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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
29 March 2018 (29.03.2018)

WIPO | PCT

(10) International Publication Number
WO 2018/056802 A1

(51) International Patent Classification:
G06T 7/55 (2017.01)

(21) International Application Number:
PCT/MY2017/050057

(22) International Filing Date:
14 September 2017 (14.09.2017)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
PI 2016703427 21 September 2016 (21.09.2016) MY

(71) Applicant: **UNIVERSITI PUTRA MALAYSIA** [MY/MY]; Putra Science Park, Pejabat Timbalan Naib Canselor (penyelidikan Dan Inovasi), Universiti Putra Malaysia, Serdang Selangor, 43400 (MY).

(72) Inventors: **O.K. RAHMAT, Rahmita Wirza**; Universiti Putra Malaysia, Upm Serdang, Selangor Darul Ehsan, 43400 (MY). **SENG BENG, Ng**; Faculty Of Computer Science And Information Technology, Universiti Putra Malaysia, Upm Serdang, 43400 (MY). **KHALID, Fatimah**; Fakulti Sains Komputer Dan Teknologi Maklumat, Universiti Putra Malaysia, Serdang Selangor, 43400 (MY).

(74) Agent: **YIP, JIUN HANN**; Trademark2u Sdn Bhd (670910-m) No. 1, Block C, Jalan Dataran Sd 1, Dataran Sd Pju 9, Bandar Sri Damansara, Kuala Lumpur, 52200 (MY).

(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, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME,

(54) Title: A METHOD FOR ESTIMATING THREE-DIMENSIONAL DEPTH VALUE FROM TWO-DIMENSIONAL IMAGES

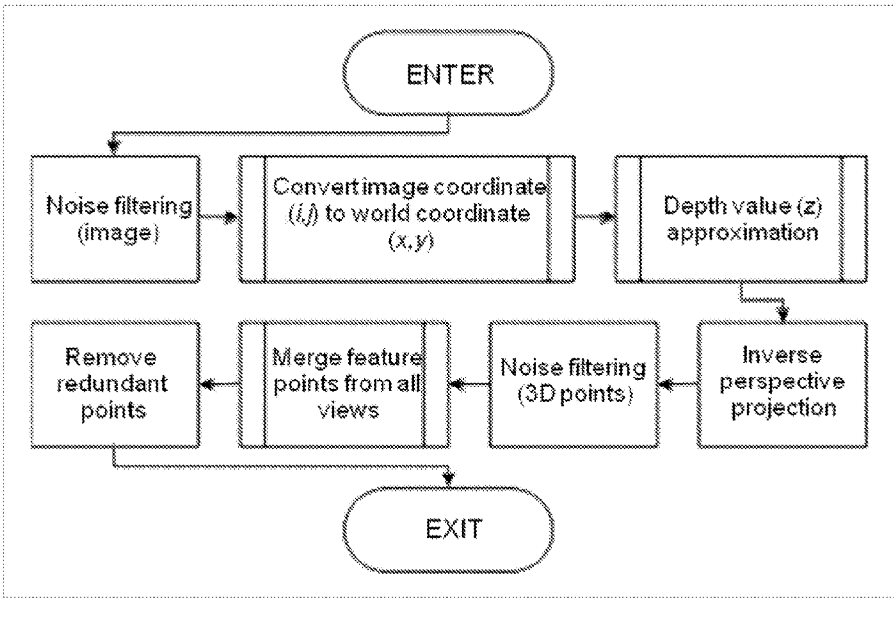

Fig. 1

(57) Abstract: The present invention relates to a method for estimating three-dimensional depth value from two-dimensional images, characterised by the steps of placing an object (102) on a rotatable plate (103); acquiring a first view of the object (102) comprising a first image and a second image of the object (102), wherein the first image of the object (102) is captured at an angle between 0° to 360°, and the second image is captured at an angle in a range of 1° to 35° relative to the first image; obtaining two-dimensional feature point coordinates by applying Good Features to Track technique and extracting colour information, simultaneously, from the first image and the second image; filtering two-dimensional feature point coordinates; applying Pyramidal Lucas-Kanade Optical Flow technique to obtain displacement magnitudes from the two-dimensional feature point coordinates; calculating World Coordinate; estimating the three-dimensional depth value of the acquired images; and implementing inverse perspective mapping algorithm on the

WO 2018/056802 A1

[Continued on next page]

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).

Declarations under Rule 4.17:
— *as to the identity of the inventor (Rule 4.17(i))*
— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*
— *of inventorship (Rule 4.17(iv))*

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

three-dimensional feature point coordinates.

A METHOD FOR ESTIMATING THREE-DIMENSIONAL DEPTH VALUE FROM TWO-DIMENSIONAL IMAGES

Background of the Invention

Field of the Invention

This invention relates to a method for estimating depth value from at least two images, and more particularly to a method of three-dimensional depth value estimation from at least two images using optical flow and trigonometry method.

Description of Related Arts

Three dimensional (3D) modeling of physical objects can be very useful in many areas, such as computer graphics and animation, robot vision, reverse engineering, and medical. 3D modeling can be represented from the scratch using modeling software, or digitized from real world objects. The basic information required for this representation is the *x*, *y*, and *z* coordinates. Further manipulation of these coordinates can deduce the objects' dimensions (width, height, and depth). Other attributes such as the models' surface colour, texture, lighting, shading and shadow contribute to a more realistic representation.

