



(51) International Patent Classification:
H04N 13/02 (2006.01)

(21) International Application Number:
PCT/CN2015/074336

(22) International Filing Date:
16 March 2015 (16.03.2015)

(25) Filing Language: English

(26) Publication Language: English

(71) Applicant: SZ DJI TECHNOLOGY CO., LTD.
[CN/CN]; 6F, HKUST SZ IER BLdg., No.9 Yuexing 1st Rd. Hi-Tech Park(South), Nanshan District, Shenzhen, Guangdong 518057 (CN).

(72) Inventors: ZHU, Zhenyu; 6F, HKUST SZ IER BLdg., No.9 Yuexing 1st Rd. Hi-Tech Park(South), Nanshan District, Shenzhen, Guangdong 518057 (CN). ZHAO, Cong; 6F, HKUST SZ IER BLdg., No.9 Yuexing 1st Rd. Hi-Tech Park(South), Nanshan District, Shenzhen, Guangdong 518057 (CN).

(74) Agent: SHENZHEN SCIENBIZIP INTELLECTUAL PROPERTY AGENCY CO.,LTD.; 9F, Rongqun build-

ing, NO.83 Longguan east Rd.Longhua new Dist., Shenzhen, Guangdong 518109 (CN).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, 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).

[Continued on next page]

(54) Title: APPARATUS AND METHOD FOR FOCAL LENGTH ADJUSTMENT AND DEPTH MAP DETERMINATION

(57) Abstract: Apparatus and methods for focal length adjustment and depth map determination are provided for making and using the same. The focal length adjustment apparatus comprises a distance assembly configured to determine a distance between an object of interest and an imaging mechanism for imaging the object of interest. The focal length adjustment apparatus likewise can include a focal length assembly configured to automatically adjust a focal length of the imaging mechanism according to the determined distance. The focal length adjustment apparatus advantageously achieves automatic focal length adjustment with lower cost compared with solutions using laser cameras.

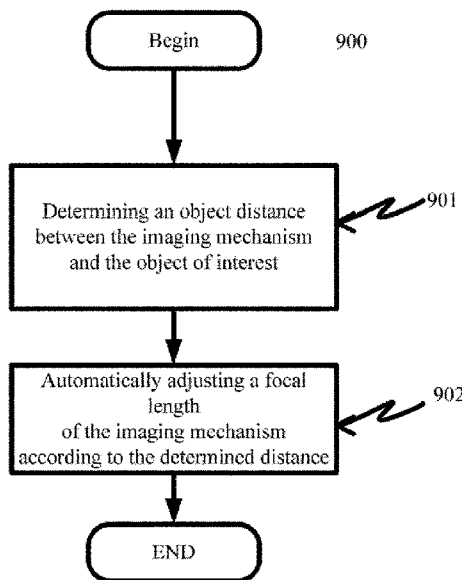


Fig. 9



Published:

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

APPARATUS AND METHOD FOR FOCAL LENGTH ADJUSTMENT AND DEPTH MAP DETERMINATION

FIELD

[0001] The disclosed embodiments relate generally to digital imaging and more particularly, but not exclusively, to apparatus and methods for adjusting a focal length automatically and/or determining a depth map for an image.

BACKGROUND

[0002] Stereoscopic imaging, a technique whereby multiple imaging devices are used to form a three dimensional image through stereopsis, is becoming increasingly common in many fields. Stereoscopic imaging is particularly useful in robotics, where it is often desirable to gather three-dimensional information about an operating environment of a machine. Stereoscopic imaging simulates the binocular visions of human eyes and applies the principle of stereopsis to achieve depth perception. This technique can be reproduced by artificial imaging devices by viewing a given object of interest using multiple imaging devices from slightly different vantage points. Differences between varying views of the object of interest convey depth information about a position of the object, thereby enabling three-dimensional imaging of the object.

[0003] During the shooting of videos or television plays, tracking focal length is a difficult and professional task. Main current focal length tracking devices require experienced focal length tracking operators to adjust focal length of the cameras based on the monitor screens and the shooting site situations in real time.

[0004] For certain imaging applications, manual focal length adjustment is cumbersome and may be impractical where the imaging device is operated remotely. Accordingly, it is desirable that a focal length adjustment system be able to automatically adjust the focal length to track a moving object of interest.

[0005] SUMMARY

[0006] In accordance with a first aspect disclosed herein, there is set forth a method for automatic focal length adjustment, comprising:

[0007] determining a distance between an object of interest and an imaging mechanism; and

[0008] automatically adjusting a focal length of the imaging mechanism according to said determining the distance.

[0009] In an exemplary embodiment of the disclosed methods, the determining comprises determining the distance between the object of interest in a scene and the imaging mechanism.

[0010] In an exemplary embodiment of the disclosed methods, the determining comprises

imaging the scene with first and second imaging devices contained in the imaging mechanism.

- [0011] In an exemplary embodiment of the disclosed methods, the determining further comprises:
- [0012] obtaining a depth map of the scene;
- [0013] selecting the object of interest in the scene; and
- [0014] calculating the distance of the object of interest according to the depth map of the scene.
- [0015] In an exemplary embodiment of the disclosed methods, the obtaining comprises calculating a disparity of scene images from the first and second imaging devices.
- [0016] In an exemplary embodiment of the disclosed methods, the calculating the disparity comprises optimizing a global energy function.
- [0017] In an exemplary embodiment of the disclosed methods, the optimizing the global energy function comprises summing a disparity energy function and a scaled smoothing term.
- [0018] In an exemplary embodiment of the disclosed methods, the disparity energy function is represented by a Birchfield-Tomasi term.
- [0019] In an exemplary embodiment of the disclosed methods, the Birchfield-Tomasi term is defined by accumulating the minimum disparity of coordinates for pixels in a first image of the scene captured by the first imaging device and a second image of the scene captured by the second imaging device.
- [0020] An exemplary embodiment of the disclosed methods further comprises, for all neighbors of a pixel, accumulating scaled trigger functions of a disparity between two neighboring pixels to obtain the smoothing term.
- [0021] In an exemplary embodiment of the disclosed methods, the accumulating comprises, for all neighbors of a pixel, accumulating scaled trigger functions of a disparity between two neighboring pixels from four domains.
- [0022] In an exemplary embodiment of the disclosed methods, the smoothing term is obtained by accumulating scaled trigger functions of disparity for all neighbors of each pixel.
- [0023] An exemplary embodiment of the disclosed methods further comprises:
- [0024] aggregating data items in a plurality of directions to obtain an energy function for each of the directions; and
- [0025] accumulating the energy functions in the directions to obtain the energy function.
- [0026] In an exemplary embodiment of the disclosed methods, the aggregating comprises obtaining energy functions in a predetermined number of directions.
- [0027] In an exemplary embodiment of the disclosed methods, the aggregating comprises obtaining energy functions in four or eight directions.

- [0028] In an exemplary embodiment of the disclosed methods, the aggregating data items comprises obtaining energy function in a direction by summing a corresponding smoothing item and a dynamic planning in the direction.
- [0029] In an exemplary embodiment of the disclosed methods, the summing the corresponding smoothing item and the dynamic planning in the direction comprises presenting the dynamic planning in the direction with a recurrence based on the energy functions of its neighbors in this direction.
- [0030] In an exemplary embodiment of the disclosed methods, the direction comprises a horizontal direction.
- [0031] In an exemplary embodiment of the disclosed methods, the aggregating data items in the horizontal direction comprises calculating the energy by a recurrence based on said energy functions of its neighbors in the horizontal direction.
- [0032] An exemplary embodiment of the disclosed methods further comprises obtaining the best depth.
- [0033] In an exemplary embodiment of the disclosed methods, the obtaining the best depth comprises seeking the disparity value that minimizes the summation of energy in a plurality of directions.
- [0034] In an exemplary embodiment of the disclosed methods, the obtaining the best depth comprises seeking the disparity value based on an energy function in one direction.
- [0035] An exemplary embodiment of the disclosed methods further comprises reducing noise by doing at least one of matching scene images from the first and second imaging devices and identifying respective unique features of the scene images while setting the disparity as -1.
- [0036] An exemplary embodiment of the disclosed methods further comprises compensating an error based on the factors at least selected from a group consisting of the distance between a central line of the two imaging devices, an actual distance of two adjacent pixels, the focal length of the two imaging devices and the depth between the object of interest and the first and second imaging devices.
- [0037] An exemplary embodiment of the disclosed methods further comprises optimizing the depth map by using a non-partial optimizing equation.
- [0038] An exemplary embodiment of the disclosed methods further comprises obtaining a Jacobi iteration of the non-partial optimizing equation by using a recurrence filtering.
- [0039] In an exemplary embodiment of the disclosed methods, the selecting the object of interest in the scene comprises receiving outside instructions to select the object of interest.
- [0040] In an exemplary embodiment of the disclosed methods, the receiving the instructions comprises identifying the object of interest selected on either of scene images from the first or the second imaging device.

- [0041] In an exemplary embodiment of the disclosed methods, the identifying the object of interest selected comprises sensing a frame on either of the scene images framing in the object of interest or sensing a click on the object of interest on either of the scene images.
- [0042] In an exemplary embodiment of the disclosed methods, the receiving the outside instructions comprises receiving vocal instructions, optionally a pre-set name of the object of interest, to determine the object of interest.
- [0043] In an exemplary embodiment of the disclosed methods, the selecting the object of interest in the scene comprises judging under at least one pre-set rule and automatically determining the object of interest based on said judging.
- [0044] In an exemplary embodiment of the disclosed methods, the judging under the at least one pre-set rule comprises judging if the object is approaching or within a certain distance of the imaging mechanism.
- [0045] In an exemplary embodiment of the disclosed methods, the automatically adjusting the focus comprises automatically adjusting the focal length of the imaging mechanism in real time with tracking learning detection based on gray level information of the object of interest.
- [0046] In accordance with another aspect disclosed herein, there is set forth a stereoscopic imaging system configured to perform automatic focal length adjustment in accordance with any one of the above methods.
- [0047] In accordance with another aspect disclosed herein, there is set forth a focal length adjustment apparatus, comprising:
- [0048] a distance assembly for determining a distance between an object of interest in a scene and an imaging mechanism for imaging the scene; and
- [0049] a focal length assembly for automatically adjusting a focal length of the imaging mechanism according to the determined distance.
- [0050] In an exemplary embodiment of the disclosed apparatus, the distance assembly is configured to determine the distance between the object of interest in a scene and the imaging mechanism.
- [0051] In an exemplary embodiment of the disclosed apparatus, the imaging mechanism comprises first and second imaging devices imaging the scene to obtain first and second scene images.
- [0052] In an exemplary embodiment of the disclosed apparatus, either of the first and second imaging devices is a camera or a sensor.
- [0053] In an exemplary embodiment of the disclosed apparatus, the first and second imaging devices are selected from a group consisting of laser cameras, infrared cameras, ultrasound cameras and Time-of-Flight cameras.
- [0054] In an exemplary embodiment of the disclosed apparatus, the first and second imaging

devices are Red-Green-Blue (RGB) cameras.

[0055] In an exemplary embodiment of the disclosed apparatus, the distance assembly comprises:

[0056] a depth estimation mechanism for obtaining a depth map of the scene;

[0057] an object determination mechanism for determining the object of interest in the scene; and

[0058] a calculating mechanism for calculating distance of the object of interest according to the depth map of the scene.

[0059] In an exemplary embodiment of the disclosed apparatus, the depth map is obtained based on a disparity of the first and the second scene images.

[0060] In an exemplary embodiment of the disclosed apparatus, the depth estimation mechanism optimizes a global energy function.

[0061] In an exemplary embodiment of the disclosed apparatus, the global energy function is defined as a sum of a disparity energy function and a scaled smoothing term.

[0062] In an exemplary embodiment of the disclosed apparatus, the disparity energy function comprises a Birchfield-Tomasi data term.

[0063] In an exemplary embodiment of the disclosed apparatus, the Birchfield-Tomasi data term is defined based on the minimum disparity of coordinates for pixels in a first image of the scene captured by the first imaging device and a second image of the scene captured by the second imaging device.

[0064] In an exemplary embodiment of the disclosed apparatus, the smoothing term adopts an energy function of differential of disparity.

[0065] In an exemplary embodiment of the disclosed apparatus, the smoothing term is, for all neighbors of a pixel with the coordination of (x,y) , a summation of scaled trigger function of a disparity between two neighboring pixels.

[0066] In an exemplary embodiment of the disclosed apparatus, the neighbors are pixels from four domains.

