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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
19 June 2014 (19.06.2014)

(10) International Publication Number
WO 2014/091307 A2

(51) International Patent Classification: *G06T 7/00* (2006.01)

(21) International Application Number: PCT/IB2013/003077

(22) International Filing Date: 27 November 2013 (27.11.2013)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

10 2012 112 322.5	14 December 2012 (14.12.2012)	DE
61/740,681	21 December 2012 (21.12.2012)	US
13/767,175	14 February 2013 (14.02.2013)	US

(71) Applicant: **FARO TECHNOLOGIES, INC.** [US/US]; 125 Technologies, Inc., Lake Mary, FL 32746-6204 (US).

(72) Inventors: **DEBITSCH, Rasmus**; Unterländer Str. 48, 70435 Stuttgart (DE). **HILLEBRAND, Gerrit**; Marienstr. 1, 71332 Waiblingen (DE). **HEIDEMANN, Rolf**; Nibelungenstr. 17, 70191 Stuttgart (DE). **OSSIG, Martin**; Wiesenstr. 23, 71732 Tamm (DE).

(74) Agent: **CHRISTENSEN, Dave, S.**; Cantor Colburn LLP, 20 Church St., 22nd Floor, Hartford, CT 06103-3207 (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, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK,

[Continued on next page]

(54) Title: METHOD FOR OPTICALLY SCANNING AND MEASURING AN ENVIRONMENT

(57) Abstract: With a method for optically scanning and measuring an environment by means of a hand-held scanner (100) for producing 3D-scans, wherein the hand-held scanner (100) produces, by means of at least one projector (121, 122), at least one pattern (X, X') onto an object (O) in the environment and records, by means of at least one camera (111, 112, 113), images of the object (O) which is provided with the pattern (X, X'), with a plurality of frames (F), for finding a ring closure in the plurality of frames (F), a frustum (V) is formed for each frame (F) in a first step, and, in a second step the latest frustum (V_n) which is assigned to the latest frame (F_n), is compared with a plurality of the last frusta (V), by means of forming the intersection, wherein the frustum (V_j), with which there is the largest intersection, is selected.

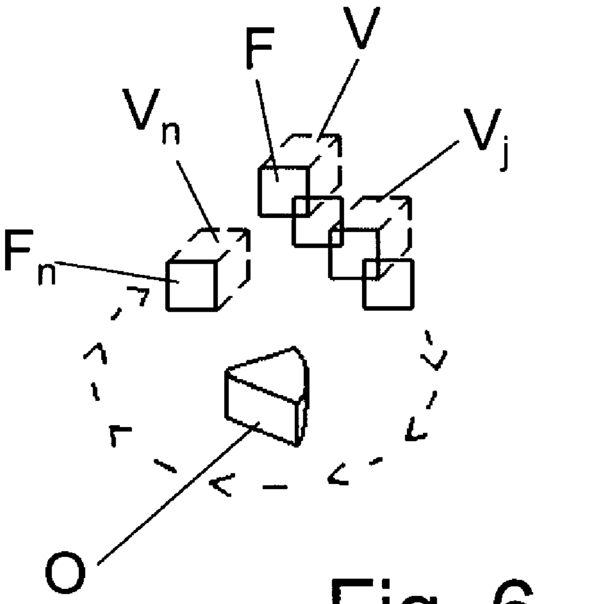

Fig. 6

WO 2014/091307 A2

EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

FARO Technologies, Inc., Lake Mary, FL, USA

Method for optically scanning and measuring an environment

The invention relates to a method having the features of the generic term of Claim 1.

A device of this kind is known from US 6,826,299 B2. A projector projects light patterns on the surface of an object to be scanned. The position of the projector is determined by means of a projected, encoded pattern. Two (or more) cameras, the relative positions and alignment of which are known or are determined, can record images of the surface with a further, uncoded pattern. The three-dimensional co-ordinates (of the points of the pattern) can be determined by means of mathematical methods which are known per se, such as epipolar geometry.

From the games sector, scanners are known as tracking devices, in which a project-or projects an encoded light pattern onto the target to be pursued, preferably the user who is playing, in order to then record this encoded light pattern with a camera and to determine the coordinates of the user.

US 8,238,611 B2 discloses a system for scanning a scene, including distance meas-uring. The system, in its most simplest form, comprises a camera unit with two cameras, optionally with filters, for the stereoscopic registration of a target area, an illumination unit for generating a pattern in the target area, preferably by means of a diffractive optical element, and a synchronizing unit, which synchronizes illumina-tion unit and camera unit. Camera unit and illumination unit can be set up in select-able relative positions. Optionally, also two camera units or two illumination units can be used.

The invention is based on the object of improving a method of the type mentioned in the introduction. This object is achieved according to the invention by means of a method having the features of Claim 1.The dependent claims relate to advantageous configurations.