Conventional digitization methods utilize Coordinate Measuring Machine (CMMs) or laser scanners to obtain coordinate for each good feature from an object to a digital data. Nevertheless, both of these devices are very costly and require a certain amount of technical knowledge during usage and maintenance.

For example, US Patent No. 7,342,669 B2 disclosed a 3D measurement system that uses laser-light and estimates depth information of an object by applying a triangulation method. The system consists of an apparatus comprising a laser projecting, an image capturing device and a computer. The cited art also provides a method for 3D measurement using said apparatus comprises projecting a line-laser to an object where the apparatus having light-emitting diodes (LEDs) attached to a line-laser projector for estimating the position and orientation of the laser projecting device; then, capturing projected line-laser light and the LEDs at the same time using the image capturing device; then, calculating, using the computer, a 3D shape of the object from the captured image using the triangulation method; and outputting the calculated 3D shape. Meanwhile, by applying the teachings in this cited art, a user can quickly acquire a precise 3D shape without using complex equipment and can also interactively check the regions whose positions are not measured during the measuring process by displaying the acquired 3D positions in real-time. Thus, measurement of 3D shape of the target object becomes efficient. However, said laser projecting device may face problems on objects with shiny surfaces or objects that do not reflect light, which include black colour and transparent surface. Thus, the user may have the difficulties to verify the depth of the object.

Another example such as US Patent No. 8,330,803 B2 disclosed a method for 3D digitization of an object, in which a plurality of camera images of the object are recorded and assembled to determine the 3D coordinates of the object. An apparatus for performing the method for 3D digitization of the object comprises a projector and one or more cameras. The projector projects a light pattern onto the object, in more particular white –light strips. In order to improve the method for 3D digitization of the object, 2D feature points from the plurality of camera images of the objects are determined without human intervention. The 2D point correspondences between the 2D feature points of a picture and the 2D feature points of another picture are determined. Several of these 2D point correspondences are selected, and an associated 3D transformation is determined. The quality of this 3D transformation is determined with reference to the transformed 3D coordinates of the 2D feature points. Valid 3D feature points are determined therefrom. For assembling the camera images of the object, the 3D coordinates of the valid 3D feature points are use. However, the cited patent does not disclosed in detail how to calculate or determine the depth value of each features across said image-set.

US Patent No. 7,573,475 B2 disclosed a method of converting 2D image to 3D image. The method includes receiving a first 2D image comprising image data, where the first 2D image is captured from a first camera location. The method also includes projecting at least a portion of the first 2D image onto computer-generated geometry. The image data has depth value associated with the computer-generated geometry. The system includes rendering, using the computer-generated geometry and a second camera location differs from the first camera location, a second 2D image that is stereoscopically complementary to the first 2D image, and infilling image data that absent from the second 2D image. The cited patent uses two different images captured from two different locations for example image from the left and the right of the object. Then, geometry map is built for both objects and compared to each other to calculate depth for the geometry map. Geometry map is accurate to be used if the object is simple but when the object is complicated, said method becomes inaccurate to determine the depth of the object.

Accordingly, it can be seen in the prior arts that there exists a need to provide a simple depth estimation technique to merge the 2D camera images into a set of 3D surface points with colour information without using high cost gadgets such as coordinate measuring machine (CMM) or laser scanner.

Summary of Invention

It is an objective of the present invention to provide to a method of three-dimensional depth value estimation from at least two images using optical flow and trigonometry method.

It is also an objective of the present invention to provide to a method of three-dimensional depth value estimation and colour information extraction from at least two images.

It is yet another objective of the present invention to provide a method of three-dimensional depth value estimation in either a control environment or open space environment.

Accordingly, these objectives may be achieved by following the teachings of the present invention. The present invention relates to a method for estimating three-dimensional depth value from two-dimensional images, characterised by the steps of placing an object on a rotatable plate; acquiring a first view of the object comprising a first image and a second image of the object, wherein the first image of the object is captured at an angle between 0° to 360°, and the second image is captured at an angle in a range of 1° to 35° relative to the first image; obtaining two-dimensional feature point coordinates by applying Good Features to Track technique and extracting colour information, simultaneously, from the first image and the second image; filtering two-dimensional feature point coordinates; applying Pyramidal Lucas-Kanade Optical Flow technique to obtain displacement magnitudes from the two-dimensional feature point coordinates; calculating World Coordinate; estimating the three-dimensional depth value of the acquired images; and implementing inverse perspective mapping algorithm on the three-dimensional feature point coordinates.

Brief Description of the Drawings

The features of the invention will be more readily understood and appreciated from the following detailed description when read in conjunction with the accompanying drawings of the preferred embodiment of the present invention, in which:

Fig. 1 is a flow chart of a method for estimating three-dimensional depth value from a two-dimensional image; and

Fig. 2 is diagram showing preferred arrangement of a kit for estimating three-dimensional depth value.

Detailed Description of the Invention

As required, detailed embodiments of the present invention are disclosed herein; however, it is to be understood that the disclosed embodiments are merely exemplary of the invention, which may be embodied in various forms. Therefore, specific structural and functional details disclosed herein are not to be interpreted as limiting but merely as a basis for claims. It should be understood that the drawings and detailed description thereto are not intended to limit the invention to the particular form disclosed, but on the contrary, the invention is to cover all modifications, equivalents and alternatives falling within the scope of the present invention as defined by the appended claims. As used throughout this application, the word "may" is used in a permissive sense (i.e., meaning having the potential to), rather than the mandatory sense (i.e., meaning must). Similarly, the words "include," "including," and "includes" mean including, but not limited to. Further, the words "a" or "an" mean "at least one" and the word "plurality" means one or more, unless otherwise mentioned. Where the abbreviations or technical terms are used, these indicate the commonly accepted meanings as known in the technical field. For ease of reference, common reference numerals will be used throughout the figures when referring to the same or similar features common to the figures. The present invention will now be described with reference to Figs. 1 and 2.

