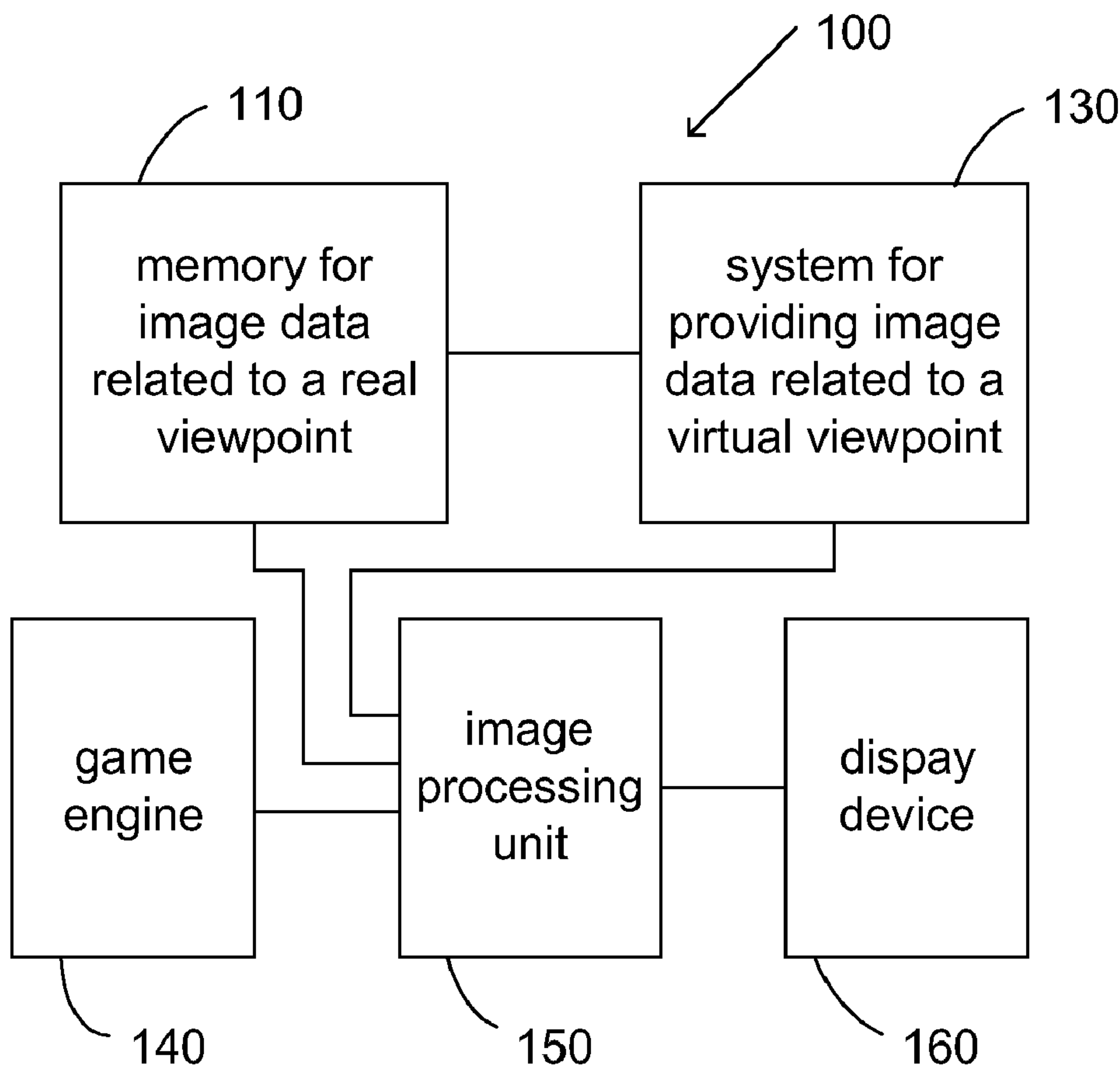




(86) **Date de dépôt PCT/PCT Filing Date:** 2008/12/19
 (87) **Date publication PCT/PCT Publication Date:** 2010/06/24
 (45) **Date de délivrance/Issue Date:** 2016/06/21
 (85) **Entrée phase nationale/National Entry:** 2011/06/17
 (86) **N° demande PCT/PCT Application No.:** SE 2008/051544
 (87) **N° publication PCT/PCT Publication No.:** 2010/071531

(51) **Cl.Int./Int.Cl. G06T 15/10** (2011.01),
H04N 13/00 (2006.01), **H04N 5/262** (2006.01)
 (72) **Inventeurs/Inventors:**
ANDERSSON, INGMAR, SE;
ISAKSSON, FOLKE, SE;
BORG, JOHAN, SE;
ERLANDSSON, ULF, SE;
TENGBLAD, ADAM, SE;
HAGSTROEM, MANS, SE
 (73) **Propriétaire/Owner:**
SAAB AB, SE
 (74) **Agent:** MARKS & CLERK

(54) **Titre : SYSTEME ET PROCEDE DE MIXAGE D'UNE SCENE ET D'UN SCENARIO VIRTUEL**
 (54) **Title: SYSTEM AND METHOD FOR MIXING A SCENE WITH A VIRTUAL SCENARIO**



(57) **Abrégé/Abstract:**

The present invention relates to a system (100) and method for mixing a scene with a virtual scenario. The system comprises image capturing means arranged to capture at least one image so as to cover the scene from a first viewpoint, and an image

(57) Abrégé(suite)/Abstract(continued):

representation generation unit arranged to generate at least one image representation based on said captured image. The system comprises further a game engine unit (140) arranged to generate a virtual scenario, an image processing unit (150) arranged to adapt the at least one image representation based on said generated virtual scenario so as to provide a virtual video sequence.

