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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
19 November 2009 (19.11.2009)

PCT

(10) International Publication Number
WO 2009/138350 A1

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

(21) International Application Number:
PCT/EP2009/055496

(22) International Filing Date:
6 May 2009 (06.05.2009)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
08305166.4 14 May 2008 (14.05.2008) EP

(71) Applicant *(for all designated States except US)*: **THOMSON LICENSING** [FR/FR]; 46 Quai Alphonse Le Gallo, F-92100 Boulogne Billancourt (FR).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **SCHLOSSER, Markus** [DE/DE]; Lister Kirchweg 98, 30177 Hannover (DE). **HERPEL, Carsten** [DE/DE]; Schwarze-Dorn-Str. 4, 30974 Wennigsen (DE).

(74) Agent: **LINDEMANN, Robert**; Deutsche Thomson OHG, Karl-Wiechert-Allee 74, 30625 Hannover (DE).

(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, 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, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, 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, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (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, SE, SI, SK, 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))*

— *with amended claims and statement (Art. 19(1))*

(54) Title: APPARATUS FOR EVALUATING IMAGES FROM A MULTI CAMERA SYSTEM, MULTI CAMERA SYSTEM AND PROCESS FOR EVALUATING

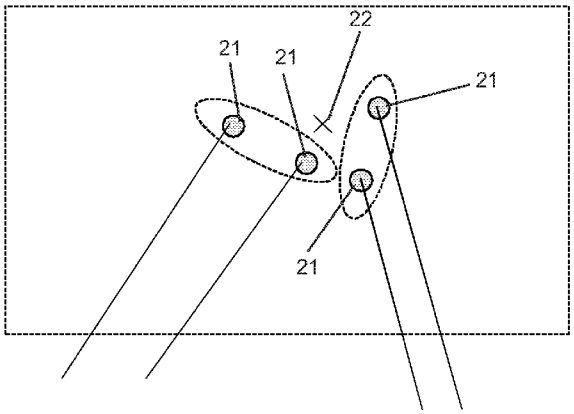


Fig. 4

(57) Abstract: An apparatus 2 for evaluating images from a multi camera system 1 is proposed, the multi camera system 1 comprising a main camera 3 for generating a main image and at least two satellite cameras 4a, 4b for generating at least a first and a second satellite image. The cameras 3, 4a, 4b can be orientated to a common observation area. The apparatus 2 is operable to estimate a combined positional data 18 of a point in the 3D-space 22 of the observation area corresponding to a pixel or group of pixels of interest 19 of the main image. The apparatus 2 comprises first disparity means 11 for estimating at least a first disparity data 12 concerning the pixel or group of pixels of interest (19) derived from the main image and the first satellite image, second disparity means 11 for estimating at least a second disparity data 13 concerning the pixel or group of pixels of interest 19 derived from the main image and the second satellite image, and positional data means 7,17 for estimating the combined positional data of the point in the 3D-space of the observation area corresponding to the pixel or group of pixels of interest 19. The positional data means 7, 14, 17 is operable to estimate first positional data 15 on basis of the first disparity data 12 and second positional data 16 on basis of the second disparity data 13, and to combine the first positional data 15 and the second positional data 16 to the combined positional data 18.

WO 2009/138350 A1

Apparatus for Evaluating Images from a Multi Camera System, Multi Camera System and Process for Evaluating

The invention relates to an apparatus for evaluating images from a multi camera system, the multi camera system comprising a main camera for generating a main image and at least two satellite cameras for generating at least a first and a second satellite image, whereby the cameras can be orientated to a common observation area, wherein the apparatus is operable to estimate combined positional data of a point in the 3D-space of the observation area corresponding to a pixel of interest, or a group of pixels of interest, or block, of the main image, the apparatus comprising first disparity means for estimating at least a first disparity data concerning the interesting pixel derived from the main image and the first satellite image, second disparity means for estimating at least a second disparity data concerning the interesting pixel derived from the main image and the second satellite image and positional data means for estimating the combined positional data of the point in the 3D-space of the observation area corresponding to the pixel of interest, the group of pixels of interest, or block. The invention furthermore relates to a multi camera system comprising said apparatus and to a process for evaluating images.

Distance measurement with the help of video cameras is well-known in the art. A common approach is to use a so-called stereo camera system, which comprises two cameras being orientated to a common observation area. As the two cameras of the stereo camera system are arranged with some distance between them the images of the observation area differ slightly. This slight difference can be used to compute a so-called disparity between the two images. On basis of this disparity the distance or the 3D-position of an object point corresponding to the pixel in the images can be computed. In order to facilitate the search for corresponding image points in the two images of the two cameras, it is known to perform a rectification in a first step, whereby the two images are transformed in such a manner that the epipolar lines of the images are parallel to each other. In general the term rectification as used in the following specification also relates to calculating correction values for images for correcting image distortions, notably

geometrical image distortions, due to the varying viewing angles of the cameras in the system. The geometrical distortions are also known as keystone distortions. The term rectification is, however, also used for calculations for correcting distortions originating from the camera itself, e.g. lens distortions, which may also be referred to by the term pincushion distortion, and shall not be construed to be limited to a single one of the calculations mentioned above.

It is furthermore known from the state of the art to use three cameras for distance measurement in order to improve the accuracy and the reliability of the resulting measurement values. From a theoretical point of view the evaluation of images from a multi camera system can be divided into an 'implicit method' and an 'explicit method'.

An approach for the 'implicit method' is described for example in the document US 2003/0072483 concerning a multi camera system with three cameras, whereby disparities are estimated jointly and consistently over all three cameras. For this purpose, triple point-correspondences between all three cameras respecting a trifocal constraint are tested against each other using an overall correlation criterion.

The 'explicit method' works the opposite way. The three cameras are separated into two independent camera pairs (e.g. middle - left and middle - right), and disparities are estimated completely independently from each other. An example for the 'explicit method' is disclosed by the document US 2006/0056727 A1. The document has the title “System for combining multiple disparity maps” and describes a process, whereby a plurality of images of a scene are provided, whereby pairs of images are selected and whereby for each pair of selected images a disparity map is created. In the next step, a reference point is selected and all disparity maps are transformed into the coordinates of the reference point. Afterwards the transformed disparity maps are merged into one single disparity map.