The present invention relates to a method for estimating three-dimensional depth value from two-dimensional images, characterised by the steps of:

placing an object (102) on a rotatable plate (103) with a distance from an image capture apparatus (101);

acquiring a first view of the object (102) comprising a first image and a second image of the object (102), wherein the first image of the object (102) is captured at an angle between 0° to 360°, and the second image is captured at an angle in a range of 1° to 35° relative to the first image;

obtaining two-dimensional feature point coordinates by applying Good Features to Track technique and extracting colour information, simultaneously, from the first image and the second image;

filtering two-dimensional feature point coordinates to eliminate noise using first filtering means;

applying Pyramidal Lucas-Kanade Optical Flow technique to obtain displacement magnitudes from the two-dimensional feature point coordinates;

calculating World Coordinate for each of the two-dimensional feature point coordinates;

estimating the three-dimensional depth value of the acquired images by calculating each of the two-dimensional feature point coordinates using equation (1), thereby producing three-dimensional feature point coordinates:

$$z_W = \frac{objdist(x_{W1}y_{W2}\cos\alpha - x_{W2}y_{W1})}{objdist(y_{W2}\sin\alpha) + x_{W1}y_{W2}\cos\alpha - x_{W2}y_{W1}} \qquad (1)$$

wherein,

z_W is the three-dimensional depth value for the two-dimensional feature point coordinate (x_{W1}, y_{W1}),

objdist is a distance of center of rotation (COR) from the image capture apparatus (101) (optic center),

(x_{W1}, y_{W1}) is a coordinate of the two-dimensional feature point from the first image,

(x_{W2}, y_{W2}) is a matching coordinate of the two-dimensional feature point in the second image,

α is a rotation angle;

implementing inverse perspective mapping algorithm on the three-dimensional feature point coordinates to remove perspective distortion effects.

In a preferred embodiment of the method for estimating three-dimensional depth value from two-dimensional images, at least two views of the object (102) are acquired to reflect all sides of the object (102).

In a preferred embodiment of the method for estimating three-dimensional depth value from two-dimensional images, said first image is captured at an angle of 0°, 60°, 120°, 180°, 240°, 360° or a combination thereof.

In a preferred embodiment of the method for estimating three-dimensional depth value from two-dimensional images, the three-dimensional feature point coordinates from all views are merged and collected into a single set of three-dimensional feature point coordinates.

In a preferred embodiment of the method for estimating three-dimensional depth value from two-dimensional images, the views are merged by using an inverse rotation matrix with equation (8):

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = R_y(-\alpha)\begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

$$= \begin{pmatrix} \cos(-\alpha) & 0 & \sin(-\alpha) \\ 0 & 1 & 0 \\ -\sin(-\alpha) & 0 & \cos(-\alpha) \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} \qquad (8)$$

wherein,

(x', y', z') are merged three-dimensional feature point coordinate after applying the inverse rotation matrix on three-dimensional feature point coordinate (x, y, z).

In a preferred embodiment of the method for estimating three-dimensional depth value from two-dimensional images, further filtering step on the single set of three-dimensional feature point coordinates is required to remove redundant points using a second filtering means.

In a preferred embodiment of the method for estimating three-dimensional depth value from two-dimensional images, the colour information is extracted using equations (2), (3) and (4):

$$Bi,j = (i \times imgW \times 3) + [(j \times 3) + 0]$$

$$= 3[i(imgW) + j] \tag{2}$$

$$Gi,j = (i \times imgW \times 3) + [(j \times 3) + 1] \tag{3}$$

$$Ri,j = (i \times imgW \times 3) + [(j \times 3) + 2] \tag{4}$$

wherein,

Bi,j is a blue colour information of pixel (*i* ,*j*),

Gi,j is a green colour information of pixel (*i,j*),

Ri,j is a red colour information of pixel (*i,j*).

In a preferred embodiment of the method for estimating three-dimensional depth value from two-dimensional images, the world coordinate is obtained using equations (5) and (6):

$$\text{Scaling Factor, k} = \frac{objheight}{detectedheight} \tag{5}$$

wherein,

objheight is an actual object (102) height measured using linear callipers,

detectedheight is a distance of highest point and lowest point of the object (102) detected in the acquired image, measured in pixels:

$$(x_{Wi}, y_{Wi}) = (kx_{Pi}, ky_{Pi}) \tag{6}$$

wherein,

(x_{Wi}, y_{Wi}) is the world coordinate for the two-dimensional feature point coordinate i,

(x_{Pi}, y_{Pi}) is a projected coordinate for the two-dimensional feature point coordinate i on the acquired image.