Compared to an encoded pattern, an uncoded pattern can be produced more easily, for example as a regular pattern of light points. In order to then record images of the object which is provided with the pattern and to get unambiguous correspondences of the light points of the pattern, two (or more) cameras are used. The two cameras and the projector are arranged in a way that is not co-linear, but in a triangle arrangement. It is thus possible to use three epipolar-geometry-relations, in order to determine the correspondence between the patterns in the camera images. When these correspondences are known, the three-dimensional coordinates of the point cloud, i.e. the 3D-scan, can be determined.

Preferably, the pattern is not produced within the visible wavelength range, but within the infrared range (700 nanometers - 1 millimeter). The two cameras have a corresponding sensitiveness in this wavelength range, while scattered light and other interferences can be filtered out in the visible wavelength range. A color camera can be provided as third camera for color information, such camera recording images of the object to be scanned, too. The 3D-scan can be colored with the color information thus obtained.

The hand-held scanner produces a plurality of 3D-scans of the same scene, from different positions. Registration of the different 3D-scans in a common coordinate system is facilitated by a stationary pattern, which can be captured by different 3D-scans. The stationary pattern rests with respect to the object, when the hand-held scanner is moved and takes the different positions. The natural texture of the surface of the object and other structures, such as edges, can be used as stationary pattern, such texture being captured by means of a color camera as third camera, or a projected pattern, which is produced by a separate (external) projector, is used (additionally or alternatively). This stationary pattern can be distinguishable in terms of geometry, time or spectrum from the pattern produced by the hand-held scanner.

A modular design with three (or more) cameras and a plurality of projectors, by means of which requirements which depend on the application are fulfilled by projecting and recording images of patterns having different point densities and lateral resolution, is conceivable.

The production of the pattern can be realized by means of deflecting methods, such as production by means of diffractive optical elements or micro-lenses (or single lasers), or by shading methods, for example the production by means of shutters, transparencies (as they would be used in a transparency projector) and other masks. The deflecting methods have the advantage of less light getting lost and consequently a higher intensity being available.

In the present invention, the hand-held scanner is designed as a portable scanner, i.e. it works at high speed and has little weight. It is, however, also possible to mount the hand-held scanner on a tripod (or on another stand), on a manually movable trolley (or another cart), or on an autonomously moving robot, i.e. that it is not carried by the user – optionally also by using another housing, for example without grip part. The notion "hand-held scanner" must consequently have a wide interpretation, so that it comprises in general scanners which are configured as compact units.

A frustum is a truncated-pyramid-shaped (or truncated-konical-shaped) area of space, which extends from the image plane (near clipping plane), in correspondence with the viewing direction, to a plane far away (far clipping plane), i.e. the frustum is a normally finite area of space. However, the far clipping plane may be in the infinity. Points inside the frustum are visible, points outside the frustum are not visible. Objects containing several points may be represented by simple bounding volumes, like spheres, cuboids, truncated pyramids or the like. Frustum culling is testing the visibilty of an object by an intersection of the volumes of the frustum and of said object (e.g. an other frustum). If the intersection of the volumes is zero, the object is not visible and will not be displayed. The intersection may be formed analytically (by known mathematical methods or approaches, e.g. intersecting planes, calculating distances of corners of said object to planes of the frustum) or numerically. In the latter case, the volume of the object may be discretized, and the field elements are checked whether they are inside the frustum (and counted) or whether they are outside the frustum (and ignored).

Operation of the hand-held scanner can, in a sequence of frames or in a video, entail a ring closure, in particular when an object O is circumnavigated. Preferably, the ring closure is recognized automatically and used for correcting potential measuring errors. For this purpose, preferably one frustum each is formed for any frame out of the plurality of frames, such frustum containing a certain part of the points of the three-dimensional point cloud which represents the 3D-scan, which is determined from the frame and assigned thereto. The intersection of the frustum of the latest frame and of a plurality of the past frames is formed, wherein the past frustum having the largest intersection is chosen. The ring closure can be recognized by means of searching, comparing and identifying features.

In order to reduce the amount of data to be saved and/or transferred by the hand-held scanner, preferably (in a post-processing) an averaging is performed over adjacent frames, preferably by dividing the two-dimensionally structured amount of data up into groups of frames and averaging by means of the frames of the group.

The invention is explained in more detail below on the basis of an exemplary embodiment illustrated in the drawing, in which

Fig. 1 shows a schematic illustration of the device,

Fig. 2 shows a schematic illustration of the production of a pattern by means of a diffractive optical element,

Fig. 3 shows a pattern and another pattern,

Fig. 4 shows a schematic illustration of projector plane, image plans and epipolar lines,

Fig. 5 shows a schematic illustration of an averaging, and

Fig. 6 shows a schematic illustration of a ring closure.

A hand-held scanner 100 is provided as portable part of a device for optically scanning and measuring an environment of the hand-held scanner 100. The hand-held scanner 100 has a base part 104, a grip part 106, which protrudes from the base part 104 - pointing upward, if properly used -, and a head end 108, which is provided on the grip part 106 - at the upper end thereof, if properly used. A user of the hand-held scanner 100 can take the hand-held scanner 100 at the grip part 106, which is configured in order to stride through the environment with the hand-held scanner 100 and to align the hand-held scanner 100 to objects O in the environment.

