



(12) **EUROPEAN PATENT APPLICATION**

(43) Date of publication:
18.04.2012 Bulletin 2012/16

(51) Int Cl.:
B63B 21/00 (2006.01) G05D 1/02 (2006.01)

(21) Application number: **11182355.5**

(22) Date of filing: **22.09.2011**

(84) Designated Contracting States:
AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR
Designated Extension States:
BA ME

(72) Inventors:
• **Colombari, Andrea**
37138 Verona (IT)
• **Marzotto, Roberto**
37057 S. Giovanni Lupatoto (VR) (IT)
• **Murino, Vittorio**
16035 Rapallo (GE) (IT)

(30) Priority: **24.09.2010 IT MI20101737**

(71) Applicant: **EVS Embedded Vision Systems s.r.l.**
37134 Verona (IT)

(74) Representative: **Mittler, Enrico et al**
Mittler & C. S.r.l.
Viale Lombardia, 20
20131 Milano (IT)

(54) **Assisting visual apparatus for mooring a ship**

(57) An assisting visual apparatus for mooring a ship (1) is described. The apparatus comprises at least one pair of shooting means (TC1-TC3) arranged at given heights (H, H3) on the ship so as to have a visual field spanning the part surrounding the ship stern and a part of the ship itself, an image control device (20) comprising memory (27), a data processing unit (28) and an application software and adapted to combine the individual images shot by said shooting means (TC1-TC3), a display (24) on which the final image (I_o) processed by the image control device is formed in real time. The application software of the image control device comprises a

calibration function (100) of the shooting means which comprises, for all shooting means, the calculation (103) of the homographies with respect to a selected reference plane (π_d), the generation of a virtual image (I_{out}) as combination of the individual images deriving from the shooting means on the basis of the calculated homographies, the formation of at least one fusion map (M) on the basis of the virtual image (I_{out}). The application software also comprises a successive rendering function (200) which includes the formation of the final image (I_o) to be displayed on the display according to said at least one fusion map (M). (Fig. 2)

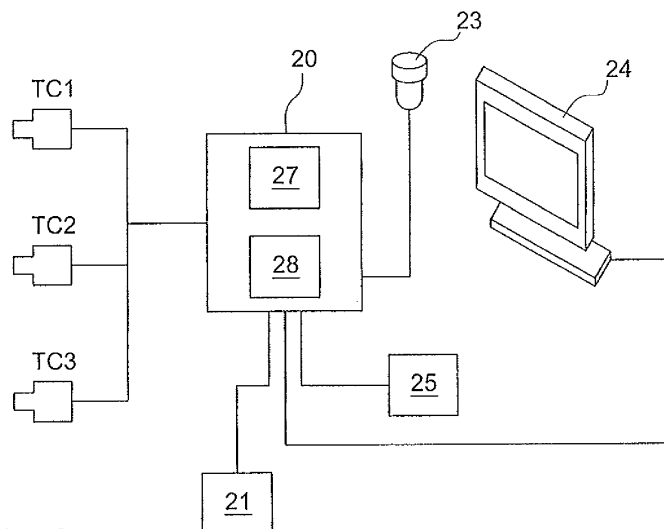


Fig.2

Description

[0001] The present invention relates to an assisting visual apparatus for mooring a ship.

[0002] The difficulties are known for performing a manoeuvre for mooring a ship, in particular the manoeuvre for stern mooring, due to the difficulties in identifying some obstacles and in estimating the volumes of the ship which is being manoeuvred due to poor visibility of the stern from the governing stations. The most difficult step in the manoeuvre for mooring is the approach and the alignment with the berth. The simple installation of a camera and associated viewer that shows the ship stern and the quay to which the ship is nearing, is not capable of displaying the ship volumes.

[0003] In view of the state of the art, it is the object of the present invention to provide an assisting visual apparatus for mooring a ship.

[0004] In accordance with the present invention, said object is achieved by means of an assisting visual apparatus for mooring a ship, said apparatus comprising at least one pair of shooting means arranged at given heights on the ship so as to have a visual field spanning the part surrounding the ship stern and a part of the same ship, an image control device comprising a memory, a data processing unit and an application software, and it being adapted to combine the single images shot by the shooting means, a display on which the final image processed by the image control device is displayed in real time, characterized in that said application software of the image control device comprises a calibration function of the shooting means which includes, for all shooting means, the calculation of the homographies with respect to a selected reference plane, the generation of a virtual image as combination of the single images deriving from the shooting means on the base of the calculated homographies, the formation of at least one fusion map on the base of the virtual image, said application software comprising a successive function of rendering including the formation of the final image to be displayed on the display as a function of said at least one fusion map.

[0005] Due to the present invention, it is possible to provide an assisting visual apparatus for mooring a ship which allows the operator to be shown the largest view possible of the stern zone, by merging the views acquired by a group of cameras. The system allows the nearing of the ship to the quay and the presence of any other ships or obstacles of another kind in the planned trajectory, to be monitored. The system helps the operator to prevent the corners of the ship from knocking against other ships and to stop at an adequate distance from the quay during the mooring.