In a preferred embodiment of the method for estimating three-dimensional depth value from a two-dimensional image, the inverse perspective mapping algorithm is implemented using equation (7):

$$(x'_{W1}, y'_{W1}) = \left(x_{W1}\left(1 - \frac{z_1}{objdist}\right), y_{W1}\left(1 - \frac{z_1}{objdist}\right)\right) \quad (7)$$

wherein,

(x'_{W1}, y'_{W1}) is a corrected coordinate for the two-dimensional feature point coordinate from the first image,

(x_{W1}, y_{W1}) is the world coordinate for the two-dimensional feature point coordinate in the first image.

z_1 is an approximated three-dimensional depth value for the two-dimensional feature point coordinate (x_{W1}, y_{W1}),

objdist is a distance of center of rotation (COR) from the image capture apparatus (101) (optic center),

In a preferred embodiment of the method for estimating three-dimensional depth value from two-dimensional images, the step of filtering two-dimensional feature point coordinates using the first filtering means is by applying Euclidean distance method.

In a preferred embodiment of the method for estimating three-dimensional depth value from two-dimensional images, the estimation of three-dimensional depth value is carried out in a controlled environment includes controlling lighting and background colour.

In a preferred embodiment of the method for estimating three-dimensional depth value from two-dimensional images, the estimation of three-dimensional depth value is carried out in an open space.

In a preferred embodiment of the method for estimating three-dimensional depth value from two-dimensional images, further filtering step on the three-dimensional feature point coordinates is carried out to eliminate noise using a third filtering means, before the merging step.

Examples

In an exemplary embodiment of the present invention as illustrated in Figure 2, a kit (100) mainly comprises a rotatable plate (103) and an image capture apparatus (101), where an object (102) is a blue toy bird which placed on the rotatable plate (103) with a distance from the image capture apparatus (101). In a preferred embodiment, the blue toy bird has a width, height and depth of 134.03 mm, 94.90 mm, 35.12 mm respectively. The distance from the image capture apparatus (101) to a centre of rotatable plate (103) is set at a length that is able to capture the whole image of the object. An embodiment shows either use of two pieces of 180° half circle protractor for forming a circle or one piece of 360° protractor, which is placed underneath the rotatable plate (103) for ease of reference when measuring the angle of rotation of the rotatable plate (103).

The kit (100) of the present invention is for estimating three-dimensional depth value of the object (102) by carrying a method as shown in Figure 1. In an exemplary embodiment, the kit (100) is placed in a controlled environment which includes controlling lighting and background colour. The background colour is preferably black colour to eliminate reflections during image capturing. A first image of the object (102) is acquired by rotating the rotatable plate (103) to a predetermined angle between 0° to 360°. In a preferred embodiment, said predetermined angle is 0°, 60°, 120°, 180°, 240° and 360°. In another preferred embodiment, a combination of predetermined angle may be selected for acquiring multiple images at different views of the object (102), for example 0° and 180° in a pair angles and 60° and 240°in a pair angles. The pair angles is preferably a vertically opposite angles. There is at least two views of the object (102) are acquired to obtain at least four images to reflect all sides of the object (102).

Then, the rotatable plate (103) is rotated to an angle in a range of 1° to 35° relative to the predetermined angle of the first image for capturing a second image. In a preferred embodiment, the first image and the second image are captured using the image capture apparatus (101), wherein the image capture apparatus (101) is preferably a camera and more preferably a webcam.

Two-dimensional feature point coordinates is obtained by applying Good Features to Track technique and the colour information is extracted from the two-dimensional feature point coordinates, simultaneously, from the first image and the second image. Said colour information is extracted using equations (2), (3), and (4) as followings:

$$Bi,j = (i \times imgW \times 3) + [(j \times 3) + 0]$$

$$= 3[i(imgW) + j] \qquad (2)$$

$$Gi,j = (i \times imgW \times 3) + [(j \times 3) + 1] \qquad (3)$$

$$Ri,j = (i \times imgW \times 3) + [(j \times 3) + 2] \qquad (4)$$

wherein,

Bi,j is a blue colour information of pixel (*i,j*),

Gi,j is a green colour information of pixel (*i,j*),

Ri,j is a red colour information of pixel (*i,j*).

The two-dimensional feature point coordinates is filtered to eliminate noise using first filtering means, where the detected two-dimensional feature point coordinates which are located further away from the centre of the images are assumed to have a higher probability of noise. The two-dimensional feature point coordinates is filtered by applying Euclidean distance method.

Pyramidal Lucas-Kanade optical Flow technique is applied to the two-dimensional feature point coordinates to obtain displacement magnitudes and to find matching two-dimensional feature point coordinates from the first image and second image. If any of the two-dimensional feature point coordinates from the first and second image do not match, the unmatched two-dimensional feature point coordinates are discarded. Then, a World Coordinate for each of the two-dimensional feature point coordinates is calculated using equations (5) and (6) as followings:

$$\text{Scaling Factor,k} = \frac{objheight}{detectedheight} \quad (5)$$

wherein,

objheight is an actual object (102) height measured using linear callipers,

detectedheight is a distance of highest point and lowest point of the object (102) detected in the acquired image, measured in pixels:

$$(x_{Wi}, y_{Wi}) = (kx_{Pi}, ky_{Pi}) \quad (6)$$

wherein,

(x_{Wi}, y_{Wi}) is the world coordinate for the two-dimensional feature point coordinate i,

(x_{Pi}, y_{Pi}) is a projected coordinate for the two-dimensional feature point coordinate i on the acquired image.