It is an object of the invention to present an apparatus, a multi camera system and a process improving the depth measurement in images of multi camera systems.

According to the invention an apparatus with the features of claim 1, a multi camera system with the features of claim 11 and a process with the features of claim 13 are proposed. Preferred or advantageous embodiments of the invention are disclosed in the dependent claims, in the description and/or in the figures as attached.

The apparatus according to the invention is adapted and/or operable to evaluate correlated images originating from a multi camera system. The apparatus itself can be embodied as a data processing unit, for example a computer running a corresponding computer program, a microcontroller, a dsp-card etc. The multi camera system comprises at least a main camera for generating a main image and at least two satellite cameras for generating at least a first and a second satellite image. In operation the cameras are orientated to observe a common observation area, so that the main image and the satellite images show at least partly overlapping areas. Although it may be preferred to use the said three cameras, it is also possible to use more than two satellite cameras accompanying the main camera.

The apparatus is operable and/or adapted to estimate a combined positional data of a point in the 3D-space of the observation area corresponding to a pixel of interest, the group of pixels of interest, or block of the main image. In the following description the point in the 3-D space may be referred to by any one or combination of the terms pixel of interest, group of pixels of interest, or block, without being limited to the term used, unless explicitly stated. It will be appreciated that any one of the terms relates to an object or point of interest located in the observation area, and the size of the object on the cameras' imagers determines whether it can be represented by one or more pixels, or a block of pixels. One task of the apparatus is thus to evaluate the images in such a way that combined positional data of the point in the 3D-space, for example 3D-

coordinates or depth or distance values, corresponding to an interesting pixel of the main image , are calculated or estimated.

The apparatus comprises first disparity means for estimating at least a first disparity data concerning the interesting pixel derived from the main image and the first satellite image. Disparity in its general definition may be explained as a spatial difference between two corresponding image points, respectively representing the same original point, in two images from two cameras viewing the same or an overlapping observation area from different positions. From the respective positions of the corresponding points in the two images or on the basis of a corresponding disparity value it is possible to calculate the distance and/or the viewing angle to the object corresponding to the image points. In case the positions of the cameras are known the 3D-position of the object can be derived, otherwise the position is determined relative to the cameras.

Furthermore, the apparatus comprises second disparity means for estimating at least a second disparity data concerning the interesting pixel derived from the main image and the second satellite image. Preferably, the first disparity data and the second disparity data are calculated separately, especially without using the second satellite image in case of the first disparity data and without using the first satellite image in case of the second disparity data. In further embodiments of the invention third, fourth etc. disparity means can be integrated to evaluate images from third, fourth etc. satellite cameras in order to estimate third, fourth etc. disparity data.

A positional data means is operable and/or adapted to estimate or determine the combined positional data of the point in the 3D-space of the observation area corresponding to the interesting pixel from the positional data previously determined or estimated.

According to the invention, the positional data means is operable to estimate a first positional data on the basis of the first disparity information and a second positional data on the basis of the second disparity information and to combine the first positional data and the second positional data into the combined

positional data. Preferably the estimation of the first positional data and the second positional data are performed separately, which means that the first positional data is estimated without using the second disparity information and the second positional data is estimated without using the first disparity data. Optionally third, fourth etc. positional data resulting from third fourth etc. disparity data may be used.

One finding of the invention is that the novel apparatus is able to determine the combined positional data with a high accuracy and high reliability. This advantage is based on the fact that the image information of the main image and the two satellite images are presented in some kind of a matrix-form or an equivalent thereof, due to the nature and finite resolution of the image sensor. As a result from the rectification performed when estimating the disparity data the input images are transformed into another matrix-form which does not necessarily fully coincide exactly with the matrix-form of the input images. In other words, the different viewing angles result in matrix distortions that have to be compensated for by rectification, even when the image resolution of the image sensors in terms of pixels per line and lines per image is the same. This may even be aggravated when the resolution of the image sensors differs. Thus, the disparity data may contain some interpolated values. In order to prevent a further transformation of the disparity data into a common disparity map as proposed by Jones et al. (US 20060056727) which is again realized as a matrix-form, it is proposed to virtually re-project the pixel of interest or group of pixels of interest from each camera into the 3-D space, or 'real world', based on the associated distance estimates, i.e. the disparity data, and combine the projected points to the point in the 3D-space corresponding to the pixel of interest or group of pixels of interest. As the image resolution in the 'real world' or the 3D-space is not limited by a matrix-form, one further interpolation step as required in Jones et al. can be avoided. Thus the resulting point in the 3D-space as determined in accordance with the invention has a higher accuracy and reliability compared to the results from apparatii or processes known from the prior art.

In a preferred embodiment of the invention the apparatus comprises a rectification means for performing the rectification of the main image with the first

satellite image and – preferably separately – the main image with the second satellite image. As mentioned further above, the task of the rectification is to project multiple images onto a common image surface. It is for example used to correct a distorted image into a standard coordinate system. In one special embodiment the transformation is realized to align the epipolar lines of the images horizontally.

In a preferred and practical embodiment of the invention, the apparatus is operable to estimate a combined positional data map comprising the positional data of the pixel of interest or the group of pixels of interest. The map may comprise the complete main image, so that for every pixel or group of pixels in the main image a combined positional data is mapped. Alternatively, only a subset or subarea of the main image is mapped with combined positional data. Furthermore, it is preferred to estimate a first disparity map comprising the first disparity data and/or to estimate a second disparity map comprising the second disparity data. The dimension, for example the number of rows and number of columns, of the first disparity map, and the second disparity map and the combined positional data map are preferably the same.