[0006] The features and advantages of the present invention will become apparent from the following detailed description of a practical embodiment thereof, shown by way of non-limiting example in the accompanying drawings, in which:

figure 1 is a diagrammatic top view of the manoeuvre for mooring a ship equipped with an assisting visual apparatus for mooring in accordance with the present invention;

figure 2 shows an assisting visual apparatus for mooring a ship in accordance with the present invention;

figures 3-4 are views of a ship equipped with the apparatus in accordance with the present invention; figure 5 is a block diagram of the various steps of the calibration and rendering procedure performed by the assisting visual apparatus for mooring a ship in accordance with the present invention;

figure 6 shows the combination of two images with the apparatus in accordance with the present invention;

figure 7 shows the combination of three images with the apparatus in accordance with the present invention.

[0007] Figure 1 diagrammatically shows a ship in mooring manoeuvre, wherein ship 1 is equipped with an assisting visual apparatus for mooring in accordance with the present invention. Ship 1 is to be arranged near quay 3 between ships 5 and 6. The assisting visual apparatus for mooring in accordance with the present invention provides the operator on the ship with a greater visual angle 2 with respect to the one of a simple camera arranged on the stern of ship 1.

[0008] The assisting visual apparatus for mooring ship 1 is best seen in figure 2. The apparatus comprises at least two shooting means, preferably three cameras TC1-TC3 arranged on ship 1, one image control device 20 powered by a power source 21, having as input the images provided by cameras TC1-TC3 by means of a video cable and that can be activated by means of an external button 23. The image control device 20 is adapted to display images on an on-board display 24. Alternatively to display 24, it is possible to use a dedicated display mounted up high, such as a rear-view mirror, to prevent hindrances during the manoeuvre for mooring.

[0009] The image control device 20 is for applying the required transformations to the images deriving from the group of cameras TC1-TC3 and is adapted to perform the transfer of the processed video image (video output) on the available on-board navigation display 24. The image control device 20 has a control 25 for allowing the operator to select one of the available operating modes. Device 20 includes a data processing unit 28 and a memory 27 on which an application software is installed and operating; device 20 is configured in such a way to perform the transformations of the images in real time.

[0010] Cameras TC1-TC3 shoot, in real time, the frame towards which they have been set up and send the shot frames to device 20, in real time.

[0011] A possible arrangement of cameras TC1-TC3 on ship 1 is shown in figures 3-4. The cameras TC1 and TC2 are mounted on ship 1 at the same height H and

with the same vertical tilt G at a distance D from each other, while camera TC3 is positioned at a lower height $H3$ and so as to directly frame platform 11 of ship 1. The arrangement of cameras TC1-TC3 ensures that an area about the ship is framed without obstructions obstructing the view of the periphery of ship 1 itself, by merging the video flows of all the cameras.

[0012] The assisting visual apparatus for mooring in accordance with the invention provides a calibration procedure 100, to be performed during the installation, and a rendering procedure 200, in real time, which may comprise three various modes together or alternatively with each other: an orthographical view or top view, fusion of all the cameras (TC1, TC2, TC3), a perspective view, merging of two cameras (TC1, TC2) and a direct view, from one camera (TC3). In particular, certain or all the steps in the calibration procedure 100 and rendering procedure 200 are performed by the application software installed in memory 27 of device 20.

[0013] The calibration procedure comprises the succession of the following steps: a step 101 for calculating the intrinsic parameters of the cameras TC1-TC3, a step 102 for removing the distortion, a step 103 for calculating the homographies of the reference plane, a step 104 for estimating the fusion maps and lastly a step 105 for estimating the maps for the photometric normalization. The calibration procedure is only performed once, in the installation step.

[0014] For the mooring assisting, the frames of interest are:

- a top view using the platform plane as a reference plane, union of all cameras (TC1, TC2 and TC3), such as to allow the rear and side distance from the other ships and from the quay to be visually estimated;
- a single perspective view given by the merging of the cameras installed at the same height and with the same vertical tilt (TC1 and TC2). In this case, the reference plane is parallel to the two cameras, at a defined distance;
- a single view (TC3) of the rear of the ship.

[0015] If on the one hand the top view allows a global view of the scene, on the other it determines a perspective deformation of the objects which may cause difficulties in interpreting the scene for the operator. The use of some other views allows this problem to be obviated in the event these difficulties arise.

[0016] Step 101 comprises the calculation of the intrinsic parameters of cameras TC1-TC3 by means of a calibration algorithm known in literature, by acquiring a series of images from a "planar grid" (known e.g. from the article "Flexible Camera Calibration by Viewing a Plane from Unknown Orientations", Zhang, in Proceedings of the International Conference on Computer Vision, pp. 666-673, 1999). The intrinsic parameters of the cameras comprise the focal length, the main point and the distortion parameters.

The calculation of the intrinsic parameters may also be performed prior to the installation of the apparatus on the ship and thus might not be performed by the application software installed in memory 27 of device 20.

