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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
16 January 2014 (16.01.2014)

WIPO | PCT

(10) International Publication Number
WO 2014/011182 A1

(51) International Patent Classification:
G01B 11/22 (2006.01) *G06K 9/20* (2006.01)

(21) International Application Number:
PCT/US2012/046557

(22) International Filing Date:
12 July 2012 (12.07.2012)

(25) Filing Language: English

(26) Publication Language: English

(71) Applicant *(for all designated States except US)*: **CALFORNIA INSTITUTE OF TECHNOLOGY** [US/US]; 1200 East California Blvd., Pasadena, CA 91125 (US).

(72) Inventor; and

(75) Inventor/Applicant *(for US only)*: **GHARIB, Morteza** [US/US]; 1200 East California Blvd., Pasadena, CA 91125 (US).

(74) Agent: **STIRRAT, Mark, A.**; ONE, LLP, 4000 MacArthur Blvd., West Tower, Suite 1100, Newport Beach, CA 92660 (US).

(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, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

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

WO 2014/011182 A1

(54) Title: CONVERGENCE/DIVERGENCE BASED DEPTH DETERMINATION TECHNIQUES AND USES WITH DEFOCUSING IMAGING

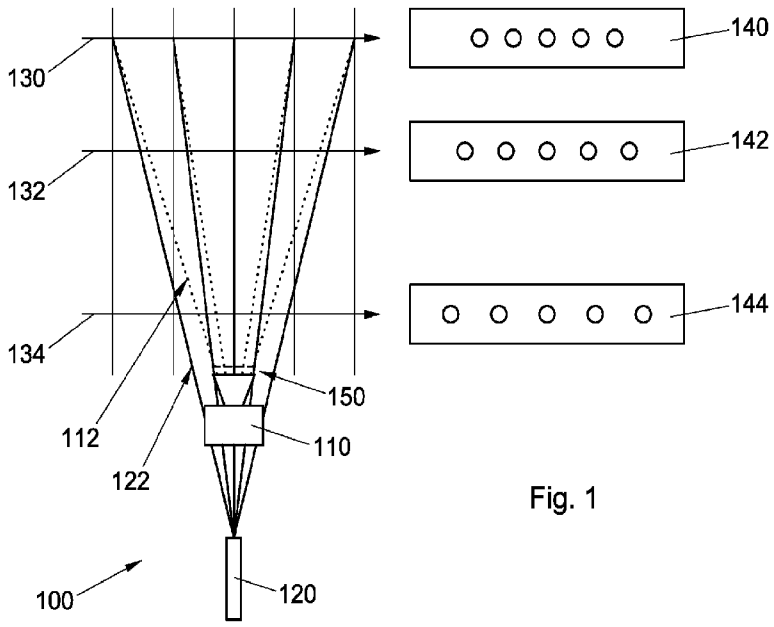


Fig. 1

(57) Abstract: Light is projected toward a surface and depth measurement at points determined by attributes of convergence or divergence of the light captured by a camera system. In defocusing applications, the optical system includes at least two off-axis apertures or camera portions, arranged to obtain an image of the projected light patterns including defocused information. The camera may be movable between different positions to image the surface from said different positions.

CONVERGENCE/DIVERGENCE BASED DEPTH DETERMINATION TECHNIQUES AND USES WITH DEFOCUSING IMAGING

BACKGROUND

[0001] Active stereo imaging is a 3D imaging method that operates by projecting a structured or random light pattern onto an object surface and capturing an image of that light pattern with a camera. The system is historically calibrated. The characteristics of the plane projection on the system are obtained for the known calibration sample. The light pattern is projected onto the surface from an angle relative to the camera such that the light appears to be at different locations based on the distance to the object or to a point on the object. Based on the calibration set, the position of the surface, and hence the contour of the object, can be determined. However, use of non-uniform patterns limit the resolution the system can achieve due to un-used space in the pattern.

[0002] A type of "defocusing" system described in US Publication No. 2009/0295908 employs a projected pattern of laser dots or "markers" for the purpose of generating high accuracy local 3D object data from the reflected light. With these, z-axis or depth image information is derived from the spacing observed between matched sets of points, each point in a given matched set (e.g., as arranged in a triangle) derived from a different aperture from a multi-aperture mask.

[0003] Knowing the position or relative "pose" of the camera, the 3D data determined at each of the different times or positions can be combined to stitch together multiple different scenes to complete a model for an object larger than each individual imaged area. The '908 application describes one manner of determining the 3D date, another is presented in PCT/US10/57532.

[0004] US Publication No. 2011/0074932 also projects a pattern of laser dots onto a surface to be imaged. In this '932 reference, the pattern is projected at an angular offset from the camera. Data recorded by the camera is used in two different modes. Information obtained through multiple offset apertures is used for defocusing. Information obtained through another aperture (e.g., a central aperture) is used for resolving depth (z) information from the deformation of the planar (x, y) coordinates of the pattern.

[0005] According to the '932 description, defocus-based imaging of the laser dots results in lower z axis accuracy than possible with the active stereo imaging due to a higher out-of-plane to in-plane error ratio generated from the off-axis apertures (as

compared to a central aperture). As such, the '932 reference uses the depth information from defocusing to identify the correspondence of dots between the deformed and original pattern for stereo imaging depth determination. This approach reportedly offers greatly increased working depth of the system and allows the active stereo imaging to be used in more applications including at lower angles and over greater working depths. Stated otherwise, the '932 system (via its processor) uses defocused information from the projected dots to determine a correspondence of the optical dots recorded on a surface to the optical dots of the projected pattern to determine and approximate z-axis position for in performing active stereo imaging.

[0006] More generally, the process of "defocus-based" imaging is one in which large data structures are created by light capture (with a CMOS, CCD or other imaging apparatus) through restricted areas positioned at different radial locations along a common optical axis. Corresponding x and y value points (or features) from a given imaged scene are matched and z-values calculated from the separation of the points (or features) such as by equations of the type described in US Patent Nos. 6,278,847 and 7,006,132.

[0007] One example of defocus-based imaging employs a technique called aperture-coded imaging. Suitable hardware for such purposes is described in PCT/US10/57532 as well as USPNs 6,278,847; 7,006,132; 7,612,869 and 7,612,870. This technique uses off-axis apertures to measure the depth and location of a scattering site in separated (color, time-wise, etc.) channels. The shifts in the images caused by these off-axis apertures are monitored to determine the three-dimensional position of the site or sites.

[0008] There are often tradeoffs in aperture-coded systems for obtaining defocused information of this type. The '908 application is intended to address some of these by separating the routines for feature-based pose determination (optionally obtained through color-coded apertures) and determination of high-resolution 3-D data with the laser dot projection. By doing so, it is possible to avoid marking the subject object with a dyed or sprayed-on contrast pattern. Additionally, the use of a regular projected pattern in the '908 reference enables extremely dense point sets (thus also addressing issues presented with the use of structured/random patterns as introduced above). The '932 reference instead utilized defocusing for gross (i.e., less accurate) depth determination of recorded x, y location dots in order to facilitate the arrival at image correspondence for active stereo determination of most accurate position. Moreover, multiple images can be combined or "stitched" together to fully image a 3D object.