In a further preferred embodiment of the invention at least one, some or each of the first disparity data comprises at least two, preferably four neighbor disparity data values for each pixel of interest or group of pixels of interest in the original image. Four neighbor disparity values may represent a tradeoff between required computing power and accuracy. Additionally or alternatively, for at least one, some or each of the second disparity data at least two, preferably four neighbor disparity data values are provided. As explained before, the transformation from the input images into the disparity maps usually leads to an interpolation, at least for parts of the image, as the input images are embodied in matrix-form and the disparity maps are also embodied in matrix-form, but having a different matrix due to geometrical distortion or the like. In order to minimize the transformation and/or interpolation inaccuracies it is preferred to provide not only one single disparity data value per disparity map but at least two for the pixel of interest or group of pixels of interest. Preferably four disparity data values resulting from matrix places located adjacent to the ideal position of the pixel of interest or group

of pixels of interest in the disparity map are provided. With the knowledge that the ideal position of the interesting pixel in the disparity maps is located between the matrix places corresponding to the two, preferably four neighbor disparity data values ideally no, or at least only a minimum of information is lost by the transformation. With this embodiment, sub-pixel accuracy can be achieved.

Consequently, it is preferred, that each of the first and/or the second positional data comprises at least two, preferably four positional data values corresponding to the neighbor disparity data values. Thus, for each interesting pixel and per each disparity map at least two, preferably four positional data values are calculated in the 3D-space or the real world.

In yet a further preferred embodiment the positional data means is operable to estimate the combined positional data on the basis of the positional data values. As a result, the combined positional data of one single pixel or group of pixels of interest is estimated by four, preferably eight positional data values in case a multi camera system with two satellite cameras is employed. With this architecture it is secured that the amount of interpolation errors introduced by the processing of the input images is kept at a minimum.

The positional data values may be combined in a weighted manner, whereby the weights depend on the distance in - for example - image coordinates between the projection of the interesting pixel and its neighbors in the disparity maps and/or on an estimated reliability of the positional data values based on the spatial distribution of the positional data values in the 3D-space.

In a possible further embodiment the positional data means is operable and/or adapted to estimate a reliability value for each combined positional data. One possible basis for the reliability value is the spatial distribution of the positional data values in the 3D-space. The distance between these two positional data values indicates the consistency between the estimates of the camera pairs formed from main camera and first satellite camera on the one hand side and from main camera and second satellite camera on the other hand side.

In a first possible realization only one positional data value of the first positional data and only one positional data value of the second positional data are used. Another possible basis for determination of the reliability value is the position of the four, preferably eight positional data values, whereby also the distance between the data values can be evaluated.

A further subject matter of the invention is a multi camera system comprising a main camera for generating a main image and at least two satellite cameras for generating at least a first and a second satellite image, whereby the cameras can be orientated to a common observation area. The multi camera system is characterized by an apparatus as described above and/or according to one of the preceding claims.

In a preferred embodiment the main camera shows a higher resolution, a higher color depth, a higher frame rate and/or a better optical setup than the satellite cameras. One finding of this embodiment is that in special applications, for example in the movie industry, a main camera is used for taking the images, which are used later on for the final movie. The satellite cameras are only auxiliary cameras, which allow determining a depth map or a 3D-model of the observation area, especially of the observation scene. Thus, the quality of the images taken by the satellite cameras may have a lower quality than the images taken by the main camera. However, in case objects the position of which is to be determined are represented by a group of pixels in the main camera it may be sufficient to determine the position or distance of this group of pixels rather than for each individual pixel. Other possible applications of the multi camera system and/or of the apparatus can be found in the automotive area, especially to support driver assistant systems by analyzing and/or observing the surrounding of a vehicle.

A further subject matter of the invention is a method for evaluating images preferably from a multi camera system as described above/and/or according to the preceding claims and/or by using the apparatus as described above and/or according to one of the preceding claims. The method of evaluating comprises the steps:

estimating at least first disparity data concerning a pixel or a group of pixels of interest in a main image derived from the main image and a first satellite image;

estimating second disparity data concerning that pixel or group of pixels derived from the main image and a second satellite image;

estimating a combined positional data of a point in the 3D-space of an observation area corresponding to the pixel or group of pixels of interest, wherein first positional data are determined on the basis of the first disparity data and second positional data are determined on the basis of the second disparity data, and wherein the first positional data and the second positional data are combined into the combined positional data.

Further advantages, features and characteristics of the invention are disclosed by the following description and the figures as attached of a preferred embodiment of the invention. In the drawing,

Fig. 1 shows a block diagram of a first embodiment of the invention;

Fig. 2 shows the block diagram of Fig. 1 with more details in order to illustrate a second embodiment of the invention;

Fig. 3 shows a schematic illustration of the distribution of special pixels in the rectified images and the 3D-space according to the second embodiment of the invention;

Fig. 4 shows a detail from the Fig. 3.

Fig. 1 shows a schematic view of multi camera system 1 coupled with an apparatus 2, which is adapted to evaluate the images from the multi camera system 1 as a first embodiment of the invention.

The multi camera system 1 comprises at least three cameras, a main camera 3 and two satellite cameras 4a and 4b. The three cameras 3, 4a, 4b may be arranged in arbitrary order but observe a common observation scene or area 5. In one special embodiment, the main camera 3 is embodied as a movie camera, the first and the second satellite camera 4a,b are realized as cameras with a lower resolution in view of the color depth, the number of pixels and/or in view of the optical setup.

The apparatus 2 comprises a disparity module 6, which is adapted to calculate a disparity map as explained below and a positional data module 7 for calculating positional data of a point in the 3D-space of the observation area 5 corresponding to an interesting pixel of a main image, resulting from the main camera 3.

The task of the apparatus 2 and the multi camera system 1 is to generate a depth map or a 3D-map for the observation area 5, so that at least one, some and/or any pixel of the main image resulting from the main camera 3 is mapped to a 3D-position of the observation area 5.

The multi camera system 1 is operable to produce at least three images, namely the main image resulting from the main camera 3, a first satellite image, resulting from the first satellite camera 4a and a second satellite image, resulting from the satellite camera 4b. These three images are applied to a rectification unit 8 being a part of the disparity module 6. In the rectification unit 8 the first satellite image and the main image are rectified to each other resulting in a first pair 9 of rectified images and the main image and the second satellite image are rectified to each other resulting in a second pair of rectified images 10. During the rectifying step the epipolar-lines of the main image and the first satellite image are adjusted parallel by a transform operation to generate the first pair of rectified images 9. The same step is processed to generate the second pair of rectified images 10. The steps of rectification for the first pair 9 and the second pair 10 of rectified images are performed separate from each other.

