(12) PATENT (11) Application No. AU 199669352 B2
(19) AUSTRALIAN PATENT OFFICE (10) Patent No. 714317

(54) Title
Method and apparatus for automatic electronic replacement of billboards in a video image

(51)[6] International Patent Classification(s)
H04N 005/272 H04N 009/75
H04N 005/262

(21) Application No: 199669352 (22) Application Date: 1996.09.09

(87) WIPO No: WO97/09823

(30) Priority Data

(31) Number	(32) Date	(33) Country
9518439	1995.09.08	GB
9601101	1996.01.19	GB

(43) Publication Date : 1997.03.27
(43) Publication Journal Date : 1997.05.22
(44) Accepted Journal Date : 2000.01.06

(71) Applicant(s)
Orad Hi-Tec Systems Limited

(72) Inventor(s)
Itzhak Wilf; Avi Sharir; Michael Tamir

(74) Agent/Attorney
GRIFFITH HACK,GPO Box 1285K,MELBOURNE VIC 3001

(56) Related Art
WO 94/05118
US 5353392
US 4393394

OPI DATE 27/03/97 APPLN. ID 69352/96
AOJP DATE 22/05/97 PCT NUMBER PCT/GB96/02226

AU9669352

(51) International Patent Classification 6 :
H04N 5/272, 9/75, 5/262

A1

(11) International Publication Number: WO 97/09823

(43) International Publication Date: 13 March 1997 (13.03.97)

(21) International Application Number: PCT/GB96/02226

(22) International Filing Date: 9 September 1996 (09.09.96)

(30) Priority Data:

9518439.6	8 September 1995 (08.09.95)	GB
9601101.0	19 January 1996 (19.01.96)	GB

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

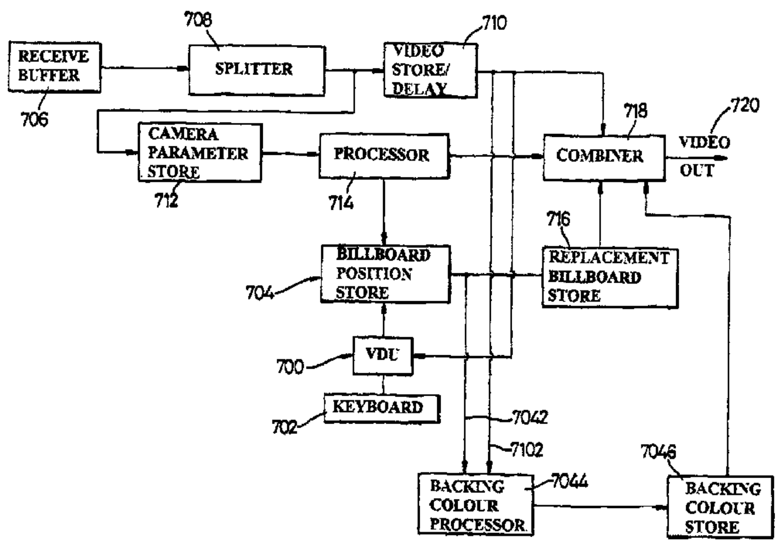
(71) Applicant *(for all designated States except US)*: ORAD HI-TEC SYSTEMS LIMITED [IL/IL]; Savir CTR, 14 Hata'as Street, P.O. Box 2177, 44425 Kfar Saba (IL).

(71) Applicant *(for KG KZ RU SI SK UZ only)*: GOODMAN, Christopher [GB/GB]; 22 Village Close, Edwalton, Nottingham NG12 4AD (GB).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: WILF, Itzhak [IL/IL]; 263 Ben-Gurion Road, 47212 Ramat-Gan (IL). SHARIR, Avi [IL/IL]; 21 Ani Maamin Street, 67727 Ramat Hasharon (IL). TAMIR, Michael [IL/IL]; 13 Beit-Tsuri Street, 69122 Tel-Aviv (IL).

(74) Agent: GILDING, Martin; Eric Potter Clarkson, St. Mary's Court, St. Mary's Gate, Nottingham NG1 1LE (GB).

(54) Title: METHOD AND APPARATUS FOR AUTOMATIC ELECTRONIC REPLACEMENT OF BILLBOARDS IN A VIDEO IMAGE



(57) Abstract

Apparatus for automatic electronic replacement of a billboard in a video image including an automatic camera orientation measurement apparatus including motion measurement means operative to measure the Field of View (FOV) of the TV camera relative to a known reference position.

METHOD AND APPARATUS FOR AUTOMATIC ELECTRONIC REPLACEMENT OF BILLBOARDS IN A VIDEO IMAGE

The present invention relates to a method and apparatus for automatically replacing billboards in a video image.

The present invention has particular use in electronic replacement of billboards in a stadium or other venue but can be used to provide accurate data relating to camera orientation for other purposes.

In previous systems it has been proposed to electronically replace billboards in a stadium which are viewed by a viewer on television. The billboards in the stadium are televised by a TV camera and the boards are electronically altered so that the TV viewer at home sees a different board to the spectator in the stadium or other venue.

The known systems such as those described in US 5,266,933, an apparatus and method for electronically altering video images is disclosed. The apparatus and method disclosed in the US patent and also in US patent 5,353,392 whilst theoretically allowing replacement of billboards do not solve the many practical problems encountered in real environments. Most of these problems are related to the recognition and replacement processes.

Relying entirely on pattern recognition techniques which utilize only the video signal to identify and localise billboards for replacement introduces major problems which affect the practical value of such a system.

Clearly, any pattern recognition scheme, including those described in US 5,264,933 and US 5,353,392 must rely on useful visible features in the image that can be compared with pre-defined descriptions. Such features should be located inside the billboard or at its neighbourhood.

In realistic situations, the visibility of these features might change, continuously or otherwise from practically zero to a some threshold visibility which allows the pattern recognition scheme to work properly. These changes can occur in the direction of growing or reducing visibility.

Such situations include :

- Acceleration or de-acceleration of camera motion introducing a huge amount of blur.
- Excessive zooming-in or zooming-out of the billboard.
- Excessive occlusion by players.
- Entering or exiting a camera's field of view by any combination of pan, tilt and zoom operations.
- Any combination of the above mentioned mechanisms.

Therefore, in practical situations, a continuous replacement of billboards, is not possible. Even if an interrupted replacement was allowed, it would require a delay of at least a few seconds to decide whether the resulting replacement interval is acceptable or not. Such a delay is usually not permitted in live broadcasting of sports events.

Replacing arbitrary billboards introduces further problems. A seamless replacement requires to identify the foreground objects occluding the billboard in order to inhibit replacement at places of occlusion. Foreground objects mainly consist of players but also the ball or other objects. Consider now a player with a red shirt, occluding a part of a similarly red portion of a billboard. Colour contrast cannot be used robustly to identify occlusion. Furthermore, since the player is a non-rigid object, motion or shape information cannot be used accurately enough to guarantee perfect replacement.

Another problem which may arise in practical situation is resolution of billboard identity. Consider two identical billboards positioned at two different locations in the arena. Suppose different replacement billboards are assigned to each of these physical billboards, then one must be able to tell which one is which. This can prove to be extremely difficult especially if no unambiguous features are visible.

WO 94/05118 discloses a system for derivation of a studio camera position and motion from a camera image. The system calculates the position of a camera in a virtual studio from a knowledge of the initial position of the camera and an analysis of the image of a checkerboard panel viewed by the camera.

The present invention provides apparatus for automatic electronic replacement of a billboard in a video image including automatic camera orientation measurement apparatus including motion measurement means operative to measure the Field of View (FOV) of a TV camera relative to a known reference position;

said apparatus further including a dynamic billboard memory said dynamic billboard memory including:

a) static image data recording means recording data indicating the identity of said billboard,

b) static set-up data recording means recording data indicating the static position of the billboard with the camera in a static position with said billboard in full view in said Field of View of said camera,

c) dynamic calibration data recording means said dynamic calibration recording means including:

i) left to right pan data recording means recording data indicating the position of the billboard in the Field of View of the camera when the camera is panned from left to right across the billboard,

ii) right to left pan data recording means recording data indicating the position of the billboard in the Field of View of the camera when the camera is panned from right to left,

iii) tilt up data recording means recording data indicating the position of the billboard in the Field of View of the camera when the camera is tilted in a direction from below the billboard to above the billboard; and

iv) tilt down recording means recording data indicating the position of the billboard in the Field of View of the camera when the camera is tilted in a direction from above the billboard to below the billboard.

The present invention also provides a method for automatic replacement of a billboard in a video image including the steps of:

a) locating said billboard in a known position relative to a TV camera,

b) determining a position for said billboard when said TV camera is static and storing said position in a first store,

c) determining a plurality of dynamic

positional data for said billboard as said TV camera is panned from left to right, right to left, tilted up and tilted down and storing said dynamic data in respective dynamic data stores,

d) storing a replacement virtual billboard in a replacement billboard store,

e) replacing said billboard in a picture taken by said TV camera using data from said static data store and from selected dynamic data stores dependent on detected movement of said TV camera.

Embodiments of the present invention will now be described, by way of example with reference to the accompanying drawings in which:-

PAGES 6 TO 12 ARE INTENTIONALLY BLANK

Figure 1 shows a stadium or other venue illustrating the apparatus according to the present invention;

Figure 2 shows the video image of the stadium as seen by the camera in a first position;

Figure 3 illustrates a stadium with billboards in several different positions;

Figure 4 illustrates a zoomed camera shot of a billboard illustrating the problem with prior art systems;

Figure 5 shows a patterned chroma-key billboard for use with the present invention;

Figure 6 shows in block diagrammatic form circuitry associated with the camera arrangement of Figure 1 for transmitting video data and camera orientation data;

Figure 7 shows in block diagrammatic form receiver circuitry for co-operation with the transmitter circuitry of Figure 6;

Figure 8 shows a flow diagram for the operation of the circuitry of Figure 7;

Figure 9 shows an arrangement for a billboard setup data store;

Figure 10 shows an arrangement for billboard setup data;

Figure 11 shows a flow diagram for perspective transformation computation;

Figure 12 shows an arrangement for camera intrinsic parameters store;

Figure 13 shows equations for dynamic recalibration;

Figure 14 shows a flow diagram for dynamic recalibration;

Figure 15 shows the process of recalibration;

Figure 16 illustrates the problems relating to siting of billboards at different locations within a stadium with differing lighting conditions;

Figure 17 shows a graph of minimum and maximum levels for U & V illustrating the operation of a chroma-keyer;

Figure 18 shows a billboard with occluding object illustrating the principle of adjustment of the chroma-key colour for a billboard;

Figure 19 shows an exemplary remote receiver for reception of billboard coordinate data and perfect chroma-key colour, occlusion being effected by chroma-key techniques;

Figure 20 shows an alternative arrangement for the billboard set up data store illustrating an alternative embodiment;

Figure 21 shows a flow diagram for dynamic set up procedure for camera panning for use with the billboard set up data store of Figure 20; and

Figure 22 shows a flow diagram for dynamic set up procedure for camera tilting for use with the billboard set up data store of Figure 20.

With reference now to Figures 1 to 4 the principle of the present invention is now explained.

In a stadium or other venue 10 billboards 14,16,18 are installed at the side of a pitch represented by markings 12. These billboards are viewable by a camera 20. Billboards 15,17,19 may be present on the opposite side of the stadium for viewing by a further camera 21. The stadium terraces/seating are shown diagrammatically by the lines 11.