[0009] Yet, there exists opportunity for other advantageous systems employing a projected pattern for 3-D imaging in connection with defocusing techniques.

SUMMARY

[0010] Accordingly, new approaches to projected grid use are desirable given the opportunity they provide for improved resolution z-axis determinations derived from x, y image data. The approaches described herein are suitable for use alone or advantageously in connection with defocusing imaging techniques. All of them employ different principles than those systems described above. As such, distinct advantages are described and may be further apparent to those with skill in the art upon review of the present description.

[0011] A first embodiment combines a diverging light projection with a defocusing camera system. As long as the origins of the two are at different distances, optionally along a common axis from the measurement volume, depth can be estimated by measuring the distance between points on the projection pattern as imaged through a single aperture. In such a setup, the distance between neighboring points on the light pattern will be linearly related to the distance from the camera. Stated otherwise, the apparent spacing between the imaged dots varies with depth, depending on the relative distance from their respective origins.

[0012] In contrast, if the origins are the same, the divergence of the projection will be canceled by the decreasing magnification of the camera system. If the projector is angled relative to the camera axis, simple measurement of distance between points will not suffice, but instead such offset can be compensated for in the initial calibration. Once estimated depth is obtained, it may then be refined for the points as imaged through a plurality of offset apertures by using defocusing to achieve final accuracy.

[0013] Granted, such a system may not offer the accuracy described in connection with the approach in the '932 publication application, which teaches away from that described here (on the basis of the noted out-of-plane to in-plane error ratio introduced by attempting to obtain final depth values from offset defocusing apertures/camera). However, the approach according to this embodiment of the invention that first approximates image depth using divergence, and then refines with defocusing, offers advantages in terms of hardware and computational simplicity. Namely, knowing – basically – *a priori* at which depth various imaged points should be located for matching offset defocusing images provides an advantage. With such an advantage, signal

separation (i.e., by color, time, polarization, pattern, etc.) with associated hardware need not be employed in some cases. Stated otherwise, image crowding problems (e.g., as commented upon in the '870 patent above) are reduced or eliminated depending on setup.

[0014] Another embodiment employs converging structured light projections that focus on one plane. In such an approach, one camera can be used to measure depth by comparing the distance between corresponding points. In a sense, the approach works as "reverse" defocusing (i.e., by using multiple light projections to produce doublets, triplets, etc., the positions of which are compared instead of using multiple apertures).

[0015] As in co-pending PCT Application No. PCT/2012/46484, filed July 12, 2012 and incorporated by reference herein in its entirety, the projections can be configured as a grid of varying intensity and/or varying color to help identify corresponding points. For example, a red and blue projection pattern - where the respective grids register at one plane - can be used to infer depth by looking at the distance between red and blue points. In any case, multiple projection systems are used that come from different origins, but focus on the same plane. The apparent position of the multiple dots relates to the distance from the focal plane.

[0016] In a simplified version of such a system reliant on geometric matching, an imaging device (lens), with three lasers placed at vertices of an equilateral triangle on the periphery of the lens, can measure exact distance between a particular point and a reference plane. The reference plane is chosen with the design as the plane at which the beams intersect, though if necessary it could be changed during use by retuning the wavelength of the laser (such that the focal distance through its lens will change slightly) or by mechanically moving the laser. Points on planes between the imaging device and the reference plane will form a pattern of an upright equilateral triangle (identical in orientation to the layout of the lasers). At a plane farther away than the reference plane, the pattern is inverted. Thus, the orientation identifies the point of interest as being behind or in front of the reference plane while the size of the pattern projected correlates directly to the distance from the plane in question to the reference plane.

[0017] In the three-laser embodiment, the three lasers are used so that it is known whether the imaged point is ahead of or behind the reference plane, and so that an X, Y coordinate relative to the center of the system can be established. An alternative is to use two light sources of different colors, or two on perpendicular planes (so that in the region closer than the reference plane the two dots make an "L" with the optical axis,

and in the region farther than the reference plane they make an upside-down "L"), or even just one laser in certain situations (in which case X, Y would not be known from the measurement so it would have to come from knowledge of the object and/or imaging system *a priori* via calibration or similar techniques). If multiple points in the XY plane are to be imaged simultaneously, then the three-source or multi-color arrangement is preferred, but for single point measurements any of these embodiments are acceptable.

[0018] In any case, such systems are adaptable to existing imagers such as cameras, photodetectors, microscopes, and others, and need not be at wavelengths in the visible range. The beams need not be separate lasers; they can be one beam split into the appropriate quantity of beams that the system requires via a diffraction grating, prism, thin-film filters, etc.

[0019] Yet another arrangement employs converging light focused at a reference plane in which depth measurement can be extrapolated from the intensity of the scattered light since such intensity will vary as the inverse of the area of the spot size. The advantage of this arrangement is that the intensity can be measured with a single photocell and thus acquisition can be very fast – no post-processing is needed, other than any desired calibration. Moreover, this arrangement has strong potential for miniaturization into all-inclusive microchip sensors to develop small, portable, hand-held devices. It also has the advantage in that a multitude of light sources can be used as a light cone that can easily be formed with simple lenses and fiber optics.

[0020] One embodiment employs a single light source projected as a ring onto a surface, and the receiving optics are tuned to be sensitive to a specific depth range by designing them with a telecentric field of view. Telecentricity is optional but has advantages as discussed below. As long as all of the light from the scattered image of the ring is received by the receptor, a depth measurement can be obtained. The intensity may be calibrated for the material being imaged simply by scanning X, Y for the reference plane, establishing a maximum value, and scanning X, Y out of range of the receptor where the received intensity should drop rapidly.

[0021] The light source is focused at the reference plane, thus it converges as it approaches it and then diverges as it propagates farther from it. With a telecentric field of view, the usable range of the sensor is the intersection of the field of view cylinder and the light source cone(s).

[0022] Capture is fast enough that detection of which region is being used (aft of or in front of the reference plane) may be confirmed by checking to see if the reference plane

is crossed at any point in time. Such a fast acquisition device may be used, for example, to measure the out-of-plane deformation of a material in a highly dynamic stress situation (such as impact) in a small region (point).

[0023] Another version of this concept discards all points measured outside of the reference plane. In such a system, detector components are be moved away from or towards the point in question until the reference plane is found. Thus, by recording the position of the device itself and knowing the distance between the device and the reference plane one can exactly ascertain the position of the point in question.

[0024] Such a "single-valued" embodiment is even simpler than the intensity-driven system in that the optical arrangement would constitute an "on-off" device (phototransistor) that may be used, for example, as part of already-present accurate measurement devices. This type of device has many applications, including but not limited to precise focusing of optical devices. By using a carefully tuned laser beam as a light source, the focal distance variation with wavelength can be measured. All these devices could be connected to a piezo-electric actuator for rapid scanning in depth (i.e., the Z direction).