In a disparity unit 11 being a part of the disparity module 6 a first disparity map 12 and a second disparity map 13 are created, whereby the first disparity map12 is

created exclusively on the basis of the first pair of rectified images 9 and the second disparity map 13 is created exclusively on the basis of the second pair of rectified images 10.

The disparity maps 12 and 13 and the rectified main image 9, 10, respectively, are transferred to the positional data module 7. In a positional data unit 14 the positional data of the pixels or groups of pixels of the first rectified main image 9 and the second rectified main image 10 are calculated, thereby creating a first and a second positional data map 15, 16. It should be noted that also the positional data maps 15, 16 are calculated independently from each other. The positional data maps 15, 16 comprise positional data about the 3D-position in the observation area 5 corresponding to the pixels or groups of pixels in the rectified main images 9 and 10.

In a subsequent combining unit 17 being a part of the positional data module 7 the positional data of the positional data maps 15 and 16 are combined to a combined positional data map 18.

Optionally, a reliability value for each pixel or group of pixels of the combined positional data map 18 may be computed by evaluating the distance of the positional data corresponding to a common pixel or group of pixels of the main image.

Fig. 2 shows a further embodiment of the present invention, which allows a higher accuracy and a better evaluation of the reliability values as compared to the embodiment shown in Fig. 1. Same parts of the embodiment in Fig. 2 and the embodiment in Fig. 1 are designated with the same reference numbers. As a difference to the previous embodiment the pixel or group of pixels of interest 19 in the main image resulting from the main camera 3 is not evaluated and processed as a single pixel (or group of pixels), since already the first transformation – the rectification in the rectification unit 8 – leads to interpolation errors. Instead of using a single pixel or matrix value for the pixel or group of pixels of interest 19, four adjacent and/or neighboring pixels 20 are processed. The neighboring pixels 20 are pixels originating from the rectified images 9, 10 which surround the ideal,

or real, position of the pixel or group of pixels of interest 19 in the rectified images 9, 10. These neighboring pixels 20 are processed by the positional data unit 14, which calculates, for each neighboring pixel 20, the 3D-positions on basis of the disparity maps 15 and 16. So – summed up – eight 3D-positions are calculated for each single pixel or group of pixels 19 of the main image from the main camera 3, which are combined by the combining unit 17 to a common 3D-position as explained below.

Fig. 3 shows a view from the top of two rectified images 9, 10 for the main camera 3, and represents the distance of the 3D point or group of 3D points to the camera and the horizontal position, or X- and Z-direction, rather than the position of the pixel or group of pixels within the image sensor's matrix of lines and columns, or X- and Y-direction. The pixel or group of pixels or interest 19 is surrounded by two neighboring pixels 20 in each rectified image 9, 10. It shall be noted that only 2x2 of the 2x4 neighboring pixels 20 are shown in Fig. 4 due to the top view presentation; the corresponding Y-direction lies perpendicular to the surface of the drawing sheet. Each neighboring pixel 20 can be projected with the help of the first and second disparity map 12, 13, respectively, to their 3D-position in the observation area 5, so that four 3D-positions 21 in the 3D-space of the observation area 5 are calculated. It shall be noted that in fact eight 3D-positions 21 are estimated.

Fig. 4 shows a detail of the observation area 5 in Fig. 3, illustrating the 3D-positions 21 of the neighboring pixels 20 and the 3D-position 22 of the ideal pixel position 19. The evaluation of the plurality of 3D-positions 21 of the neighboring pixels 20 can be computed in different ways: In a simple implementation, an average value of the eight 3D-positions 21 is calculated. The same limitation as to the reproduction of the Y-direction as mentioned for figure 3 applies.

In more sophisticated implementations a weighted average can be used, whereby the weight may be dependent from (i) - the spatial distribution - the distance in image coordinates between the ideal position 19 of the pixel or group of pixels of interest and the respective neighboring pixels or groups of pixels 20 in the rectified images 9, 10 of the main camera 3 (or an equivalent thereof) and/or (ii) –

the confidence - the estimated reliability of the distance associated with the respective pixel or group of pixels 20 for example derived when calculating the disparity data and/or (iii) – the consistence - the distance between the 3D-positions, optionally also in order to eliminate inconsistent outliers.

In yet a further implementation the following steps are proposed: Defining sub-clusters with the 3D-positions of the neighbouring pixel or group of pixels 20, whereby each sub-cluster comprises 3D-positions resulting only from one disparity map 15 or 16. Calculating of an optionally weighted centre of mass and of the corresponding standard deviation for each sub-cluster. Optionally elemination of outliers or blips, for example elimination of pixels or group of pixels 20 with 3D-positions having a distance to the centre of mass which is larger than three times of the standard deviation (three sigma test). Recalculating the centre of mass and standard deviation of the remaining pixels or group of pixels 20 of the sub-clusters. Weighted fusion of the centre of masses of the sub-classes for example weighted under consideration of the corresponding standard deviations. Optionally using the standard deviation and/ or the weights as used for defining the reliability value. One finding of this implementation is that an inconsistency of the positional data values of one pixel or group of pixels of interest is often based on the inconsistency of the positional data values of different camera pairs, and thus of different disparity maps, and not based on the inconsistency of the positional data values of one disparity map. By the fusion of the centre of masses resulting from different camera pairs and/or different disparity maps, centre of masses, which are not reliable are considered with a lesser weight.

Optionally, a reliability measure for each of these distance estimates, i.e. the combined positional data, is determined. It does not only depend on individual reliability estimates for each positional data, but also on the distance between the 3D-points of the positional data values or the centre of mass of different sub-clusters as explained below for one point of interest. This distance is a measure for the consistency of the different estimates – mainly between the different camera pairs but also within one camera pair. A large distance indicates a mismatch so that the reliability estimate is decreased accordingly.

In a further development of the invention, the consistency of the positional data values and/or the reliability value is considered or weighted in respect to the distance of the point in 3D-space to the camera or cameras. In general and due to optical reasons, the inconsistency should decrease along with the distance. Thus a limit of allowable inconsistency should also decrease with the distance. On the other hand it must be respected that in small distances distortions arise due to perspective effects. Thus the curve progression of a curve concerning the allowable inconsistency vs. the camera distance should increase again below a pre-defined camera distance.