A first camera 111 and a second camera 112 are arranged in the head end 108 - spaced apart at a defined distance to each other. The alignments of the first camera 111 and of the second camera 112 to each other are adjusted or adjustable in such a way that the fields of view overlap and stereoscopic images of the objects O are possible. If the alignments are fixed, there is an optimum overlapping range - depending on the application. With regard to precision, an overlapping range similar to the projector - camera distances would be favorable. Depending on typical environment situations, also a range of several decimeters or meters may be preferred. Alternatively, the alignments can be adjusted by the user, for example by pivoting the cameras 111 and 112 in opposite sense, about axes of rotation that are parallel to the grip part 106. The alignment can be known to the hand scanner 100 at any time, if the adjusting process of the user is tracked, or the alignment is initially at random (and unknown), and is then made known to the hand-held scanner 100 by calibration.

The first camera 111 and the second camera 112 are preferably monochrome, i.e. sensitive to a narrow wavelength range, for example by being provided with corresponding filters, which then filter out other wavelength ranges, including scattered light. This narrow wavelength range is preferably within the infrared range. In order to nevertheless obtain color information on the objects O, preferably a color camera 113 is arranged in the head end 108, preferably symmetrically aligned to the first camera 111 and to the second camera 112, and arranged in the center thereof. The color camera 113 is then sensitive in the visible wavelength range.

The hand-held scanner 100 has a display and control unit 115, which is configured preferably as a touch screen. The display and control unit 115 is preferably arranged at the head end 108, on the side facing away from the cameras 111, 112 and, if available, 113. The display and control unit 115 can be configured to be detachable. The cameras 111, 112 and, if available, 113, as well as the display and control unit 115 are connected to a control and evaluation unit 118, which is arranged preferably in the head end 108. The control and evaluation unit 118 can preferably pre-process the data of the cameras 111, 112 and, if available 113, optionally already produce the 3D-scans and give suitable views onto the display and control unit 115. Alternatively, there is no display and control unit 115 at all, but the hand-held scanner 100 is operated by means of a remote control, for example from a stationary or from a portable computer (PC, tablet, smartphone or the like), which is in continuous connection (cabled or wireless) with the control and evaluation unit 118.

Unless the control and evaluation unit 118 transfers the 3D-scans or the data of the cameras 111, 112 and, if available 113, by means of radio communication (for example by means of WLAN to a stationary computer), the hand-held scanner 100 is provided with a data connection, preferably on the base part 104 (or, alternatively, at another point of the hand-held scanner 100). Said data connection can be, for example, a standardized interface for LAN, USB or the like, or a special interface, as is described in DE 10 2009 010 465 B3. If appropriate, the data connection can be configured also for introducing a portable storage medium (SD-card, USB-stick etc.). For power supply, preferably an accumulator is provided in the base part 104. For charging the same, a power supply outlet can be provided, preferably on the base part 104, too. Alternatively, the accumulator is exchangeable.

Generally, from the images recorded by the first camera 111 and by the second camera 112, already three-dimensional data can be determined (in the control and evaluation unit 118), i.e. 3D-scans of the objects O can be produced, for example by means of photogrammetry. The objects O, however, often have few structures and many smooth surfaces, so that generation of 3D-scans from the scattered light of the objects O is difficult.

A (first) projector 121 is therefore provided, which is configured preferably in the base part 104 (or alternatively in the head end 108) and aligned in correspondence with the two cameras 111 and 112. The relative distance and the relative alignment are pre-set or can be set by the user. The (first) projector 121 projects a pattern X onto the objects O to be scanned. The pattern X does not need to be encoded (that is to say single-valued), but it is preferably uncoded, for example periodically, that is to say multivalued. The multivaluedness is resolved by the use of the two cameras 111 and 112.

The uncoded pattern X is preferably a point pattern, comprising a regular arrangement of points in a grid. In the present invention, for example, one hundred times one hundred points are projected at an angle of approximately 50° to a distance of approx. 0.5 m to 5 m. The pattern X can also be a line pattern or a combined pattern of points and lines, each of which is formed by tightly arranged light points. The two cameras 111 and 112 project the pattern X in their respective image planes B_{111} and B_{112}, in each of which one photo sensor (for example CMOS or CCD) is arranged, in order to record the pattern X.

There is a relationship between the point density, the distance between the (first) projector 121 and the object and the resolution that can be obtained with the produced pattern X. If only single images are available, fine structures of the object O can be examined with a higher point density, and only coarse structures with low point densities. It therefore seems helpful to be able to produce, in addition to pattern X, at least one other pattern X'. Depending on the production of the patterns X, X', a dynamic transition between the patterns and/or a spatial intermingling is possible, in order to adapt the point density to the structures of the object O.

In an embodiment for special examinations a second projector 122, which is aligned correspondingly and can produce the other pattern X', is therefore configured in addition to the (first) projector 121. Alternatively, the (first) projector 121 can produce, in addition to pattern X, also the other pattern X', for example offset to each other with respect to time and/or in another wavelength range. The other pattern X' preferably is a pattern which deviates from pattern X, in the present invention an uncoded pattern, in the present invention a point pattern with a regular arrangement of points having another distance (grid length) to each other.