[0025] If the position of device can also be known, then the Z-scanning of a point will increase the accuracy of the measurement. Rapid Z-scanning could also expand the depth range of the devices which is a factor of the design of the light source, and, in the case of the on-off device, could convert it into a full depth measuring device in which a surface would be scanned in Z by the piezo-actuator and the depths at which the reference plane coincides with the object are recorded. The piezo device could be replaced with a tunable light source, as by changing the wavelength of the light the focal distance through its optics will change in a predictable or measurable manner.

[0026] Moreover, these single-point measurement devices can be extended into array form where multiple points in the X, Y plane can be imaged simultaneously by multiplying the number of devices (as in the case of the light-intensity based unit) or by multiplying the apparent number of light sources (in the tri-laser or the multi-color/intensity, etc. arrangements already described).

[0027] Multiplying the apparent number of light sources can be done by using beam splitters and/or holographic lenses or diffraction gratings so that a single beam can be split into multiple beams and thus be emitted from several different points. The same holds true for generating the light patterns in any of the embodiments herein.

[0028] Likewise, in certain embodiments it should be understood that the illumination and the sensor need not be on the same side of the surface, for example, if measuring on a semi-opaque film. It should also be clear from the embodiments where the depth is extracted from the position of projected dots that the detector need not be concentric with the central axis of the illumination, as any offset can be compensated for in the initial calibration. Further, the approaches described may be used alone or in combination with offset multi-aperture or multi-camera defocusing techniques (as variously stated or otherwise possible as may be appreciated by those with skill in the art). Aspects of the invention include the subject hardware, software, and related methods of operation and also methods of manufacture.

BRIEF DESCRIPTION OF THE DRAWINGS

[0029] The figures provided are diagramatic and not drawn to scale. Variations of the invention from the embodiments pictured are contemplated. Accordingly, depiction of aspects and elements of the inventions in the figures are not intended to limit the scope of the inventions.

[0030] Fig. 1 illustrates a light beam divergence-based imaging system; Fig. 2 illustrates an imaging system employing converging light beams; Fig. 3 illustrates a system employing light beams in a pattern converging and diverging across a focal plane; Fig. 4 illustrates a system employing “ring” or “flood” lighting converging and diverging across a focal plane; and Figs. 5A and 5B are flowcharts illustrating methods of system operation.

DETAILED DESCRIPTION

[0031] Various exemplary embodiments of the inventions are described below. Reference is made to these examples in a non-limiting sense. They are provided to illustrate more broadly applicable aspects of the present inventions. Various changes may be made to the inventions described and equivalents may be substituted without departing from the true spirit and scope of the inventions. In addition, many modifications may be made to adapt a particular situation, material, composition of matter, process, process act(s) or step(s) to the objective(s), spirit or scope of the present inventions. All such modifications are intended to be within the scope of the claims made herein.

[0032] Turning to FIG. 1, the divergence-based depth determination approach is described. System 100 combines the concepts of a diverging structured light projection with a defocusing camera system. Along with a computer processor and such other

connective hardware or housing features (not shown), such a system includes a camera 110 and a projector 120. The camera view is taken as lines 112. The projector provides a structured light pattern illustrated with lines 122 incident upon an object. At an object plane 130, the apparent spacing between the dots as imaged by the camera is illustrated as image capture 140; at object position 132, the apparent spacing between the dots is illustrated as image capture 142; and at object position 134, the apparent spacing between the dots is illustrated as image capture 144.

[0033] As shown, the apparent spacing between the imaged dots varies with depth, depending on the relative distance from their respective origins. In connection with a sensor in a camera and processing by a computer, a complex mapping of an object and/or object image tracking can be performed in real time. By "real time," what is meant – in connection with this system and others – is that a user experiences no appreciable or apparent delay as a computer system provides display or responsiveness to action.

[0034] As long as the origins of the camera and projector are offset from one another by different distances from the measurement volume (as shown – aligned along the camera axis, but not necessarily so), depth can be estimated by measuring the distance between points on the projected light pattern. The distance between neighboring points on the light pattern will be linearly related to the distance from the camera in the setup shown. The estimated depth is then advantageously refined by using defocusing to achieve final position accuracy.

[0035] When followed by calculations for defocusing with images taken through at least one additional (offset) aperture (optionally provided in connection with an aperture mask 150), the divergence-based calculations regarding point depth (z-axis) can be very useful. The divergence-based calculations are very simple and thus fast. As a "starting point" they offer advantage in terms of completing the more complex defocusing refinement calculations while still maintaining activity in real time (for display purposes, active imagine, etc.) even with a relatively simple or low-powered processor. Once the first set of calculations (i.e., the divergence based calculations) have run, data is provided that can prove useful for point match-up between images captured with an offset for defocusing. Thus, such a system can operate very effectively, even at higher point densities and/or without coding the apertures.

[0036] Notably, these and other embodiments may use a laser, an LED or any other kind of other light generator for the light source. The light patterns may be provided by multiplying these entities or using beam splitters and/or holographic lenses or diffraction

gratings (i.e., as in a diffractive optical element (DOE)) as mentioned above. Also, while shown with projected patterns and described essentially with respect to a 2-D format with a line of points captured in each of Figs. 1 and 2, it is to be appreciated that a projected "grid" may be used in different embodiments that extends (in reference to the noted figures) into and out of the page, such that the captured images and associated data structures that are processed measure between parts vertically (as well as horizontally as shown and described in connection with the image capture screens 140, 142, etc.). Alternatively, if the camera system is of low enough frame rate, one can also devise a system using a scanning laser, similar to CRT's, so that successive line-by-line capture and comparison is made. As yet another option, comparison across the captured lines in a scan-type system may be employed. In any case, the examples are provided in a non-limiting sense as should be understood with respect to the other inventive variations described.

[0037] Fig. 2 illustrates a system 200 in which a converging structured light projection 202 from two different projection sources 202, 204 is focused on a plane 210. One camera 220 (optionally with a single aperture) can then measure depth by looking at the distance between the points, in the same way as defocusing with matched pairs (triplets, etc.) in the projection pattern using multiple light projections instead of multiple apertures. The projection can have a grid of varying intensity or varying color to help identify corresponding points – for example as described in above-referenced co-pending PCT Application No. PCT/US12/46484, filed July 12, 2012 and incorporated by reference in its entirety.

[0038] In one example, red and blue projections (as shown), may be provided where the respective grids register at one plane 210. Then depth can be calculated in connection with images captured for an object (e.g., by pixel-by-pixel comparison and data transformation) as related to the distance measured between red and blue points in different depth (z-axis) planes 212, 214. While such activity will typically be populated across an entire x,y grid (optionally as noted above) a linear representation of the corresponding image frames 222, 224 captured by the camera are shown. In this case, the relative position of the multiple dots relates to the distance from the focal plane as determined by a computer running appropriate software or by an application specific integrated circuit (ASIC) or chip set in reference to an image calibration set. Profilometry of an entire surface, or tracking the changing shape of a surface or body, can be accomplished in this fashion.