As a consequence, the proposed approach does not only produce a combined distance map that is more accurate but the reliability of every distance is estimated, too. This additional information is very valuable in its own right as e.g. some applications may need sparse, but highly reliable and accurate distance estimates. Furthermore, this information may be advantageously be used during a subsequent post-processing step. Disparity estimation is intrinsically difficult at object borders due to the resulting occlusions. Therefore, distance maps are commonly refined using segmentation information. During this process, the reliability estimates provide additional information for the performed smoothing and interpolation.

Finally and in a further possible implementation, the computational requirements many be reduced by avoiding the combination of the neighbouring distance estimates in the combining unit 17, as detailed above. The nearest neighbour to the ideal position of the pixel or group of pixels of interest may simply be used instead, without too much performance degradation.

The invention advantageously reduces the number of transformations required during rectification of images and consequently reduces errors introduced by interpolation.

Claims

1. Apparatus (2) for evaluating correlated images originating from a multi camera system (1), the correlated images at least partly covering the same 3-D observation area captured from respective different viewing positions, the correlated images comprising a main image and at least a first and a second satellite image, wherein the apparatus (2) is operable to estimate combined positional data (18) of a point in the 3D-space (22) of the observation area corresponding to at least one pixel of interest (19) of the main image, the apparatus (2) comprising:
- first disparity means (11) for estimating at least first disparity data (12) concerning the at least one pixel of interest (19) derived from the main image and the first satellite image,
- second disparity means (11) for estimating at least second disparity data (13) concerning the at least one pixel of interest (19) derived from the main image and the second satellite image,
- positional data means (7,17) for estimating the combined positional data of the point in the 3D-space of the observation area corresponding to the at least one pixel of interest (19) in the main image,
characterized in that
the positional data means (7, 14, 17) is operable to estimate first positional data (15) on basis of the first disparity data (12) and second positional data (16) on basis of the second disparity data (13), and to combine the first positional data (15) and the second positional data (16) to the combined positional data (18), and that each of the first and/or the second disparity data (12, 13) comprises at least two neighbor disparity data values for each observation area corresponding to at least one pixel of interest (19) in the main image.

2. Apparatus (2) according to claim 1, characterized by rectification means (8) for performing a rectification of the main image with respect to the first satellite image and/or of the main image with respect to the second satellite image, and/or vice versa.

3. Apparatus (2) according to claim 1, characterized in that the apparatus (2) is operable to estimate a combined positional data map (18) comprising the combined positional data of the at least one pixel of interest and/or to estimate a first disparity map (12) comprising the first disparity data and/or to estimate a second disparity map (13) comprising the second disparity data.

4. Apparatus (2) according to claim 1, characterized in that the neighbor disparity data values are values of the disparity map (12,13) located adjacent to the ideal position (19) of the at least one pixel of interest in the disparity map (12,13).

5. Apparatus (2) according to claim 1, characterized in that each of the first and/or the second positional data (15,16) comprises at least two positional data values (21) corresponding to the neighbor disparity data values.

6. Apparatus (2) according to claim 5, characterized in that the positional data means (7) is operable to estimate the combined positional data (18) on the basis of the positional data values (21).

7. Apparatus (2) according to claim 1, characterized in that the positional data means (7) is operable to estimate a reliability value for each combined positional data.

8. Apparatus (2) according to claim 7, characterized in that the reliability value is estimated by evaluating the spatial distribution of positional data values (21) in the 3D-space.

9. Apparatus (2) according to claim 7 or 8, characterized in that the reliability value is estimated by evaluating the distance between the ideal position of the at least one pixel of interest and the location of the neighbor disparity data values in the disparity map and/or in the rectified images and/or in the 3D-space.

10. Multi camera system (1) comprising a main camera (3) for generating a main image and at least two satellite cameras (4a, 4b) for generating at least a first and a second satellite image, whereby the cameras (3, 4a, 4b) can be orientated to a common observation area (5), the multi camera system further being characterized by an apparatus (2) according to one of the preceding claims.

11. Multi camera system (1) according to claim 10, characterized in that the main camera (3) shows a higher resolution, a higher color depth, a higher frame rate and/or a better optical set-up than the satellite cameras.

12. Method for evaluating images of a multi camera system (1), preferably according to one of the preceding claims 10 or 11, the method comprising the steps of:
- estimating at least first disparity data (12) concerning an observation area corresponding to at least one pixel of interest in a main image, derived from the main image and a first satellite image;
- estimating second disparity data (13) concerning the observation area corresponding to at least one pixel of interest, derived from the main image and a second satellite image;
- estimating a combined positional data (18) of a point in the 3D-space of the observation area (5) corresponding to at least one pixel of interest, wherein first positional data on the basis of the first disparity data and second positional data on the basis of the second disparity data are estimated, wherein the first positional data and the second positional data are combined into the combined positional data; and
- providing, for each of the first and/or the second disparity data (12, 13), at least two neighbor disparity data values for the observation area corresponding to the at least one pixels of interest (19) in the main image.

AMENDED CLAIMS
received by the International Bureau on 29 October 2009 (29.10.2009)

Claims

1. Apparatus (2) for evaluating correlated images originating from a multi camera system (1), the correlated images at least partly covering the same 3-D observation area captured from respective different viewing positions, the correlated images comprising a main image and at least a first and a second satellite image, wherein the apparatus (2) is operable to estimate combined positional data (18) of a point in the 3D-space (22) of the observation area corresponding to at least one pixel of interest (19) of the main image, the apparatus (2) comprising:
- first disparity means (11) for estimating at least first disparity data (12) concerning the at least one pixel of interest (19) derived from the main image and the first satellite image,
- second disparity means (11) for estimating at least second disparity data (13) concerning the at least one pixel of interest (19) derived from the main image and the second satellite image,
- positional data means (7,17) for estimating the combined positional data of the point in the 3D-space of the observation area corresponding to the at least one pixel of interest (19) in the main image,
characterized in that
the positional data means (7, 14, 17) is operable to estimate first positional data (15) on basis of the first disparity data (12) and second positional data (16) on basis of the second disparity data (13), and wherein the positional data means is operable to combine the first positional data (15) and the second positional data (16) to the combined positional data (18) in accordance with respective reliability values for the first and the second positional data, wherein the reliability values for the first and the second positional data are based on the spatial distribution of the positional data values in the 3D-space.