[0067] In an exemplary embodiment of the disclosed apparatus, the smoothing term is defined based on scaled trigger functions of a disparity for all neighbors of each pixel.

[0068] In an exemplary embodiment of the disclosed apparatus, the global energy function is optimized by

[0069] aggregating data terms in a plurality of directions to obtain a directional energy function for each of the directions; and

[0070] accumulating the directional energy functions in the directions to obtain the energy function.

[0071] In an exemplary embodiment of the disclosed apparatus, the directions comprise a predetermined number of directions.

[0072] In an exemplary embodiment of the disclosed apparatus, the predetermined number

of directions comprise four or eight directions.

- [0073] In an exemplary embodiment of the disclosed apparatus, the energy function in one direction is based on a dynamic planning in the direction.
- [0074] In an exemplary embodiment of the disclosed apparatus, the energy function in one direction is obtained by summing a corresponding smoothing term and the dynamic planning in this direction.
- [0075] In an exemplary embodiment of the disclosed apparatus, the dynamic planning in the direction is a recurrence based on the energy functions of its neighbors in this direction.
- [0076] In an exemplary embodiment of the disclosed apparatus, the direction comprises a horizontal direction.
- [0077] In an exemplary embodiment of the disclosed apparatus, the energy function in horizontal direction is obtained by a recurrence based on said energy functions of its neighbors the horizontal direction.
- [0078] In an exemplary embodiment of the disclosed apparatus, the best depth is obtained by seeking the disparity value that minimizes the summation of energy in a plurality of directions.
- [0079] In an exemplary embodiment of the disclosed apparatus, the best depth is obtained based on an energy function in one direction.
- [0080] In an exemplary embodiment of the disclosed apparatus, noise is reduced by matching the first and second scene images and/or identifying respective unique features of the first and second scene images while setting the disparity as -1.
- [0081] In an exemplary embodiment of the disclosed apparatus, an error is compensated based on the factors at least selected from a group consisting of the distance between the central lines of the two imaging devices, actual distance of two adjacent pixels, the focal length of the two imaging devices and the depth between the object of interest and the first and second imaging devices.
- [0082] In an exemplary embodiment of the disclosed apparatus, the depth map is optimized by using a non-partial optimizing equation.
- [0083] An exemplary embodiment of the disclosed apparatus comprises obtaining a Jacobi iteration of the non-partial optimizing equation by a recurrence filtering.
- [0084] In an exemplary embodiment of the disclosed apparatus, the object determination mechanism receives outside instructions to determine the object of interest.
- [0085] In an exemplary embodiment of the disclosed apparatus, the object determination mechanism enables to identify the object of interest selected on either of the first and the second scene images.
- [0086] In an exemplary embodiment of the disclosed apparatus, the object determination mechanism enables to identify the object of interest by at least one of sensing a frame

on either of the first and the second scene images to frame in the object of interest and sensing a click on the object of interest on either of the first and the second scene images.

[0087] In an exemplary embodiment of the disclosed apparatus, the object determination mechanism receives outside vocal instructions, optionally a pre-set name of the object of interest, to determine the object of interest.

[0088] In an exemplary embodiment of the disclosed apparatus, the object determination mechanism automatically determines the object of interest based on a judgment under at least a pre-set rule.

[0089] In an exemplary embodiment of the disclosed apparatus, the pre-set rule comprises if the object of interest is approaching the first and the second imaging devices or within a certain distance.

[0090] In an exemplary embodiment of the disclosed apparatus, the focal length assembly automatically adjusts the focal length of the imaging mechanism in real time with tracking learning detection based on gray level information of the object of interest.

[0091] In accordance with another aspect disclosed herein, there is set forth a mobile platform performing in accordance with any one of the above methods.

[0092] In accordance with another aspect disclosed herein, there is set forth a mobile platform comprising any one of the above apparatus.

[0093] In accordance with another aspect disclosed herein, the above mobile platform is an unmanned aerial vehicle (UAV).

[0094] In accordance with another aspect disclosed herein, the above mobile platform is a self-stabilizing platform.

[0095] In accordance with another aspect disclosed herein, there is set forth a method for obtaining a depth map of a scene, comprising:

[0096] capturing a plurality of scene images of the scene; and

[0097] calculating a disparity of the plurality of scene images.

[0098] In an exemplary embodiment of the disclosed methods, the capturing the plurality of scene images comprises capturing the plurality of scene images via first and second imaging devices.

[0099] In an exemplary embodiment of the disclosed methods, the calculating the disparity comprises optimizing a global energy function.

[0100] In an exemplary embodiment of the disclosed methods, the optimizing the global energy function comprises summing a disparity energy function and a scaled smoothing term.

[0101] In an exemplary embodiment of the disclosed methods, the disparity energy function is represented by a Birchfield-Tomasi term.

[0102] In an exemplary embodiment of the disclosed methods, wherein the Birchfield-

Tomasi term is defined by accumulating a minimum disparity of coordinates for pixels in a first image of the scene captured by the first imaging device and a second image of the scene captured by the second imaging device.

- [0103] An exemplary embodiment of the disclosed methods further comprises, for all neighbors of a selected pixel, accumulating scaled trigger functions of a disparity between two neighboring pixels to the selected pixel to obtain the smoothing term.
- [0104] In an exemplary embodiment of the disclosed methods, the accumulating comprises, for all neighbors of the selected pixel, accumulating scaled trigger functions of a disparity between two neighboring pixels from four domains.
- [0105] In an exemplary embodiment of the disclosed methods, the smoothing term is obtained by accumulating scaled trigger functions of a disparity for all neighbors of each pixel.
- [0106] An exemplary embodiment of the disclosed methods further comprises:
- [0107] aggregating data items in a plurality of directions to obtain a directional energy function for each of the directions; and
- [0108] accumulating the directional energy functions in the directions to obtain the energy function.
- [0109] In an exemplary embodiment of the disclosed methods, the aggregating comprises obtaining energy functions in a predetermined number of directions.
- [0110] In an exemplary embodiment of the disclosed methods, the aggregating comprises obtaining energy functions in four or eight directions.
- [0111] In an exemplary embodiment of the disclosed methods, the aggregating data items comprises obtaining the energy function in a selected direction by summing a corresponding smoothing item and a dynamic planning in the selected direction.
- [0112] In an exemplary embodiment of the disclosed methods, the summing the corresponding smoothing item and the dynamic planning in the direction comprises presenting the dynamic planning in the direction with a recurrence based on the energy functions of its neighbors in the direction.
- [0113] In an exemplary embodiment of the disclosed methods, the direction comprises a horizontal direction.
- [0114] In an exemplary embodiment of the disclosed methods, the aggregating data items in the horizontal direction comprises calculating the energy by a recurrence based on the energy functions of its neighbors in the horizontal direction.
- [0115] An exemplary embodiment of the disclosed methods further comprises obtaining a best depth.
- [0116] In an exemplary embodiment of the disclosed methods, the obtaining the best depth comprises seeking the disparity value that minimizes the summation of energy in a plurality of directions.

- [0117] In an exemplary embodiment of the disclosed methods, the obtaining the best depth comprises seeking the disparity value based on an energy function in one direction.
- [0118] An exemplary embodiment of the disclosed methods further comprises reducing noise by doing at least one of matching scene images from the first and second imaging devices and identifying respective unique features of the scene images while setting the disparity as -1.
- [0119] An exemplary embodiment of the disclosed methods further comprises compensating an error based on factors at least selected from a group consisting of a distance between central lines of the two imaging devices, an actual distance of two adjacent pixels, a focal length of the two imaging devices and a depth between the object of interest and the first and second imaging devices.
- [0120] An exemplary embodiment of the disclosed methods further comprises optimizing the depth map by using a non-partial optimizing equation.
- [0121] An exemplary embodiment of the disclosed methods further comprises obtaining a Jacobi iteration of the non-partial optimizing equation by using a recurrence filtering.
- [0122] In accordance with another aspect disclosed herein, there is set forth an apparatus for obtaining depth map of scene, comprising:
- [0123] an imaging system for capturing a plurality of scene images; and
- [0124] a depth assembly for calculating a disparity of the plurality of scene images.
- [0125] In an exemplary embodiment of the disclosed apparatus, the imaging system comprises first and second imaging devices.
- [0126] In an exemplary embodiment of the disclosed apparatus, the depth assembly is configured to optimize a global energy function.
- [0127] In an exemplary embodiment of the disclosed apparatus, the global energy function is defined as a sum of a disparity energy function and a scaled smoothing term.
- [0128] In an exemplary embodiment of the disclosed apparatus, the disparity energy function comprises a Birchfield-Tomasi data term.
- [0129] In an exemplary embodiment of the disclosed apparatus, the Birchfield-Tomasi data term is defined based on a minimum disparity of coordinates for pixels in a first image of the scene captured by the first imaging device and a second image of the scene captured by the second imaging device.
- [0130] In an exemplary embodiment of the disclosed apparatus, the smoothing term adopts an energy function of differential of disparity.
- [0131] In an exemplary embodiment of the disclosed apparatus, the smoothing term is, for all neighbors of a selected pixel with the coordination of (x,y), a summation of scaled trigger functions of a disparity between two neighboring pixels.
- [0132] In an exemplary embodiment of the disclosed apparatus, the neighbors are pixels from four domains.

- [0133] In an exemplary embodiment of the disclosed apparatus, the smoothing term is defined based on scaled trigger functions of a disparity for all neighbors of each pixel.
- [0134] In an exemplary embodiment of the disclosed apparatus, the global energy function is optimized by
- [0135] aggregating data terms in a plurality of directions to obtain an energy function for each of the directions; and
- [0136] accumulating the energy functions in the directions to obtain the energy function.
- [0137] In an exemplary embodiment of the disclosed apparatus, the directions comprise a predetermined number of directions.
- [0138] In an exemplary embodiment of the disclosed apparatus, the predetermined number of directions comprise four or eight directions.
- [0139] In an exemplary embodiment of the disclosed apparatus, the energy function in one direction is based on a dynamic planning in the direction.
- [0140] In an exemplary embodiment of the disclosed apparatus, the energy function in one direction is obtained by summing a corresponding smoothing term and the dynamic planning in the direction.
- [0141] In an exemplary embodiment of the disclosed apparatus, the dynamic planning in the direction is a recurrence based on the energy functions of its neighbors in the direction.
- [0142] In an exemplary embodiment of the disclosed apparatus, the direction comprises a horizontal direction.
- [0143] In an exemplary embodiment of the disclosed apparatus, the energy function in horizontal direction is obtained by a recurrence based on the energy functions of its neighbors the horizontal direction.
- [0144] In an exemplary embodiment of the disclosed apparatus, the depth assembly is configured to obtain a best depth by seeking the disparity value that minimizes the summation of energy in a plurality of directions.
- [0145] In an exemplary embodiment of the disclosed apparatus, the best depth is obtained based on an energy function in one direction.
- [0146] In an exemplary embodiment of the disclosed apparatus, the depth assembly is configured to reduce noise by matching the plurality of images and/or identifying respective unique features of the plurality of images while setting the disparity as -1.
- [0147] In an exemplary embodiment of the disclosed apparatus, the depth assembly is configured to compensate an error based on factors at least selected from a group consisting of a distance between central lines of the two imaging devices, an actual distance of two adjacent pixels, a focal length of the two imaging devices and a depth between the object of interest and the first and second imaging devices.
- [0148] In an exemplary embodiment of the disclosed apparatus, the depth map is optimized by using a non-partial optimizing equation.

[0149] In an exemplary embodiment of the disclosed apparatus, the depth assembly is configured to obtain a Jacobi iteration of the non-partial optimizing equation by a recurrence filtering.

brief description of the drawings

[0150] Fig. 1 is an exemplary top-level block diagram illustrating an embodiment of a focal length adjustment apparatus and an imaging mechanism with a first imaging device and a second imaging device.

[0151] Fig. 2 and Fig. 3 show examples of two images of a scene obtained by the first imaging device and second imaging device of Fig. 1.

[0152] Fig. 4 is an exemplary block diagram showing an embodiment of the first imaging device and the second imaging device.

[0153] Fig. 5 is a detail drawing showing an alternative embodiment of the first imaging device and the second imaging device of Fig. 1, in which the first imaging device and the second imaging device are installed on an unmanned aerial vehicle (UAV).

[0154] Fig. 6 schematically illustrates the process of computing a distance between an object of interest and the first and the second imaging devices of Fig. 1 via triangulation.

[0155] Fig. 7 is an exemplary top-level block diagram illustrating an embodiment of a system for adjusting focal length, wherein the system includes a distance assembly.