[0017] Once the intrinsic parameters are known, the distortion may be removed in such a way that the relation linking points of the three-dimensional world and image points may be expressed, according to the "pinhole geometrical model" (known e.g. from the book "Multiple View Geometry in Computer Vision", Hartley and Zisserman, pp. 153-156, Cambridge University Press, 2004), by means of a matrix known as "of projection".

[0018] In step 102, the distortion is modelled by means of a non-linear transformation of the ideal coordinates (not distorted) into the notable true coordinates (distorted). By means of a known system such as "backward mapping" (e.g. known from the book "Digital Image Warping", Wolberg, p. 43, Wiley-IEEE Computer Society Press, 1990), a scan of the non-distorted destination image allows the distortion of the image acquired by cameras TC1-TC3 to be compensated for. Step 102 may also be performed also prior to the installation of the apparatus on the ship and thus might not be performed by the application software installed in memory 27 of device 20.

[0019] Step 103 comprises the calculation of the homographies of the reference plane; step 103 is performed by the application software installed in memory 27 of device 20. According to the pinhole geometrical model, it is known from literature that a plane in space and the perspective image thereof are linked by a linear relation called homography. Given the homography, it is possible to generate a new summary view of the plane from an arbitrary point of view. The result is equivalent to placing a virtual camera in that point, thus generating a virtual image or view.

[0020] In accordance with an embodiment of the invention, the calculation of the homographies with respect to the reference plane requires the assistance of a planar chessboard (calibration grid) that is large enough to be seen in all the cameras. The homography of one spatial plane is calculated with a group of relations (at least 4) of three-dimensional points (3D) of the plane and points in the image. The grid should be placed on the spatial plane intended to be used (hence the plane of platform 25, for example). Due to auto-obstructions, the grid might not be viewed by all the cameras. To obviate this problem, the following solution is implemented: the calibration grid is positioned on a planes π_r which is visible to all cameras TC1-TC3 and the homography H_r of that plane is estimated for all cameras.

[0021] The 3×4 projection matrix of the camera indicates how the 3D points are projected in the image. According to the pinhole model, said matrix has the shape $P = K [R \ t]$, where K is the 3×3 matrix of the intrinsic parameters and $R = [r_1 \ r_2 \ r_3]$ and t are the 3×3 rotation matrix and the 3×1 translation vector, respectively, which indicate the position of the camera with respect to the

reference system. By fixing the reference on the plane and assuming that said plane has the equation $Z=0$, homography H_r of plane π_r estimated using the calibration grid may hence be expressed according to the equation $H_r = K [r_1 \ r_2 \ t]$.

[0022] Writing $H_r = [h_1 \ h_2 \ h_3]$ obtains $r_1 = K^{-1}h_1$ and $r_2 = K^{-1}h_2$; r_3 is estimated while considering the constraint of orthogonality of r_3 with respect to the plane determined by r_1 and r_2 (R is composed of orthonormal columns as it is a rotation matrix,), that is $r_3 = (r_1 \times r_2)$. Given H_r , the projection matrix P is hence now capable of being obtained.

[0023] Knowing the spatial relation (the roto-translation in the 3D space) between π_r and the plane π_d of interest, it is possible to transform these homographies so that they are mapped on π_d instead of on π_r .

[0024] Being

$$G_e = \begin{bmatrix} R_e & t_e \\ 0 & 1 \end{bmatrix}$$

the 4x4 roto-translation matrix that relates π_r to π_d . By applying this transformation to matrix P , the position of the camera is obtained with respect to a reference arranged on plane π_d instead of π_r , i.e. $P_d = P \cdot G_e$ is calculated. From here, H_d — the homography of plane π_d — is obtained from P_d by taking the first two columns and the last column of the matrix, i.e. $H_d = [p_1 \ p_2 \ p_4]$ (as noted in "Multiple View Geometry in Computer Vision", Hartley and Zisserman, p. 196, Cambridge University Press, 2004).

[0025] This allows the calculation of the homography to be performed for one individual plane, on which the calibration grid is clearly visible from all the cameras, and then to transform them into homographies with respect to any other spatial plane. Based on the required points of view, the reference planes are the plane of the platform, the common image plane of the cameras arranged parallel and at the same height (TC1 and TC2) and the image plane itself of the individual camera (TC3), respectively.

[0026] In accordance with a variant of the embodiment of the invention, if the superimposition between views is not such to be able to display a same common grid for all the cameras TC1-TC3, it is possible to extend the above-described system thus determining the homography of all cameras with respect to a different convenient spatial plane to be estimated and i.e. the planes π_{TC1} , π_{TC2} , π_{TC3} . Knowing the spatial relation of these planes with respect to the reference plane π_d , the homographies are transformed so that they are directly mapped on plane π_d .

[0027] In accordance with another variant of the embodiment of the invention, the use of the calibration grid is entirely eliminated. A "structure-and-motion" algorithm

is used which allows the position to be obtained of the cameras with respect to the ship (motion) and a three-dimensional reconstruction by points (structure) of the framed scene. The homographies may be directly calculated with this information and if the position of the planes of interest with respect to the ship is known.