The three-dimensional depth value of the acquired images is estimated by calculating each of the two-dimensional feature point coordinates using equation (1), thereby producing three-dimensional feature point coordinates:

$$z_W = \frac{objdist(x_{W1}y_{W2}\cos\alpha - x_{W2}y_{W1})}{objdist(y_{W2}\sin\alpha) + x_{W1}y_{W2}\cos\alpha - x_{W2}y_{W1}} \quad (1)$$

wherein,

z_W is the three-dimensional depth value for the two-dimensional feature point coordinate (x_{W1}, y_{W1}),

objdist is a distance of center of rotation (COR) from the image capture apparatus (101) (optic center),

(x_{W1}, y_{W1}) is a coordinate of the two-dimensional feature point from the first image,

(x_{W2}, y_{W2}) is a matching coordinate of the two-dimensional feature point in the second image,

α is a rotation angle.

Then, an inverse perspective mapping algorithm is implemented on the three-dimensional feature point coordinates to remove perspective distortion effects using equation (7):

$$(x'_{W1}, y'_{W1}) = \left(x_{W1}\left(1 - \frac{z_1}{objdist}\right), y_{W1}\left(1 - \frac{z_1}{objdist}\right)\right) \quad (7)$$

wherein,

(x'_{W1}, y'_{W1}) is a corrected coordinate for the two-dimensional feature point coordinate from the first image,

(x_{W1}, y_{W1}) is the world coordinate for the two-dimensional feature point coordinate in the first image.

z_1 is an approximated three-dimensional depth value for the two-dimensional feature point coordinate (x_{W1}, y_{W1}),

objdist is a distance of center of rotation (COR) from the image capture apparatus (101) (optic center).

The three-dimensional feature point coordinates from all individual views collected are merged into a single set of three-dimensional feature point coordinates. The merging step is performed using an inverse rotation matrix with equation (8):

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = R_y(-\alpha)\begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

$$= \begin{pmatrix} \cos(-\alpha) & 0 & \sin(-\alpha) \\ 0 & 1 & 0 \\ -\sin(-\alpha) & 0 & \cos(-\alpha) \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad (8)$$

wherein,

(x', y', z') are merged three-dimensional feature point coordinate after applying the inverse rotation matrix on three-dimensional feature point coordinate (x, y, z).

The single set of three-dimensional feature point coordinates is filtered to remove redundant points using a second filtering means.

In another exemplary embodiment of the present invention, the kit (100) is placed in an open space for carrying out the estimation of three-dimensional depth value of the object (102). If the kit (100) is placed in the open space, a further filtering step is required on the three-dimensional feature point coordinates to eliminate noise using a third filtering means, before the merging step of the present invention.

<u>Example of calculation for three-dimensional depth value approximation</u>

User input (known values):

$objdist = 300$ mm

$\alpha = 3°$

$objectheight = 94.90$ mm

Detected two-dimensional feature point pair:

Point 1 (x_{P1}, y_{P1}) = (12.3456, 123.4567)

Point 2 (x_{P2}, y_{P2}) = (15.1234, 120.3333)

*Point 2 is a matching two-dimensional feature point of Point 1 in the second image after the α° rotation.

A minimum value and a maximum value of y is extracted from the two-dimensional feature points collected to calculate a detected height. For example,

$\min(y) = 50.1932$

$\max(y) = 313.4567$

$detectedheight = 313.4567 - 50.1932$

$= 263.2635$

$$\text{Scaling Factor, } k = \frac{objheight}{detectedheight}$$

$$= \frac{94.90}{263.2635}$$

$$= 0.3605$$

Calculation for world coordinate:

$$(x_{W_i}, y_{W_i}) = (kx_{p_i}, ky_{p_i})$$

For point 1,

$$(x_{W1}, y_{W1}) = (0.3605 \times 12.3456, 0.3605 \times 123.4567)$$
$$= (4.4506, 44.5061)$$
$$(x_{W2}, y_{W2}) = (0.3605 \times 15.1234, 0.3605 \times 120.3333)$$
$$= (5.4520, 43.3802)$$

Depth value estimation:

$$z_W = \frac{objdist(x_{W1}y_{W2}\cos\alpha - x_{W2}y_{W1})}{objdist(y_{W2}\sin\alpha) + x_{W1}y_{W2}\cos\alpha - x_{W2}y_{W1}}$$
$$= \frac{300 \times [(4.4506 \times 43.3802 \times \cos 3°) - (5.4520 \times 44.5061)]}{300 \times 43.3802 \times (\sin 3°) + (4.4506 \times 43.3802 \times \cos 3°) - (5.4520 \times 44.5061)}$$
$$= \frac{(57920.3754 \times \cos 3°) - 242.6473}{(13014.06 \times \sin 3°) + (57920.3754 \times \cos 3°) - 242.6473}$$
$$= \frac{57840.9975 - 242.6473}{681.1033 + 57840.9975 - 242.6473}$$
$$= \frac{57598.3502}{58279.4535}$$
$$= 0.9883$$

So, the 3D coordinate before inverse perspective projection for Point 1 is (4.4506, 44.5061, 0.9883).

By applying the perspective distortion, inverse perspective projection is applied on the point (x_{w1}, y_{w1}) to find the corrected coordinates, (x'_{w1}, y'_{w1})

$$(x'_{w1}, y'_{w1}) = (x_{w1}\left(1 - \frac{z_1}{objdist}\right), y_{w1}\left(1 - \frac{z_1}{objdist}\right))$$

$$= (4.4506 \times \left(1 - \frac{0.9883}{300}\right), 44.5061 \times \left(1 - \frac{0.9883}{300}\right))$$

$$= (4.4506 \times 0.9967, 44.5061 \times 0.9967)$$

$$= (4.4359, 44.3592)$$

So, the three-dimensional coordinate for Point 1 after inverse perspective projection is (4.4359, 44.3592, 0.9883).