[0156] Fig. 8 is an exemplary depth map obtained by the focal length adjustment apparatus of Fig. 1.

[0157] Fig. 9 is an exemplary top-level flow chart illustrating an embodiment of a method for focal length adjustment.

[0158] Fig. 10 is an exemplary flow chart illustrating a process of determining the distance between the imaging mechanism and the object of interest.

[0159] Figs. 11-13 are exemplary schematic charts for showing errors of the focal length adjustment apparatus of Fig. 1 and the compensation effect for the errors.

[0160] It should be noted that the figures are not drawn to scale and that elements of similar structures or functions are generally represented by like reference numerals for illustrative purposes throughout the figures. It also should be noted that the figures are only intended to facilitate the description of the preferred embodiments. The figures do not illustrate every aspect of the described embodiments and do not limit the scope of the present disclosure.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

[0162] Since currently-available focal length adjustment systems are incapable of providing an automatic focal length adjustment for imaging systems, a focal length adjustment apparatus and method are provided for adjusting focal length automatically and serving as a basis for a wide range of applications, such as applications on unmanned aerial vehicles (UAVs) and other mobile platforms. This result can be achieved, according to

one embodiment disclosed herein, by a focal length adjustment apparatus 110 as illustrated in Fig. 1.

- [0163] Fig. 1 depicts an illustrative embodiment of a focal length adjustment apparatus 110. As shown in Fig. 1, the focal length adjustment apparatus 110 can be coupled with an imaging mechanism 130. The imaging mechanism 130 can generate one or more images of a scene 100 where an object 120 of interest is positioned.
- [0164] An example of images 199 of the scene 100 obtained by the imaging mechanism 130 are shown in Fig. 2 and Fig. 3. The generated images of the scene 100 can be processed by the focal length adjustment apparatus 110 to generate a signal for adjusting a focal length of the imaging mechanism 130. The focal length of the imaging mechanism 130 can be adjusted in any suitable manner, preferably in real time.
- [0165] Although shown and described with reference to Fig. 1 as comprising two imaging devices 131, 132 for purposes of illustration only, the imaging mechanism 130 can comprise any suitable number of imaging devices 133. For example, the imaging mechanism 130 can have 2, 3, 4, 5, 6, or even a greater number of imaging devices. For an imaging mechanism 130 having more than two imaging devices, the automatic focal length adjustment illustrated herein can be applied as to any pair of the imaging devices.
- [0166] The imaging devices 131, 132 of Fig. 1 can be arranged in any desired manner in the imaging mechanism 130. The specific arrangement of the imaging devices 131, 132 can depend on a relevant imaging application. In some embodiments, for example, the imaging devices 131, 132 can be positioned side-by-side so that the imaging devices 131, 132 have parallel optical axes. In other embodiments, the imaging devices 131, 132 can be positioned such that the optical axes of the imaging devices 131, 132 are not parallel.
- [0167] Each of the imaging devices 131, 132 can sense light and convert the sensed light into electronic signals that can be ultimately rendered as an image. Exemplary imaging devices 131, 132 suitable for use with the focal length adjustment apparatus 110, include, but are not limited to, commercially-available cameras (color and/or monochrome) and camcorders. Suitable imaging devices 131, 132 can include analog imaging devices (for example, video camera tubes) and/or digital imaging devices (for example, charge-coupled device (CCD), complementary metal-oxide-semiconductor (CMOS), N-type metal-oxide-semiconductor (NMOS) imaging devices, and hybrids/variants thereof). Digital imaging devices, for example, can include a two-dimensional array of photosensor elements (not shown) that can each capture one pixel of image information. Either of the imaging devices 131, 132 can be, for example, an electro-optical sensor, a thermal/infrared sensor, a color or monochrome sensor, a multi-

spectral imaging sensor, a spectrophotometer, a spectrometer, a thermometer, and/or an illuminometer. Furthermore, either of the imaging devices 131, 132 can be, for example, an Red-Green-Blue (RGB) camera, an ultrasonic camera, a laser camera, an infrared camera, an ultrasound camera or a Time-of-Flight camera. However, the imaging devices 131, 132 can be alternatively of the same type. Similarly, the focal lengths of the imaging devices 131, 132 can be the same and/or different without limitation to the scope of the present disclosure.

[0168] An exemplary first imaging device 131 and second imaging device 132 are shown in Fig. 4. A distance D between the first imaging device 131 and the second imaging device 132 can be adjustable depending on an object distance Z (shown in Fig. 6) between the imaging devices 131, 132 and the object 120 of interest. In an embodiment, the first imaging device 131 and the second imaging device 132 can be installed on a portable cradle head 150. Once the object 120 of interest is determined, the focal lengths of the first imaging device 131 and the second imaging device 132 can be adjusted automatically based on the object distance Z . By adjusting the focal lengths of the imaging devices 131, 132, the object 120 of interest can be made clearly visible.

[0169] In some embodiments, the focal length adjustment apparatus 110 (shown in Fig. 1) can be physically located adjacent to the imaging mechanism 130 (shown in Fig. 1), in which case data between the focal length adjustment apparatus 110 and the imaging mechanism 130 can be communicated locally. An advantage of local communication is that transmission delay can be reduced to facilitate real-time focal length adjustment, image processing, and parameter calibration. In other embodiments, the focal length adjustment apparatus 110 can be located remotely from the imaging mechanism 130. Remote processing may be preferable, for example, because of weight restrictions or other reasons relating to an operational environment of the focal length adjustment apparatus 110. As a non-limiting example, if the imaging mechanism 130 is mounted aboard a mobile platform, such as an unmanned aerial vehicle (UAV) (shown in Fig. 5), it may be desirable to convey imaging data to a remote terminal (not shown) for centralized processing, such as a ground terminal or base station. Centralized processing may be desirable, for example, where multiple UAVs are imaging a given object 120 of interest in a coordinated fashion. Fig. 4 illustrates an exemplary embodiment of the focal length adjustment apparatus 110, the imaging devices 131, 132 are installed on a UAV 400.

[0170] Although shown and described as being the UAV 400 in Fig. 5 for exemplary purpose only, the mobile platform can be any kind of such platforms, including but not limited to any self-stabilizing mobile platforms.

[0171] Various communication methods can be used for remote communication between the

imaging mechanism 130 and the focal length adjustment apparatus 110. Suitable communication methods include, for example, radio, Wireless Fidelity (Wi-Fi), cellular, satellite, and broadcasting. Exemplary wireless communication technologies include, but are not limited to, Global System for Mobile Communications (GSM), General Packet Radio Service (GPRS), Code Division Multiple Access (CDMA), Wideband CDMA (W-CDMA), CDMA2000, IMT Single Carrier, Enhanced Data Rates for GSM Evolution (EDGE), Long-Term Evolution (LTE), LTE Advanced, Time-Division LTE (TD-LTE), High Performance Radio Local Area Network (HiperLAN), High Performance Radio Wide Area Network (HiperWAN), High Performance Radio Metropolitan Area Network (HiperMAN), Local Multipoint Distribution Service (LMDS), Worldwide Interoperability for Microwave Access (WiMAX), ZigBee, Bluetooth, Flash Orthogonal Frequency-Division Multiplexing (Flash-OFDM), High Capacity Spatial Division Multiple Access (HC-SDMA), iBurst, Universal Mobile Telecommunications System (UMTS), UMTS Time-Division Duplexing (UMTS-TDD), Evolved High Speed Packet Access (HSPA+), Time Division Synchronous Code Division Multiple Access (TD-SCDMA), Evolution-Data Optimized (EV-DO), Digital Enhanced Cordless Telecommunications (DECT), and others.

[0172] Alternatively and/or additionally, the imaging mechanism 130 can be at least partially incorporated into the focal length adjustment apparatus 110. The imaging mechanism 130 thereby can advantageously serve as a component of the focal length adjustment apparatus 110.

[0173] As shown in Fig. 1, the imaging mechanism 130 can interface with the focal length adjustment apparatus 110. For example, the imaging devices 131, 132 of the imaging mechanism 130 can acquire respective images 199 (shown in Fig. 2 and Fig. 3) of the scene 100 and relay the acquired images to the focal length adjustment apparatus 110 locally and/or remotely via a data communication system (not shown). The focal length adjustment apparatus 110 can be configured, for example, to reconstruct a three-dimensional depiction of the object 120 of interest using the two images via stereopsis. The focal length adjustment apparatus 110 thereby can determine whether a focal length adjustment would be advantageous based on the object distance Z between the imaging mechanism 130 and the object 120 of interest and/or to convey calibrating signals to the imaging mechanism 130 for a focal length adjustment. Additionally and/or alternatively, the focal length adjustment apparatus 110 can be advantageously configured to automatically calibrate one or more extrinsic parameters for stereoscopic imaging.

[0174] Referring now to Fig. 6, the images 199 acquired by the imaging devices 131, 132 can include images 199A and 199B. The images 199A (left, indicated as l in the following equations) and 199B (right, indicated as r in the following equations) can be

compared to ascertain the object distance Z between the imaging devices 131, 132 and the object 120 of interest. A method of triangulation can be used to ascertain the object distance Z using a binocular disparity d between the two images 199A and 199B.

Specifically, coordinates (X_i, Y_i, Z_i) of a pixel i in the image 199A(left) can be given as follows:

$$[0175] \quad X_i = \frac{T}{d} (x_i^l - c_x),$$

(Equation 1)

$$[0176] \quad Y_i = \frac{T}{d} (y_i^l - c_y),$$

(Equation 2)

$$[0177] \quad Z_i = \frac{T}{d} f$$

(Equation 3)

[0178] where c_x and c_y represent respective center coordinates of the imaging devices 131, 132, x_i and y_i represent the coordinates of the object 120 of interest in one or both of the images 199A (left) and 199B (right), T is the baseline (in other words, the distance between the center coordinates of the imaging devices 131, 132), and f is a rectified focal length of the imaging devices 131, 132, i is an index over multiple objects 120 of interest and/or over multiple selected points of an object 120 of interest that can be used to determine the object distance Z , and d is the binocular disparity between the images 199A(l) and 199B(r), represented here as:

$$[0179] \quad d_i = x_i^l - x_i^r$$

(Equation 4)

[0180] The focal length adjustment apparatus 110 can include any processing hardware and/or software needed to perform image acquisition, focal length adjustment, calibration, and any other functions and operations described herein. Without limitation, the focal length adjustment apparatus 110 can include one or more general purpose microprocessors (for example, single or multi-core processors), application-specific integrated circuits, application-specific instruction-set processors, graphics processing units, physics processing units, digital signal processing units, coprocessors, network processing units, audio processing units, encryption processing units, and the like. In certain embodiments, the focal length adjustment apparatus 110 can include an image processing engine or media processing unit, which can include specialized hardware for enhancing the speed and efficiency of image capture, filtering, and processing operations. Such operations include, for example, Bayer transformations, demosaicing operations, noise reduction operations, and/or image sharpening/softening operations.

[0181] In certain embodiments, the focal length adjustment apparatus 110 can include spe-

cialized hardware and/or software for performing focal length adjustment and parameter calibration. For example, specialized hardware and/or software can be provided for functions including, but are not limited to, reconstructing a three-dimensional depiction of the object 120 of interest using the two-dimensional images via stereopsis, determining whether a focal length adjustment is needed based on a distance between the imaging mechanism 130 and the object 120 of interest, determining an optimal focal length, conveying control signals to any components of the focal length adjustment apparatus 110 for focal length adjustment.

[0182] In some embodiments, the focal length adjustment apparatus 110 can include one or more additional hardware components (not shown), as desired. Exemplary additional hardware components include, but are not limited to, memories (for example, a random access memory (RAM), static RAM, dynamic RAM, read-only memory (ROM), programmable ROM, erasable programmable ROM, electrically erasable programmable ROM, flash memory, secure digital (SD) card, etc.), and/or one or more input/output interfaces (for example, universal serial bus (USB), digital visual interface (DVI), display port, serial ATA (SATA), IEEE 1394 interface (also known as FireWire), serial, video graphics array (VGA), super video graphics array (SVGA), small computer system interface (SCSI), high-definition multimedia interface (HDMI), audio ports, and/or proprietary input/output interfaces). Without limitation, one or more input/output devices (for example, buttons, a keyboard, keypad, trackball, displays, and a monitor) can also be included in the focal length adjustment apparatus 100, as desired.