It is also conceivable that the other pattern X' constantly interferes with pattern X, for example with a different intensity. Pattern X which is thus produced then has, for example, light points having a higher intensity at larger distances and, between them, light points having a lower intensity with smaller distances. With pattern X having different intensities, the limited camera dynamics (if the exposure time is given, the light points are visible without overexposure / underexposure only in a limited, combined distance and reflectivity area) can be overcome, and a larger dy-namics range for depth and intensity can be covered. It is true that said pattern X has a higher periodicity, but it is still uncoded in the sense of the present invention.

Besides, it is further conceivable that more than two patterns X, X' are used, for ex-ample a defined sequence of a plurality of patterns, which are produced, for ex-ample, subsequently with regard to time.

The - preferably monochromatic - pattern X (and X') is produced preferably by means of a diffractive optical element 124, which divides a light beam produced by a laser in the wavelength range (infrared) of the two cameras 111 and 112 in corres-pondence with the pattern X, without losing intensity. The lateral resolution is then limited only by the beam diameter, i.e. the size of the points. Since the pattern X (and X') is produced within the infrared range, it is possible to both, record the im-ages of the color camera 113 without interference, and to avoid safety measures to protect eyes or the like. For the same purpose, the pattern X (and X') could alternat-ively be produced in the ultraviolet range.

Two patterns X and X' can be produced with two diffractive optical elements, which are screened at different times or with different wavelengths. With a time-variable diffractive optical element, it is possible to quickly (i.e. with approximately each frame) or slowly (for example manually controlled) change between the patterns X and X', or pattern X can be adapted dynamically to the changing facts (with regard to the density of the light points and the reach of the projected pattern X). A gradual transition between the patterns X and X' is conceivable as well (fade-over). As an alternative to diffractive optical elements, arrays of microlenses or of single lasers can be used. Optionally, also a classical imaging by means of a mask, in particular of a transparency, is possible.

For reasons of energy efficiency and eye safety, the (first) projector 121 produces the pattern X on the objects O only, when the cameras 111 and 112 (and if available 113) record images of the objects O which are provided with the pattern X. For this purpose, the two cameras 111, 112 and the projector 121 (and if available the second projector 122) are synchronized, i.e. coordinated internally with each other, with regard to both, time and the pattern X used (and, if available, X'). Each recording process starts by the (first) projector 121 producing the pattern X, similar to a flash in photography, and the cameras 111 and 112 (and, if available 113) following with their records, more precisely their pairs of records (frames), i.e. one image each from each of the two cameras 111, 112. The recording process can comprise one single frame (shot), or a sequence of a plurality of frames (video). A trigger switch 126, by means of which such a shot or such a video can be triggered, is provided preferably at the grip part 106. After processing of the data, each frame then constitutes a 3D-scan, i.e. a point cloud in the three-dimensional space, in relative coordinates of the hand-held scanner 100. Alternatively to the trigger switch 126, the recording process can be triggered by means of the above-named remote control of the hand-held scanner 100.

The (first) projector 121 and the second projector 122, which is optionally available, are not arranged co-linear to the two cameras 121 and 122, but in a triangle arrangement. This arrangement of the two cameras 111 and 112, as well as of the (first) projector 121 (and optionally of the second projector 122) makes the use of mathematic methods of optics, which are known per se, possible, in particular of epipolar geometry, according to which one point in the image plane B_{112} of the second camera 112 can be observed on a (known) line, namely the epipolar line e, in the image plane B_{111} of the first camera 111, and vice versa, and/or a point which is produced by the (first) projector 121 from a projector level P_{121} can be observed on one epipolar line e each, in the image planes B_{111}, B_{112} of the two cameras 111 and 112.

In the present invention (at least) three units (projector 121 and the two cameras 111 and 112) are involved, i.e. proceeding from each of the units, two stereo geometries each (with plenty of epipolar lines e each) can be defined with the two other units. Unambigous triangle relations of points and epipolar lines e, from which the correspondence of projections of the pattern X (and optionally X') in the two image levels B_{111}, B_{112} can be determined, consequently result for the present arrangement. Due to the additional stereo geometry (compared to a pair of cameras), considerably more of the points of the pattern, which otherwise cannot be distinguished, can be identified on an epipolar line e. The density of features can thus simultaneously be high, and the size of the feature can be kept very low. With alternative methods with encoded patterns (having features consisting, for example, of a plurality of points), the size of the feature has a lower limit, limiting the lateral resolution. If said correspondence has been determined, the three-dimensional coordinates of the points on the surface of the object O are determined for the 3D-scan by means of triangulation.