Camera 21 may in a preferred example be a normal TV video camera and will transmit its output video signal directly to a first feed which may serve the local population. Although we refer to camera 20 or 21, it may be clearly understood that there could be a plurality of cameras on each side of the stadium providing differing views.

Camera 21 in a preferred embodiment will televise boards 15,17,19 which will be transmitted to the local population in an unchanged manner.

Camera 20 will in this preferred embodiment transmit a feed to an international audience. Camera 20 is equipped with orientation sensing means which preferably comprises one or more of the following :

pan measurement means 24;
tilt measurement means 25;
zoom measurement means 26; and
focus measurement means 28.

Suitable sensors may comprise the Virtual Reality Encoder from RADAMEC EPO, Bridge Road, Chertsey, Surrey KT16, 8LJ, England.

Dependent on the allowed mobility of the camera only one, several or all of these may be required. For example, if camera 20 is fixed in pan and tilt and focus and can only zoom, as in the case of some remotely controlled unmanned cameras then only the zoom parameter need be measured.

Most cameras in sports stadiums can zoom, tilt and pan and it is assumed that these parameters are measured for each camera as now explained. The focus is assumed to be fixed but in similar manner the parameter could be added if required.

Figure 2 shows the video image as seen by a viewer and in particular by the operator of the equipment. The camera 20 is zoomed, panned and/or tilted to "centre" the billboard 14 in a suitable position and at a reasonable size. With reference to Figure 7 each billboard is then viewed at a receiver and its position is marked preferably by using a touch screen 700, or keyboard mouse 702 and marking the four corners. The positions are stored in a store 704.

For billboards higher in the stadium such as 30 (Figure 3) a correction factor for the camera may be stored dependent on the tilt position of the camera.

Each billboard position is stored in store 704 together with the camera parameter information at the reference position for the camera 20 obtained from the camera parameter information which is correct at the time that the billboard position is stored.

The following procedure is preferably repeated for each of the cameras and for each of the target billboard :

1. Point the camera at the target to obtain a stable unoccluded view of the target. Adjust the zoom to get a large view of the target yet keeping the whole target within the field of view.
2. While the camera is not moving trigger an acquisition device, to grab a picture of the target, as well as the corresponding readings of the sensors.
3. Mark the corners of the target, on the video image.

Preferably a corner detector is used to pinpoint the corners of the target at sub-pixel precision.

This camera parameter information is obtained (Figure 6) from the sensors mounted on the camera and the camera movement is referenced to a first or fixed reference position for each parameter. The movements of the camera are sensed and the signals are fed into a combiner circuit 24 and then to a transmit buffer 36 from which the combined video and position data signals are transmitted.

During set up, at the receiver (Figure 7) the receive buffer 706 receives the signals and feeds these to a splitter 708. The video signal is stored and delayed in a suitable store 710 and the camera parameter data is extracted and stored in store 712.

In set up the VDU 700 is used to mark each billboard that may require replacement. The camera 20 is panned etc to move each billboard into a suitable position on the screen and its position is recorded in the billboard store 704 together with the camera parameters obtained from store 712 via processor 714.

A replacement billboard store 716 stores a plurality of replacement billboards and these are selectable to be able to replace the original billboard.

The replacement billboard is in operation inserted into the video signal in a combiner 718 to provide a modified output video signal 720.

The setting up procedure can also identify billboard locations and camera parameters for several cameras by storing a camera ID from a source 30 (figure 6). Thus, billboard position store 704 will store separate lists of billboard data for each camera.

The operation of the system will now be described with reference to a single billboard and a single camera 20.

With reference to Figure 4, it is assumed that billboard 14 enters the field of view in an enlarged form on the left-hand side of the screen as camera 20 pans following zooming from the Figure 3 position.

The camera orientation data is constantly being received by the receiver and the processor 714 will constantly match on a pixel by pixel basis the video image with the known billboard position stored in store 704. As soon as the billboard appears in the video image the pixels representing the billboard will be identified and the replacement billboard pixels which relate to those pixels will be substituted in the combiner 718. The delay will be minimal since the identification of the pixels is by an address correlation process which will be virtually instantaneous.

After a period of time the camera sensors may drift and in this case the replacement billboard may not exactly align with the original. This may only be by one or two pixels and may not be discernible to the viewer. To correct this two solutions are possible. Firstly, the billboard position can be manually restored periodically at a suitable time, for example, when a camera is not active. This requires the co-operation of the operator.

Secondly, a comparison can be made on a pixel by pixel basis of the billboard against an original stored billboard and an adjustment of the reference camera parameters can be made in billboard position store 704. This process can be done automatically at either set intervals or when the processor 714 has a suitable time slot.

The essential steps of a preferred recalibration process are to perspectively transform the current video image using the camera data to provide an estimated transformed model. A stored image of the billboard is then compared with the transformed model to provide a residual video field. The residual distortion between the transformed model and the residual video field is resolved to provide updating information for updating the estimated transformation and to thereby provide a calibration correction factor for recalibrating the position of each billboard in the store in accordance with the camera sense information.

The replacement of each billboard is accomplished by use of the processor 714 (Fig. 7) and the various parameter and billboard stores using appropriate software programmes as now described in more detail.

Figure 8 describes the complete process which allows to determine the position of each billboard in the camera's field of view, and render the corresponding part of the billboard into the frame buffer. Since the rendering and later the compositing of the graphics buffer with the video buffer by means of chroma-key are known art, we will concentrate on the billboard position determination with reference also to Figures 6 and 7.

At the beginning of each video field, the pan, tilt, zoom and focus sensors (24,25,26) are read 800. These values, combined with billboard data from billboard setup data store 704 and camera data from camera intrinsic parameters store 712, enable the detection and recognition of all billboards in the camera's FOV, independently of the video signal. The processing of Fig. 1 consists of a loop on all billboards (m) 802,804. For each billboard, its setup data is retrieved 806 from billboard setup data store 704 and used with camera intrinsic parameters 808 to compute the perspective transformation 810 from billboard m to current field. The replacement billboard information is then stored (812) in a frame buffer.

Fig. 9 describes the billboard setup data store 900 which consists of a separate record 902..904 for each billboard in the arena. Such a record consists of a static image 906 grabbed in favourable conditions and of the corresponding static setup data 908. The record also consist of dynamic setup data 910 which is computed using the image processing means in a process known as dynamic re-calibration which has been briefly described above and will be further described with reference to Figure 11. An alternative procedure providing static and dynamic calibration is described with reference to Figures 20, 21 and 22.

Fig. 10 describes the setup data (either static or dynamic) 1000 for a single billboard. It consists of the sensors' readings 1002 at the setup instance, the billboard quadrilateral vertex 1004 coordinates and the time-code of the setup instance 1006.

The method of dynamic re-calibration can be explained as follows:

Due to sensors' drift and inaccuracies, a final calibration table and other practical reasons, it is impossible to predict the exact location of all visible billboards at a given instance. However, at many video fields, a billboard's visibility may be such that an exact geometric position correction can be performed. Since that position is closer both temporally and spatially to the subsequent video fields, it is preferable to rely on that "luck shot" by predicting the billboard position relative to its sensors' readings and exact quad coordinates. Consider for example a billboard which exits the field of view due to camera panning. Having a luck shot while it is still highly visible, allows the smooth tracking of the billboard by sensors only, when its visibility does not allow any image processing means to be applied.

Fig. 11 depicts the flow diagram 1100 for perspective transformation computation. A setup data selection logic 1102 selects either the static 1103 or the dynamic 1105 setup data from the setup data store 806 as described above. This setup data, together with camera intrinsic parameter is used to compute a sensor-based prediction of the perspective transformation 1104, independently of the video signal.

A dynamic re-calibration 1106 based on image processing means is then applied to the prediction. It utilizes the video 1108 and chroma-key 1110 signals as well as the billboards model image 1112 from the setup data store 806 (Fig. 8). Based on a quality factor derived from the image processing means, either the sensors-based 1118 or the corrected transformation 1116 are output. If the estimated quality of the geometric correction is high, then the dynamic setup data is updated 1114.

Figs. 12,13,14 describe the sensor-based prediction of billboard coordinates in the video field. Such a prediction utilizes the sensors reading as well as the camera intrinsic parameters. These parameters are described in Fig. 12 and have to be tabulated for a dense sampling of the (zoom,focus) space. The meaning of these parameters is clear from Figure 13 to which reference is now made.

Let the set of measurements given by the pan, tilt, zoom and focus sensors be represented by the vector (P,T,Z,F). The tilt angle is assumed to be relative to the horizon.

Consider an object point whose image at some setup instance is, at frame-buffer coordinates (x_S, y_S). Let also the sensor measurement vector at that instance be (P_S, T_S, Z_S, F_S).

At another instance, the prediction instance, let the sensor measurements vector be (P_P, T_P, Z_P, F_P). It is required to predict the location of the object point in frame-buffer coordinates (possibly out of the actual frame-buffer), (x_P, y_P).

To enable the procedure we define the setup rotation matrix as shown at 600 and the prediction rotation matrix is defined as shown at 602.

Then, the perspective transformation matrix between the two image-plane coordinate systems is given as shown at 604 and 1402 (Fig. 14).

R_{SP} is a 3*3 matrix with row and column indices ranging from 0 to 2. $R_{SP}[i][j]$ denotes the term in row i, column j in the matrix. Thus, given the setup image-plane coordinates of the object point (u_S, v_S), the predicted location of the object point in image-plane coordinates (u_P. v_P) is given as shown at 606, 1404.

Image-plane to frame-buffer coordinate transformation is achieved as shown at 608, 1406. Aberration compensation is achieved as shown at 608, 1406 (Fig. 14) to provide predicted frame buffer billboard coordinates and perspective transformation data.

An effective way of deriving these parameters for a specific (zoom,focus) pair is described in [J. Weng et al., Calibration of stereo cameras using a non-linear distortion model, IEEE 10th Intl. Conf. Pattern Recognition (1990), pp. 246-253]. The image processing means for geometric correction, which allow also the process of re-calibration, is now described with reference to Fig. 15.

The image processing means for geometric correction of sensors-based prediction are based on the differential method for motion estimation [C. Cafforio and F. Roca, The differential method for motion estimation, in: T.S. Huang, eg., Image sequence processing and dynamic scene analysis, Spring, Berlin, 1983, pp. 104-124]. Let C be the current video field and let M be the static billboard setup image, perspectively transformed according to the sensors-based prediction. Here we consider only luminance images. Ideally, M and C are identical within the support of the billboard quadrilateral. Actual differences may include:

- Occlusion present in C but not in M.
- Geometric errors due to sensors and intrinsic camera parameters errors.
- Luminance changes.