[0028] Similarly to the formation of planar mosaics, the homographies of a plane itself in space deriving from various cameras may be used for mapping various views in a single view (planar mosaic), thus obtaining an enlargement of the global visual field from the desired point of view. Since the cameras are fixed, once the homographies of the plane have been calculated and how the views merge to form the final view has been established, i.e. the fusion map has been established, the relation that maps the individual views in the single view remains constant.

[0029] Step 104 comprises the estimation of the fusion maps which, in the rendering step 200, allow the summary views to be generated in real time; step 104 is performed by the application software installed in memory 27 of device 20.

[0030] There is a need to establish the position of the virtual camera with respect to the reference plane. The chosen configuration results in the virtual view being exactly parallel to the plane at a distance Dh for the plane of the platform, for the common image plane of the cameras arranged parallel and at the same height (TC1 and TC2) and the image plane itself of the individual camera (TC3). The orthographical view is the top view and merges together the informative contents of all the cameras. The desired specifications to be displayed relating to the area about the ship and the distance between the stern end of ship 1 and the halfway point of the distance between camera TC3 and the straight line joining cameras TC1 and TC2 allow distance Dh to be defined to arrange the virtual camera to display the virtual visual field. In the case of image plane of the individual camera, distance $Dh=0$.

[0031] Given the dimension $w \times h$ of the virtual view I_{out} (the planar mosaic):

- I_{out} is divided into regions $Rtc1$, $Rtc2$, $Rtc3$ thus establishing for all regions, on the basis of the position of the cameras, which video flow to show. Figures 6 and 7 show the division in the case of the perspective view and of the orthographical view;
- an edge of pixel dimension b is defined, at the periphery of the regions, which depicts the fusion area between various views. In figures 6 and 7, the dotted edge shows the periphery zone;
- a weight mask F is formed, which assigns variable values between 0.5 and 1 to the edge zones and value equal to 1 to the remaining zones.

[0032] For all x, y positions in the planar mosaic or virtual view I_{out} , the following is registered in the fusion map M :

— the region to which the pixel belongs, i.e. camera TC_i that sees that area;
 — coordinates u, v in the image deriving from camera TC_i, corresponding to x, y : first, the inverted homography is applied to determine position u', v' in the non-distorted image deriving from camera TC_i, then the distortion to determine position u, v in the original distorted image deriving from camera TC_i;
 — if $F(x, y)$ is other than 1, i.e. we are on the edge at the periphery between the images deriving from two different cameras TC_i and TC_j, the coordinates h, k in the image deriving from camera TC_j corresponding to x, y are to be calculated as previously described.

[0033] Lastly, the virtual view I_{out} is reflected with respect to the central axis of the image, to facilitate the reversing of the ship.

[0034] This information constitutes the fusion map M , which will be specific for all rendering modes. The map allows the summary views to be generated for the different views by directly using the video flows (hence without having to remove the distortion).

[0035] Step 105, that is the estimation of the fusion maps for the photometric normalization, allows any brightness differences to be corrected between the video flows; step 105 is performed by the application software installed in memory 27 of device 20. For the photometric normalization, it is required to determine the multiplicative gain factors α_i to be applied to the images so that these appear with the same brightness. This estimation is to be performed in the rendering step, since the brightness varies based on the lighting conditions. To avoid estimating the brightness of an image using all the pixels in that image, the coordinates are registered of a common periphery area between all the various cameras, i.e. for all points of coordinates x, y in aforesaid area, the corresponding distorted coordinates u, v are registered for all images deriving from cameras TC1-TC3. The obtained map allows the pixels for all images belonging to this area of intersection to be quickly recovered in the rendering step, to be used for estimating the gain factors.

[0036] The rendering procedure 200, successive to the calibration procedure 100, is performed by the application software installed in memory 27 of device 20 and comprises the following steps: a step 201 for the synchronized acquisition of the video flows from all the cameras, a step 202 in which, by using the fusion map for the photometric normalization, the average brightness is determined of all cameras and the gain factors α_i are calculated which allow the cameras to be arranged at the same brightness by using one as a reference, a step 203 in which, on the basis of the selected rendering mode, the corresponding fusion map is selected and a step 204 in which image I_o is formed by considering the values of the pixels from the corresponding video flows, on the basis of what is registered in the fusion map and of the gain factors α_i .