Additional three-dimensional data can be gained by means of photogrammetry from several frames with different camera positions, for example from the color camera 113 or from the part of the signal of the cameras 111 and 112, which comes from the ambient light, i.e. from the natural texture of the environment. It can also be advantageous, if the hand-held scanner 100 or another unit can illuminate the object O and optionally the background, for example with white light or infrared light, so that not only the parts of the object O which are illuminated by the pattern O and optionally the background are visible, but also areas which are between those. Such illumination is particularly suitable, if the data of the color camera 113 shall be used already for making the 3D-scans (and not only for the coloration thereof), and for calibrating the cameras 111 and 112, if filters allow only a limited spectral range to penetrate.

The scanning process also shows an aspect of time. Whereas, with stationary devices, a whole sequence of patterns can be projected and images be recorded in order to determine one single 3D-scan, one 3D-scan is produced with each shot of the hand-held scanner 100 in the present invention. If a second projector 122 or a further diffractive optical element 124 or at least a second pattern X' in addition to pattern X is provided for special examinations, it is possible, by means of a suitable switching over, to also record, with one shot, images with different patterns X and X' consecutively, so that the 3D-scan will then get a higher resolution.

In order to capture the complete scene, the 3D-scans which are produced with the shot must be registered, i.e. the three-dimensional point clouds of each frame must be inserted in a common coordinate system. Registration is possible, for example, by videogrammetry, i.e., for example, "structure from motion" (SFM) or "simultaneous localisation and mapping" (SLAM). The natural texture of the objects O can be used for common points of reference, or a stationary pattern Y can be produced. The natural texture can be captured by the color camera 113, which thus gets a second function (in addition to obtaining the color information). As additional part of the device for optically scanning and measuring the environment of the hand-held scanner 100, at least a separate projector 130 is preferably provided, however.

The separate projector 130 projects the stationary pattern Y onto the objects to be scanned, i.e. a pattern similar to pattern X (and optionally X'), nonetheless distinguishable, and preferably produced in the same manner. While pattern X and optionally X' moves with the hand-held scanner 100, the stationary pattern Y rests - also if the hand-held scanner 100 is moved, and if shots are made from different positions - in a common coordinate system. The stationary pattern Y is then visible in a plurality of images (frames) of the cameras 111 and 112, so that the 3D-scans determined thereof can be brought into relation to each other by means of the stationary pattern Y. The stationary pattern Y differs from pattern X and optionally from X' with regard to geometry or time or spectrum (or a combination thereof). If it differs with regard to time, the stationary pattern Y is produced at least in intervals of time, in which the pattern X and optionally X' is not produced (alternating or overlapping). If it differs with regard to spectrum, the stationary pattern Y is within another wavelength range as pattern X and optionally X', so that the cameras 111 and 112 must be sensitive for this, too, i.e. they should be provided with corresponding filters. The separate projector 130 can be synchronized with the hand-held scanner 100, i.e. time and kind of the projected stationary pattern Y are known to the hand-held scanner 100.

Depending on the object O to be scanned, it might be appropriate, after a plurality of 3D scans have been made, to take the separate projector 130 to the other side of the object O, in order to project a stationary pattern Y onto the surface there. Shadings can thus be avoided. The separate projector 130 is therefore preferably portable or movable and is correspondingly mounted, for example, on a tripod or on a trolley (or another cart) or can be mounted thereon. Alternatively, a plurality of separate projectors 130 is used in order to avoid shadings. A corresponding building-block system is possible.

Basically, an automation is possible, too, i.e. the hand-held scanner 100 is mounted on a manually movable trolley (or on another cart), or on an autonomously moving robot, or can be mounted thereon. The hand-held scanner 100, which is no longer carried by the user, scans its environment in a defined manner, preferably rather by producing a video than by producing a sequence of shots. Cameras and projectors are not arranged in a co-linear manner.

The hand-held scanner 100 can produce a video with a high density of frames, for example seventy frames per second. Since the held-scanner 100, however, does hardly move between two frames, the video contains very much redundant information: two frames which are adjacent with regard to time, differ only very slightly. In order to reduce the amount of data to be saved and/or to be transferred, suitable averagings in a post-processing are therefore recommendable (Fig. 5). First of all, in a first averaging step, the frames F are divided into groups $[F]_i$, with a plurality of frames per group $[F]_i$ around one key frame F_i each.

So-called voxels, which completely fill space as the sum of single volume elements, are known from 3D-computer graphics. Such structures are frequently used in order to unite three-dimensional data from different perspectives in one point cloud. A disadvantage when recording surface data are the many remaining empty voxels, which must be processed in terms of data in some way.

In the present invention, data structures which are adapted to the problem and optimized, are used. Within a group $[F]_i$ of considerably overlapping frames F, single measuring points still can be summarized very well and efficiently in a common two-dimensional data structure (grid structure), i.e. optimized for surface data and very similar to a two-dimensional image. The smaller storage capacity required permits to initially save all captured measured values as a vector in the two-dimensional data structure, i.e. gray-tone value/color and distance to the hand-held scanner 100 for each of the pixels of the frames F of the group $[F]_i$.