[0183] In some embodiments, the image acquisition, focal length adjustment, calibration, and any other functions and operations described herein for the focal length adjustment apparatus 110 can be achieved by software running on a conventional processor or a general purpose computer, such as a personal computer. The software can be operated with suitable hardware discussed above as desired. The software, for example, can take any form of source code, object code, executable code and machine readable code. The source code can be written in any form of high-level programming languages, including but not limited to, C++, Java, Pascal, Visual B and the like.

[0184] Turning now to Fig. 7, an exemplary block diagram illustrating an alternative embodiment of the focal length adjustment apparatus 110 of Fig. 1 is shown. The focal length adjustment apparatus 110 of Fig. 7 comprises a distance assembly 701 for determining the object distance Z , that is, a distance between the object 120 of interest in the scene 100 and the imaging mechanism 130. The focal length adjustment apparatus 110 further includes a focal length assembly 702 configured to automatically adjust the focal length of the imaging mechanism 130 according to the determined distance by the distance assembly 701.

- [0185] As seen from Fig. 7, the distance assembly 701 is shown as comprising a depth estimation mechanism 7011 for obtaining a depth map of the scene 100, an object determination mechanism 7012 for determining the object 120 of interest in the scene 100 and a calculating mechanism 7013 for calculating distance of the object 120 of interest according to the depth map of the scene 100 from depth estimation mechanism 7011.
- [0186] In an embodiment of the present disclosure, the depth estimation mechanism 7011 receives a first image 199A (shown in Fig. 2) of the scene 100 from the first imaging device 131 and a second image 199B (shown in Fig. 3) of the scene 100 from the second imaging device 132. Based on the first image 199A and the second image 199B of the scene 100 as shown in Figs. 2 and 3, the depth estimation mechanism 7011 obtains their disparity, based on which a depth map 800 (shown in Fig. 8) is acquired. The specific operation of the depth estimation mechanism 7011 for obtaining the depth map will be described in detail below with reference to Fig. 8.
- [0187] An exemplary depth map 800 is depicted in Fig. 8. Each pixel (not shown) in the depth map 800 is associated with a value that represents a distance between the point corresponding to the pixel in the scene 100 (shown in Fig. 1) and the imaging mechanism 130 (shown in Fig. 1). For example, in certain embodiments, a brightness value is utilized to represent a distance between a point in the scene (either on the object 120 of interest or not on the object 120 of interest) and an imaging device which images the scene. Alternatively and/or additionally, different color values can be assigned to pixels to represent the distance. In a brightness value example, as seen from Fig. 8, a brighter area 810 indicates points in the scene 100 with nearer distances to the imaging mechanism 130 (shown in Fig. 1), a darker area 820 indicates points in the scene 100 with further distances to the imaging mechanism 130, and a grey area 830 indicates points in the scene 100 with distances between the near and far distance. If the object 120 of interest moves in the scene 100, a brightness of the pixels for the object 120 of interest can vary based upon the distance between the object 120 of interest and the imaging mechanism 130. The selected pixels for the object 120 of interest can become brighter when the distance between the object 120 of interest and the imaging mechanism 130 decreases and can become dimmer when the distance increases as shown in Fig. 8.
- [0188] Returning to Fig. 7, the object determination mechanism 7012 can receive outside instructions to determine the object 120 (shown in Fig. 1) of interest by identifying the object 120 of interest selected in either of the first and the second images of the scene 100 (shown in Fig. 1). Outside instructions can be given by, for example, operators of the focal length adjustment apparatus 110. The selection of the object 120 of interest can be done by framing in the object 120 of interest in either of the first and the second images of the scene 100. The first and the second images of the scene 100 can be

displayed on a display screen (not shown) and/or on a display screen (not shown) to the operator of the focal length adjustment apparatus 110 for selection. Alternatively or additionally, the selection of the object 120 of interest can be performed by clicking the display screen(s) on the object 120 of interest in either of the first and the second images of the scene 100 displayed on the display screen(s). The object determination mechanism 7012 can sense the framing and/or the clicking operation to identify the object 120 of interest being selected.

[0189] In another embodiment, the object determination mechanism 7012 enables to receive outside oral instructions from, for example, an operator of the focal length adjustment apparatus 110. Optionally, the oral instructions can be a pre-set name of the object 120 of interest.

[0190] Alternatively and/or additionally, the object determination mechanism 7012 can be enabled to automatically determine the object 120 of interest based on a judgment under at least a pre-set rule. Any rule for the judgment can be set as desired. For example, the pre-set rule may comprise that the object 120 of interest is determined if the object 120 of interest is approaching the first and the second imaging devices 131, 132 and/or if the object 120 of interest is within a certain distance from the first and the second imaging devices 131, 132.

[0191] Based on the depth map from the depth estimation mechanism 7011 and the information about the object 120 of interest from the object determination mechanism 7012, the calculating mechanism 7013 can be enabled to calculate the distance between the imaging mechanism 130 and the object 120 of interest. The calculating mechanism 7013 preferably calculates the distance in real time.

[0192] Based on the calculated distance, the focal length assembly 702 can be enabled to automatically adjust the focal length of the imaging mechanism 130, preferably in real time, with a tracking learning detection method based on gray level information of the object 120 of interest serving as initial values.

[0193] If a user wants to focus the imaging mechanism 130 on a particular object, for example, the object determination mechanism 7012 can enable the user to draw a frame on the display showing the images of the scene 100 to frame in the object 120 of interest to be tracked. The frame can be in any suitable dimension, size or shape, including, but not limited to, a shape of a rectangular, a square or a circle or even an irregular shape. Optionally, the user can be enabled to click the display screen(s) to confirm the selection. By using the depth map 800 (shown in Fig. 8) obtained by the depth estimation mechanism 7011, the calculating mechanism 7013 can be enabled to calculate the distance of the object being selected by the user and the focal length assembly 702 enables to adjust the focal length of the focal length adjustment apparatus 110 automatically according to the object distance Z (shown in Fig. 6),

which distance preferably is acquired in real time.

[0194] Referring now to Fig. 9, one embodiment of a method 900 for focal length adjustment is illustrated. At 901, an object distance between the imaging mechanism 130 (shown in Fig. 1) and the object 120 (shown in Fig. 1) of interest is determined. The object distance Z (shown in Fig. 6) can be determined by using any of several various methods, as desired. In some embodiments, the object distance Z can be determined by using a plurality of imaging devices 133 (shown in Fig. 1) in the imaging mechanism 130 via stereopsis. For example, two imaging devices 131, 132 (shown in Fig. 1) of the imaging mechanism 130 each can acquire an image (shown in Figs. 2 and 3) of the object 120 of interest, and overlapping portions of the acquired images can be analyzed to assess the depth of the object 120 of interest. Alternatively and/or additionally, the object distance can be acquired using one or more non-stereopsis methods, such as by using a laser and/or using ultrasound. At 902, the focal length of the imaging mechanism 130 is automatically adjusted according to the object distance Z determined in step 901.

[0195] A detailed process 901 of determining the distance between the imaging mechanism 130 and the object 120 of interest is illustrated in Fig. 10. At 9011, a depth map of the scene 100 is obtained. Any process of obtaining a depth map may be applied here without limitation. An exemplary process is illustrated in details below.

[0196] In an embodiment, a depth map is obtained by calculating a disparity of scene images from the first and second imaging devices 131, 132. In the field of obtaining a depth map, an energy function is usually computed over a subset of the whole image. In one embodiment, a global energy function is optimized to obtain a disparity global energy. Specifically, the disparity global energy can be calculated by summing a disparity energy function and a scaled smoothing term.

[0197] An exemplary optimizing can be illustrated by the following equation:

$$[0198] \quad E(d) = E_d(d) + pE_s(S)$$

(Equation 5)

[0199] Wherein d indicates the disparity between the first and the images of the scene, $E_d(d)$ is a data term, which indicates the disparity energy function and $E_s(d)$ indicates a smoothing term.

[0200] The data term $E_d(d)$ comprises a Birchfield-Tomasi data term which can be obtained in accordance with the equation:

$$[0201] \quad E_d(d) = \sum_{x,y} E_{d_{BT-SAD}}(x, y, d(x, y) = d)$$

(Equation 6)

[0202] wherein,

$$[0203] \quad E_{d_{BT}}(x, y, d(x, y) = d) = \min\{C_1, C_2\}$$

(Equation 7)

$$[0204] \quad C_1 = \min_{x-d-0.5 \leq x' \leq x-d+0.5} |I_L(x) - I_R(x')|$$

(Equation 8)

$$[0205] \quad C_2 = \min_{x-0.5 \leq x' \leq x+0.5} |I_L(x') - I_R(x-d)|$$

(Equation 9)

$$[0206] \quad E_{d_{BT-SAD}}(x, y, d(x, y) = d) = \sum \min\{C_1, C_2\}$$

(Equation 10)

[0207] wherein, I_L represents the first image of the scene 100 captured by the first imaging device 131 and I_R represents the second image of the scene 100 captured by the second imaging device 132 respectively; x in $I_L(x)$ represents a horizontal coordinate of a pixel in the first image and x' in $I_L(x')$ represents a horizontal coordinate of the pixel in the second image; (x, y) represents coordinates of a pixel in the scene 100.

[0208] The introducing of the Birchfield-Tomasi data term which is a data term often used in image sampling/matching resolves a problem of incorrect image matching by utilizing a matching precision of sub pixels, which is a pixel dissimilarity measure that is insensitive to image sampling. The contents of IEEE Transactions on Pattern Analysis and Machine Intelligence (1998) explaining the Birchfield-Tomasi data term is incorporated here by reference. Other data terms may be adopted as the data term in Equation 5 when calculating the disparity global energy.

[0209] The smoothing term $E_s(d)$ can be presented by an energy function of differential of disparity, which can be obtained by, for all neighbors of a pixel with the coordination of (x, y) , summing scaled trigger functions of a disparity between two neighboring pixels. Specifically, the smoothing term $E_s(d)$ can be obtained in accordance with the equation as follows:

$$[0210] \quad E_s(\nabla d) = \sum p_1 T(|d(x, y) - d(x', y')| = 1) + p_2 T(|d(x, y) - d(x', y')| > 1)$$

(Equation 11)

[0211] wherein, (x, y) represents the coordinates of the pixel of the first image and (x', y') represents the coordinates of the pixel of the second image; p_2 and p_1 are two adjustable weights, and usually $p_2 \geq p_1$ and the summation sign is for all neighbors of (x, y) in which four domains are usually used, T is a trigger function which is triggered when the conditions in the parentheses are true.

[0212] In order to optimize the smoothing term $E_s(d)$, a fastest dynamic planning is utilized by aggregating data terms in a plurality of directions to obtain a directional energy function for each of the directions and accumulating the directional energy functions in the directions to obtain the energy function. In an embodiment, four or eight directions

are selected to aggregate data terms. Other numbers of directions, such as three, five or ten directions, can be selected without limitation.

[0213] In an embodiment, an energy function in one direction can be obtained by summing its corresponding smoothing term and the dynamic planning in this direction. A dynamic planning in one direction can be presented by a recurrence based on the energy functions of its neighbors in this direction. For example, energy functions of one pixel's neighbors in the horizontal direction are presented as $L(x-1, y, d)$, $L(x-1, y, d+1)$, $L(x-1, y, d-1)$, and $L(x-1, y, d')$. Specifically, an energy function in horizontal direction is obtained in accordance with the equation:

$$[0214] \quad L(x, y, d) = E_s(x, y, d) + \min\{L(x-1, y, d), L(x-1, y, d-1) + p_1, L(x-1, y, d+1) + p_1, \min_{d'} L(x-1, y, d') + p_2\} - \min_{d'} L(x-1, y, d')$$

(Equation 12)

[0215] wherein, the energy received at the coordinates (x, y, d) is defined as $L(x, y, d)$, which is presented as a recurrence based on the energy of neighbors $L(x-1, y, d)$, $L(x-1, y, d+1)$, $L(x-1, y, d-1)$, and $L(x-1, y, d')$.

[0216] Furthermore, a best depth is obtained by seeking the disparity value that minimizes the summation of energy in a plurality of directions. In an embodiment, the best depth is obtained in accordance with the equation:

$$[0217] \quad d^* = \operatorname{argmin}_d \sum L(x, y, d)$$

(Equation 13)

[0218] wherein, d^* indicates the best depth, $L(x, y, d)$ indicates an energy function in one direction.

[0219] As an embodiment, some noise is reduced by matching the first and second scene images and/or identifying respective unique features of the first and second scene images while setting the disparity as -1.