Claims

1. Assisting visual apparatus for mooring a ship (1), said apparatus comprising at least one pair of shooting means (TC1-TC3) arranged at given heights (H_1, H_2) on the ship so as to have a visual field spanning the part surrounding the ship stern and a part of the same ship, an image control device (20) comprising a memory (27), a data processing unit (28) and an application software and it being adapted to combine the single images shoot by the shooting means (TC1-TC3), a display (24) on which the final image (I_o) processed by the image control device is displayed in real time, **characterized in that** said application software of the image control device comprises a calibration function (100) of the shooting means which includes, for each shooting means, the calculation (103) of the homographies with respect to the selected reference plane (π_d), the generation of a virtual image (I_{out}) as combination of the single images deriving from the shooting means on the base of the calculated homographies, the formation of at least one fusion map (M) on the base of the virtual image (I_{out}), said software comprising a successive function of rendering (200) including the formation of the final image (I_o) which has to be displayed on the display as a function of said at least one fusion map (M).
2. Apparatus according to claim 1, **characterised in that** said shooting means of said pair of shooting means (TC1-TC3) are arranged at equal height on the ship.
3. Apparatus according to claim 1, **characterised by** comprising at least three shooting means, two shooting means (TC1-TC2) are arranged at equal height on the ship and the other shooting means (TC3) is arranged at a different height on the ship.
4. Apparatus according to claim 1, **characterised in that** said calibration function (100) comprises, after the calculation of the intrinsic parameters of the shooting means (101) and the removal of the distortion in the images deriving from the shooting means (102), said calculation (103) of the homographies with respect to the selected reference plane (π_d), the formation of at least one fusion map (M) and the estimation of the maps for the photometric normalization (105).
5. Apparatus according to claim 4, **characterised in that** said calculation (103) of the homographies with respect to the selected reference plane (π_d) comprises the assistance of a calibration grid arranged on a further plane (π_r) which is visible to all the shooting means, the calculation of the homographies with respect to said further plane (π_r) in which the calibration

grid is visible and the transformation of the calculated homographies into homographies with respect to the reference plane (π_d).

6. Apparatus according to claim 4, **characterised in that** said calculation (103) of the homographies with respect to the selected reference plane (π_d) comprises the determination of the homography of each shooting means with respect to a respective spatial plane (π_{TC1} , π_{TC2} , π_{TC3}) and the transformation of all the homographies so as to map them on the selected reference plane (π_d). 5

7. Apparatus according to claim 4, **characterised in that** the formation of a fusion map (M) comprises the determination of the position of the virtual shooting means with respect to the reference plane, the division of the virtual image (I_{out}), which derives from the virtual shooting means, into regions (Rtc1, Rtc2, Rtc3) by stabilizing for each region, on the base of the position of the shooting means (TC1-TC3), which image deriving from the shooting means (TC1-TC3) must be showed, the definition of an edge (b) in the periphery of said regions which represents the fusion area of the different images deriving from the shooting means and the formation of a weight mask (F) which assigns values comprises between 0.5 and 1 to the zones next to the edge and value 1 to the remaining zones, the registration of the region to which the pixel belongs for each planar position of the pixel (x, y) of the virtual image (I_{out}), the determination of the coordinates (u, v), in the image deriving from the shooting means, which correspond to the planar position of the pixel (x, y) by applying first the inverted homography for determining the position (u', v') in the non-distorted image and then applying the distortion for determining the position (u, v) in the original distorted image, the application to the virtual image of the reflection with respect to the central axis of the virtual image. 10 15 20 25 30 35 40

8. Apparatus according to claim 7, **characterised in that** the estimation of the maps for the photometric normalization comprises the determination of a superimposition area of regions belonging to images deriving from different shooting means and the registration, for each pixel of this area, of the corresponding coordinates (u, v) of each distorted image deriving from the respective shooting means, said estimation being used for the calculation of the gain factors (α_i) for each shooting means. 45 50

9. Apparatus according to claim 1, **characterised in that** said rendering function comprises the synchronized acquisition of the images deriving from the shooting means, the determination of the medium brightness of each shooting means in the base of the fusion map for the photometric normalization 55

which has been used and the calculation of the multiplicative gain factors (α_i) which allow to arrange the shooting means (TC1-TC3) at the same brightness, the formation of the final image (I_o) as a function of the values of the pixels deriving from the shooting means on the base of the fusion map (M) and said calculated gain factors (α_i).