In a second averaging step, an averaging is performed within each group $[F]_i$, in order to very simply eliminate faulty measurements. For such averaging (with regard to gray tones/colors and/or distances), only a defined part of the vector within the central range of the sorted measured values is taken. The central range can be distinguished by means of threshold values. Such averaging corresponds to a replacement of the group $[F]_i$ by a key frame F_i with averaged measured values, wherein the key frames F_i still show considerable overlapping. Each measuring point which is gained in such a way is then carried on as a point (corresponding to a three-dimensional vector) of the three-dimensional overall point cloud.

In an optional third step, the measuring points gained by averaging can be brought together with data from another group $[F]_i$, for example by Cartesian averaging.

Operation of the hand-held scanner 100 entails, in particular when an object O is circumnavigated, that a ring closure might occur, i.e. after quite a lot of frames, the video (or the sequence of shots) shows the same view into the room or at least a very similar view. The ring closures could be recognized immediately, if it were possible to look at all available data, at any time during the production of the overall point cloud. The amount of data and the computing time resulting therefrom, however, don't allow this. A method is required, by means of which it can be seen very quickly, which data from earlier sequences must also be looked at, due to the ring closure. If all measurements were completely without faults (and the movement of the hand-held scanner 100 were sufficiently regular), the ring closure would immediately result from the registration of the 3D-scan in the common coordinate system. In reality, a fault will, however, sum up, leading to an offset of two similar frames F and the resulting 3D scans. A possibility of automatically recognizing the ring closure nevertheless (and to correct the fault), is described in the following (Fig. 6).

A frustum (more precisely: viewing frustum) usually is a truncated-pyramid-shaped area of space, which extends from the image plane, in correspondence with the viewing direction, into the infinite. In the present invention, a frustum V is formed for each frame in a first step, such frustum comprising (at least approximately), 80 % of the captured points from the three-dimensional point cloud, i.e. a finite part of said area of space of the assigned 3D scan, which is determined from the frame F. The latest frustum V_n is assigned to the latest frame F_n, which was recorded last. In a second step the latest frustum V_n is then compared to the past frusta V by forming the intersection. The frustum out of the past frusta Vj, with which there is the largest intersection, is selected for carrying out an exacter analysis.

In a third step, within the latest frustum V_n and the selected frustum Vj each, features, for example edges and corners, are looked for in a manner known per se. In a fourth step, the detected features are compared to each other, for example with regard to their embedded geometry, and the coinciding features are identified. Depending on the degree of coincidence, it is decided in a fifth step, whether there is a ring closure or not.

In order to profit from the knowledge of the ring closure, common features are generated from the identified, coinciding features. By means of methods known under the denomination "bundle adjustment", said error of measurement can be corrected in a sixth step, i.e. the 3D scans are corrected up to a defined depth of penetration into space, i.e. the three-dimensional point cloud is - in some places and to a certain degree - displaced, so that said offset is eliminated in the frames, 3D scans and frusta which are per se identical. If correction is not completely possible, i.e. after the sixth step (with the "bundle adjustment"), a certain deviation of data and consequently a certain error of measurement which cannot be corrected, still remains, this certain deviation (i.e. the error which cannot be corrected) preferably is a measure for the quality of the measurements and of the data as a whole.

The movement of the hand-held scanner 100 and processing of the produced frames can be treated also as tracking, i.e. the hand-held scanner 100 tracks the relative movement of its environment with the methods used during tracking. If tracking gets lost, for example, if the hand-held scanner 100 has been moved too fast, there is a simple possibility of re-assuming tracking. For this purpose, the latest video image, as it is provided by the color camera 113, and the last video still image from tracking provided by it, are represented side by side (or one above the other) on the display and control unit 115 for the user. The user then must move the hand-held scanner 100 until the two video images get to coincide. A support, for example acoustic or optical, due to a processing of the video images and a comparison thereof is helpful and therefore preferably implemented.

List of Reference Numerals

100	hand-held scanner
104	base part
106	grip part
108	head end
111	first camera
112	second camera
113	color camera
115	display and control unit
118	control and evaluation unit
121	(first) projector
122	second projector
124	diffractive optical element
126	trigger switch
130	separate projector
B_{111}	image plane of the first camera
B_{112}	image plane of the second camera
e	epipolar line
F, F_n	frame
F_i	key frame
$[F]_i$	group
O	object
P_{121}	projector plane
V, V_j, V_n	frustum
X, X'	pattern
Y	stationary pattern

Patent Claims

1. Method for optically scanning and measuring an environment by means of a hand-held scanner (100) for producing 3D-scans, wherein the hand-held scanner (100) produces, by means of at least one projector (121, 122), at least one pattern (X, X') onto an object (O) in the environment and records, by means of at least one camera (111, 112, 113) images of the object (O) which is provided with the pattern (X, X'), with a plurality of frames (F), characterized in that, for finding a ring closure in the plurality of frames (F), a frustum (V) is formed for each frame (F) in a first step, and that, in a second step the latest frustum (V_n) which is assigned to the latest frame (F_n), is compared with a plurality of the last frusta (V), by means of forming the intersection, wherein the frustum (V_j), with which there is the largest intersection, is selected.

2. Method according to Claim 1, characterized in that, in a third step, features are looked for within the latest frustum (V_n) and the selected frustum (V_n) each.