[0220] The present disclosure also optimizes the depth map obtained by the depth estimation mechanism 7011. As described above, the global energy function is optimized by compensating errors of the depth estimation mechanism 7011. Specifically, the errors can be compensated by following equations:

$$[0221] \quad \frac{b}{c_1} = \frac{f}{f+D} \Rightarrow c_1 = \frac{b(f+D)}{f}$$

(Equation 14)

$$[0222] \quad \frac{c_1}{a} = \frac{x}{x+D+f} \Rightarrow x = \frac{c_1(D+f)}{a-c_1}$$

(Equation 15)

[0223] by combining Equations 14 and 15, we get:

[0224]
$$+x = \frac{b(f+D)^2}{af-b(f+D)} \approx \frac{D^2}{a \times \frac{f}{b} - D}$$

(Equation 16)

[0225] on the other hand, we have:

[0226]
$$\frac{b}{c_2} = \frac{f}{f+D-x} \Rightarrow c_2 = \frac{b(f+D-x)}{f}$$

(Equation 17)

[0227]
$$\frac{c_2}{a} = \frac{x}{D+f} \Rightarrow x = \frac{c_2(D+f)}{a}$$

(Equation 18)

[0228] by combining Equations 17 and 18, we get:

[0229]
$$-x = \frac{b(f+D)^2}{af+b(f+D)} \approx \frac{D^2}{a \times \frac{f}{b} + D}$$

(Equation 19)

[0230] Wherein

[0231] f = focal length (mm)

[0232] D = depth (depth between the object and the imaging plane, mm)

[0233] a = baseline (the distance between the central lines of the two imaging devices, mm)

[0234] b = actual distance of two adjacent pixels (mm)

[0235] d = disparity (measured in pixel)

[0236] Therefore, the depth estimation errors are within a range of [-x, x], and the estimated depth is in a range of

$$[D - x, D + x]$$

. The above error analysis and the compensation are based on following assumptions. For example, the error estimation is based on an assumption that camera calibration parameters are completely correct and an average error of the disparity map is a theoretical value for 1 pixel. An actual calibration can introduce errors, and a depth estimation error may exceed 1 pixel. Therefore, the above data of errors reflects a trend only. As a reference: an average error of a depth estimation in four testing maps by using the first ranked stereo matching process, Middelbury, is $1.29+0.14+6.47+5.70/4=3.4$ (pixel). Even under a situation with no noise, no alteration of light beams and calibration parameters are correct, and the current best process presents an average error for all points as 3.4 pixels.

[0237] A few schematic charts showing the errors of the focal length adjustment apparatus 110 and the compensation effect for the errors are depicted in Figs. 11-13.

[0238] Fig. 11 illustrates an exemplary relationship between the estimated error of the focal length adjustment apparatus 100 and the measured depth D of the scene 100 when only

the measured depth D varies. In other words, the baseline a, the distance of two adjacent pixels b and the focal length f of the equations 16, 17 remain constant as the measured depth D changes. As shown in Fig. 11, a horizontal axis represents a measured depth D in mm, and a vertical axis represents the estimated error of the focal length adjustment apparatus 110. As Fig. 11 shows, when the measured depth D changes, the estimated error of the focal length adjustment apparatus 100 also changes in a non-linear relationship with reference to the measured depth D. For example, when the measured depth D is 3000mm, the estimated error according to Fig. 11 is about 5mm, but when the measured depth D increases to 6000mm, the corresponding estimated error increases to over 20mm.

[0239] Although shown and described as being the non-linear relationship in Fig. 11 for exemplary purposes only, the relationship of the estimated error and the measured depth D can be any linear and/or non-linear relationships.

[0240] Fig. 12 illustrates an exemplary relationship between the estimated error of the focal length adjustment apparatus 100 and the baseline when only the baseline a varies. Here, the measured depth D, the distance of two adjacent pixels b and the focal length f of the equations 16, 17 remain constant when the baseline a changes. In Fig. 12, a horizontal axis represents the baseline a in mm, and a vertical axis represents the estimated error in mm. As Fig. 12 shows, when the baseline a increases, the estimated error can decrease by in accordance with a non-linear relationship between the estimated error and the baseline a. For example, when the baseline a is 500mm, the estimated error according to Fig. 12 can be as high as 35mm, but when the baseline a increases to 2000mm, the estimated error decreases to about 8mm.

[0241] Although shown and described as being the non-linear relationship in Fig. 12 for exemplary purposes only, the relationship of the estimated error and the baseline can be any linear and/or non-linear relationships.

[0242] Fig. 13 is an illustrative example for showing representative corresponding relationships among the image representation symbols and the variables contained in the Equations 14-19. In other words, Equations 14-19 can be deducted from the represented relationships illustrated in Fig. 13. Here, the imaging devices 131, 132 (shown in Fig. 1) are represented with two cameras, Cam1 and Cam2, which have a baseline “a” and a focus length “f”.

[0243] As shown in Fig. 13, triangles ABO₂ and CBO₂ are similar because AB is parallel to CD. Therefore, we can get Equation 14:

$$\frac{b}{c_s} = \frac{f}{f+D} \Rightarrow c_s = \frac{b(f+D)}{f}$$

. In addition, triangles ADE and O₁O₂E are similar because CD is parallel to O₁O₂. Therefore, we can get Equation 15:

$$\frac{c_1}{a} = \frac{x}{x+D+f} \Rightarrow x = \frac{c_1(D+f)}{a-c_1}$$

. By combining Equations 14 and 15, we can reach Equation 16:

$$+x = \frac{b(f+D)^2}{af-b(f+D)} \approx \frac{D^2}{a\frac{f}{b}-D}$$

. For same reasons, Equation 19 can be deduced from a combination of a similarity relationship of triangles AHO₂ and a similarity relationship of FGO₂ and O₁O₂C. In both Equations 16 and 19, D is the actual depth between the scene 100 and the imaging plane, a is the baseline between the two image devices 131, 132, b is the distance of two adjacent pixels and f is the focal length of the imaging devices 131, 132, as shown in Fig. 13.

[0244] Although shown and described as being the same absolute value for -x and +x in Fig. 13 for exemplary purposes only, the estimated errors -x and +x can be different absolute values as results of equations 16 and 19, in which case, b can carry different values in each equation.

[0245] Based on the characteristics of the estimated errors shown in Figs. 11 and 12, the depth map can be further optimized by applying a non-partial optimizing equation whose Jacobi iteration can be obtained by a recurrence filtering. Specifically, the non-partial optimizing equation is defined in accordance with the following equation:

$$E(d) = \sum |d(x,y) - d^*(x,y)|^2 + \sum \exp (|I_L(x,y) - I_L(x',y')| + |x' - x| + |y' - y|) |d(x,y) - d(x'y')|$$

(Equation 20)

[0247] in which, d* (x, y) indicates the optimal depth map and d(x,y) is the estimated depth map; I(x,y) represents the intensity of the image; x,y are the coordinates of the pixel in an image coordinate; x',y' are the coordinates of an adjacent pixel of x,y in the same image.

[0248] Similarly to the operations performed by the object determination mechanism 7012 and the calculating mechanism 7013, in step 9012 of Fig. 10, the object 120 of interest is determined and the distance between the object 120 of interest and the imaging mechanism 130 is calculated in step 9013 of Fig. 10, which in turn serves as a basis for automatically adjusting the focal length of the imaging mechanism 130.

[0249] A stereoscopic imaging system configured to conduct the aforementioned operations to perform automatic focal length adjustment can be obtained according to any embodiment of the present disclosure.

[0250] Furthermore, a computer program product comprising instructions for automatically adjusting focal length of a stereoscopic imaging system having at least two imaging devices in accordance with the aforementioned operations can be obtained according to an embodiment of the present disclosure. Preferably, the method for automatically

adjusting focal length according to the present disclosure can be achieved by an ordinary computing device, such as a personal computer and/or a microcomputer.

[0251] The disclosed embodiments are susceptible to various modifications and alternative forms, and specific examples thereof have been shown by way of example in the drawings and are herein described in detail. It should be understood, however, that the disclosed embodiments are not to be limited to the particular forms or methods disclosed, but to the contrary, the disclosed embodiments are to cover all modifications, equivalents, and alternatives.

Claims

- [Claim 1] A method for focal length adjustment, comprising:
determining a distance between an object of interest and an imaging mechanism; and
automatically adjusting a focal length of the imaging mechanism according to said determining the distance.
- [Claim 2] The method of claim 1, wherein said determining comprises determining the distance between the object of interest in a scene and the imaging mechanism.
- [Claim 3] The method of claim 2, wherein said determining comprises imaging the scene with first and second imaging devices contained in the imaging mechanism.
- [Claim 4] The method of claim 3, wherein said determining further comprises:
obtaining a depth map of the scene;
selecting the object of interest in the scene; and
calculating the distance of the object of interest according to the depth map of the scene.
- [Claim 5] The method of claim 4, wherein said obtaining comprises calculating a disparity of scene images from the first and second imaging devices.
- [Claim 6] The method of claim 5, wherein said calculating the disparity comprises optimizing a global energy function.
- [Claim 7] The method of claim 6, wherein said optimizing the global energy function comprises summing a disparity energy function and a scaled smoothing term.
- [Claim 8] The method of claim 7, wherein said disparity energy function is represented by a Birchfield-Tomasi term.
- [Claim 9] The method of claim 8, wherein the Birchfield-Tomasi term is defined by accumulating the minimum disparity of coordinates for pixels in a first image of the scene captured by the first imaging device and a second image of the scene captured by the second imaging device.
- [Claim 10] The method of any one of claims 7-9, further comprising, for all neighbors of a pixel, accumulating scaled trigger functions of a disparity between two neighboring pixels to obtain the smoothing term.
- [Claim 11] The method of claim 10, wherein said accumulating comprises, for all neighbors of a pixel, accumulating scaled trigger functions of a disparity between two neighboring pixels from four domains.
- [Claim 12] The method of any one of claims 7-11, wherein said smoothing term is

- obtained by accumulating scaled trigger functions of a disparity for all neighbors of each pixel.
- [Claim 13] The method of any one of claims 1-12, further comprising:
aggregating data items in a plurality of directions to obtain a directional energy function for each of the directions; and
accumulating the directional energy functions in the directions to obtain the energy function.
- [Claim 14] The method of claim 13, wherein said aggregating comprises obtaining energy functions in a predetermined number of directions.
- [Claim 15] The method of claim 13 or claim 14, wherein said aggregating comprises obtaining energy functions in four or eight directions.
- [Claim 16] The method of any one of claims 13-15, wherein said aggregating data items comprises obtaining energy function in a direction by summing a corresponding smoothing item and a dynamic planning in the direction.
- [Claim 17] The method of claim 16, wherein said summing the corresponding smoothing item and the dynamic planning in the direction comprises presenting the dynamic planning in the direction with a recurrence based on the energy functions of its neighbors in this direction.
- [Claim 18] The method of claim 17, wherein said direction comprises a horizontal direction.
- [Claim 19] The method of claim 18, wherein said aggregating data items in the horizontal direction comprises calculating the energy by a recurrence based on said energy functions of its neighbors in the horizontal direction.
- [Claim 20] The method of any one of claims 13-19, further comprising obtaining the best depth.
- [Claim 21] The method of claim 20, wherein said obtaining the best depth comprises seeking the disparity value that minimizes the summation of energy in a plurality of directions.
- [Claim 22] The method of claim 20, wherein said obtaining the best depth comprises seeking the disparity value based on an energy function in one direction.
- [Claim 23] The method of any one of claims 3-22, further comprising reducing noise by doing at least one of matching scene images from the first and second imaging devices and identifying respective unique features of the scene images while setting the disparity as -1.
- [Claim 24] The method of any one of claims 3-23, further comprising compensating an error based on the factors at least selected from a group