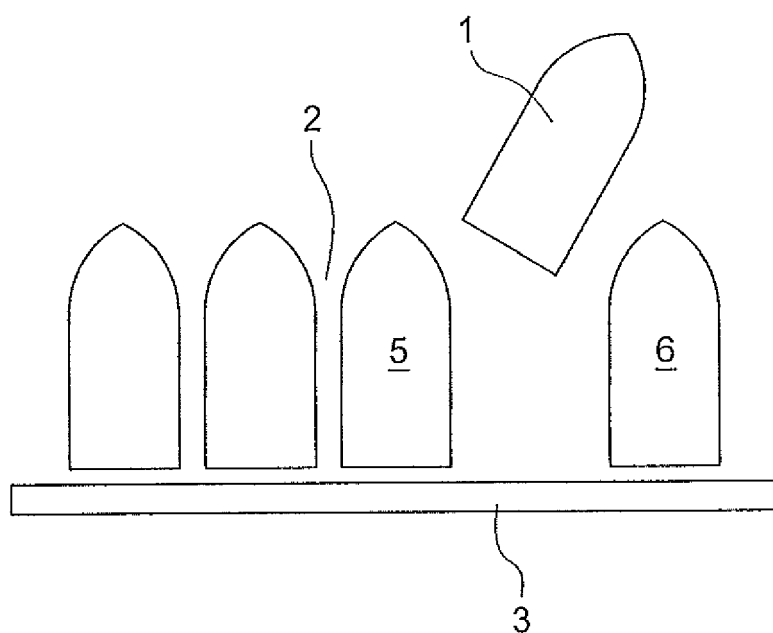


Fig.1

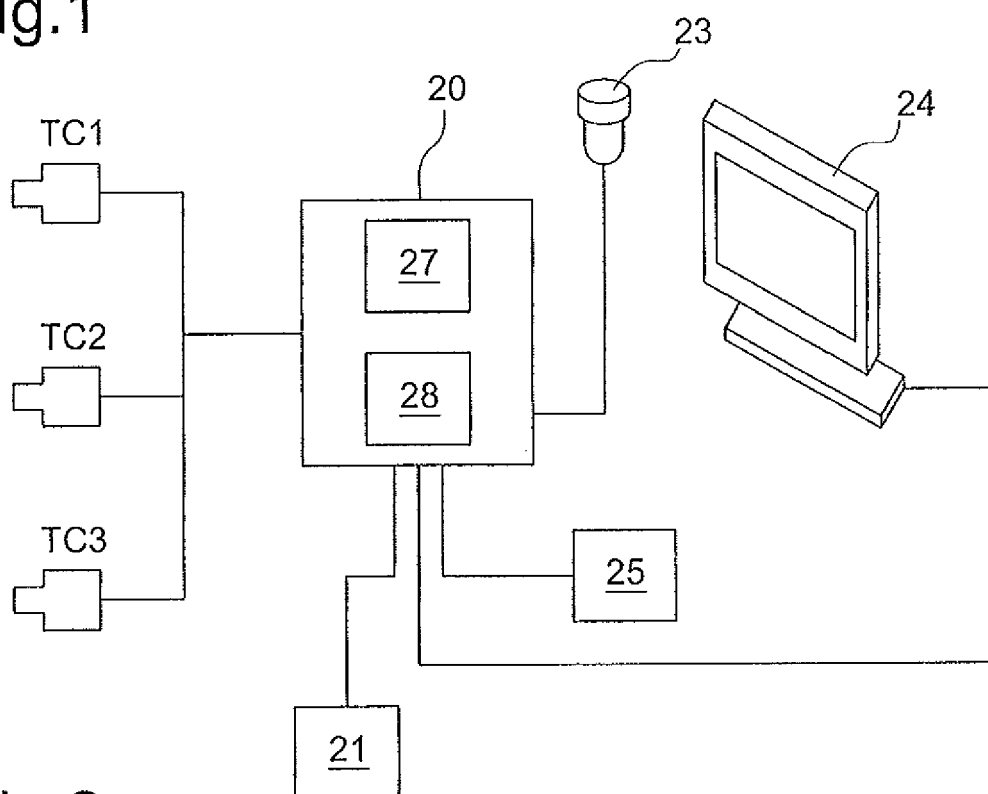


Fig.2

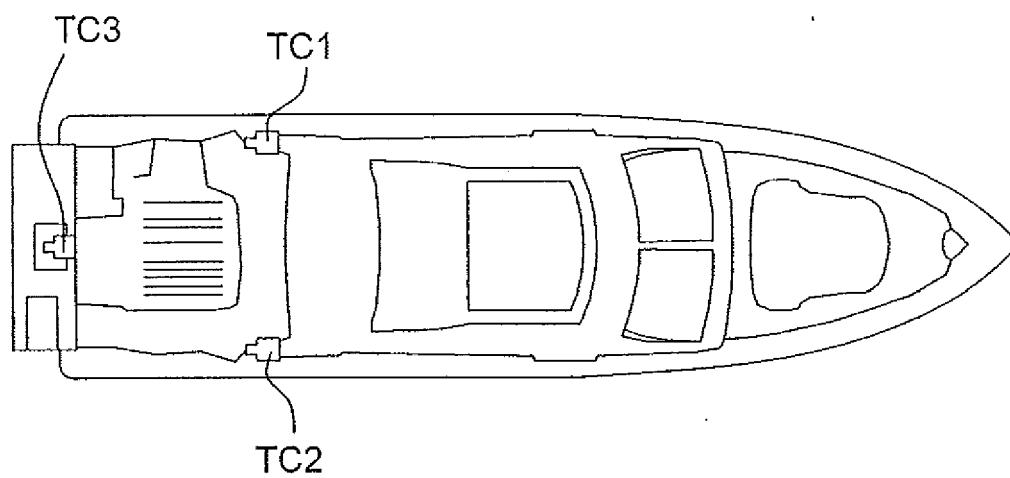


Fig.3

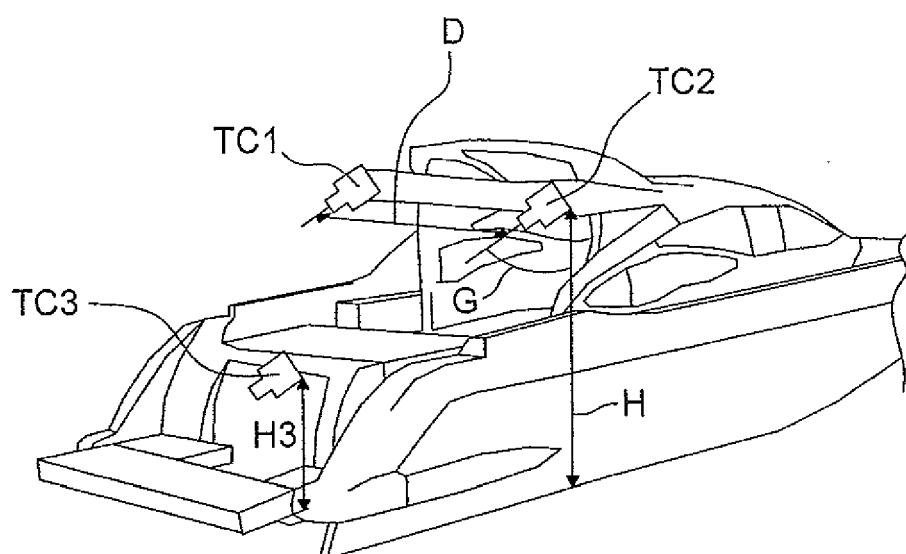


Fig.4

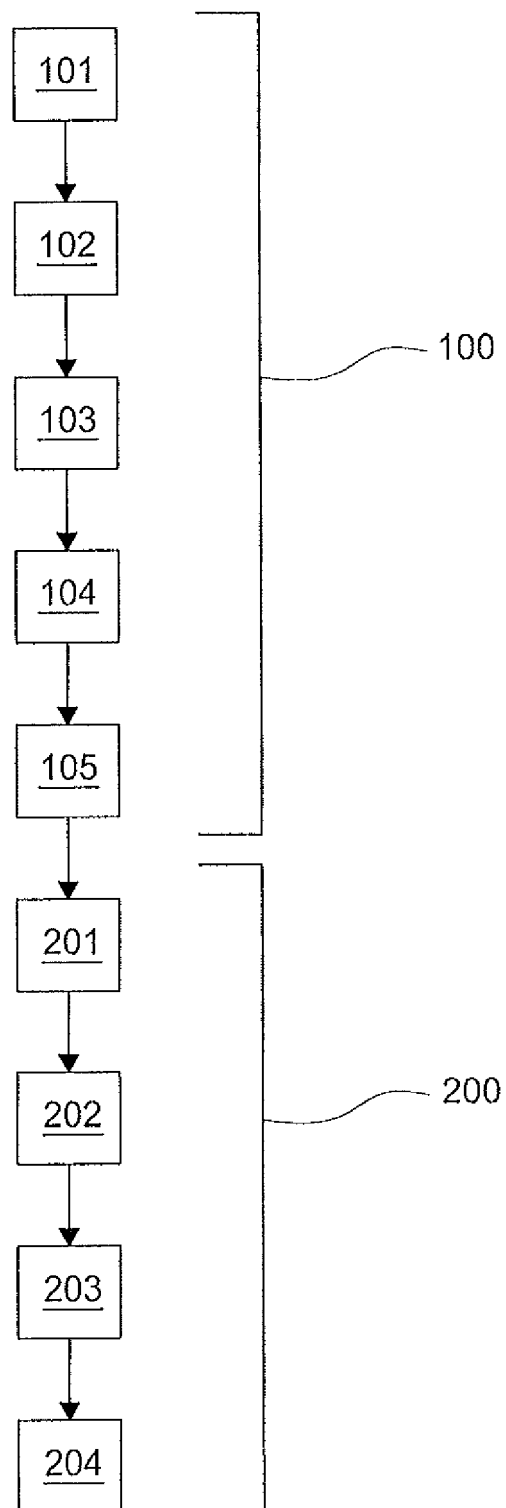


Fig.5

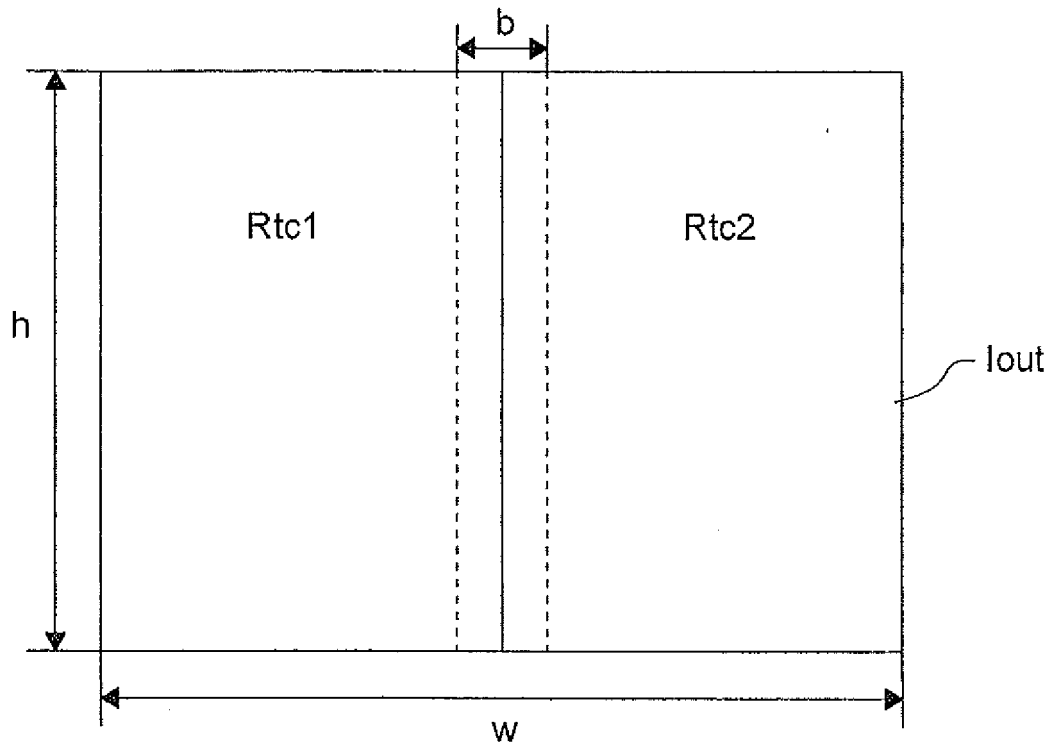


Fig. 6

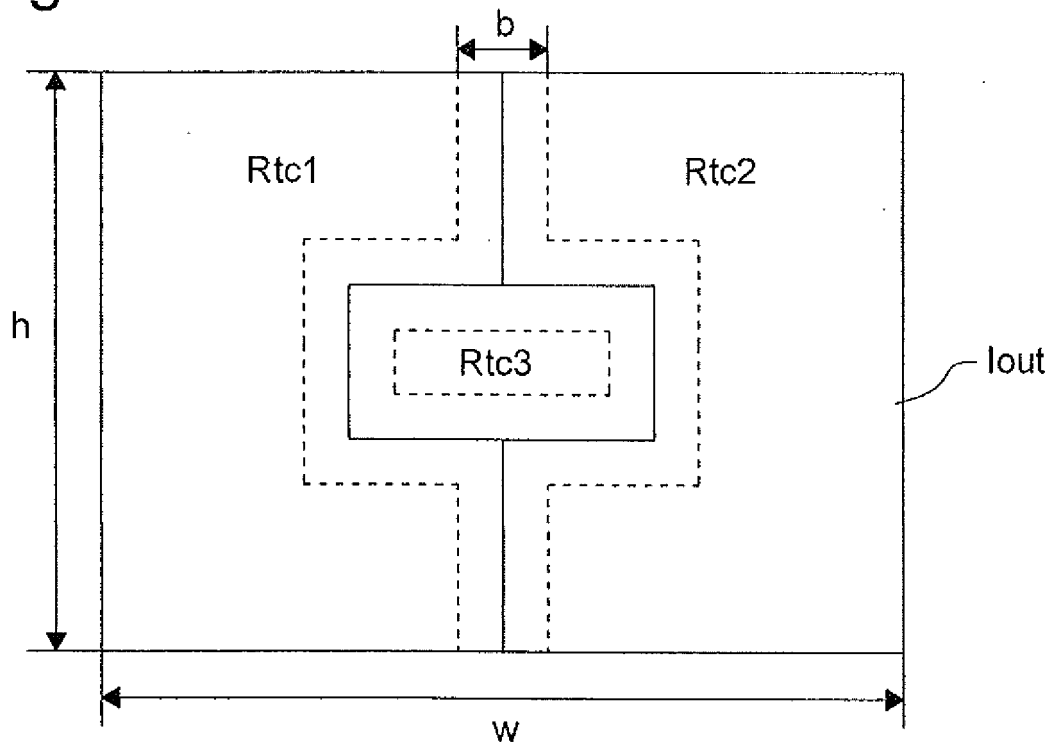


Fig. 7



EUROPEAN SEARCH REPORT

Application Number
EP 11 18 2355

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (IPC)
A	US 2007/073454 A1 (KAJI HIROTAKA [JP] ET AL) 29 March 2007 (2007-03-29) * the whole document * -----	1	INV. B63B21/00 G05D1/02
			TECHNICAL FIELDS SEARCHED (IPC)
			B63B G05D
The present search report has been drawn up for all claims			
Place of search The Hague		Date of completion of the search 6 March 2012	Examiner De Sena Hernandorena
CATEGORY OF CITED DOCUMENTS X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons & : member of the same patent family, corresponding document			

1
EPO FORM 1503 03.82 (P04C01)

06-03-2012

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2007073454 A1	29-03-2007	NONE	

EPO FORM P0459

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Non-patent literature cited in the description

- **ZHANG.** Flexible Camera Calibration by Viewing a Plane from Unknown Orientations. *Proceedings of the International Conference on Computer Vision*, 1999, 666-673 **[0016]**
- **HARTLEY ; ZISSERMAN.** Multiple View Geometry in Computer Vision. Cambridge University Press, 2004, 153-156 **[0017]**
- **WOLBERG.** Digital Image Warping. Wiley-IEEE Computer Society Press, 1990, 43 **[0018]**
- **HARTLEY ; ZISSERMAN.** Multiple View Geometry in Computer Vision. Cambridge University Press, 2004, 196 **[0024]**