Neglecting for the moment any difference which is not due to geometric errors, consider a point (x,y) inside the support of the billboard quadrilateral. Let (p,q) be the local geometric error then we may write for the luminance signals of the respective images:

$$M(x + p, y + q) = C(x,y)$$

Under the assumption that the error is small, one may write a Tailor series expansion:

$$M(x + p, y + q) = M(x,y) + p\frac{dM}{dx} + q\frac{dM}{dy} + (\textit{Second order terms})$$

Neglecting the second order terms and denoting the spatial derivatives

$$dM/dx = H$$
$$dM/dy = V$$

we obtain

$$C(x,y) - M(x,y) = pH + qV$$

Also denoting the differences C(x,y) - M(x,y) by D we obtain

$$D = pH + qV$$

The equation above holds, locally. For a global billboard solution, and small error assumption we may use the perspective model [G. Adiv, Determining Three-Dimensional Motion and Structure from Optical Flow Generated by several moving objects, IEEE Trans. Pattern Analysis and Machine intelligence, 7, pp. 384 - 401, 1985].

$$p(x,y) = \frac{a_1x + a_2y + a_3}{a_7x + a_8y + 1}$$

$$q(x,y) = \frac{a_4x + a_5y + a_6}{a_7x + a_8y + 1}$$

The coefficients a1,..,a8 are computed by minimizing the following expression:

$$\sum_{(x,y)} (D(x,y)-p(x,y)H(x,y)-q(x,y)\,V(x,y))^2$$

Now, the perspective transformation matrix (based on sensors' prediction) is multiplied by:

$$\begin{bmatrix} a_1 & a_4 & a_7 \\ a_2 & a_5 & a_8 \\ a_3 & a_6 & 1 \end{bmatrix}$$

The matrix obtained can be considered to be the updated prediction of billboard perspective.

In a practical environment the following considerations may apply.

- Occlusion may cause major problem in this formulation, since if pixels from occluding and moving objects participate in the minimization of the expression above they might bias the solution significantly. Preferably, such pixels are discarded from processing by using chroma-key panels. A key signal output by a chroma-keyer, is preferably utilized to discard these pixels.

- Luminance variations can be minimized by pre-processing the current video field, using histogram matching techniques.
- The prediction-correction process may require 2-3 iterations to converge.
- Noise immunity and convergence can both be enhanced by pre-smoothing the images.

Thus the billboards 14 etc are in accordance with the present invention chroma-key boards and occlusion is by colour discrimination using the normal chroma-key techniques. These techniques will enable perfect occlusion providing that the players do not wear any colour which is the same as the board. This may not always be possible and it is proposed in accordance with a particular embodiment of the present invention to use boards which can rotate or otherwise change to a second or third colour. For example, three colours may be blue, green and red which may be selected when the colour of the players strips are known.

Alternatively if it is required to display a billboard in an area of the pitch or a surrounding area then such an area must be selected to be of a known colour which can then be recorded in the chroma-keyer as a chroma-key colour.

In a preferred embodiment chroma-key apparatus can comprise the ULTIMATTE-7 digital video image compositing device from ULTIMATTE Corp., 20554 Plummer St., Chatsworth, CA 91311, USA.

The backing colour can be selected between blue, green and red. In order for the chroma-keyer to calculate all parameters necessary to perform proper image compositing, the system requires a sample of the background colour as reference. This step can be done automatically by scanning the image and detecting the purest and brightest colour. Advanced chroma-keyers enable the user to manually select the area to be sampled.

In a particular embodiment it is proposed to use a patterned chroma-key panel. Calibration of the camera sensors can then readily be accomplished by comparison of the pattern on a pixel by pixel basis. The pattern on the billboard panel should preferably have critical dimensions less than anticipated sensors error (projected to world coordinates).

In summary, the above system can operate even in extremely poor weather conditions since the electronic processing circuitry knows exactly where each billboard is situated and does not rely on any analysis of the video image to detect the billboard. In the event that the video image is so distorted that recalibration cannot be carried out with reasonable certainty, then the original camera parameter settings can continue to be used since the video image as viewed will be of poor quality and thus the viewer will not notice an error of one or two pixels in the positioning of the replacement billboard which will require to be displayed in an equivalent quality which matches the poor quality video image.

In a further preferred embodiment of the present invention, the problem addressed is that of having billboards situated in different positions in a stadium as shown in Figure 16.

In such conditions the lighting of billboards, 1,3 and 5 will be different because of the location of lights 7, 9 and 13. Also this lighting can change all the time during the game.

If such billboards are chroma-key boards all of the same colour then the billboards will all appear to be slightly different colours due to the different lighting conditions.

A fixed adjustment of a global backing colour might result in partial object background separation by the chroma-keyer.

In the present invention it is proposed to provide spatial adaptation of the backing colour map so that the chroma-keyer can correctly recognise each billboard. This can be provided by storing in store 704 (Figure 7) a spatial map providing information relating to the colour of each chroma key board.

Thus, the chroma-keyer will compare the colour in each video location with a specific colour associated with the billboard in that location.

In a preferred embodiment the locations of the billboards may be identified by "painting" a slightly enlarged box surrounding the billboard to identify the location. Such boxes are identified as 1′, 3′ and 5″ by dotted lines in Figure 16.

The system will track the backing colours over time and therefore will continuously update, to ensure correct identification, once correctly set up.

The operation of the system is as follows.

Firstly with reference to Figure 17, minimum and maximum levels are set for U and V. These should be wide enough to encompass all billboards which are reasonably lit.

Then for each billboard, as lighting conditions change, an adjustment can be made to its stored values, as shown in Figure 18, which assumes occlusion of billboard 1 by an object 13. An inner box 1″ is defined to ensure only pixels from within 1 are considered. Most occluding pixels can be discarded as these will be of a different colour. Then all pixels (YuV) inside FOV and billboard quad 1″ are measured and an addition to the average (which is backing colour average UV over billboard) is made if :-

$$U_{min} \leq U \geq U_{max}$$

$$\text{or} \quad V_{min} \leq V \geq V_{max}$$

The inventors have recognised a further problem which arises from the use of chroma-key billboards in a stadium. Due to the variable lighting as described above, each billboard will appear on the video image as a slightly different colour. In order to transmit correct occlusion information it is necessary to transmit an occlusion map for each billboard.

In accordance with a preferred embodiment of the present invention it is proposed to transmit, for each billboard, a perfect background colour and to then allow a chroma-keyer in each receiving station to introduce the occluded portions by normal chroma-key procedures.

Consider now the billboard arrangement as shown in Figure 16. Each billboard 1,3 and 5 will, because of its different lighting conditions, appear to be a different colour even though this colour may be within the maximum and minimum limits as set out in Figure 17.

In accordance with this preferred embodiment of the present invention, the transmitting apparatus (see Figure 6) will transmit a perfect chroma-key colour within the area of the billboard and will also transmit the coordinates of the quadrilateral formed by the billboard.

In this way the receiving station only has to decode/extract the quadrilateral coordinates of the billboard and then within that quadrilateral replace those pixels which are the perfect chroma-key colour by the replacement billboard. Those pixels which are not a perfect chroma-key colour are not replaced.

In accordance with this system it is not necessary for the chroma-keyer at the remote receiving station to be able to recognise different billboards and to have to store different chroma-key values for each billboard. Also it is not necessary to transmit any occlusion information since occlusion by the chroma-keyer will be relatively simple at each remote location.

With reference to Figure 7, the billboard position and backing colour store 704 knows the position of each billboard and a control output 7042 from the store is used, in combination with the video output 7102 to provide inputs for a backing colour processor 7044 which can change the colour of the billboard within the coordinates provided by store 704. The output of processor 7044 is used to control a backing colour store 7046 which changes the colour of the billboard within the required coordinates and can also provide the coordinates to the video output 720 for the remote receiver. These may be transmitted by a standard video data transmission system.

An exemplary remote receiver is shown in Figure 19. Video data is received at receiver buffer 1900 and split and delayed 1902, 1904.

Billboard coordinate store 1906 stores the transmitted billboard coordinates and in combination with graphics generator 1908 and replacement billboard image store 1910 provides an output signal to a combiner/chroma-keyer 1912 to produce the desired, occluded billboard on the screen.

With reference now to Figures 20 to 22 in a further embodiment the set up data stored in store 704 is modified prior to any event being televised.

The modification comprises the addition of dynamic set up data as well as the static image and static set up data shown in Figure 9.

The additional data may be used instead of the dynamic recalibration set up data 910 shown in Figure 9 or could be used in addition.

In a preferred embodiment it is assumed that the additional data is used instead of the dynamic recalibration procedure and this is now described.

As an introduction, the problems associated with replacing billboards with virtual billboards are discussed. The same problem is identifying the position, size and perspective of the original billboard and then replacing this with the virtual or replacement billboard.

In a static camera situation there is no real problem once the original co-ordinates have been recorded providing that the camera sensors do not drift substantially over time.

However, the inventors have found that during rapid panning or tilting of the camera the co-ordinates 908 of the replacement billboard as recorded in the store 900 do not coincide with the actual position of the billboard in the stadium or venue. This is because the camera sensors exhibit a degree of hysteresis. This can be compensated for by the dynamic recalibration process already described but this may not be practical in some circumstances such as during rapid panning with substantial occlusion of the target billboard.

The hysteresis could possibly be countered by a simple percentage error built into the movement of the camera but this does not produce very good results because it does not take into consideration the camera angle with respect to each billboard nor does it take into account the variability in the camera parameter sensors with angle.

In the present invention therefore in an alternative embodiment, in addition to static billboard set up data the billboard set up data store 900 stores data for each billboard for each camera at least in relation to left-right pan 2002, right-left pan 2004, up-down tilt 2008 and down-up tilt 2006. The data is obtained and stored as now herein described with reference to Figures 20 to 22.

Figure 20 shows the store 900 modified to provide, in addition to the static image data for billboards 1 to M and the static set up data for billboards I to M four further sets of data for each billboard 1 to M and these are multiplied to provide this data for each camera.

For each camera the position of each billboard is recorded with the camera panning from left to right 2002 and for right to left 2004. The panning speed may be selected as the normal speed for the event being televised. Thus for example for horse racing it could be low but for motor racing it could be higher. The position of each billboard is then recorded with the camera tilting upwards 2006 and then downwards 2008 across each billboard.

For each measurement the camera zoom and focus are preferably set at a known level at which the billboard being analysed is in a reasonable view at a reasonable size. The zoom and focus could for example be the same as that during the acquisition of the static set up data for each billboard so that a direct comparison with the static set up data can be made. In that case only an error correction figure may need to be recorded.

It is preferable not to select too high a panning or tilting speed because at very high speeds the billboards will in any case be blurred and therefore accuracy of replacement will not be an issue. This procedure is followed for each billboard for each camera and the data is then used during the event to correct the position of each billboard as the camera pans or tilts.

It may be seen that each billboard will be viewed at a different angle by each camera and also that the output of each of the sensors on each camera may vary dependent on the angle through which the camera must turn to view the billboard. By recording the static data and data relating to panning and tilting in both directions the replacement billboard will be accurately positioned in the exact position of the original or real billboard both for static shots and when the camera is moving.

Dynamic recalibration during the event as previously described will ensure, except during very fast camera movements with large occlusion, that the replacement billboard is correctly positioned but the use of static and dynamic set up data will also ensure this unless the camera sensors drift substantially during an event. Thus providing that the camera sensors are of a reasonable quality from the point of view of draft they can be of a variable quality with respect to accuracy during panning and tilting. By careful selection of camera sensors extremely accurate sensors are therefore not required since any variation with respect to camera movement is compensated for by the storage of dynamic set up data.

The data is obtained as described with reference to Figures 21 and 22 as follows.

Once the static data relating to each billboard image and its static set up data (906, 908, Figure 9) has been obtained, sequence 2100 is started and the camera is panned 2102 by the operator at a desired speed relative to the normal panning speed.