2. Apparatus (2) according to claim 1, characterized by rectification means (8) for performing a rectification of the main image with respect to the first

satellite image and/or of the main image with respect to the second satellite image, and/or vice versa.

3. Apparatus (2) according to claim 1, characterized in that each of the first and/or the second disparity data (12, 13) comprises at least two neighbor disparity data values for each observation area corresponding to at least one pixel of interest (19) in the main image.

4. Apparatus (2) according to claim 1, characterized in that the neighbor disparity data values are values of the disparity map (12,13) located adjacent to the ideal position (19) of the at least one pixel of interest in the disparity map (12,13).

5. Apparatus (2) according to claim 1, characterized in that each of the first and/or the second positional data (15,16) comprises at least two positional data values (21) corresponding to the neighbor disparity data values.

6. Apparatus (2) according to claim 5, characterized in that the positional data means (7) is operable to estimate the combined positional data (18) on the basis of the positional data values (21).

7. Apparatus (2) according to claim 1, characterized in that the positional data means (7) is operable to estimate a reliability value for each combined positional data.

8. Apparatus (2) according to claim 7, characterized in that the reliability value is estimated by evaluating the spatial distribution of positional data values (21) in the 3D-space.

9. Apparatus (2) according to claim 7 or 8, characterized in that the reliability value is estimated by evaluating the distance between the ideal position of the at least one pixel of interest and the location of the neighbor disparity

data values in the disparity map and/or in the rectified images and/or in the 3D-space.

10. Multi camera system (1) comprising a main camera (3) for generating a main image and at least two satellite cameras (4a, 4b) for generating at least a first and a second satellite image, whereby the cameras (3, 4a, 4b) can be orientated to a common observation area (5), the multi camera system further being characterized by an apparatus (2) according to one of the preceding claims.

11. Multi camera system (1) according to claim 10, characterized in that the main camera (3) shows a higher resolution, a higher color depth, a higher frame rate and/or a better optical set-up than the satellite cameras.

12. Method for evaluating images of a multi camera system (1), preferably according to one of the preceding claims 10 or 11, the method comprising the steps of:
- estimating at least first disparity data (12) concerning an observation area corresponding to at least one pixel of interest in a main image, derived from the main image and a first satellite image;
- estimating second disparity data (13) concerning the observation area corresponding to at least one pixel of interest, derived from the main image and a second satellite image;
- estimating a combined positional data (18) of a point in the 3D-space of the observation area (5) corresponding to at least one pixel of interest, wherein first positional data on the basis of the first disparity data and second positional data on the basis of the second disparity data are estimated, wherein the first positional data and the second positional data are combined into the combined positional data in accordance with respective reliability values for the first and the second positional data, wherein the reliability values for the first and the second positional data are

based on the spatial distribution of the positional data values in the 3D-space.

13. The method of claim 12, further comprising providing, for each of the first and/or the second disparity data (12, 13), at least two neighbor disparity data values for the observation area corresponding to the at least one pixels of interest (19) in the main image.

International Application No. PCT/EP2009/055496
Applicant: Thomson Licensing

Statement under Art. 19 (1) PCT

Claim 1 has been amended to improve upon the clarity of the claim. Notably, the operation of the positional data means has been further clarified. In accordance with amended claim 1 the positional data means combine the first and second positional data into combined positional data based upon the reliability values associated with the first and second positional data, respectively, only after the reliability values have been assigned to the respective first and second positional data. In other words, each of the first and second positional data has individual reliability values associated with it, which is used for combining the first and second positional data into combined positional data. This is unseen in any of the prior art documents.