[0039] Another option is to employ such a projection system 200 in connection with a multiple-aperture camera employing an aperture mask 230 (or independent camera system aligned to achieve the same basic optical result). In this case, a first depth measurement can be made employing the offset beams (optionally through a central camera or aperture – though only two suitably-used offset apertures are shown in the figure), then subsequent refinement performed comparing the offset of image points within each color. For this purpose, only the red points can be used or only the blue points can be used. Under such circumstances, one or more of the offset aperture(s) may include a color filter. Otherwise, a Bayer or other filter associated with the camera CMOS or CCD can be employed to eliminate the color channel not used for defocusing purposes.

[0040] Regardless, as with the system in Fig. 1, defocusing imaging and calculations can be performed very efficiently, with respect to matching, etc., by virtue of the information provided by the beam-dot based depth calculations. Also, it is to be appreciated that the activity shown in front of focal plane 210 in Fig. 2 can occur behind the same. Fig. 3 illustrates a related system operating as such.

[0041] In Fig. 3, system 300 uses multiple angles of illumination in conjunction with an imaging device to extrapolate the third dimension of measurement in connection with a computer processor. In this case, an imaging device 302 (including optics and a sensor, etc.), with three lasers 310, 312, 314 placed at vertices of an equilateral triangle on the periphery of a lens, is used to extract distance between one or more points and a reference plane 320. The reference plane is chosen in the design as the plane at which the beams (or multiple projected beams) meet at a focal point 340 along an optical axis 350, although the location of this plane can be changed during use by retuning the wavelength of the laser(s) - such that the focal distance through any associated lens, grating, etc. will change slightly - or mechanically by moving the laser(s).

[0042] In the simplified version of the system 300 shown, points on a plane 322 between the imaging device and the reference plane will form a pattern of an upright equilateral triangle 330 (identical in orientation to the layout of the lasers). At a plane 324 farther away than the reference plane, the pattern 330' is inverted. Thus, the orientation identifies the point of interest as being behind or in front of the reference plane while the size of the pattern projected correlates directly to the distance from the plane in question to the reference plane. By analyzing images captured by the camera system and determining the orientation of the triplets (or relative position of doublets, etc. as noted

option in connection with this embodiment of the invention above) and distance between the imaged points forming the triangles, depth (z-axis) position is determined quickly using an associated computer processor.

[0043] As above, the referenced approach can be multiplied/multiplexed and/or used in conjunction with defocusing techniques, especially as a preceding/precedent matter, to subsequently simplify defocus imaging feature (be they dot, point, SIFT or SURF resolved, etc. features) match-up for subsequent calculations.

[0044] Fig. 4 illustrates a related imaging or depth-finding system 400 in which a camera or simplified photo-intensity detector 402 is associated with a light 404 and an associated optics focusing beam 406 so it converges to a focal point 410 at a reference plane 412 and diverges beyond that point. With telecentric optics, a field of view 420 for the sensor is (at least) substantially cylindrical. The depth of the field 422 which may be imaged is then the intersection of the bi-cylindrical beam and cylinder 420.

[0045] With telecentric optics, the vision system's image magnification is (at least substantially) independent of the distance or position in the field of view of the object reflecting light. As such, a measure of depth can be obtained simply by measuring reflected light intensity where the intensity of the light is expected to vary in inverse-square relation to distance from the source (i.e., to follow "Lambert's Law" as qualified/compared against calibration results for such an object or surface). Without telecentric optics, magnification-dependent intensity differences can be accounted for with system calibration. In any case, operation in a zone at our about (i.e., scanning around to identify as discussed as an option above) the converging/diverging light focal point can be limited.

[0046] Moreover, such a system 400 with defocusing hardware and software control as described above can be incorporated such that by using captured images from multiple offset apertures (e.g., as by using an optional aperture plate 430 with offset apertures 432 for collecting defocusing information and a central 434 aperture used for the intensity-based depth determination), distances between features in the captured image identified on an object (e.g., as applied as laser dots, by way of contrast medium, etc.) can be resolved for more accurate depth measurement. Of course, various "active shutter" arrangements could instead be used. In one example, an LCD rapidly switches between collecting intensity data in a fully "open" configuration and defocusing data by "blacking out" all but at least two offset windows for additional image capture. In any case, the intensity-based depth measurement can be used to inform and speed, and/or

improve the accuracy of the defocusing calculations – especially in connection with feature matching.

[0047] Still further, it is contemplated that the system in Fig. 4 may be practiced in connection with a projection grid of laser (or otherwise provided) beam dots converging/diverging in such a pattern. In that case changes in intensity for the purpose of calculating depth for a given plan are registered as added or "dropped" points/dots from the field of view. As such, simply counting the number of such features recorded in an image can offer an approximation of depth.

[0048] Figs. 5A and 5B are flowcharts illustrating operation possibilities for the subject imaging systems. The systems above may operate according to these methods or the methods described can implicate alternative hardware options.

[0049] In the example in Fig. 5A, at 500, the system obtains image information as reflected light from a target object or environment via electronic imaging components – optionally, in connection with any of systems 100, 200, 300 or 400. Then at 510, a computer processor utilizes the image data captured as associated with the selected mode of converging/diverging light projection to make a determination of depth or distance of the point or points. The sub-process outputs an initial depth determination 520 for the feature(s) of interest. With this initial depth determination and imaging data already obtained at 500, or with image data acquired specifically for defocusing at 530 (such as with a system 400 collecting different types of image data through its various apertures), a more refined depth determination 540 is made based on defocusing principles. The refined depth determination results 550 of the defocusing sub-process may be accomplished as described in connection with the teachings of any of the above-referenced patents or applications, as in employing defocusing equations and/or in reference to a calibration set generated for the system optics (especially as further described in PCT/US10/57532).

[0050] In another example illustrated in Fig. 5B, also usable with any of the systems described herein, at 560 the system obtains the various image data to process. At 570 depth determinations are made from the same based on the converging/diverging projections. At 580, potential point matches are identified for defocusing. At 590 when a potential defocusing point match yields multiple results, the projection based depth determination is used as an estimate of expected neighbor point position depth to select which one is the most likely defocusing point pair match. So-matched, a defocusing-based depth determination is output at 600 in reference to a calibration set and/or

defocusing equations per the references cited above.

[0051] In any case, feature matching for defocusing is aided by the information from the initial depth determination. Possession of such information *a priori* limits the range where expected matching points should be found within the possibilities for defocusing processing – thus reducing the computational intensity of the problem to be solved. In other words, with a given range or set of possible locations of z-position for an identified point in x, y space from one aperture/channel, its match from another aperture/channel is more easily identified. Once so-matched, extremely accurate defocusing-based imaging proceeds apace with reduced possibilities or even the elimination of inaccurate/inappropriate "matches."