The sensors indicate the direction of pan 2106 and dependent on the direction the dynamic data is stored in store 2002 or 2004 in steps 2106 or 2108 by selection of that store. The sequence for both L-R and R-L stores 2002 and 2004 is similar and will be described for the L-R store but using reference numerals for both stores.

As the camera pans L-R each billboard is identified from data stored in store 906 step 2110; 2112. The system asks if the billboard has previously been recoded dynamically step 2114, 2116 and if so it returns to the start of the sequence and repeats steps 2104-2110 until it finds a billboard that has not been dynamically scanned. Once a new billboard has been found the position (co-ordinates) of the billboard during panning is recorded and compared with the static billboard parameters previously stored (908) in step 2118, 2120. Any error is computed (step 2122, 2124) and the errors are stored in L-R and R-L pan stores 2126, 2128 for the billboard. The system asks if all billboards recorded in store 908 have been dynamically scanned both for L-R and R-L (steps 2130, 2132). If not the sequence is continued until the last billboard has been dynamically scanned and then the program is terminated 2134, 2136.

A similar program sequence shown in Figure 22 is provided for tilting of each camera. Obviously if the cameras are not either allowed to tilt or are unlikely to be tilted to any extent then this sequence and the recordal of data in stores 2006, 2008 may not be necessary.

The sequence is started 2200 and each camera in turn is tilted 2202 and the direction of tilt determined 2204 by the camera sensors. Dependent on whether the camera is tilting up or down dynamic set up data is stored in stores 2006 or 2008 in steps 2206, 2208. Both sequences are similar and only the sequence tilting the camera down will be described with reference then to both sequences.

Each billboard is identified 2210, 2212 from the static image data and also from the camera parameters especially where all real billboards are the same. The program interrogates the billboard data step 2214, 2216 to see if the billboard has already been interrogated. If it has the program restarts but if not the co-ordinate data of the billboard during tilting is compared with the static data step 2218, 2220. The error, if any is computed step 2122, 2124 and stored in the stores 2006, 2008 (Figure 20) step 2226, 2228.

The program then interrogates stores 2006, 2008 to see if all billboards have been dynamically interrogated for tilt errors in both up (step 2230) and down (step 2232) and if so ends the program steps 2234, 2236. If not the program continues by commencing at the start of the sequence until all boards have been interrogated.

Normally camera zoom and focus will not require the same type of dynamic set up data to be stored. However, if particular camera aberrations are known then these may be compensated for by use of similar dynamic set up data.

The dynamic data stored in stores 2002-2008 may be used instead of or in conjunction with the dynamic recalibration data obtained as described with reference to Figure 9. Usually however the dynamic set up data will obviate the need for recalibration during most types of event.

During use the system knows by reading the camera sensors whether the billboard is being viewed in a static manner or is being panned past L-R or R-L or tilted past UP or DOWN. In such cases the position of the billboard is taken from the static data store and then if panning or tilting is occurring, the necessary error corrections are applied. Once camera movement ceases the static billboard parameters are reverted to.

THE CLAIMS DEFINING THE INVENTION ARE AS FOLLOWS:

1. Apparatus for automatic electronic replacement of a billboard in a video image including automatic camera orientation measurement apparatus including motion measurement means operative to measure the Field of View of a TV camera relative to a known reference position;

said apparatus further including a dynamic billboard memory said dynamic billboard memory including:

a) static image data recording means recording data indicating the identity of said billboard,

b) static set-up data recording means recording data indicating the static position of the billboard with the camera in a static position with said billboard in full view in said Field of View of said camera,

c) dynamic calibration data recording means said dynamic calibration recording means including:

i) left to right pan data recording means recording data indicating the position of the billboard in the Field of View of the camera when the camera is panned from left to right across the billboard,

ii) right to left pan data recording means recording data indicating the position of the billboard in the Field of View of the camera when the camera is panned from right to left,

iii) tilt up data recording means recording data indicating the position of the billboard in the Field of View of the camera when the camera is tilted in a direction from below the billboard to above the billboard, and

iv) tilt down recording means recording data indicating the position of the billboard in the Field of View of the camera when the camera is tilted in a direction from above the billboard to below the billboard.

2. Apparatus as claimed in claim 1 including a plurality of billboards situated at different positions within a stadium.

3. Apparatus as claimed in claim 1 or claim 2 in which the static, pan left to right, pan right to left tilt up and tilt down data are all recorded at the same zoom and focus of the camera.

4. Apparatus as claimed in claim 3 including processing means for comparison of the static position data of said billboard with the dynamic position data during panning and for computing the error in billboard position as the camera pans.

5. Apparatus as claimed in claim 3 including processing means for comparison of the static position data of said billboard with the dynamic position data during tilting and for computing the error in billboard position as the camera tilts.

6. A method for automatic replacement of a billboard in a video image including the steps of:

a) locating said billboard in a known position relative to a TV camera;

b) determining a position for said billboard when said TV camera is static and storing said position in a first store;

c) determining a plurality of dynamic

positional data for said billboard as said TV camera is panned from left to right, right to left, tilted up and tilted down and storing said dynamic data in respective dynamic data stores;

d) storing a replacement virtual billboard in a replacement billboard store;

e) replacing said billboard in a picture taken by said TV camera using data from said static data store and from selected dynamic data stores dependent on detected movement of said TV camera.

7. Apparatus for automatic electronic replacement of a billboard in a video image as claimed in claim 1 and substantially as herein described with reference to the accompanying drawings.

8. A method for automatic replacement of a billboard in a video image as claimed in claim 6 and substantially as herein described with reference to the accompanying drawings.

Dated this 27th day of October 1999
<u>ORAD HI-TEC SYSTEMS LIMITED</u>
By their Patent Attorneys
GRIFFITH HACK
Fellows Institute of Patent and
Trade Mark Attorneys of Australia