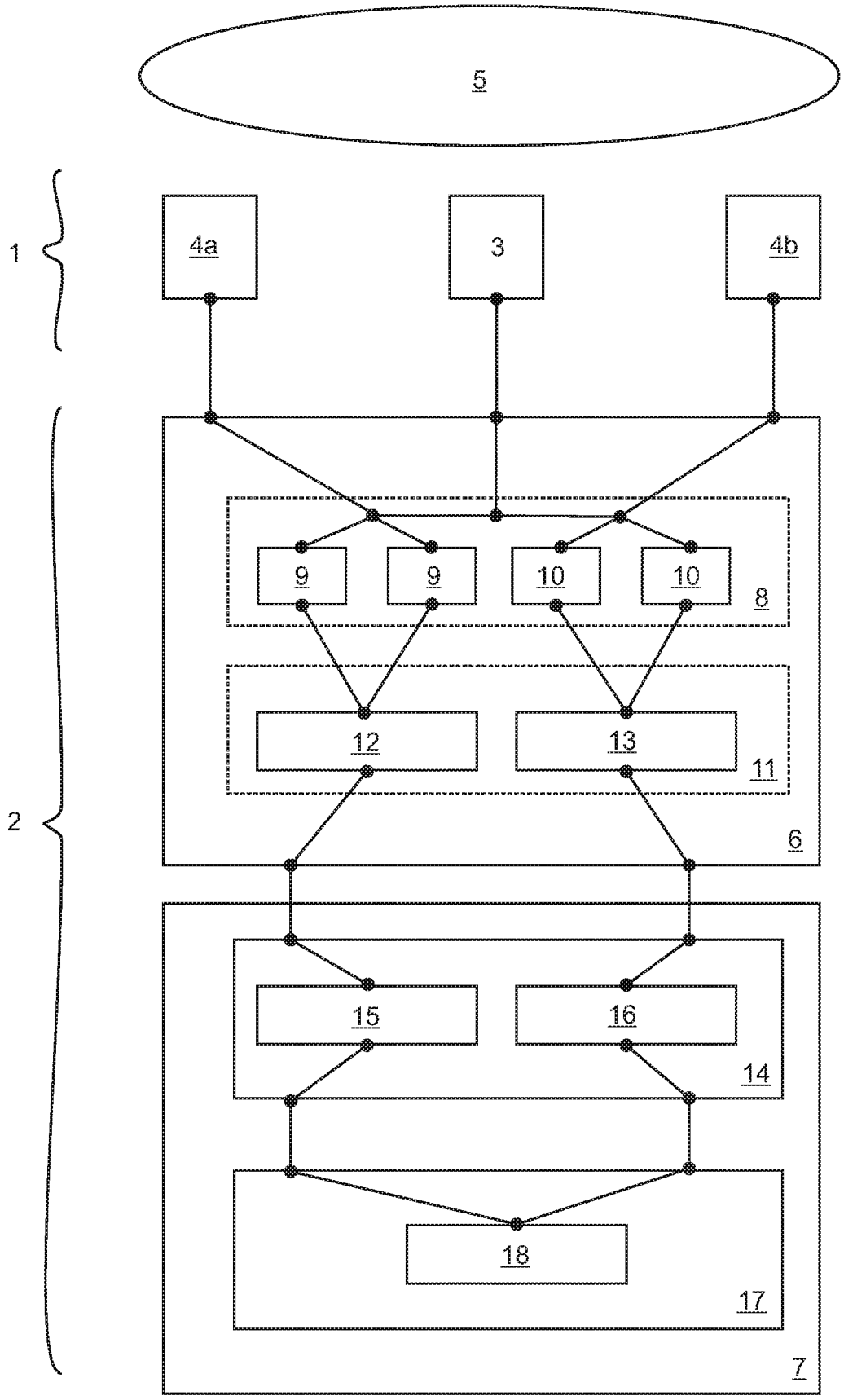


Fig. 1

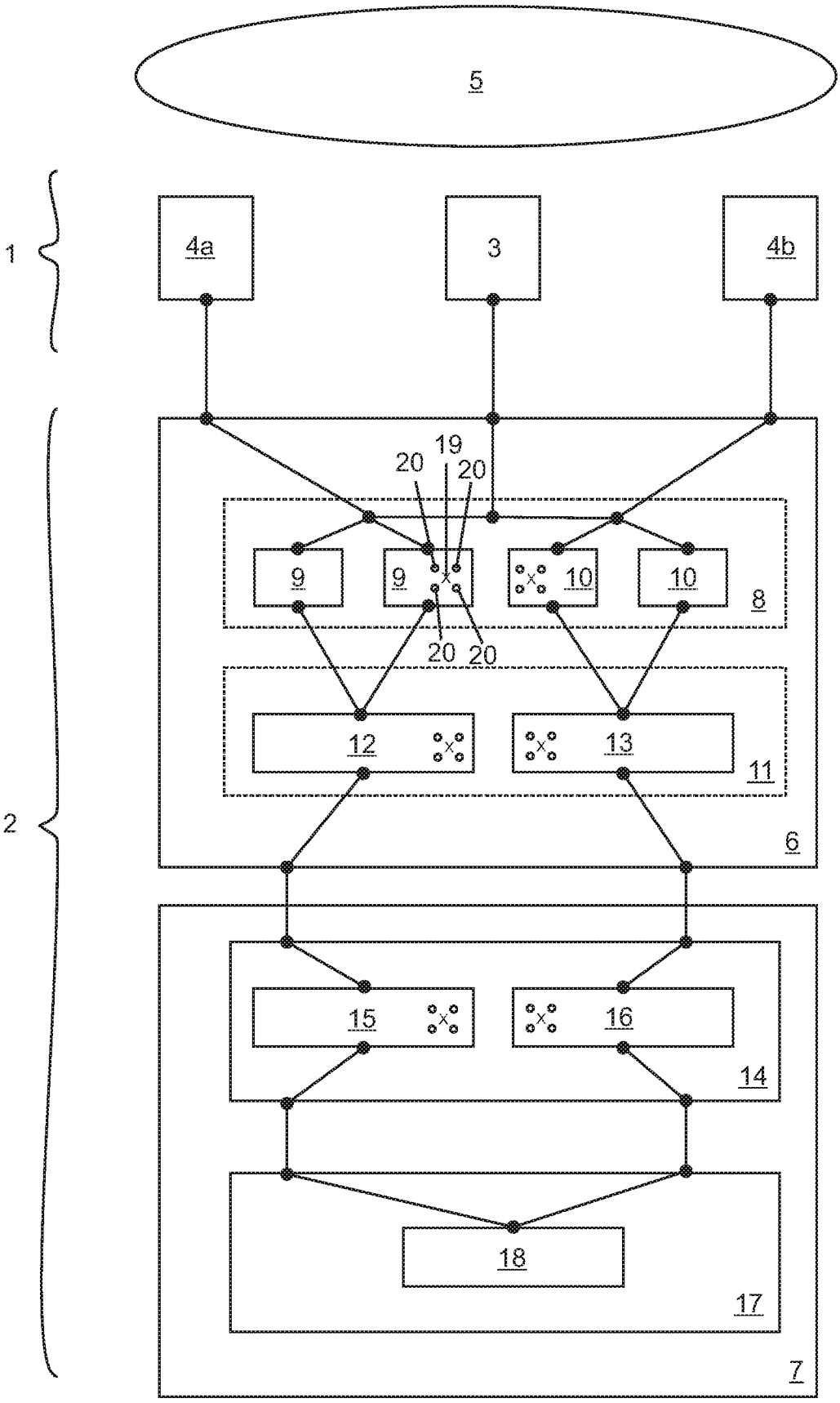


Fig. 2

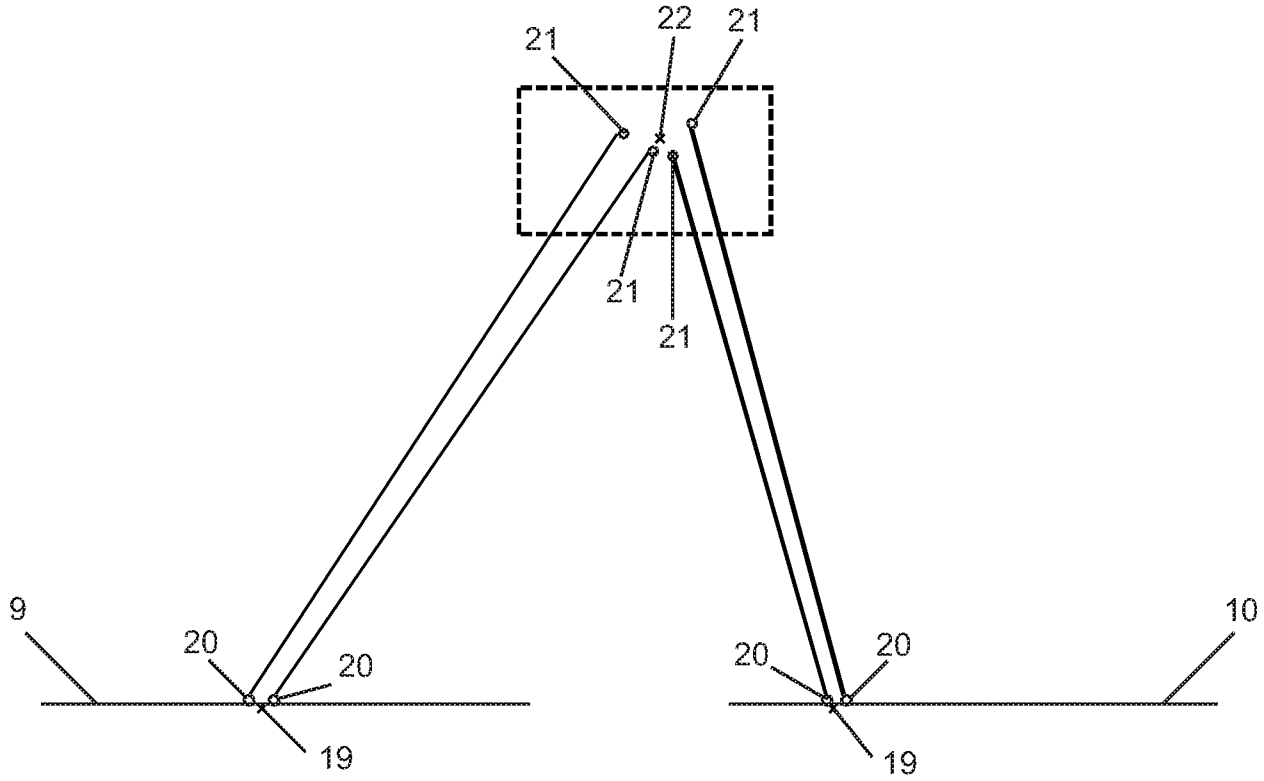


Fig. 3

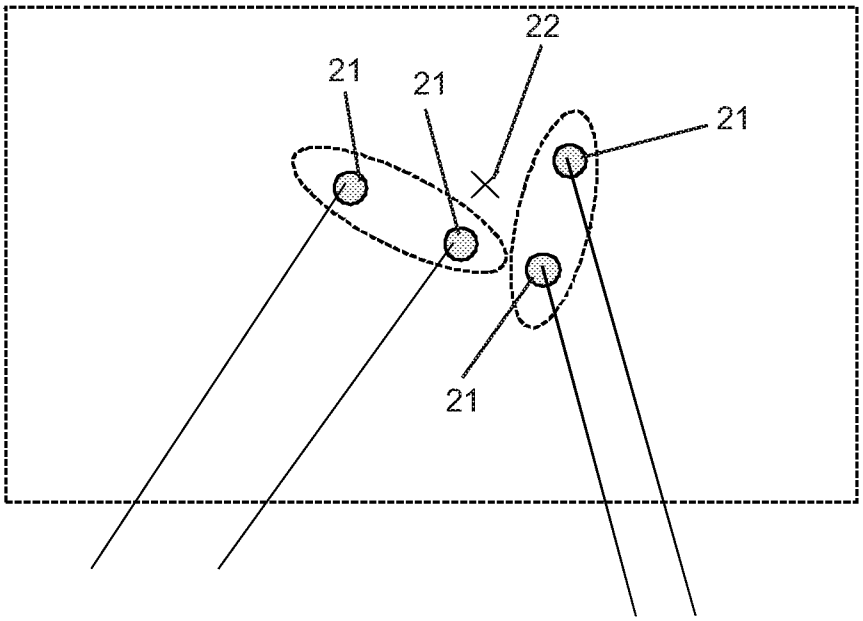


Fig. 4

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2009/055496

A. CLASSIFICATION OF SUBJECT MATTER
INV. G06T7/00

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)
G06T

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 practical, search terms used)
EPO-Internal, COMPENDEX, INSPEC, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	FUA P: "Reconstructing complex surfaces from multiple stereo views" COMPUTER VISION, 1995. PROCEEDINGS., FIFTH INTERNATIONAL CONFERENCE ON CAMBRIDGE, MA, USA 20-23 JUNE 1995, LOS ALAMITOS, CA, USA,IEEE COMPUT. SOC, US, 20 June 1995 (1995-06-20), pages 1078-1085, XP010146947 ISBN: 978-0-8186-7042-8	1,3-6, 10,12
Y	the whole document	2,7-9,11
	-/--	

[X] Further documents are listed in the continuation of Box C.

[ ] See patent family annex.

* Special categories of cited documents :

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