[0052] In certain examples, the process stops at the initial depth determination in Fig. 5A and feeds into another process. In other examples, any or all of the depth determination for a given imaged "scene" may be aggregated with other related image scene data taken for an object larger that the imager field of view. In that case, the teachings for approaches to calculating camera pose, transforming and aggregating image data as presented in PCT/US10/57532 may be applied or others as may be apparent to those with skill in the art.

[0053] Although several embodiments have been disclosed in detail above, other embodiments are possible and the inventors intend these to be encompassed within this specification. The specification describes specific examples to accomplish a more general goal that may be accomplished in another way. This disclosure is intended to be exemplary, and the claims are intended to cover any modification or alternative which might be predictable to a person having ordinary skill in the art. For example, other forms of processing can be used. Any camera type can be used, including a CCD camera, active pixel, or any other type. Also, other shapes of apertures can be used, including round, oval, triangular, and/or elongated. The above devices can be used with color filters for coding different apertures, but can also be used with polarization or other coding schemes.

[0054] The cameras described herein can be handheld portable units, or machine vision cameras, or underwater units. Or the camera may be mounted in a stationary position an object moved relative to them or otherwise configured. Still further, the camera may be worn by a user to record facial expressions or gestures to be blended with animation. Other possibilities exist as well.

[0055] Those of skill would further appreciate that the various illustrative logical blocks, modules, circuits, and algorithm steps described in connection with the embodiments disclosed herein may be implemented as electronic hardware, computer software, or combinations of both. Indeed, given the type of pixel-to-pixel matching and associated calculations required with the data structures recorded and manipulated, computer use is necessary. In imaging any object, vast sets of data are collected and stored in a data structure requiring significant manipulation in accordance with imaging principles – including defocusing principles/equations – as noted herein and as incorporated by reference. Likewise, the response time of computer control is required for effective operation of the scanning and "single-point" embodiments discussed above.

[0056] To clearly illustrate this interchangeability of hardware and software, various illustrative components, blocks, modules, circuits, and steps have been described above generally in terms of their functionality. Whether such functionality is implemented as hardware or software depends upon the particular application and design constraints imposed on the overall system. Skilled artisans may implement the described functionality in varying ways for each particular application, but such implementation decisions should not be interpreted as causing a departure from the scope of the exemplary embodiments of the invention.

[0057] The various illustrative logical blocks, modules, and circuits described in connection with the embodiments disclosed herein, may be implemented or performed with a general purpose processor, a Digital Signal Processor (DSP), an Application Specific Integrated Circuit (ASIC), a Field Programmable Gate Array (FPGA) or other programmable logic device, discrete gate or transistor logic, discrete hardware components, or any combination thereof designed to perform the functions described herein. A general purpose processor may be a microprocessor, but in the alternative, the processor may be any conventional processor, controller, microcontroller, or state machine. The processor can be part of a computer system that also has a user interface port that communicates with a user interface, and which receives commands entered by a user, has at least one memory (e.g., hard drive or other comparable storage, and random access memory) that stores electronic information including a program that operates under control of the processor and with communication via the user interface port, and a video output that produces its output via any kind of video output format, e.g., VGA, DVI, HDMI, displayport, or any other form.

[0058] A processor may also be implemented as a combination of computing devices, e.g., a combination of a DSP and a microprocessor, a plurality of microprocessors, one or more microprocessors in conjunction with a DSP core, or any other such configuration. These devices may also be used to select values for devices as described herein.

[0059] The steps of a method or algorithm described in connection with the embodiments disclosed herein may be embodied directly in hardware, in a software module executed by a processor, or in a combination of the two. A software module may reside in Random Access Memory (RAM), flash memory, Read Only Memory (ROM), Electrically Programmable ROM (EPROM), Electrically Erasable Programmable ROM (EEPROM), registers, hard disk, a removable disk, a CD-ROM, or any other form of storage medium known in the art. An exemplary storage medium is coupled to the processor such that the processor can read information from, and write information to, the storage medium. In the alternative, the storage medium may be integral to the processor. The processor and the storage medium may reside in an ASIC. The ASIC may reside in a user terminal. In the alternative, the processor and the storage medium may reside as discrete components in a user terminal.

[0060] In one or more exemplary embodiments, the functions described may be implemented in hardware, software, firmware, or any combination thereof. If implemented in software, the functions may be stored on or transmitted over as one or more instructions or code on a computer-readable medium. Computer-readable media includes both computer storage media and communication media including any medium that facilitates transfer of a computer program from one place to another. A storage media may be any available media that can be accessed by a computer. By way of example, and not limitation, such computer-readable media can comprise RAM, ROM, EEPROM, CD-ROM or other optical disk storage, magnetic disk storage or other magnetic storage devices, or any other medium that can be used to carry or store desired program code in the form of instructions or data structures and that can be accessed by a computer. The memory storage can also be rotating magnetic hard disk drives, optical disk drives, or flash memory based storage drives or other such solid state, magnetic, or optical storage devices. Also, any connection is properly termed a computer-readable medium. For example, if the software is transmitted from a website, server, or other remote source using a coaxial cable, fiber optic cable, twisted pair, digital subscriber line (DSL), or wireless technologies such as infrared, radio, and

microwave, then the coaxial cable, fiber optic cable, twisted pair, DSL, or wireless technologies such as infrared, radio, and microwave are included in the definition of medium. Disk and disc, as used herein, includes compact disc (CD), laser disc, optical disc, digital versatile disc (DVD), floppy disk and blu-ray disc where disks usually reproduce data magnetically, while discs reproduce data optically with lasers. Combinations of the above should also be included within the scope of computer-readable media. The computer readable media can be an article comprising a machine-readable non-transitory tangible medium embodying information indicative of instructions that when performed by one or more machines result in computer implemented operations comprising the actions described throughout this specification.

[0061] Operations as described herein can be carried out on or over a website. The website can be operated on a server computer, or operated locally, e.g., by being downloaded to the client computer, or operated via a server farm. The website can be accessed over a mobile phone or a PDA, or on any other client. The website can use HTML code in any form, e.g., MHTML, or XML, and via any form such as cascading style sheets ("CSS") or other.

[0062] Also, the inventors intend that only those claims which use the words "means for" are intended to be interpreted under 35 USC 112, sixth paragraph. Moreover, no limitations from the specification are intended to be read into any claims, unless those limitations are expressly included in the claims. The computers described herein may be any kind of computer, either general purpose, or some specific purpose computer such as a workstation. The programs may be written in C, or Java, Brew or any other programming language. The programs may be resident on a storage medium, e.g., magnetic or optical, e.g. the computer hard drive, a removable disk or media such as a memory stick or SD media, or other removable medium. The programs may also be run over a network, for example, with a server or other machine sending signals to the local machine, which allows the local machine to carry out the operations described herein.