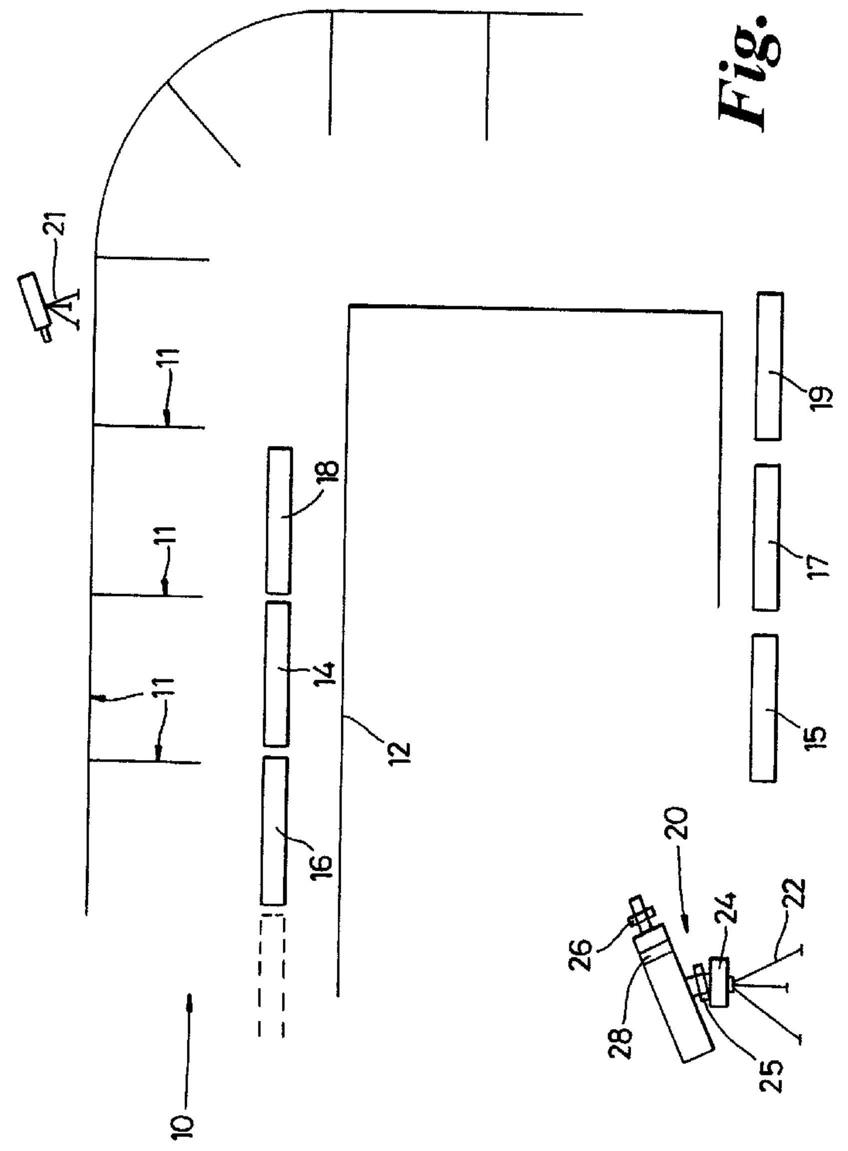


Fig. 1

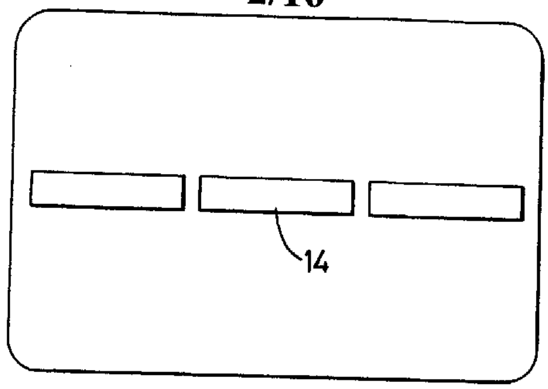


Fig. 2

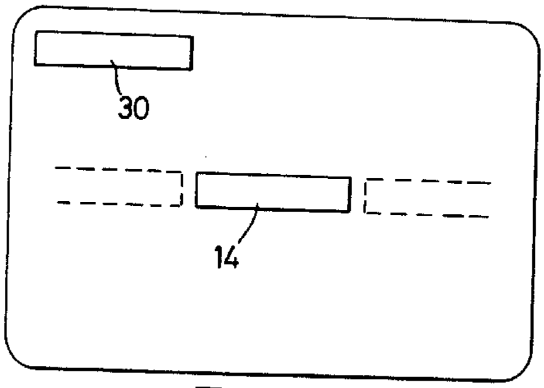


Fig. 3

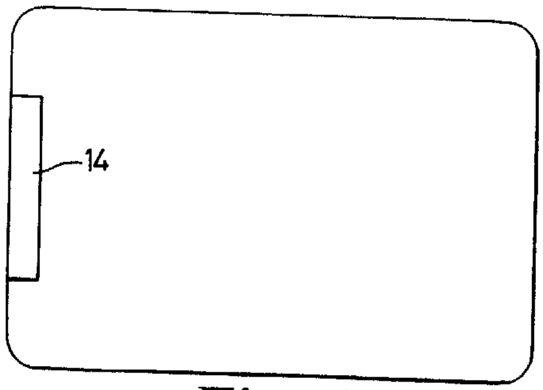


Fig. 4

Fig. 5

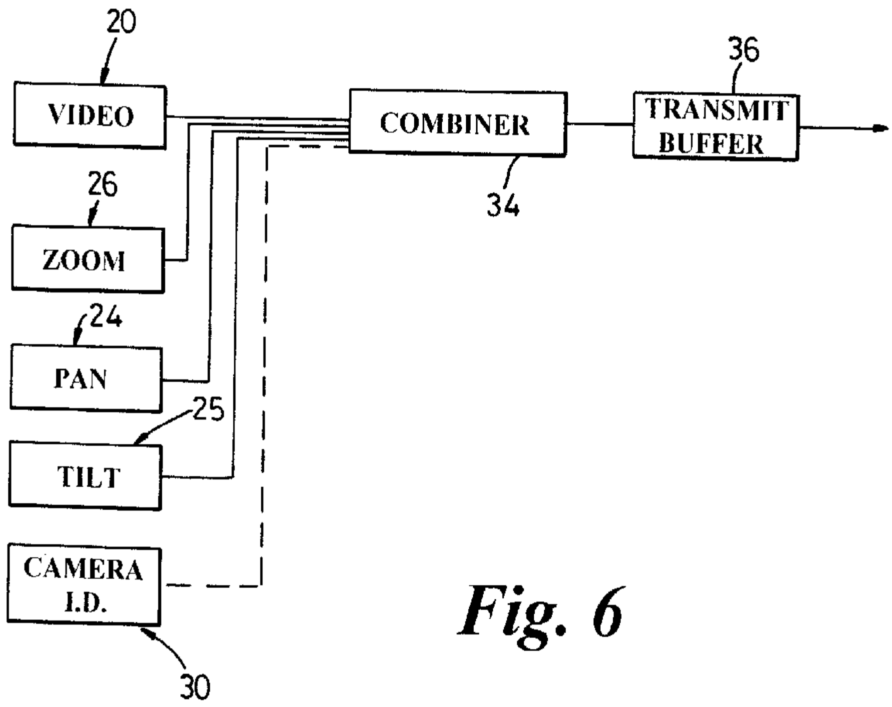

20
VIDEO
26
ZOOM
24
PAN
25
TILT
CAMERA
I.D.
30
COMBINER
34
36
TRANSMIT
BUFFER

Fig. 6

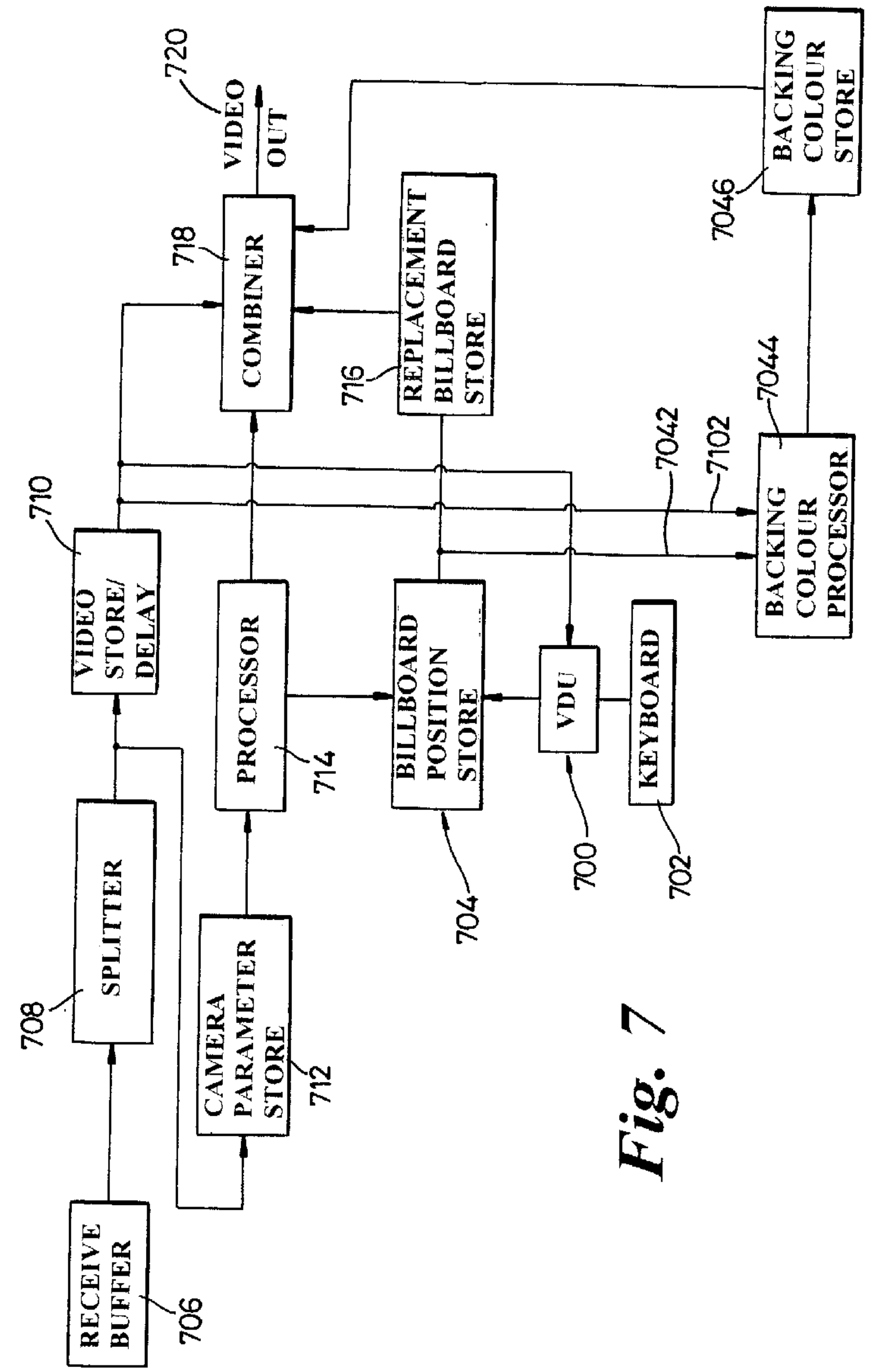

RECEIVE BUFFER
706
SPLITTER
708
VIDEO STORE/ DELAY
710
CAMERA PARAMETER STORE
712
PROCESSOR
714
BILLBOARD POSITION STORE
704
VDU
700
KEYBOARD
702
REPLACEMENT BILLBOARD STORE
716
COMBINER
718
VIDEO OUT
720
BACKING COLOUR PROCESSOR
7042
7102
7044
BACKING COLOUR STORE
7046

Fig. 7

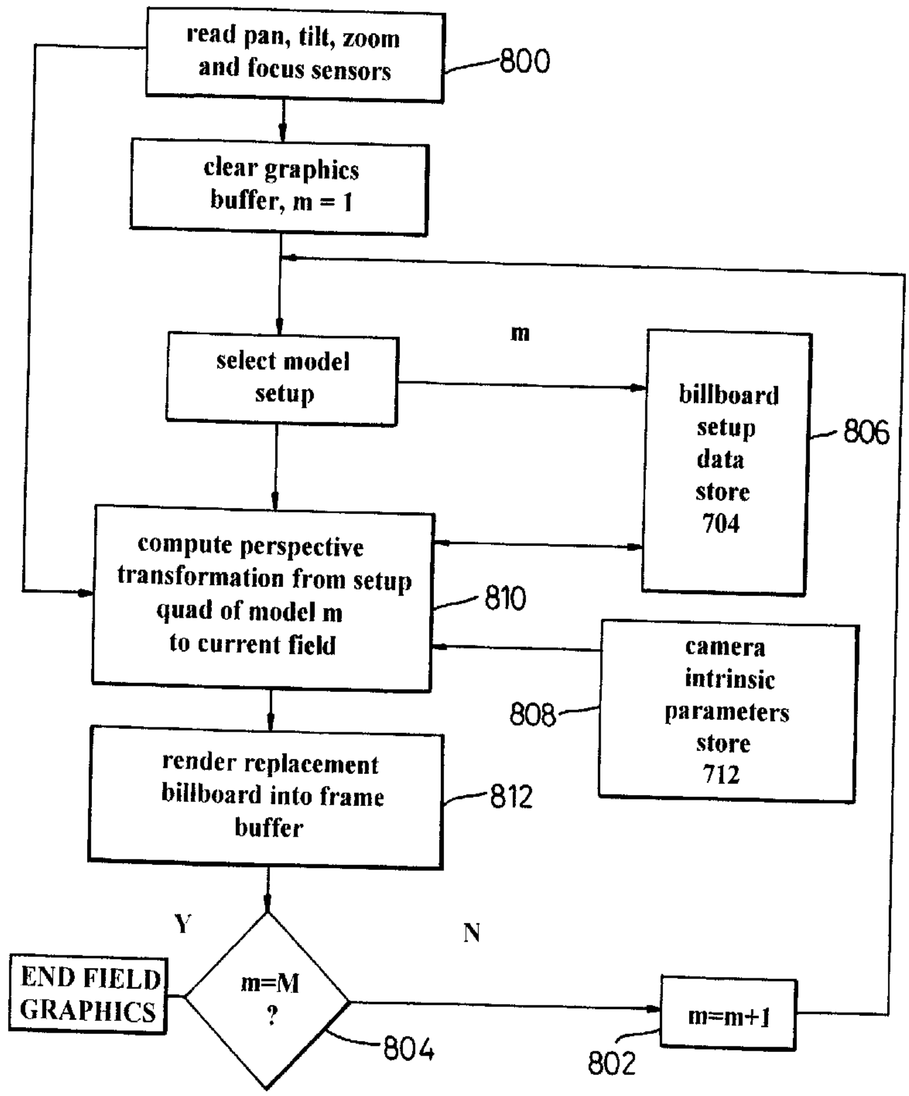

read pan, tilt, zoom and focus sensors
800
clear graphics buffer, m = 1
select model setup
m
billboard setup data store 704
806
compute perspective transformation from setup quad of model m to current field
810
camera intrinsic parameters store 712
808
render replacement billboard into frame buffer
812
Y
N
END FIELD GRAPHICS
m=M ?
804
m=m+1
802