- consisting of the distance between a central line of the two imaging devices, an actual distance of two adjacent pixels, the focal length of the two imaging devices and the depth between the object of interest and the first and second imaging devices.
- [Claim 25] The method of any one of claims 4-24, further comprising optimizing the depth map by using a non-partial optimizing equation.
- [Claim 26] The method of claim 25, further comprising obtaining a Jacobi iteration of the non-partial optimizing equation by using a recurrence filtering.
- [Claim 27] The method of one of claims 4-26, wherein said selecting the object of interest in the scene comprises receiving outside instructions to select the object of interest.
- [Claim 28] The method of claim 27, wherein said receiving the instructions comprises identifying the object of interest selected on either of scene images from the first or the second imaging device.
- [Claim 29] The method of claim 28, wherein said identifying the object of interest selected comprises sensing a frame on either of the scene images framing in the object of interest or sensing a click on the object of interest on either of the scene images.
- [Claim 30] The method of claim 29, wherein said receiving the outside instructions comprises receiving vocal instructions, optionally a pre-set name of the object of interest, to determine the object of interest.
- [Claim 31] The method of any one of claims 4-30, wherein said selecting the object of interest in the scene comprises judging under at least one pre-set rule and automatically determining the object of interest based on said judging.
- [Claim 32] The method of claim 31, wherein said judging under the at least one pre-set rule comprises judging if the object is approaching or within a certain distance of the imaging mechanism.
- [Claim 33] The method of any one of claims 1-32, wherein said automatically adjusting the focus comprises automatically adjusting the focal length of the imaging mechanism in real time with tracking learning detection based on gray level information of the object of interest.
- [Claim 34] A stereoscopic imaging system configured to perform automatic focal length adjustment in accordance with any one of claims 1-33.
- [Claim 35] A focal length adjustment apparatus, comprising:
a distance assembly for determining a distance between an object of interest in a scene and an imaging mechanism for imaging the scene;
and

- a focal length assembly for automatically adjusting a focal length of the imaging mechanism according to the determined distance.
- [Claim 36] The apparatus of claim 35 wherein said distance assembly is configured to determine the distance between the object of interest in a scene and the imaging mechanism.
- [Claim 37] The apparatus of claim 36, wherein said imaging mechanism comprises first and second imaging devices imaging the scene to obtain first and second scene images.
- [Claim 38] The apparatus of claim 37, wherein either of the first and second imaging devices is a camera or a sensor.
- [Claim 39] The apparatus of claim 37 or claim 38, wherein the first and second imaging devices are selected from a group consisting of laser cameras, infrared cameras, ultrasound cameras and Time-of-Flight cameras.
- [Claim 40] The apparatus of any one of claims 37-39, wherein the first and second imaging devices are Red-Green-Blue (RGB) cameras.
- [Claim 41] The apparatus of any one of claims 37-40, wherein the distance assembly comprises:
a depth estimation mechanism for obtaining a depth map of the scene;
an object determination mechanism for determining the object of interest in the scene; and
a calculating mechanism for calculating distance of the object of interest according to the depth map of the scene.
- [Claim 42] The apparatus of claim 41, wherein the depth map is obtained based on a disparity of the first and the second scene images.
- [Claim 43] The apparatus of claim 42, wherein the depth estimation mechanism optimizes a global energy function.
- [Claim 44] The apparatus of claim 43, wherein the global energy function is defined as a sum of a disparity energy function and a scaled smoothing term.
- [Claim 45] The apparatus of claim 44, wherein said disparity energy function comprises a Birchfield-Tomasi data term.
- [Claim 46] The apparatus of claim 47, wherein said Birchfield-Tomasi data term is defined based on the minimum disparity of coordinates for pixels in a first image of the scene captured by the first imaging device and a second image of the scene captured by the second imaging device.
- [Claim 47] The apparatus of any one of claims 44-46, wherein said smoothing term adopts an energy function of differential of disparity.
- [Claim 48] The apparatus of any one of claims 44-46, wherein said smoothing term

is, for all neighbors of a pixel with the coordination of (x,y) , a summation of scaled trigger function of a disparity between two neighboring pixels.

- [Claim 49] The apparatus of claim 48, wherein the neighbors are pixels from four domains.
- [Claim 50] The apparatus of any one of claims 44-49, wherein said smoothing term is defined based on scaled trigger functions of a disparity for all neighbors of each pixel.
- [Claim 51] The apparatus of any of claims 43-50, wherein the global energy function is optimized by aggregating data terms in a plurality of directions to obtain a directional energy function for each of the directions; and accumulating the directional energy functions in the directions to obtain the energy function.
- [Claim 52] The apparatus of claim 51, wherein the directions comprise a predetermined number of directions.
- [Claim 53] The apparatus of claim 51, wherein said predetermined number of directions comprise four or eight directions.
- [Claim 54] The apparatus of any one of claims 51-53, wherein said energy function in one direction is based on a dynamic planning in the direction.
- [Claim 55] The apparatus of claim 54, wherein said energy function in one direction is obtained by summing a corresponding smoothing term and the dynamic planning in this direction.
- [Claim 56] The apparatus of claim 54, wherein the dynamic planning in the direction is a recurrence based on the energy functions of its neighbors in this direction.
- [Claim 57] The apparatus of claim 56, wherein said direction comprises a horizontal direction.
- [Claim 58] The apparatus of claim 57, wherein said energy function in horizontal direction is obtained by a recurrence based on said energy functions of its neighbors the horizontal direction.
- [Claim 59] The apparatus of any of claims 51-58, wherein the best depth is obtained by seeking the disparity value that minimizes the summation of energy in a plurality of directions.
- [Claim 60] The apparatus of any of claims 51-58, wherein the best depth is obtained based on an energy function in one direction.
- [Claim 61] The apparatus of any one of claims 35-60, wherein noise is reduced by matching the first and second scene images and/or identifying re-

- spective unique features of the first and second scene images while setting the disparity as -1.
- [Claim 62] The apparatus of any one of claims 43-61, wherein an error is compensated based on the factors at least selected from a group consisting of the distance between the central lines of the two imaging devices, actual distance of two adjacent pixels, the focal length of the two imaging devices and the depth between the object of interest and the first and second imaging devices.
- [Claim 63] The apparatus of any one of claims 41-62, wherein the depth map is optimized by using a non-partial optimizing equation.
- [Claim 64] The apparatus of claim 63, comprising obtaining a Jacobi iteration of the non-partial optimizing equation by a recurrence filtering.
- [Claim 65] The apparatus of claim 41, wherein the object determination mechanism receives outside instructions to determine the object of interest.
- [Claim 66] The apparatus of claim 65, wherein said object determination mechanism enables to identify the object of interest selected on either of the first and the second scene images.
- [Claim 67] The apparatus of claim 66, wherein the object determination mechanism enables to identify the object of interest by at least one of sensing a frame on either of the first and the second scene images to frame in the object of interest and sensing a click on the object of interest on either of the first and the second scene images.
- [Claim 68] The apparatus of claim 65, wherein said object determination mechanism receives outside vocal instructions, optionally a pre-set name of the object of interest, to determine the object of interest.
- [Claim 69] The apparatus of claim 41, wherein the object determination mechanism automatically determines the object of interest based on a judgment under at least a pre-set rule.
- [Claim 70] The apparatus of claim 65, wherein the pre-set rule comprises if the object of interest is approaching the first and the second imaging devices or within a certain distance.
- [Claim 71] The apparatus of any one of claims 37-70, wherein said focal length assembly automatically adjusts the focal length of the imaging mechanism in real time with tracking learning detection based on gray level information of the object of interest.
- [Claim 72] A mobile platform performing the method of any one of claims 1-33.
- [Claim 73] A mobile platform comprising the apparatus of any one of claims

35-71.

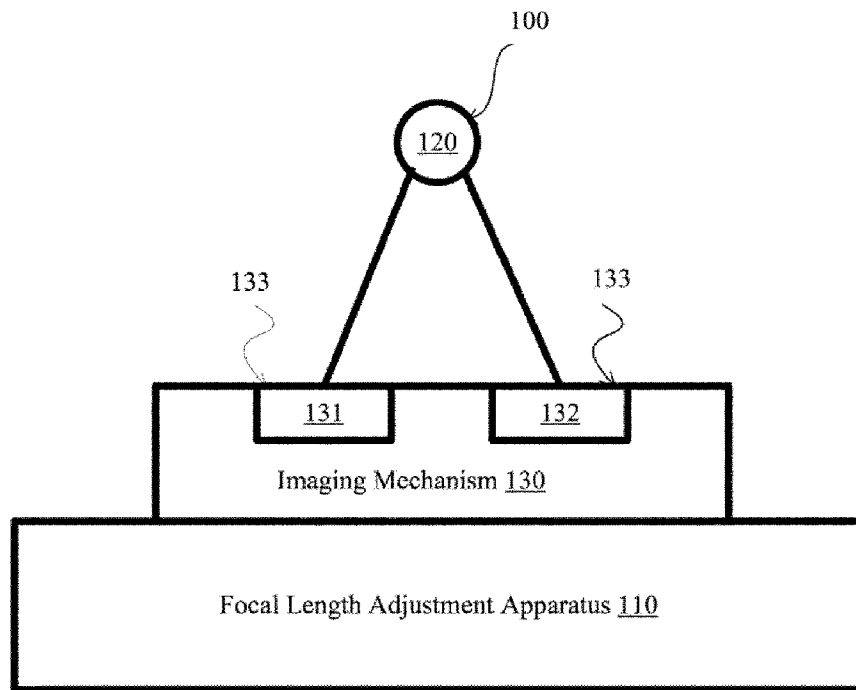
- [Claim 74] The mobile platform of claim 73, wherein the mobile platform is an unmanned aerial vehicle (UAV).
- [Claim 75] The mobile platform of claim 73 or claim 74, wherein the mobile platform is a self-stabilizing platform.
- [Claim 76] A method for obtaining a depth map of a scene, comprising:
capturing a plurality of scene images of the scene;
calculating a disparity of the plurality of scene images; and
obtaining the depth map based on said disparity.
- [Claim 77] The method of claim 76, wherein said capturing the plurality of scene images comprises capturing the plurality of scene images via first and second imaging devices.
- [Claim 78] The method of claim 76 or claim 77, wherein said calculating the disparity comprises optimizing a global energy function.
- [Claim 79] The method of claim 78, wherein said optimizing the global energy function comprises summing a disparity energy function and a scaled smoothing term.
- [Claim 80] The method of claim 79, wherein said disparity energy function is represented by a Birchfield-Tomasi term.
- [Claim 81] The method of claim 80, wherein the Birchfield-Tomasi term is defined by accumulating a minimum disparity of coordinates for pixels in a first image of the scene captured by the first imaging device and a second image of the scene captured by the second imaging device.
- [Claim 82] The method of any one of claims 79-81, further comprising, for all neighbors of a selected pixel, accumulating scaled trigger functions of a disparity between two neighboring pixels to the selected pixel to obtain the smoothing term.
- [Claim 83] The method of claim 82, wherein said accumulating comprises, for all neighbors of the selected pixel, accumulating scaled trigger functions of a disparity between two neighboring pixels from four domains.
- [Claim 84] The method of any one of claims 79-83, wherein said smoothing term is obtained by accumulating scaled trigger functions of a disparity for all neighbors of each pixel.
- [Claim 85] The method of any one of claims 78-83, further comprising:
aggregating data items in a plurality of directions to obtain a directional energy function for each of the directions; and
accumulating the directional energy functions in the directions to obtain the energy function.

- [Claim 86] The method of claim 85, wherein said aggregating comprises obtaining energy functions in a predetermined number of directions.
- [Claim 87] The method of claim 85 or claim 86, wherein said aggregating comprises obtaining energy functions in four or eight directions.
- [Claim 88] The method of any one of claims 85-87, wherein said aggregating data items comprises obtaining the energy function in a selected direction by summing a corresponding smoothing item and a dynamic planning in the selected direction.
- [Claim 89] The method of claim 88, wherein said summing the corresponding smoothing item and the dynamic planning in the direction comprises presenting the dynamic planning in the direction with a recurrence based on the energy functions of its neighbors in the direction.
- [Claim 90] The method of claim 89, wherein said direction comprises a horizontal direction.
- [Claim 91] The method of claim 90, wherein said aggregating data items in the horizontal direction comprises calculating the energy by a recurrence based on said energy functions of its neighbors in the horizontal direction.
- [Claim 92] The method of any one of claims 85-91, further comprising obtaining a best depth.
- [Claim 93] The method of claim 92, wherein said obtaining the best depth comprises seeking the disparity value that minimizes the summation of energy in a plurality of directions.
- [Claim 94] The method of claim 92, wherein said obtaining the best depth comprises seeking the disparity value based on an energy function in one direction.
- [Claim 95] The method of any one of claims 77-94, further comprising reducing noise by doing at least one of matching scene images from the first and second imaging devices and identifying respective unique features of the scene images while setting the disparity as -1.
- [Claim 96] The method of any one of claims 77-95, further comprising compensating an error based on factors at least selected from a group consisting of a distance between central lines of the two imaging devices, an actual distance of two adjacent pixels, a focal length of the two imaging devices and a depth between the object of interest and the first and second imaging devices.
- [Claim 97] The method of any one of claims 76-96, further comprising optimizing the depth map by using a non-partial optimizing equation.