[0063] Where a specific numerical value is mentioned herein, it should be considered that the value may be increased or decreased by 20%, while still staying within the teachings of the present application, unless some different range is specifically mentioned. Where a specified logical sense is used, the opposite logical sense is also intended to be encompassed.

[0064] The previous description of the disclosed exemplary embodiments is provided to enable any person skilled in the art to make or use the present inventions. Various

modifications to these exemplary embodiments will be readily apparent to those skilled in the art, and the generic principles defined herein may be applied to other embodiments without departing from the spirit or scope of the inventions. Thus, the present inventions are not intended to be limited to the embodiments shown herein but is to be accorded the widest scope consistent with the principles and novel features disclosed herein.

CLAIMS

1. An imaging method for depth determination of an object, the method comprising:

projecting light in a converging or diverging manner,

capturing at least one image of reflected light from an object with a camera to provide image data,

processing the image data to make an initial depth determination; and

processing the image data using the initial depth determination to make a refined depth determination based on matched image features from the data captured from offset positions.

2. The method of claim 1, wherein the light is projected in a diverging pattern of beams.

3. The method of claim 2, wherein the pattern is a grid pattern.

4. The method of claim 1, wherein the light is projected in a converging manner.

5. The method of claim 4, wherein the light is provided by a ring light.

6. The method of claim 4, wherein the capturing is performed with a telecentric optical setup.

7. The method of claim 4, wherein the light is provided by a plurality of sources with the camera set between the sources.