Fig. 8

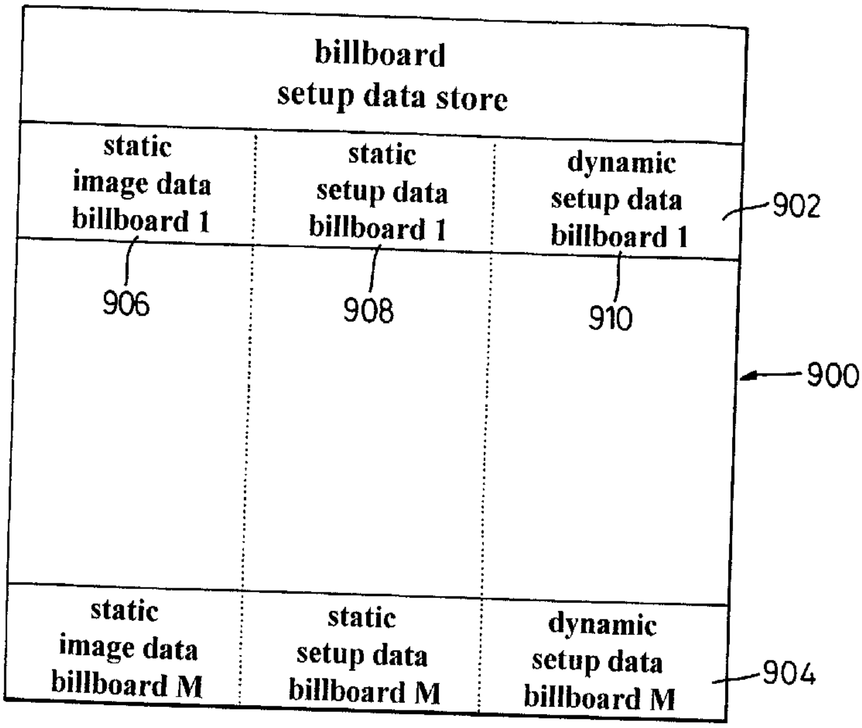
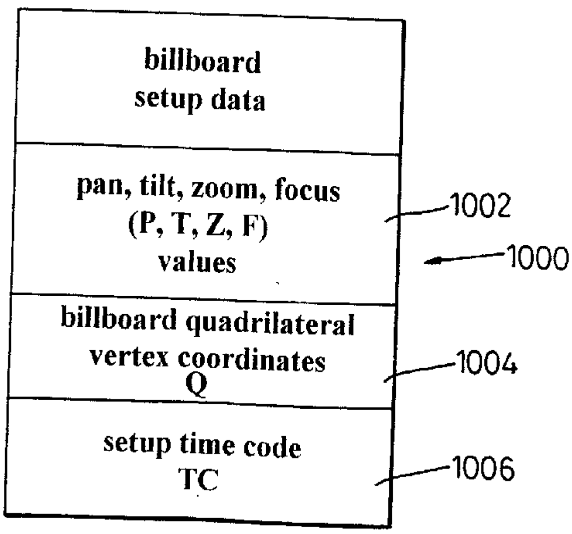
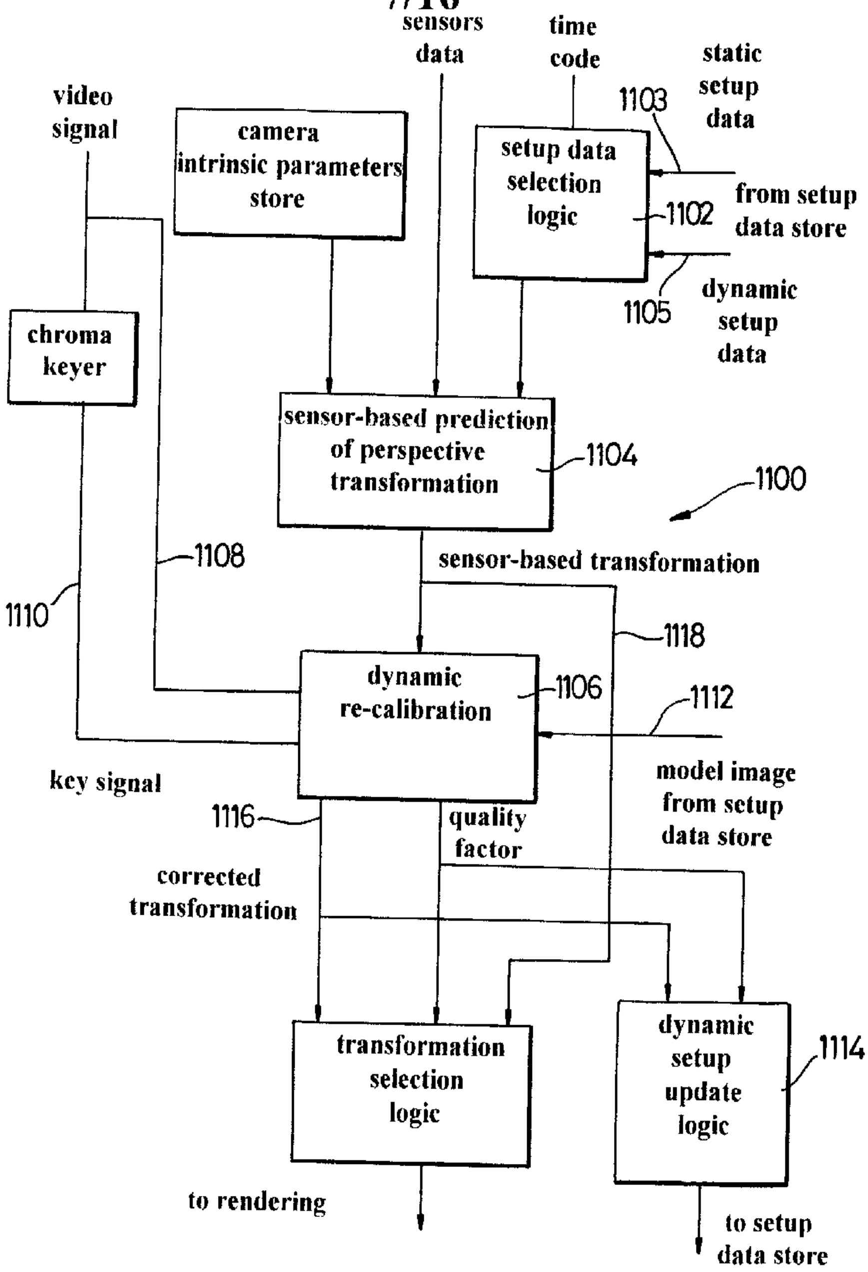

billboard
setup data store
static
image data
billboard 1
static
setup data
billboard 1
dynamic
setup data
billboard 1
902
906
908
910
900
static
image data
billboard M
static
setup data
billboard M
dynamic
setup data
billboard M
904

Fig. 9

Fig. 10


billboard
setup data
pan, tilt, zoom, focus
(P, T, Z, F)
values
1002
1000
billboard quadrilateral
vertex coordinates
Q
1004
setup time code
TC
1006



sensors data
time code
static setup data
video signal
camera intrinsic parameters store
setup data selection logic
1103
1102
from setup data store
1105
dynamic setup data
chroma keyer
sensor-based prediction of perspective transformation
1104
1100
1108
sensor-based transformation
1110
1118
dynamic re-calibration
1106
1112
key signal
model image from setup data store
1116
quality factor
corrected transformation
transformation selection logic
dynamic setup update logic
1114
to rendering
to setup data store

Fig. 11

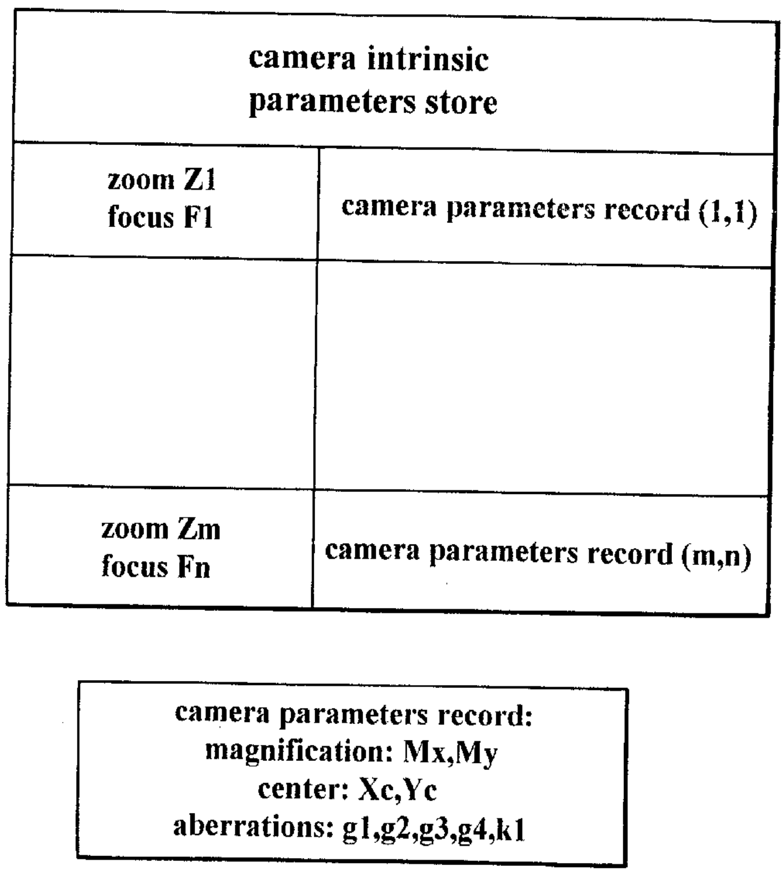

camera intrinsic
parameters store
zoom Z1
focus F1
camera parameters record (1,1)
zoom Zm
focus Fn
camera parameters record (m,n)
camera parameters record:
magnification: Mx,My
center: Xc,Yc
aberrations: g1,g2,g3,g4,k1

Fig. 12

1. setup rotation matrix

$$R_s = \begin{bmatrix} \cos(P_s) & 0 & -\sin(P_s) \\ 0 & 1 & 0 \\ \sin(P_s) & 0 & \cos(P_s) \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(T_s) & \sin(T_s) \\ 0 & -\sin(T_s) & \cos(T_s) \end{bmatrix}$$

600

2. prediction rotation matrix

$$R_p = \begin{bmatrix} \cos(P_p) & 0 & -\sin(P_p) \\ 0 & 1 & 0 \\ \sin(P_p) & 0 & \cos(P_p) \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(T_p) & \sin(T_p) \\ 0 & -\sin(T_p) & \cos(T_p) \end{bmatrix}$$

602

3. setup to prediction perspective transformation

$$R_{sp} = R_s^{-1} R_p$$

604

4. setup to prediction point transformation

$$u_p = \frac{R_{sp}[0][0]u_s + R_{sp}[1][0]v_s + R_{sp}[2][0]}{R_{sp}[0][2]u_s + R_{sp}[1][2]v_s + R_{sp}[2][2]}$$

$$v_p = \frac{R_{sp}[0][1]u_s + R_{sp}[1][1]v_s + R_{sp}[2][1]}{R_{sp}[0][2]u_s + R_{sp}[1][2]v_s + R_{sp}[2][2]}$$

606

608

5. image plane to frame-buffer

$$x = (u - A_u(u,v))M_X + X_C$$
$$y = (v - A_v(u,v))M_Y + Y_C$$

Fig. 13

6. aberrations

$$A_u(u,v) = (g_1 + g_3)u^2 + g_4 uv + g_1 v^2 + k_1 u(u^2 + v^2)$$
$$A_v(u,v) = g_2 u^2 + g_3 uv + (g_2 + g_4)v^2 + k_1 v(u^2 + v^2)$$