"E" earlier document 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
3 August 2009	14/08/2009

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Rockinger, Oliver

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2009/055496

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	GOESELE M ET AL: "Multi-View Stereo Revisited" COMPUTER VISION AND PATTERN RECOGNITION, 2006 IEEE COMPUTER SOCIETY CO NFERENCE ON NEW YORK, NY, USA 17-22 JUNE 2006, PISCATAWAY, NJ, USA,IEEE, vol. 2, 17 June 2006 (2006-06-17), pages 2402-2409, XP010923149 ISBN: 978-0-7695-2597-6 * sections 1 and 2 *	7-9
Y	TANGER R ET AL: "TRINOCULAR DEPTH ACQUISITION" SMPTE JOURNAL, SMPTE INC. SCARSDALE, N.Y, US, vol. 116, no. 5, 1 May 2007 (2007-05-01), pages 206-211, XP001541217 ISSN: 0036-1682 page 206 - page 209	2,7,11
P,X	BRADLEY D ET AL: "Accurate multi-view reconstruction using robust binocular stereo and surface meshing" 26TH IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, CVPR 26TH IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, CVPR 2008 INST. OF ELEC. AND ELEC. ENG. COMPUTER SOCIETY US, 2008, XP002501124 * sections 1 - 4 *	1-3,7-12