Although the present invention has been described with reference to specific embodiments, also shown in the appended figures, it will be apparent for those skilled in the art that many variations and modifications can be done within the scope of the invention as described in the specification and defined in the following claims.

Description of the reference numerals used in the accompanying drawings according to the present invention:

Reference Numerals	Description
100	Kit
101	Image capture apparatus
102	Object
103	Rotatable plate

Claims

I/We claim:

1. A method for estimating three-dimensional depth value from two-dimensional images, characterised by the steps of:

placing an object (102) on a rotatable plate (103) with a distance from an image capture apparatus (101);

acquiring a first view of the object (102) comprising a first image and a second image of the object (102), wherein the first image of the object (102) is captured at an angle between 0° to 360°, and the second image is captured at an angle in a range of 1° to 35° relative to the first image;

obtaining two-dimensional feature point coordinates by applying Good Features to Track technique and extracting colour information, simultaneously, from the first image and the second image;

filtering two-dimensional feature point coordinates to eliminate noise using first filtering means;

applying Pyramidal Lucas-Kanade Optical Flow technique to obtain displacement magnitudes from the two-dimensional feature point coordinates and to find matching two-dimensional feature point coordinates from the first image and second image;

calculating World Coordinate for each of the two-dimensional feature point coordinates;

estimating the three-dimensional depth value of the acquired images by calculating each of the two-dimensional feature point coordinates using equation (1), thereby producing three-dimensional feature point coordinates:

$$z_W = \frac{objdist(x_{W1}y_{W2}\cos\alpha - x_{W2}y_{W1})}{objdist(y_{W2}\sin\alpha) + x_{W1}y_{W2}\cos\alpha - x_{W2}y_{W1}} \qquad (1)$$

wherein,

z_W is the three-dimensional depth value for the two-dimensional feature point coordinate (x_{W1}, y_{W1}),

objdist is a distance of center of rotation (COR) from the image capture apparatus (101) (optic center),

(x_{W1}, y_{W1}) is a coordinate of the two-dimensional feature point from the first image,

(x_{W2}, y_{W2}) is a matching coordinate of the two-dimensional feature point in the second image,

α is a rotation angle;

implementing inverse perspective mapping algorithm on the three-dimensional feature point coordinates to remove perspective distortion effects.

2. The method for estimating three-dimensional depth value from two-dimensional images according to claim 1, wherein at least two views of the object (102) are acquired to reflect all sides of the object (102).

3. The method for estimating three-dimensional depth value from two-dimensional images according to claim 1 and claim 2, wherein said first image is captured at an angle of 0°, 60°, 120°, 180°, 240°, 360° or a combination thereof.

4. The method for estimating three-dimensional depth value from two-dimensional images according to claim 2, wherein the three-dimensional feature point coordinates from all views are merged and collected into a single set of three-dimensional feature point coordinates.

5. The method for estimating three-dimensional depth value from two-dimensional images according to claim 4, wherein the views are merged by using an inverse rotation matrix with equation (8):

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = R_y(-\propto)\begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

$$= \begin{pmatrix} \cos(-\alpha) & 0 & \sin(-\alpha) \\ 0 & 1 & 0 \\ -\sin(-\alpha) & 0 & \cos(-\alpha) \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad (8)$$

wherein,

(x', y', z') are merged three-dimensional feature point coordinate after applying the inverse rotation matrix on three-dimensional feature point coordinate (x, y, z).

6. The method for estimating three-dimensional depth value from two-dimensional images according to claim 4, wherein further filtering step on the single set of three-dimensional feature point coordinates is required to remove redundant points using a second filtering means.

7. The method for estimating three-dimensional depth value from two-dimensional images according to claim 1, wherein the colour information is extracted using equations (2), (3) and (4):

$$Bi,j = (i \times imgW \times 3) + [(j \times 3) + 0]$$

$$= 3[i(imgW) + j] \quad (2)$$

$$Gi,j = (i \times imgW \times 3) + [(j \times 3) + 1] \quad (3)$$

$$Ri,j = (i \times imgW \times 3) + [(j \times 3) + 2] \quad (4)$$

wherein,

Bi,j is a blue colour information of pixel (i,j),

Gi,j is a green colour information of pixel (i,j),

Ri,j is a red colour information of pixel (i,j).

8. The method for estimating three-dimensional depth value from two-dimensional images according to claim 1, wherein the world coordinate is obtained using equations (5) and (6):

$$\text{Scaling Factor, k} = \frac{objheight}{detectedheight} \quad (5)$$

wherein,

objheight is an actual object (102) height measured using linear callipers,

detectedheight is a distance of highest point and lowest point of the object (102) detected in the acquired image, measured in pixels:

$$(x_{Wi}, y_{Wi}) = (kx_{Pi}, ky_{Pi}) \quad (6)$$

wherein,

(x_{Wi}, y_{Wi}) is the world coordinate for the two-dimensional feature point coordinate i,

(x_{Pi}, y_{Pi}) is a projected coordinate for the two-dimensional feature point coordinate i on the acquired image.