610

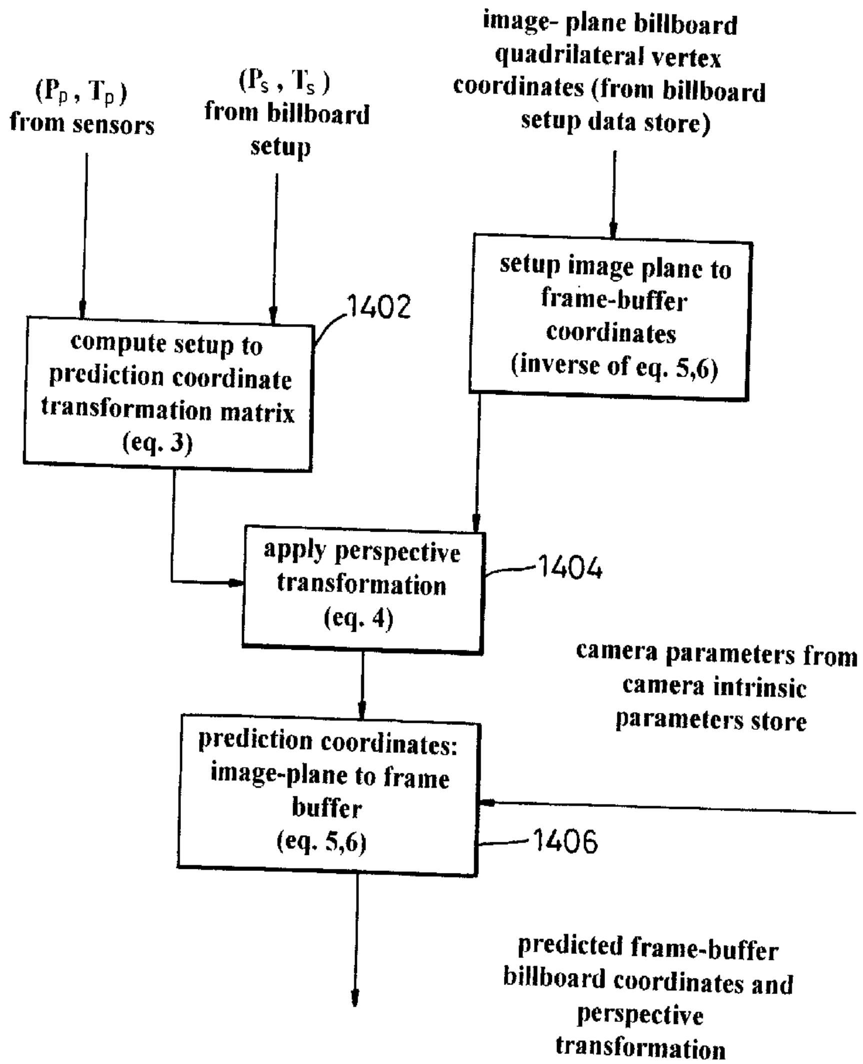


image- plane billboard quadrilateral vertex coordinates (from billboard setup data store)
(P_p, T_p) from sensors
(P_s, T_s) from billboard setup
setup image plane to frame-buffer coordinates (inverse of eq. 5,6)
1402
compute setup to prediction coordinate transformation matrix (eq. 3)
apply perspective transformation (eq. 4)
1404
camera parameters from camera intrinsic parameters store
prediction coordinates: image-plane to frame buffer (eq. 5,6)
1406
predicted frame-buffer billboard coordinates and perspective transformation

Fig. 14

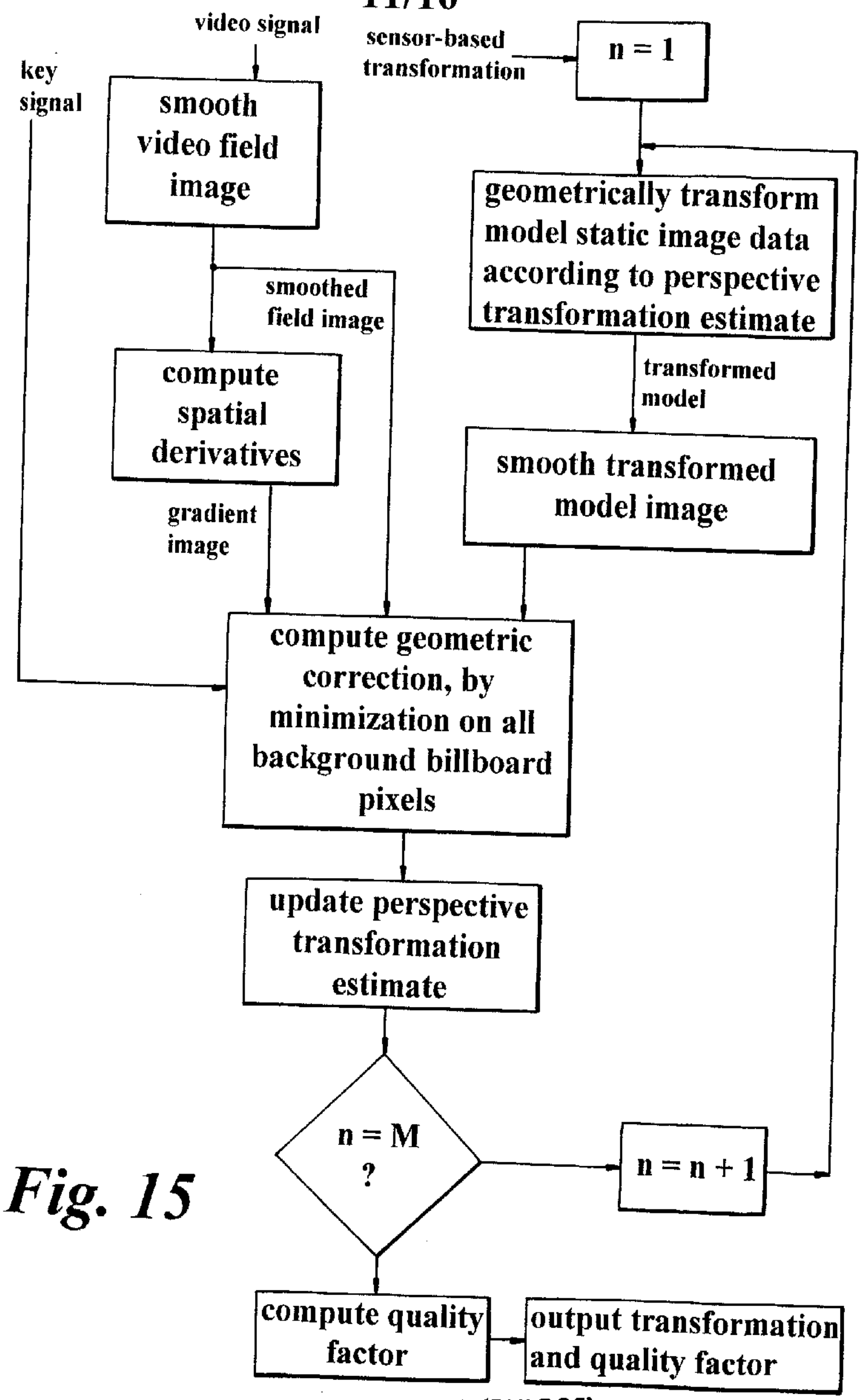

video signal
sensor-based transformation
n = 1
key signal
smooth video field image
geometrically transform model static image data according to perspective transformation estimate
smoothed field image
compute spatial derivatives
transformed model
smooth transformed model image
gradient image
compute geometric correction, by minimization on all background billboard pixels
update perspective transformation estimate
n = M ?
n = n + 1
compute quality factor
output transformation and quality factor

