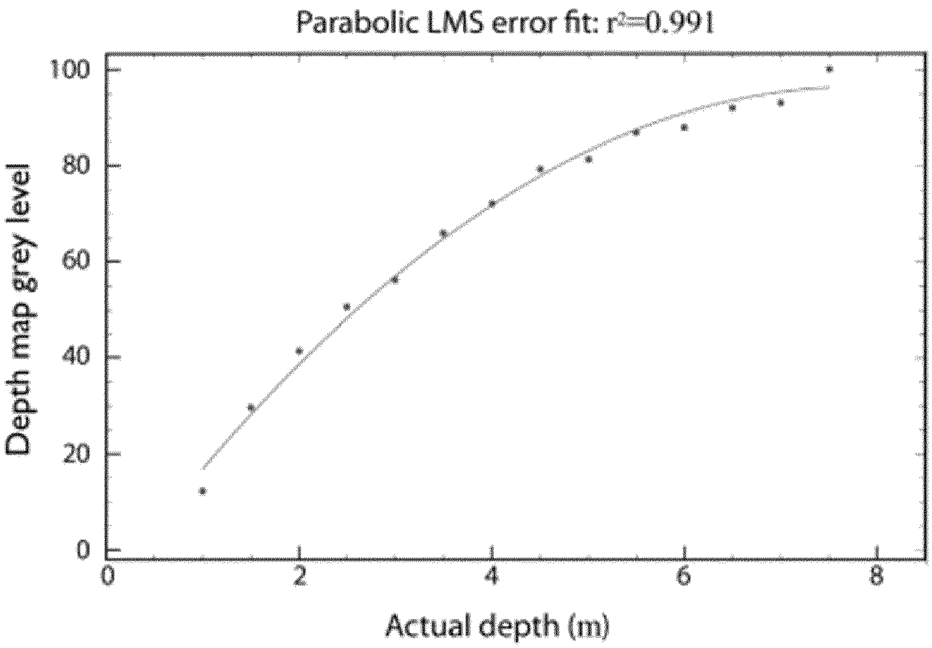

Parabolic LMS error fit: r²=0.991
Depth map grey level
Actual depth (m)
0
20
40
60
80
100
0
2
4
6
8

Fig. 8

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- US 5522789 A **[0009]**
- US 2013321790 A **[0010]**