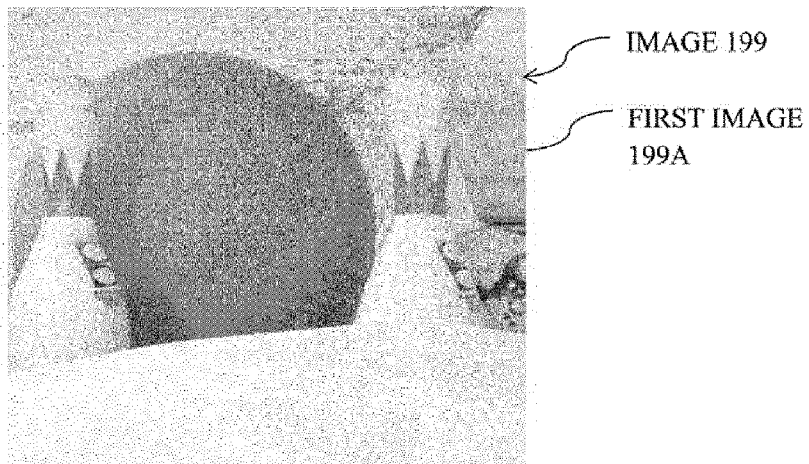
- [Claim 98] The method of claim 97, further comprising obtaining a Jacobi iteration of the non-partial optimizing equation by using a recurrence filtering.
- [Claim 99] An apparatus for obtaining depth map of scene, comprising:
an imaging system for capturing a plurality of scene images; and
a depth assembly for obtaining the depth map based on a disparity calculated from the plurality of scene images.
- [Claim 100] The apparatus of claim 99, wherein said imaging system comprises first and second imaging devices.
- [Claim 101] The apparatus of claim 100, wherein said depth assembly is configured to optimize a global energy function.
- [Claim 102] The apparatus of claim 101, wherein the global energy function is defined as a sum of a disparity energy function and a scaled smoothing term.
- [Claim 103] The apparatus of claim 102, wherein said disparity energy function comprises a Birchfield-Tomasi data term.
- [Claim 104] The apparatus of claim 103, wherein said Birchfield-Tomasi data term is defined based on a minimum disparity of coordinates for pixels in a first image of the scene captured by the first imaging device and a second image of the scene captured by the second imaging device.
- [Claim 105] The apparatus of any one of claims 102-104, wherein said smoothing term adopts an energy function of differential of disparity.
- [Claim 106] The apparatus of any one of claims 102-104, wherein said smoothing term is, for all neighbors of a selected pixel with the coordination of (x,y) , a summation of scaled trigger functions of a disparity between two neighboring pixels.
- [Claim 107] The apparatus of claim 106, wherein the neighbors are pixels from four domains.
- [Claim 108] The apparatus of any one of claims 102-107, wherein said smoothing term is defined based on scaled trigger functions of a disparity for all neighbors of each pixel.
- [Claim 109] The apparatus of any of claims 110-108, wherein the global energy function is optimized by
aggregating data terms in a plurality of directions to obtain an energy function for each of the directions; and
accumulating the energy functions in the directions to obtain the energy function.
- [Claim 110] The apparatus of claim 109, wherein the directions comprise a predetermined number of directions.

- [Claim 111] The apparatus of claim 110, wherein the predetermined number of directions comprise four or eight directions.
- [Claim 112] The apparatus of any one of claims 109-111, wherein the energy function in one direction is based on a dynamic planning in the direction.
- [Claim 113] The apparatus of claim 112, wherein said energy function in one direction is obtained by summing a corresponding smoothing term and the dynamic planning in the direction.
- [Claim 114] The apparatus of claim 113, wherein the dynamic planning in the direction is a recurrence based on the energy functions of its neighbors in the direction.
- [Claim 115] The apparatus of claim 114, wherein the direction comprises a horizontal direction.
- [Claim 116] The apparatus of claim 115, wherein said energy function in horizontal direction is obtained by a recurrence based on said energy functions of its neighbors the horizontal direction.
- [Claim 117] The apparatus of any of claims 109-116, wherein said depth assembly is configured to obtain a best depth by seeking the disparity value that minimizes the summation of energy in a plurality of directions.
- [Claim 118] The apparatus of any of claims 109-117, wherein the best depth is obtained based on an energy function in one direction.
- [Claim 119] The apparatus of any one of claims 101-118, wherein the depth assembly is configured to reduce noise by matching the plurality of images and/or identifying respective unique features of the plurality of images while setting the disparity as -1.
- [Claim 120] The apparatus of any one of claims 101-119, wherein the depth assembly is configured to compensate an error based on factors at least selected from a group consisting of a distance between central lines of the two imaging devices, an actual distance of two adjacent pixels, a focal length of the two imaging devices and a depth between the object of interest and the first and second imaging devices.
- [Claim 121] The apparatus of any one of claims 101-120, wherein the depth map is optimized by using a non-partial optimizing equation.
- [Claim 122] The apparatus of claim 121, wherein the depth assembly is configured to obtain a Jacobi iteration of the non-partial optimizing equation by a recurrence filtering.

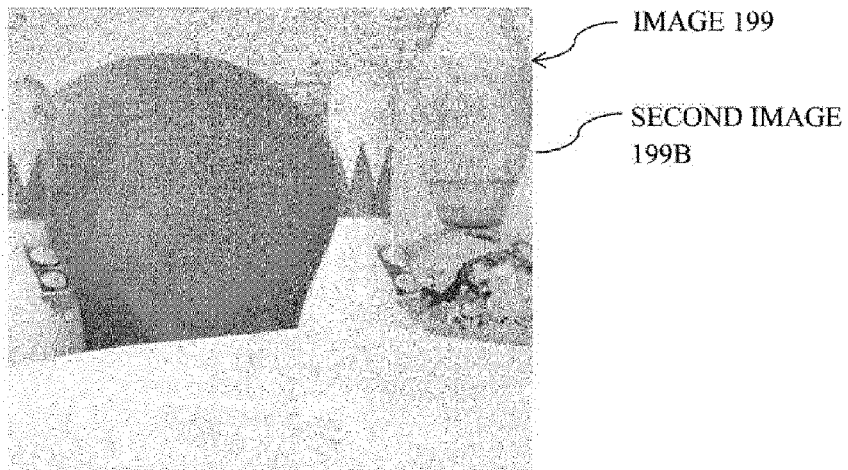
[Fig. 1]



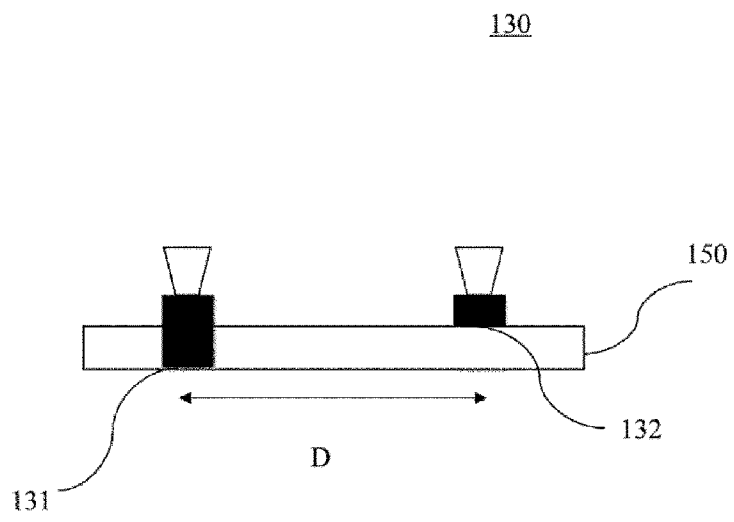
[Fig. 2]



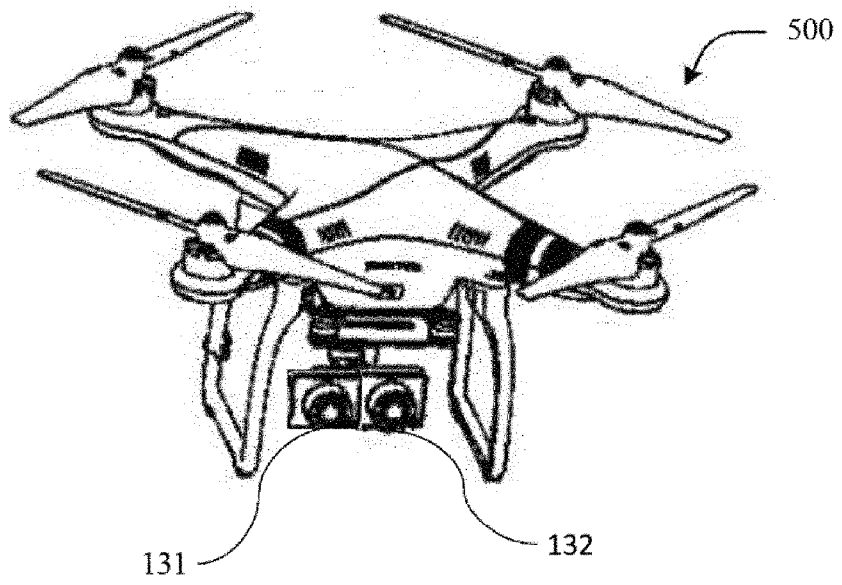
[Fig. 3]



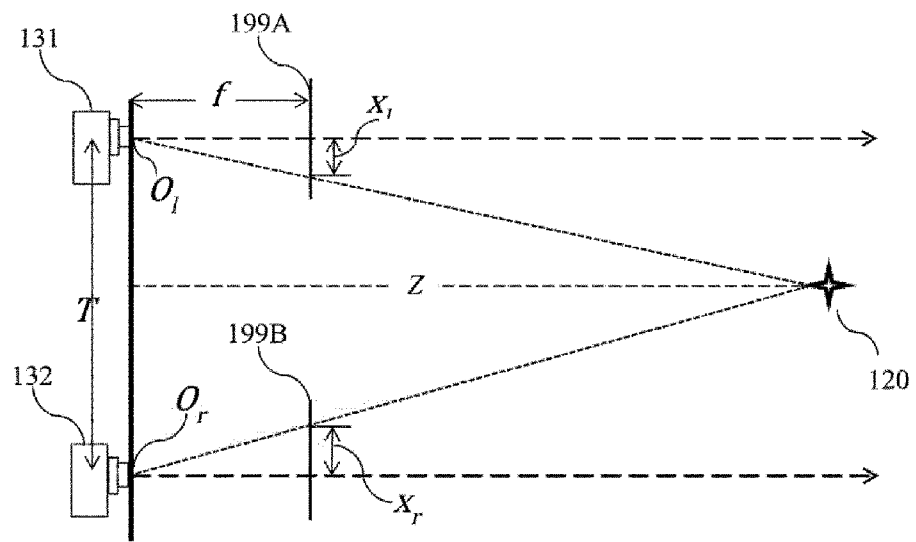
[Fig. 4]



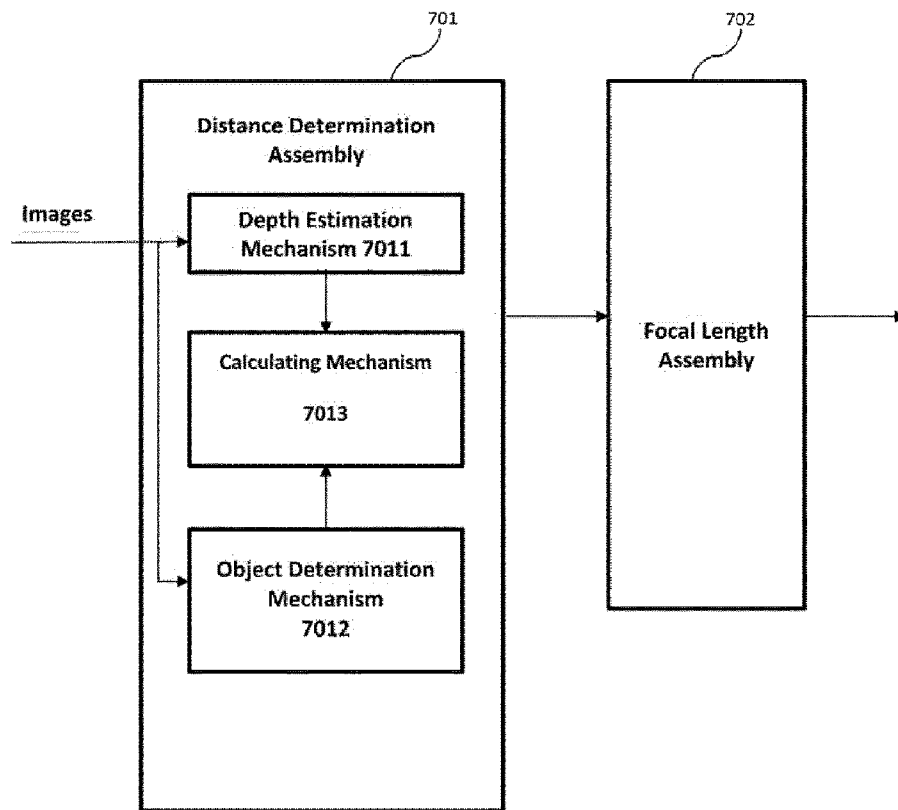
[Fig. 5]