Fig. 15

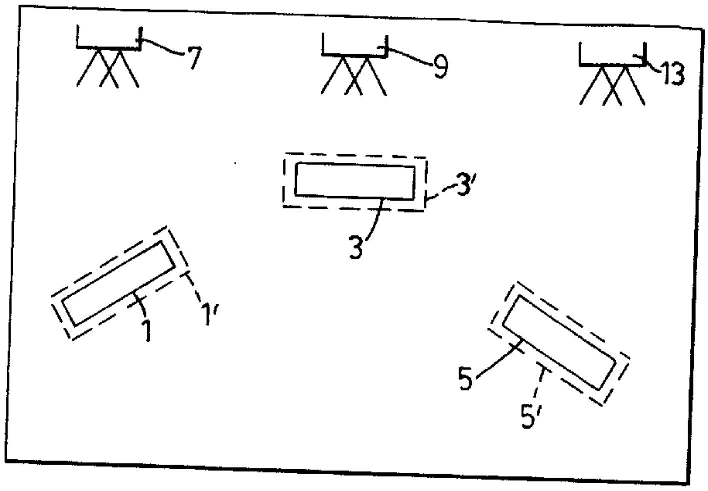


Fig. 16

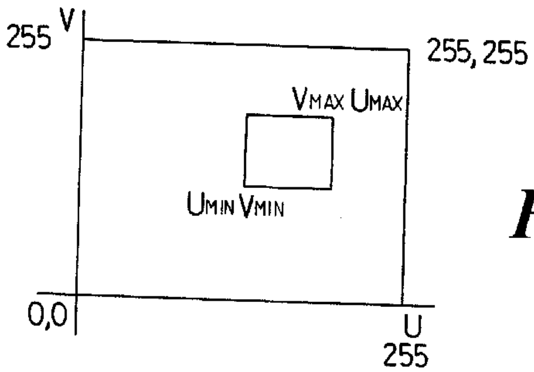


Fig. 17

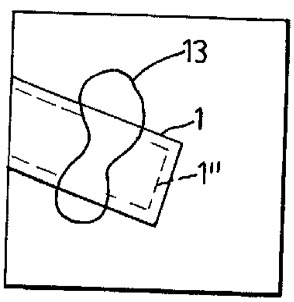


Fig. 18

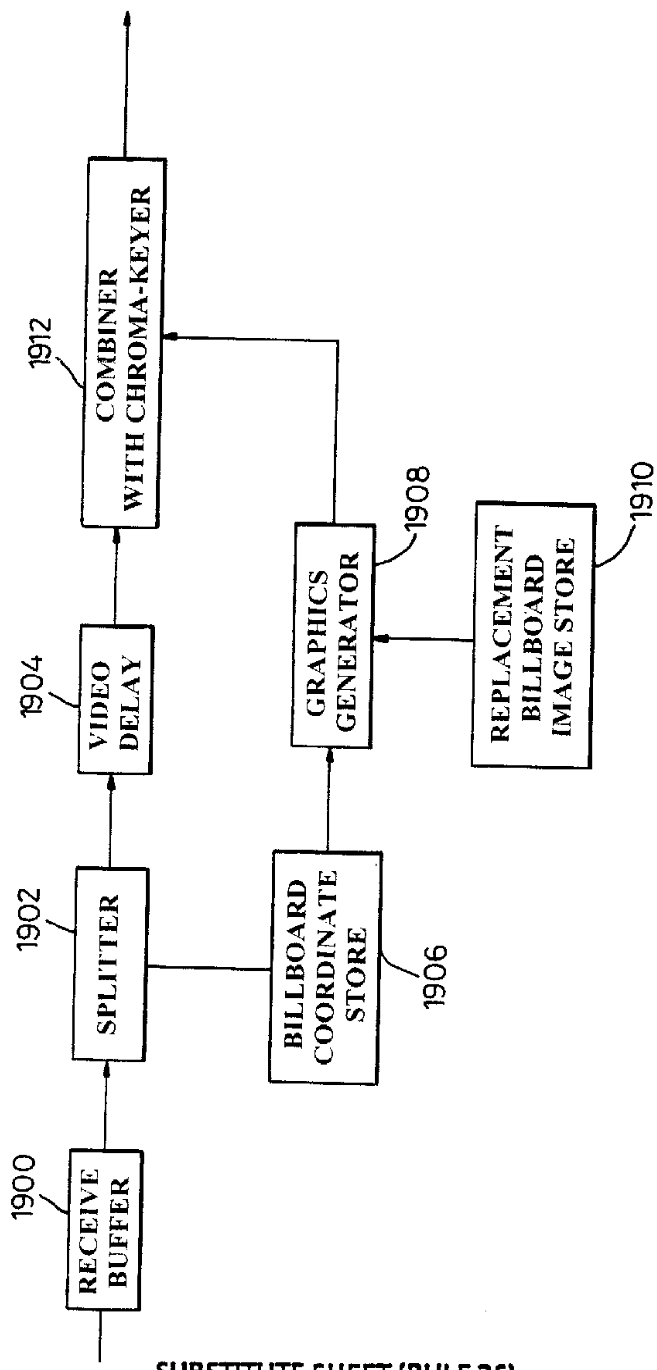
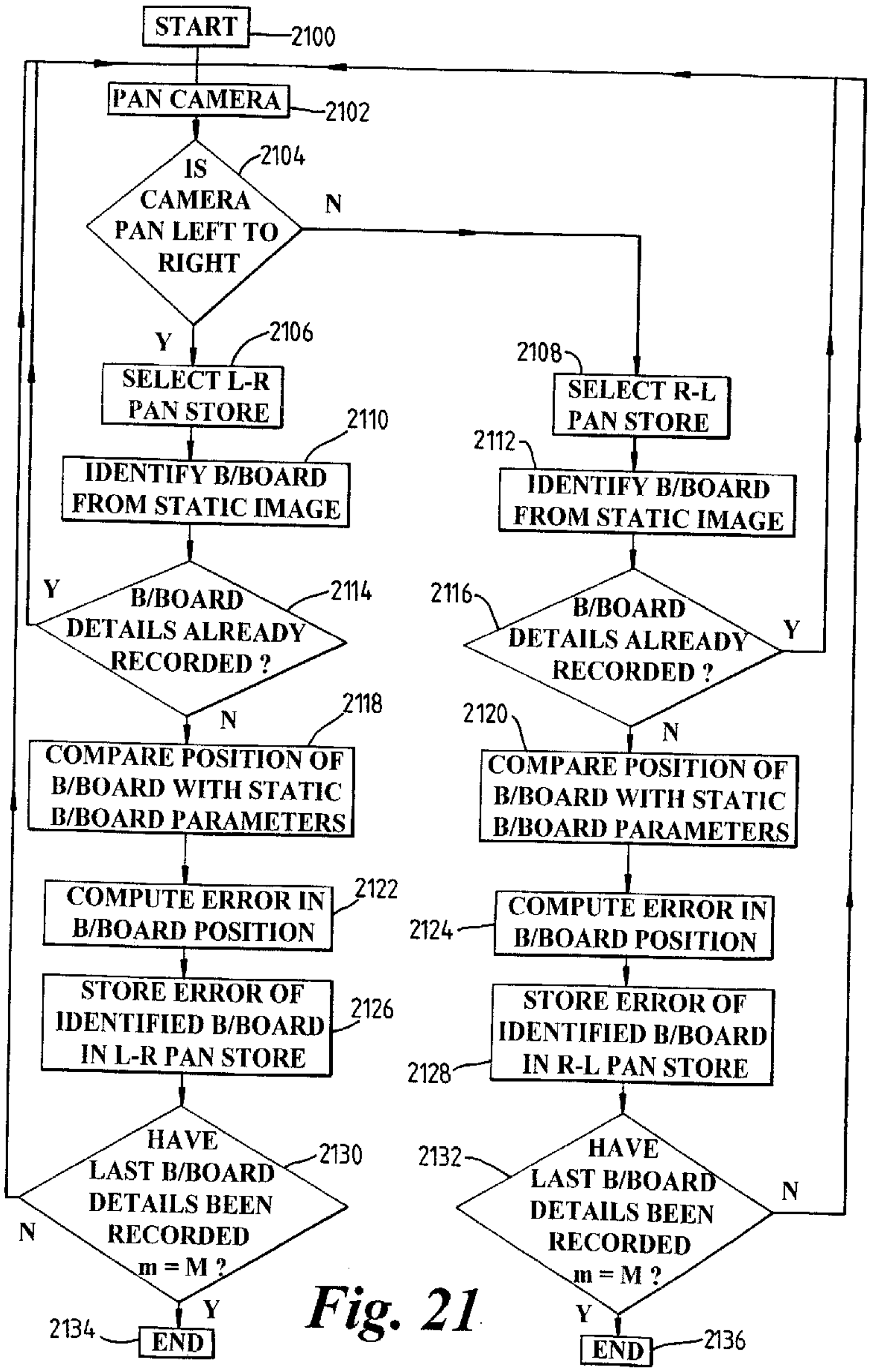

RECEIVE BUFFER
1900
SPLITTER
1902
VIDEO DELAY
1904
COMBINER WITH CHROMA-KEYER
1912
BILLBOARD COORDINATE STORE
1906
GRAPHICS GENERATOR
1908
REPLACEMENT BILLBOARD IMAGE STORE
1910

Fig. 19

BILLBOARD STATIC AND DYNAMIC SET UP DATA STORE						
STATIC IMAGE DATA	STATIC SET UP DATA	DYNAMIC CALIBRATION DATA			
		PAN L-R	PAN R+L	TILT UP	TILT DOWN
B/BOARD 1	B/BOARD 1	B/B 1	B/B 1	B/B 1	B/B 1
906	908	2002	2004	2006	2008
B/BOARD M	B/BOARD M	B/B M	B/B M	B/B M	B/B M

Fig. 20



START
2100
PAN CAMERA
2102
IS CAMERA PAN LEFT TO RIGHT
2104
N
Y
SELECT L-R PAN STORE
2106
SELECT R-L PAN STORE
2108
IDENTIFY B/BOARD FROM STATIC IMAGE
2110
IDENTIFY B/BOARD FROM STATIC IMAGE
2112
B/BOARD DETAILS ALREADY RECORDED ?
2114
Y
N
B/BOARD DETAILS ALREADY RECORDED ?
2116
Y
N
COMPARE POSITION OF B/BOARD WITH STATIC B/BOARD PARAMETERS
2118
COMPARE POSITION OF B/BOARD WITH STATIC B/BOARD PARAMETERS
2120
COMPUTE ERROR IN B/BOARD POSITION
2122
COMPUTE ERROR IN B/BOARD POSITION
2124
STORE ERROR OF IDENTIFIED B/BOARD IN L-R PAN STORE
2126
STORE ERROR OF IDENTIFIED B/BOARD IN R-L PAN STORE
2128
HAVE LAST B/BOARD DETAILS BEEN RECORDED m = M ?
2130
N
Y
HAVE LAST B/BOARD DETAILS BEEN RECORDED m = M ?
2132
N
Y
END
2134
END
2136

Fig. 21

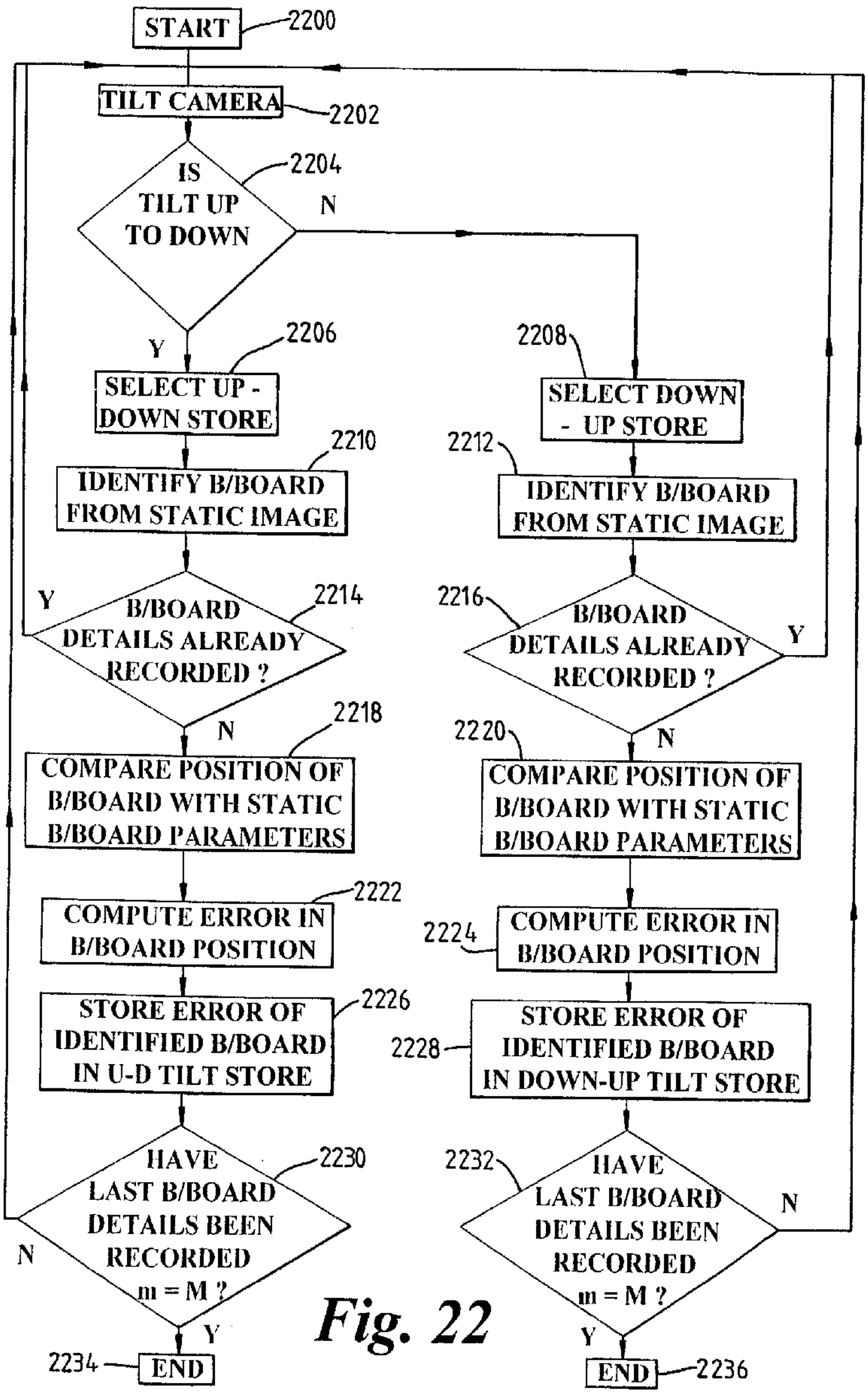

START
2200
TILT CAMERA
2202
IS TILT UP TO DOWN
2204
N
Y
SELECT UP - DOWN STORE
2206
2208
SELECT DOWN - UP STORE
IDENTIFY B/BOARD FROM STATIC IMAGE
2210
2212
IDENTIFY B/BOARD FROM STATIC IMAGE
Y
B/BOARD DETAILS ALREADY RECORDED ?
2214
2216
B/BOARD DETAILS ALREADY RECORDED ?
Y
N
2218
2220
N
COMPARE POSITION OF B/BOARD WITH STATIC B/BOARD PARAMETERS
COMPARE POSITION OF B/BOARD WITH STATIC B/BOARD PARAMETERS
COMPUTE ERROR IN B/BOARD POSITION
2222
2224
COMPUTE ERROR IN B/BOARD POSITION
STORE ERROR OF IDENTIFIED B/BOARD IN U-D TILT STORE
2226
2228
STORE ERROR OF IDENTIFIED B/BOARD IN DOWN-UP TILT STORE
HAVE LAST B/BOARD DETAILS BEEN RECORDED m = M ?
2230
2232
HAVE LAST B/BOARD DETAILS BEEN RECORDED m = M ?
N
N
Y
Y
2234
END
END
2236

Fig. 22