9. The method for estimating three-dimensional depth value from two-dimensional images according to claim 1, wherein inverse perspective mapping algorithm is implemented using equation (7):

$$(x'_{W1}, y'_{W1}) = \left(x_{W1}\left(1 - \frac{z_1}{objdist}\right), y_{W1}\left(1 - \frac{z_1}{objdist}\right)\right) \quad (7)$$

wherein,

(x'_{W1}, y'_{W1}) is a corrected coordinate for the two-dimensional feature point coordinate from the first image,

(x_{W1}, y_{W1}) is the world coordinate for the two-dimensional feature point coordinate in the first image.

z_1 is an approximated three-dimensional depth value for the two-dimensional feature point coordinate (x_{w1}, y_{w1}),

objdist is a distance of center of rotation (COR) from the image capture apparatus (101) (optic center).

10. The method for estimating three-dimensional depth value from two-dimensional images according to claim 1, wherein the step of filtering two-dimensional feature point coordinates using the first filtering means is by applying Euclidean distance method.

11. The method for estimating three-dimensional depth value from two-dimensional images according to claim 1, wherein the estimation of three-dimensional depth value is carried out in a controlled environment includes controlling lighting and background colour.

12. The method for estimating three-dimensional depth value from two-dimensional images according to claim 1, wherein the estimation of three-dimensional depth value is carried out in an open space.

13. The method for estimating three-dimensional depth value from two-dimensional images according to claim 4 and claim 12, wherein further filtering step on the three-dimensional feature point coordinates is carried out to eliminate noise using a third filtering means, before the merging step.

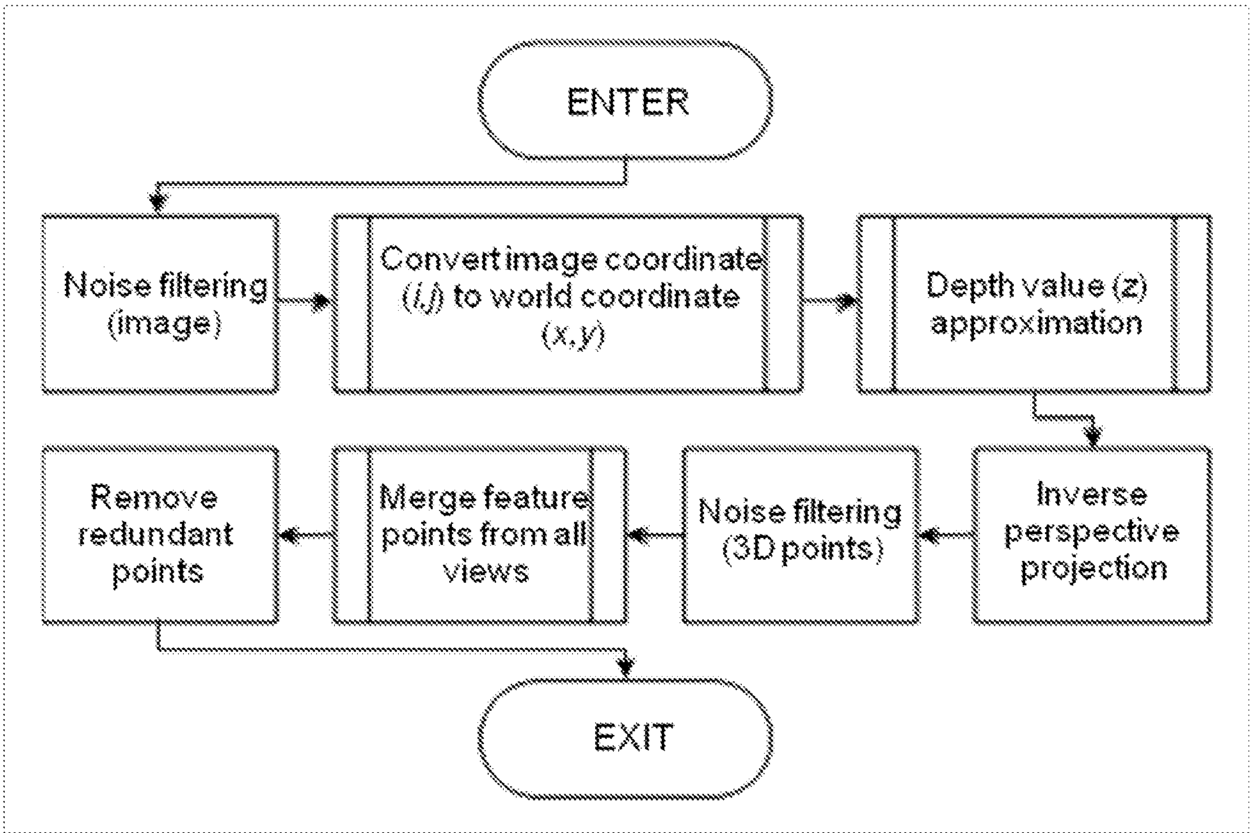
ENTER
Noise filtering (image)
Convert image coordinate (i,j) to world coordinate (x,y)
Depth value (z) approximation
Inverse perspective projection
Noise filtering (3D points)
Merge feature points from all views
Remove redundant points
EXIT