8. The method of claim 7, wherein the light is projected in points forming a triangular pattern.

9. The method of claim 8, wherein the light is not differentiated between the sources.

10. The method of claim 7, wherein the light is differentiated by color between the sources.

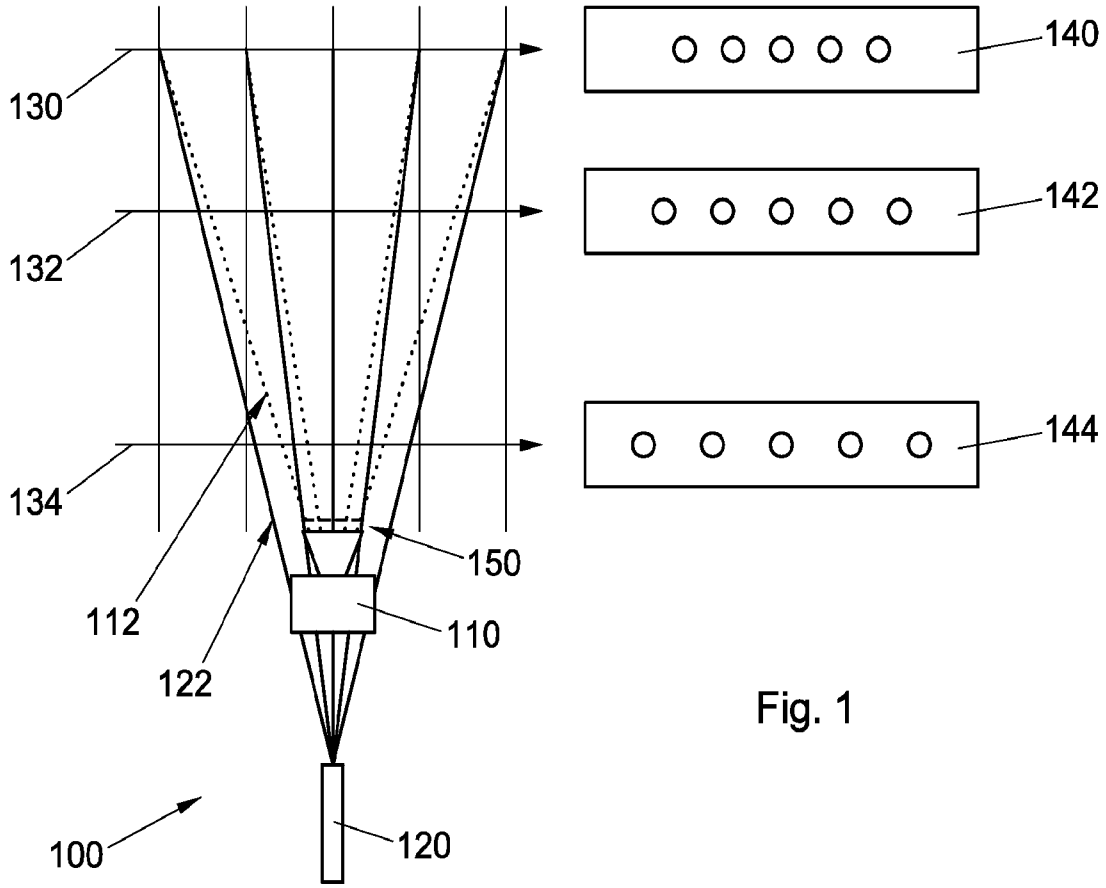

130
132
134
140
142
144
150
112
110
122
100
120

Fig. 1

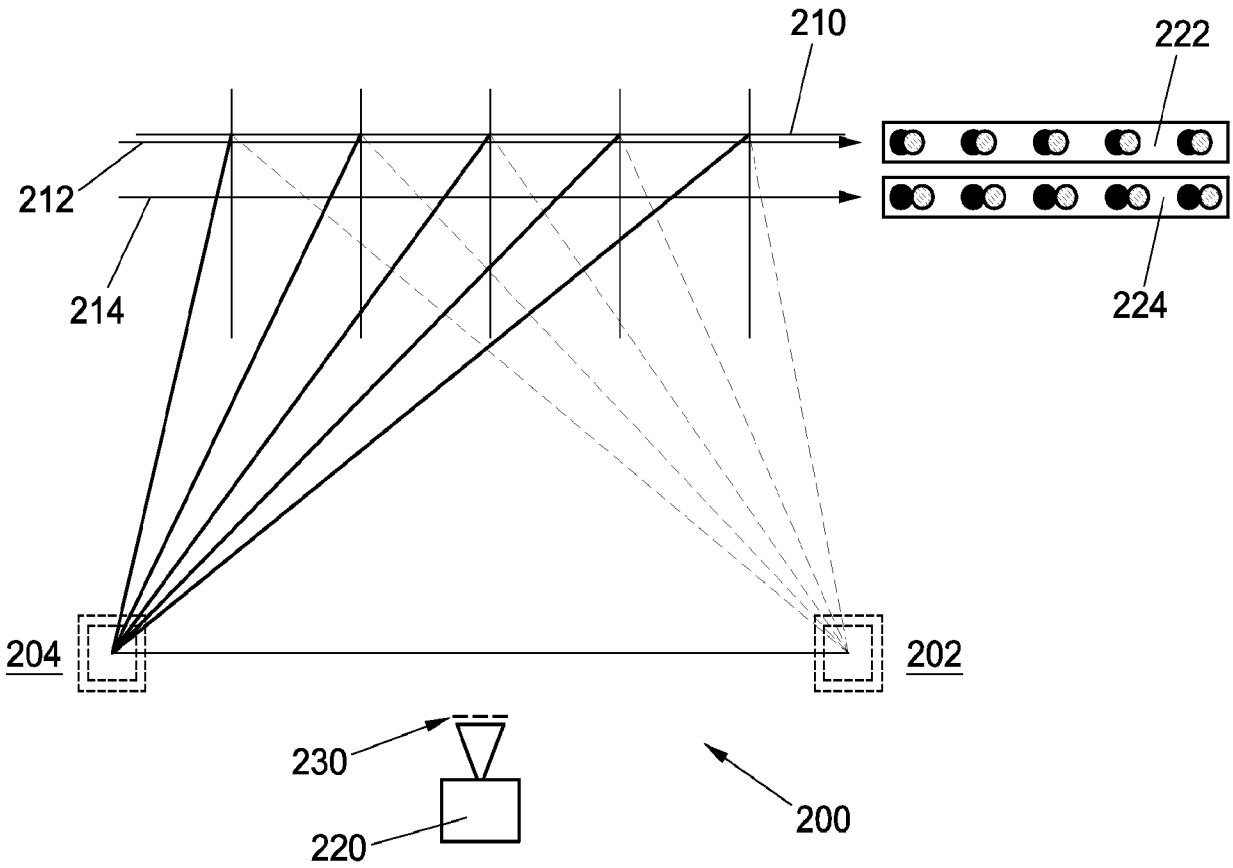

210
222
212
214
224
204
202
230
220
200

Fig. 2

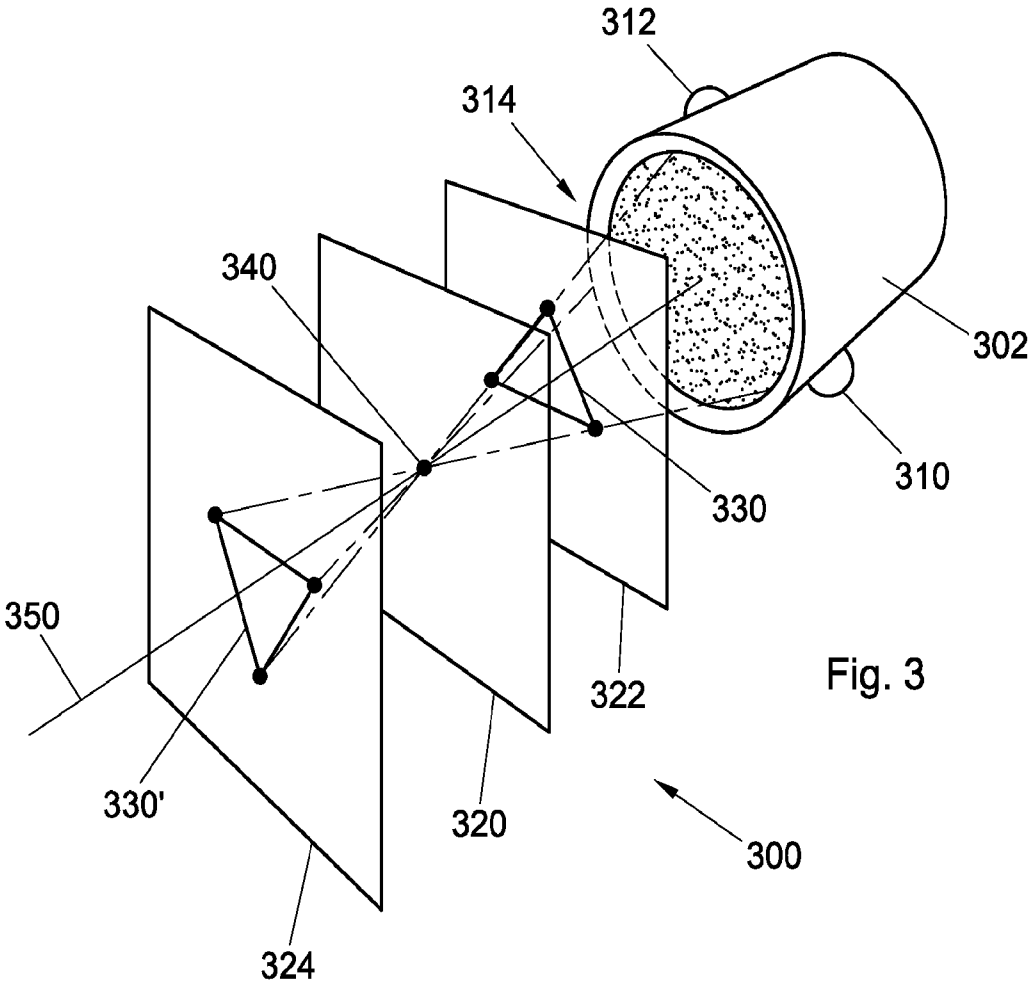

312
314
340
302
310
330
350
322
330'
320
300
324

Fig. 3

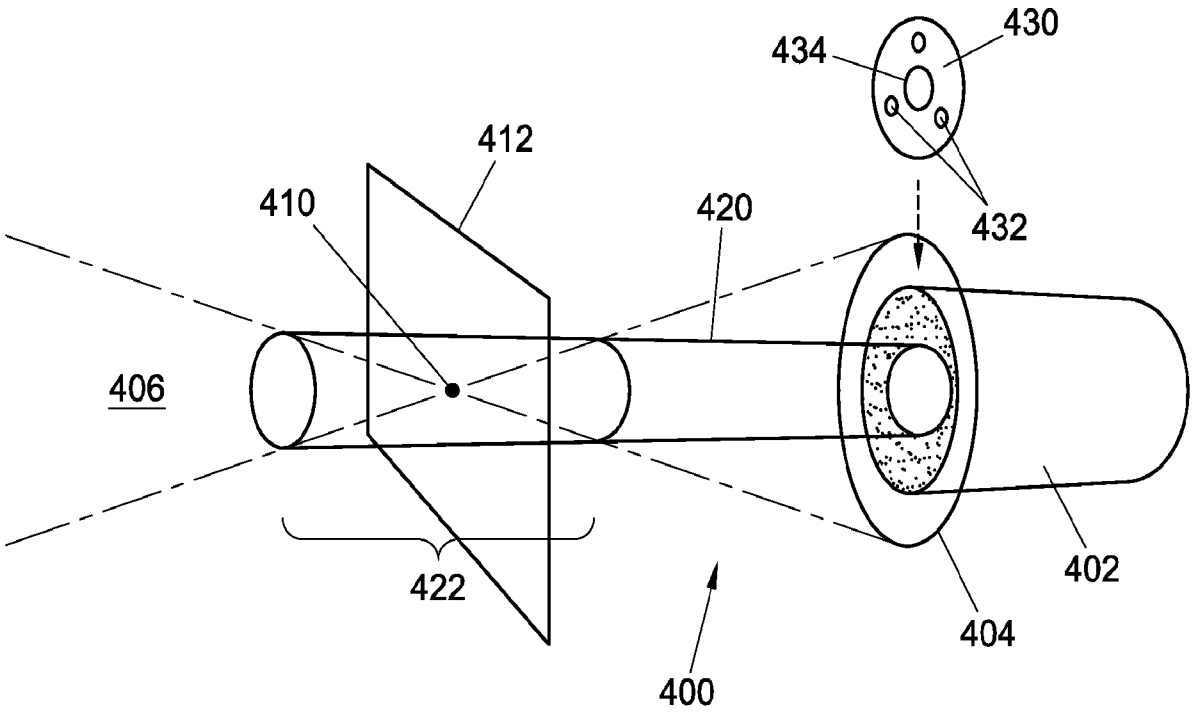
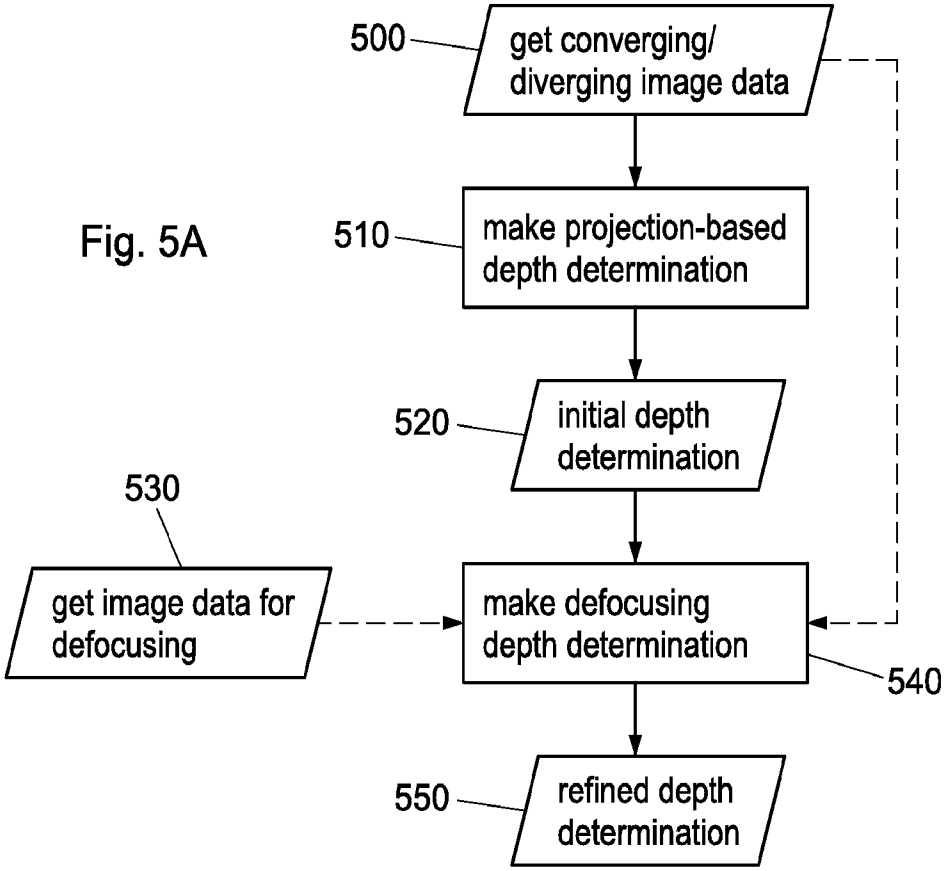

430
434
412
410
420
432
406
422
402
404
400

Fig. 4

Fig. 5A


500
get converging/ diverging image data
510
make projection-based depth determination
520
initial depth determination
530
get image data for defocusing
make defocusing depth determination
540
550
refined depth determination

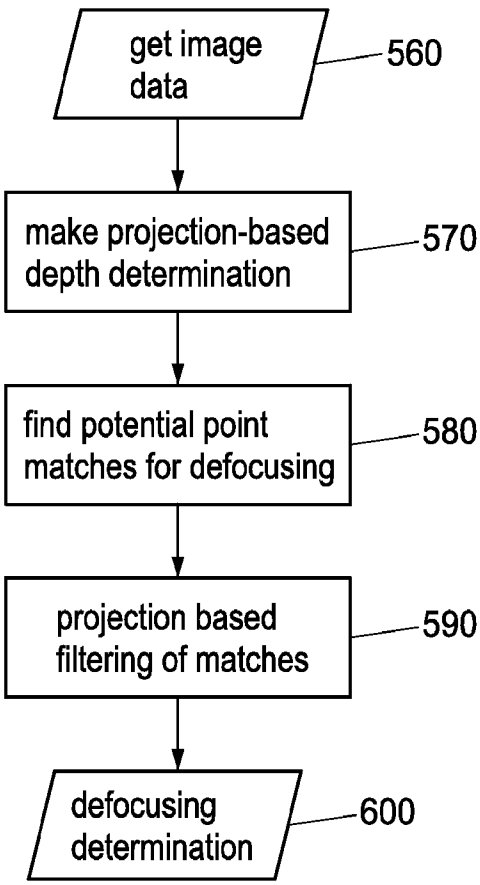

get image data
560
make projection-based depth determination
570
find potential point matches for defocusing
580
projection based filtering of matches
590
defocusing determination
600

Fig. 5B

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US12/46557

A. CLASSIFICATION OF SUBJECT MATTER
IPC(8) - G01B 11/22; G06K 9/20 (2013.01)
USPC - 356/603; 348/E13.057