3. Method according to Claim 2, characterized in that, in a fourth step, the detected features are compared with each other and that the coinciding features are identified.

4. Method according to Claim 3, characterized in that, in a fifth step, depending on the degree of correspondence of the corresponding features, the decision is taken, whether a ring closure is present or not.

5. Method according to Claim 4, characterized in that, from the identified, corresponding features, common features are generated, by means of which, in a sixth step, errors of measurement are corrected, in particular up to a certain penetration depth.

6. Method according to Claim 5, characterized in that the errors which cannot be corrected in the sixth step build a measure for the quality of measurements.

7. Method according to one of the preceding claims, characterized in that each of the formed frusta (V) contains at least approximately 80 % of the registered points from the three-dimensional point cloud of the corresponding 3D-scan.

8. Method according to the generic term of Claim 1, in particular according to one of the preceding claims, characterized in that an averaging is performed over adjacent frames (F).

9. Method according to Claim 8, characterized in that, in a first averaging step, the frames (F) are divided into groups ($[F]_i$) and are saved within each group ($[F]_i$) as a vector in a common two-dimensional data structure.

10. Method according to Claim 9, characterized in that, for averaging in a second averaging step, averaging is performed within each group ($[F]_i$) via a defined share of the vector, in particular that the averaging for each vector is performed with regard to the gray tones/colors and/or the distances.