Fig. 1

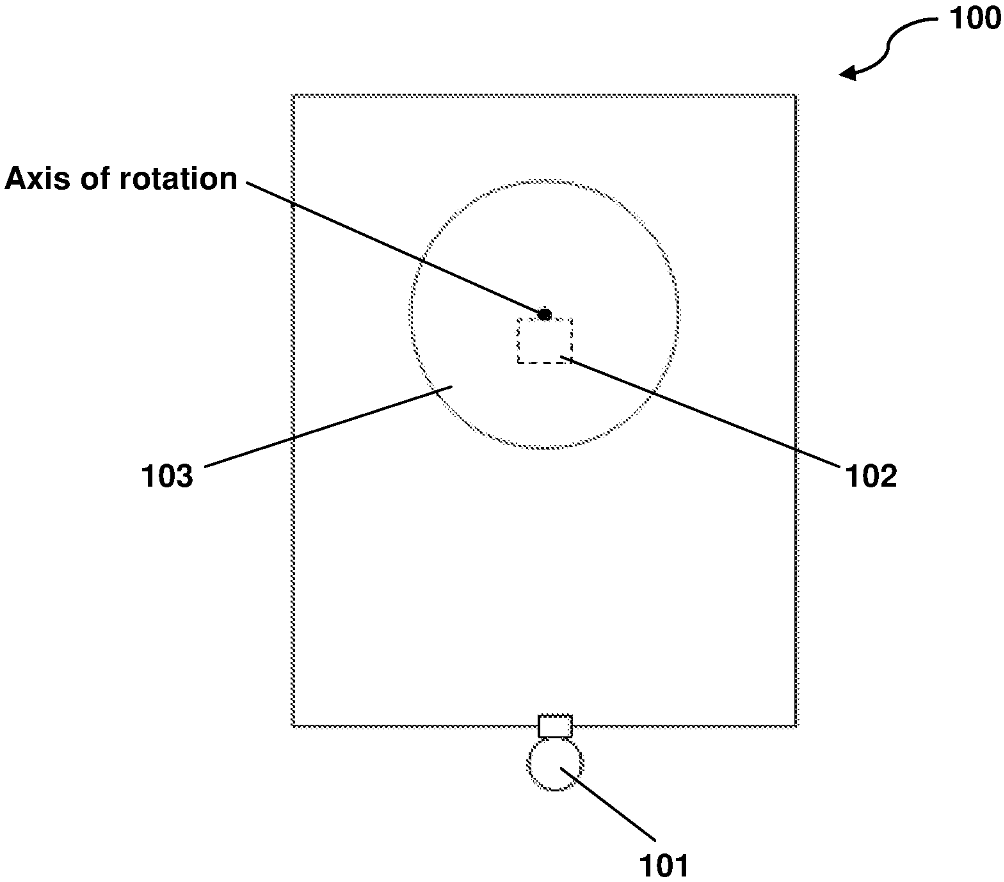

Fig. 2

INTERNATIONAL SEARCH REPORT

International application No.
PCT/MY2017/050057

A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl. G06T7/55(2017.01)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)

Int.Cl. G06T7/55

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

Published examined utility model applications of Japan 1922-1996
Published unexamined utility model applications of Japan 1971-2017
Registered utility model specifications of Japan 1996-2017
Published registered utility model applications of Japan 1994-2017

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

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2009/0226094 A1 (PIONEER CORPORATION) 2009.09.10, Paragraphs [0072]-[0114], [0133], [0134], Figs.1-6, 12 & WO 2008/032375 A1, Paragraphs [0044]-[0086], [0110], [0111], Figs.1-6, 12	1-13
A	JP 2009-222568 A (KONICA MINOLTA SENSING, INC.) 2009.10.01, Whole document (No Family)	1-13
A	US 4969036 A (BHANU, Bir and BURGER, Wilhelm) 1990.11.06, Whole document & JP 3-48977 A, Whole document & EP 390051 A2	1-13

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

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
24.11.2017	05.12.2017

Name and mailing address of the ISA/JP	Authorized officer
Japan Patent Office	KUWA, Toshitaka — 5H 5094
3-4-3, Kasumigaseki, Chiyoda-ku, Tokyo 100-8915, Japan	Telephone No. +81-3-3581-1101 Ext. 3531

Claim 1 describes a method defined by equation (1), whereas the description only refers to the equation itself. In other words, the description refers to neither derivation process of the equation nor definition of coordinate systems (the place of the origin, the direction of x axis, y axis and z axis, the relationship between the three axes and the rotation axis, the relationship between the world coordinate system and the camera coordinate system etc.).

In the technical field of image processing, it is common general knowledge at the time of filing that it is difficult to derive an equation for three-dimensional reconstruction without detailed definition of coordinate systems.
In addition, it is difficult to understand equation (1) without disclosed approximation etc., if any approximation is used in the deriving process.

In light of the common general knowledge stated above, it is evident that a "method" defined only by the vague equation, without any description to derive it, is not sufficiently specified from a technical perspective. Therefore, the invention cannot be clearly identified from the statement of claim 1, even by taking into account the description and drawings.
Thus, claim 1 is not clear, as required by PCT Article 6.

This reason also applies, mutatis mutandis, to claims 5, 8 and 9 which describe methods defined by equations (5)-(8).

Consequently, search has been carried out on methods as below in following claims, since no meaningful search can be carried out.

(Claim 1)
A method ... point coordinates;
estimating the three-dimensional depth value of the acquired images by calculating each of the two-dimensional feature point coordinates;
implementing inverse ... distortion effects.

(Claim 5)
The method for estimating three-dimensional depth value from two-dimensional images according to claim 4.

(Claim 8)
The method for estimating three-dimensional depth value from two-dimensional images according to claim 1.

(Claim 9)
The method for estimating three-dimensional depth value from two-dimensional images according to claim 1.