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC(8) - G01B 11/22; G06K 9/20 (2013.01)
USPC - 356/601, 603, 3.1; 382/154, 108; 348/E13.057

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)
MicroPatent (US Granted, US Applications, EP-A, EP-B, WO, JP, DE-G, DE-A, DE-T, DE-U, GB-A, FR-A); Google Patents; Google Scholar; IP.COM, IEEE Xplore; Search Terms Used: Depth, deep, height, topography, light, patterns, structure, triang*, camera, imaging sensor, telecentric, image features, initial, refining, accurate, wavelength, color

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 2009/0295908 A1 (GHARIB, M et al.) December 03, 2009, Figure 3, Paragraphs [0041] [0051].	1-10
Y	US 2009/0129667 A1 (HO, Y et al.) May 21, 2009, Figure 1, Paragraphs [0064] and [0065].	1-10
Y	US 2003/0160970 A1 (BASU, A et al.) August 28, 2003, Figure 2, Paragraphs [0040], [0075].	4-10
Y	US 6,229,913 B1 (NAYAR, S et al.) May 08, 2001, column 1, Lines 21-26.	6
Y	US 4,879,664 A (SUYAMA, T et al.) November 07, 1989, Abstract, Figure 1, Column 2, Lines 5-13.	8, 9

☐ Further documents are listed in the continuation of Box C. ☐

* 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
30 August 2012 (30.08.2012)	02 OCT 2012
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-3201	Authorized officer: Shane Thomas PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

Form PCT/ISA/210 (second sheet) (July 2009)