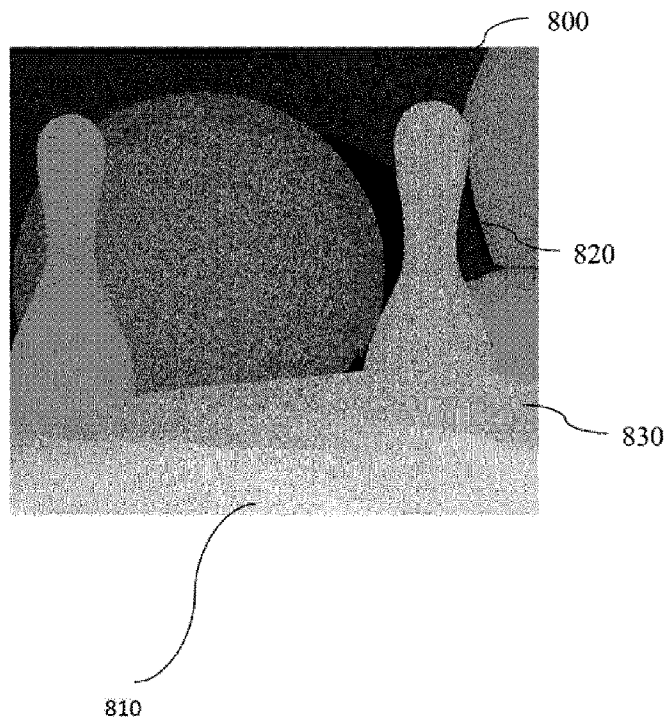
[Fig. 6]



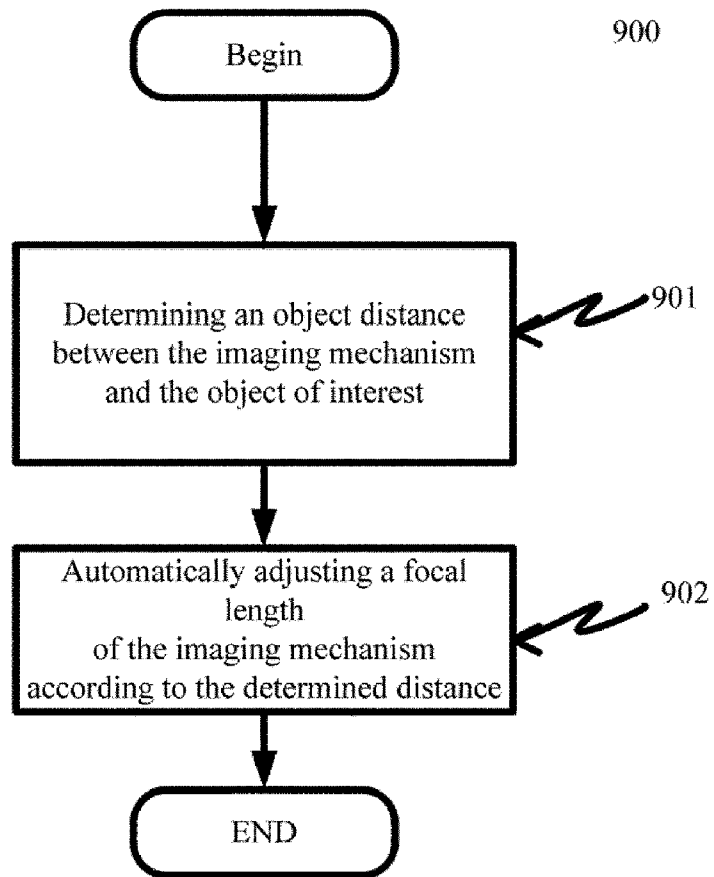
[Fig. 7]



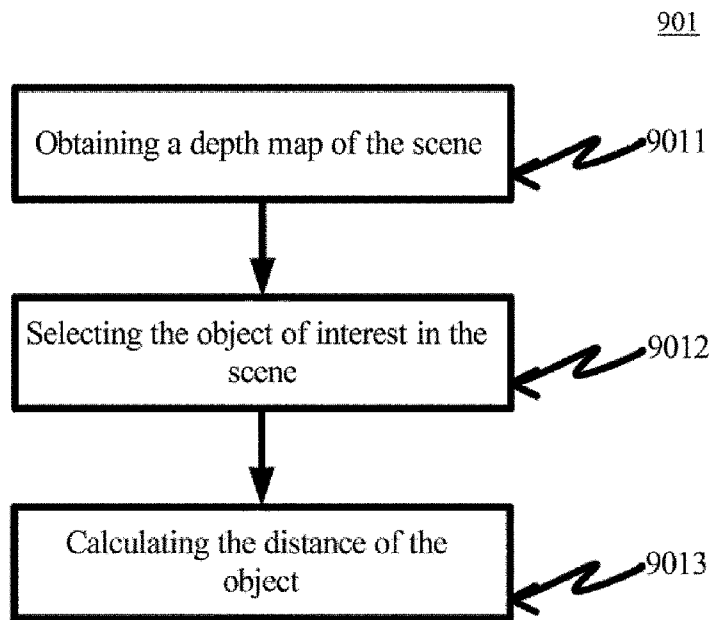
[Fig. 8]



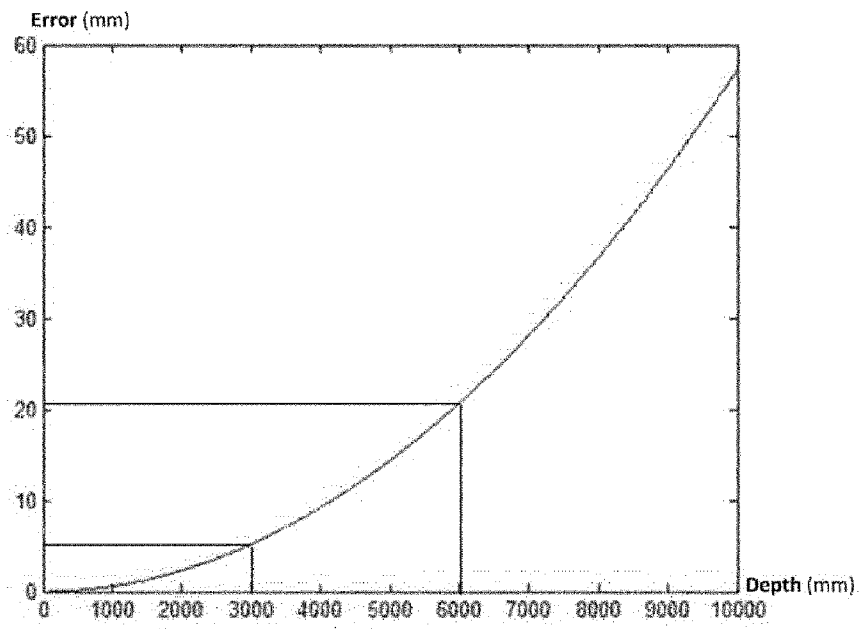
[Fig. 9]



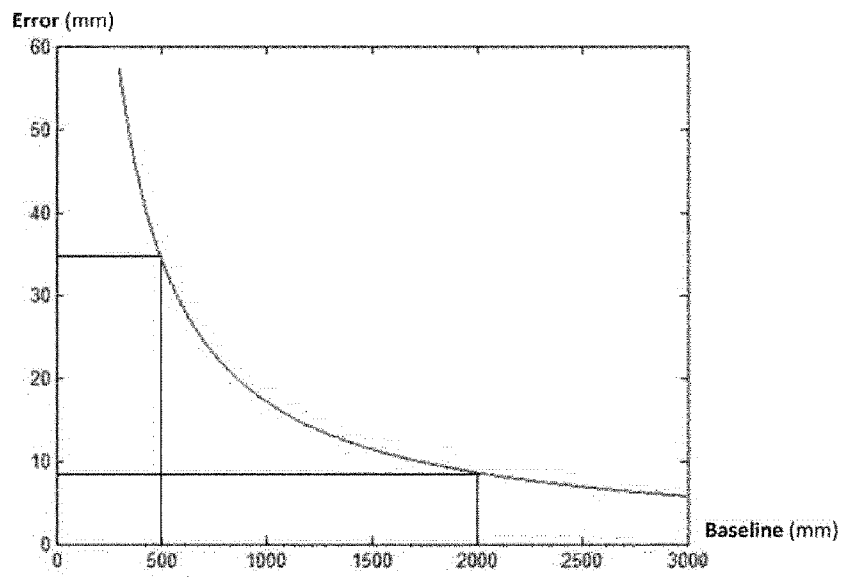
[Fig. 10]



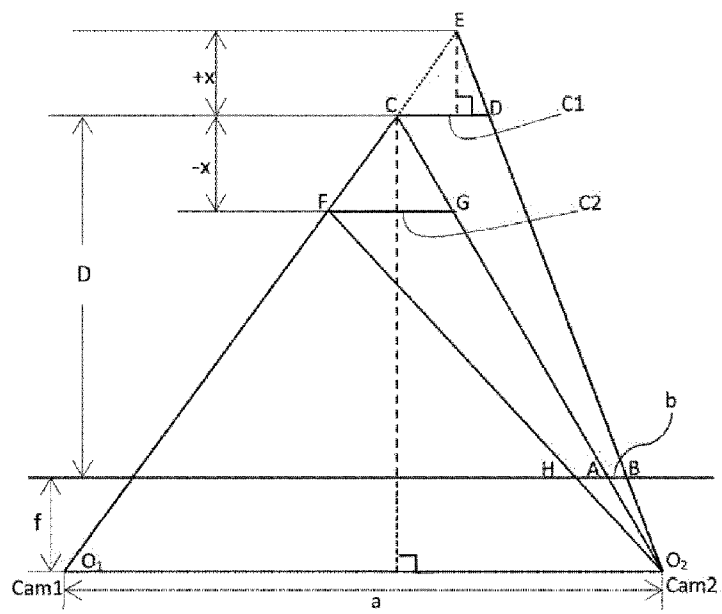
[Fig. 11]



[Fig. 12]



[Fig. 13]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2015/074336

A. CLASSIFICATION OF SUBJECT MATTER

H04N 13/02(2006.01)i

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

H04N

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

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

CNPAT, EPODOC, WPI, CNKI, IEEE: focal, length, adjust, auto, distance, 3d, three dimension, object

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 101840146 A (XIA, JIALIANG) 22 September 2010 (2010-09-22) the abstract	1-5, 27-42, 72-75
X	CN 101243458 A (SONY ELECTRONICS INC.) 13 August 2008 (2008-08-13) the abstract	76, 77, 99, 100
X	CN 103350281 A (SHENZHEN HANS LASER SCI.&TECHNOLOGY CO., LTD. ET AL.) 16 October 2013 (2013-10-16) the abstract	1
A	US 2014347456 A1 (PANASONIC CORP.) 27 November 2014 (2014-11-27) the whole document	1-122

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

* Special categories of cited documents:

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

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

04 December 2015

Date of mailing of the international search report

21 December 2015

Name and mailing address of the ISA/CN

STATE INTELLECTUAL PROPERTY OFFICE OF THE
P.R.CHINA
6, Xitucheng Rd., Jimen Bridge, Haidian District, Beijing
100088, China

Authorized officer

YANG, Shuangyi

Facsimile No. (86-10)62019451

Telephone No. (86-10)62413437

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2015/074336

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	101840146	A	22 September 2010	None			
CN	101243458	A	13 August 2008	JP	2009505104	A	05 February 2009
				WO	2007022329	A2	22 February 2007
				JP	5161087	B2	13 March 2013
				CN	101243458	B	27 June 2012
				US	2007036427	A1	15 February 2007
				US	7929801	B2	19 April 2011
				WO	2007022329	B1	15 November 2007
CN	103350281	A	16 October 2013	CN	103350281	B	29 July 2015
US	2014347456	A1	27 November 2014	JP	2014230019	A	08 December 2014