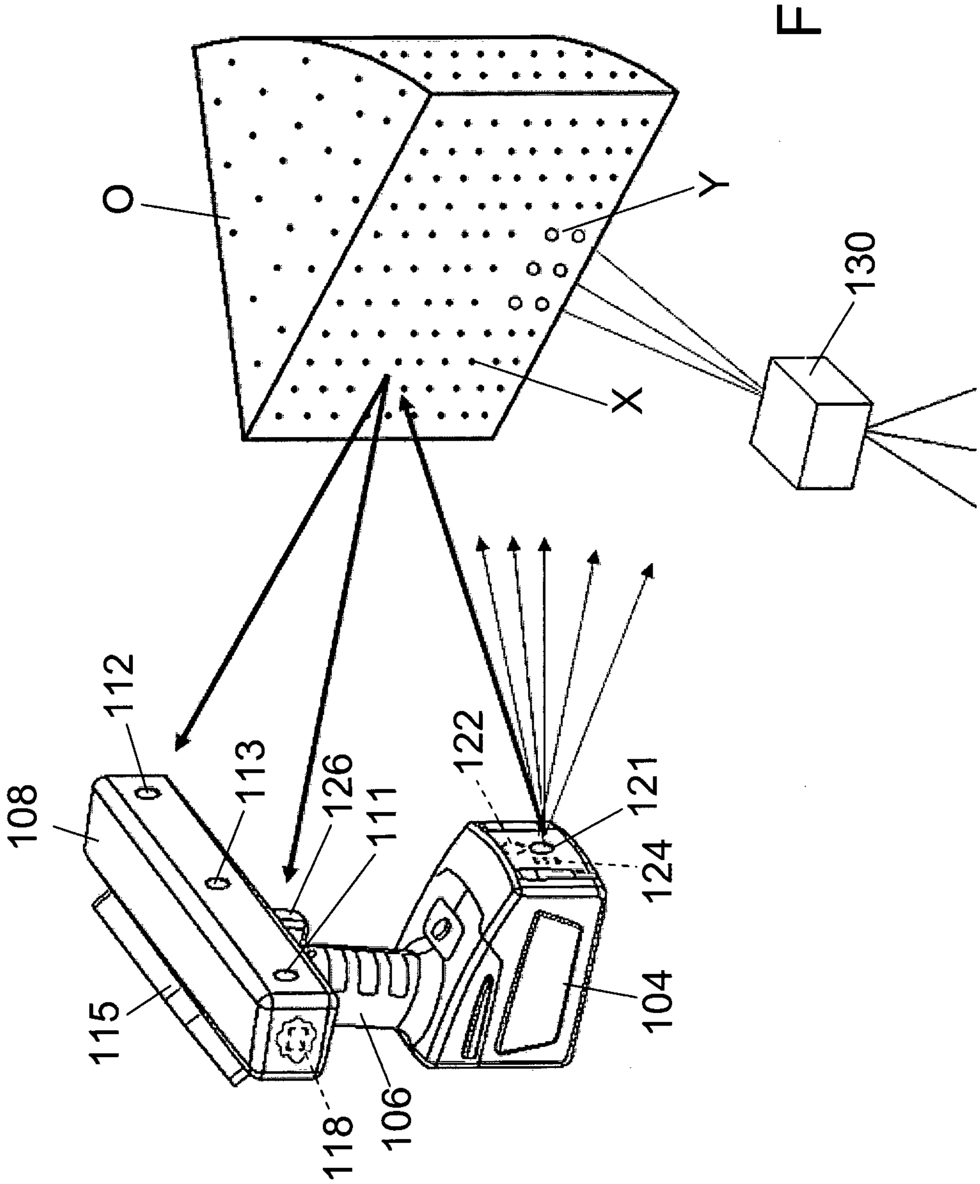

Fig. 1

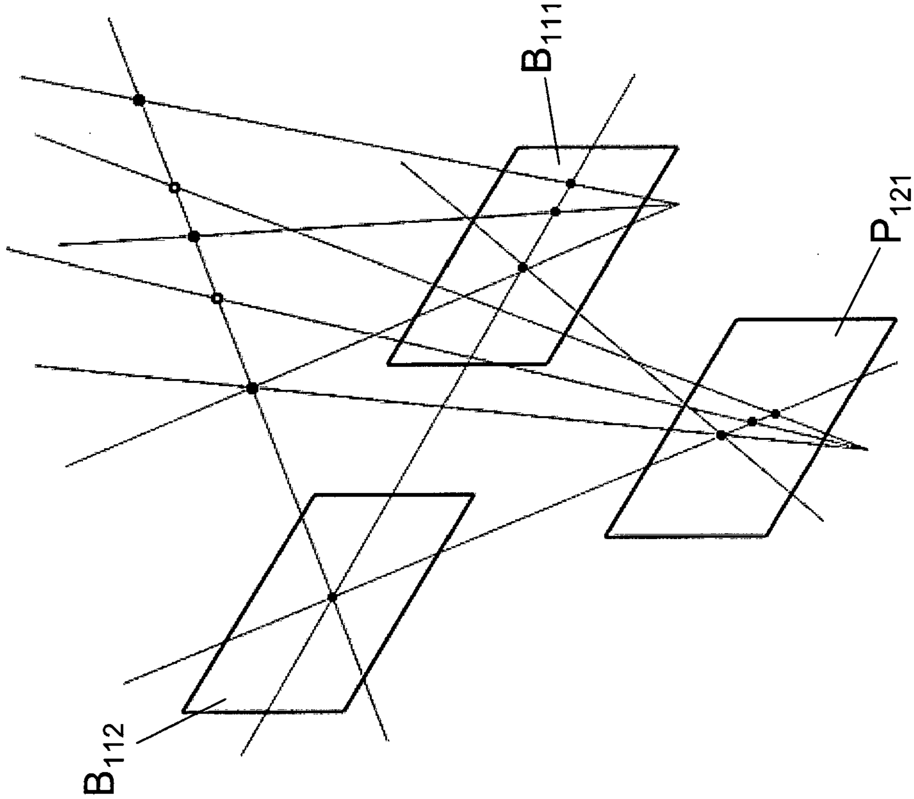

Fig. 4

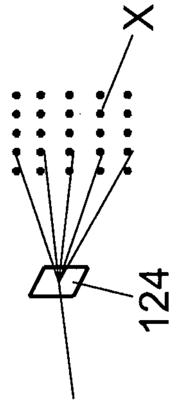

Fig. 2

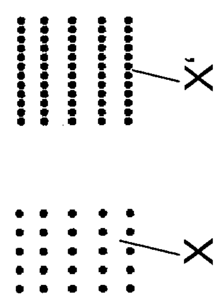

Fig. 3

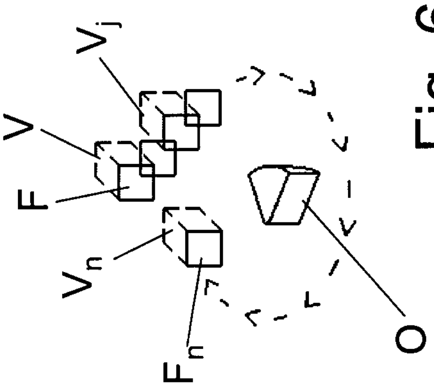

Fig. 6

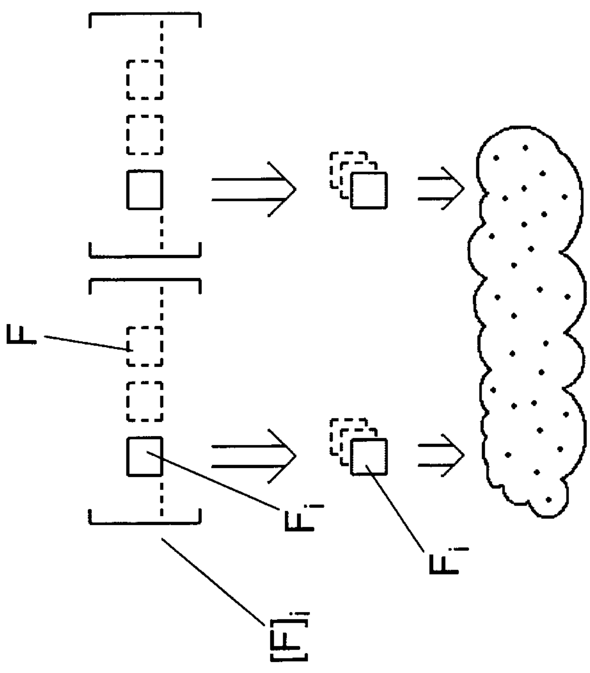

Fig. 5