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

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
24 June 2010 (24.06.2010)(10) International Publication Number
WO 2010/071531 A1

(51) International Patent Classification:

G06T 15/10 (2006.01) *H04N 5/262* (2006.01)
H04N 13/00 (2006.01)Linköping (SE). **HAGSTRÖM, Måns** [SE/SE]; Nettovägen 6, S-175 88 Järfälla (SE).

(21) International Application Number:

PCT/SE2008/051544

(74) Agent: **ALBIHNS AB**; Linnégatan 2, P.O. Box 5581, S-114 85 Stockholm (SE).

(22) International Filing Date:

19 December 2008 (19.12.2008)

(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.

(25) Filing Language:

English

(26) Publication Language:

English

(71) Applicant (*for all designated States except US*): **SAAB AB** [SE/SE]; S-581 88 Linköping (SE).

(72) Inventor; and

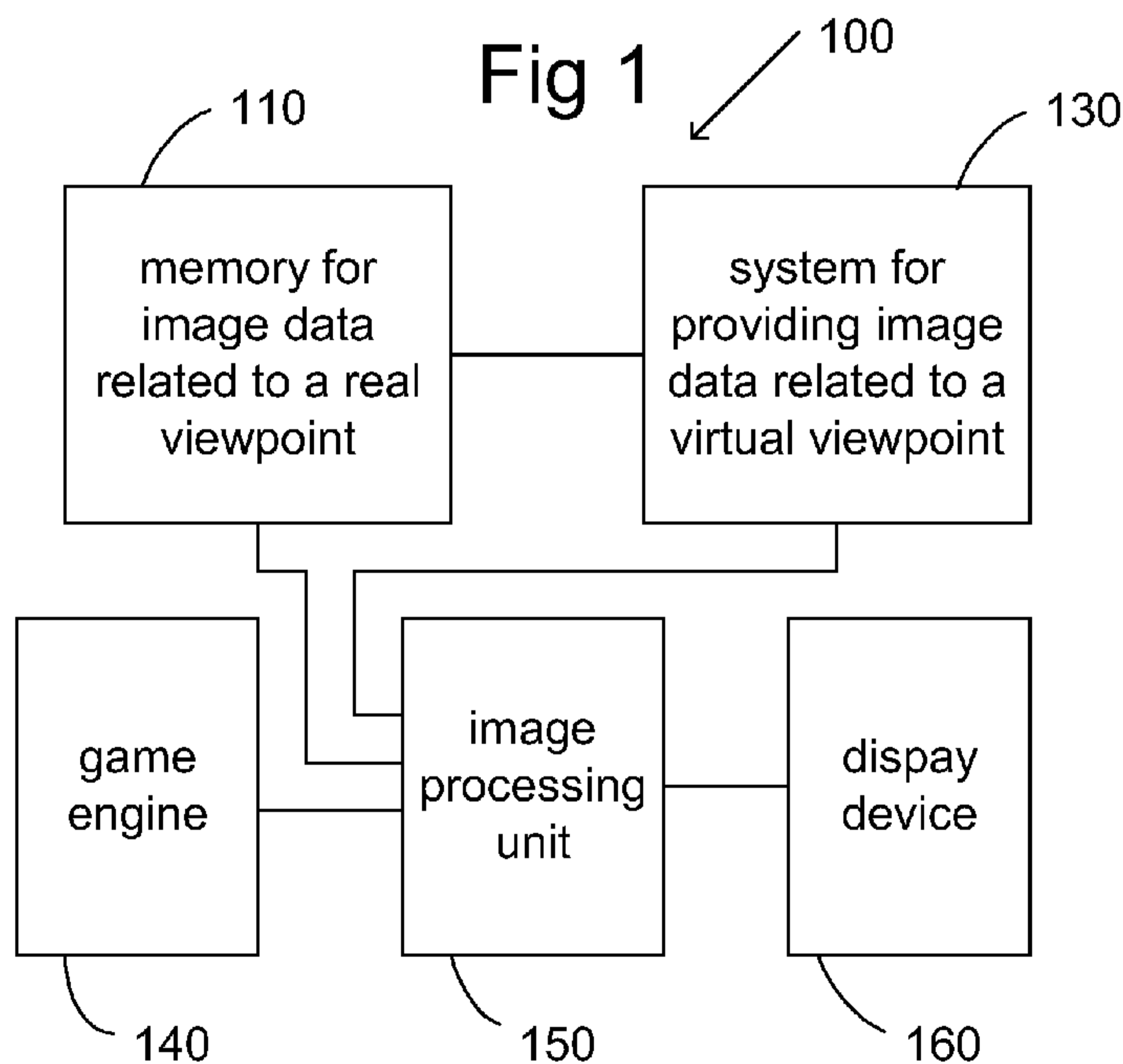
(71) Applicant : **TENGBLAD, Adam** [SE/SE]; Stenholmsvägen 44, S-561 39 Huskvarna (SE).

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): **ANDERSSON, Ingemar** [SE/SE]; Tegskiftesgatan 275, S-583 34 Linköping (SE). **ISAKSSON, Folke** [SE/SE]; Tegskiftesgatan 275, S-583 34 Linköping (SE). **BORG, Johan** [SE/SE]; Storhagsgatan 87, S-590 49 Vikingstad (SE). **ERLANDSSON, Ulf** [—/SE]; Repslagaregatan 10, S-582 22(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, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR), OAPI

[Continued on next page]

(54) Title: SYSTEM AND METHOD FOR MIXING A SCENE WITH A VIRTUAL SCENARIO



(57) Abstract: The present invention relates to a system (100) and method for mixing a scene with a virtual scenario. The system comprises image capturing means arranged to capture at least one image so as to cover the scene from a first viewpoint, and an image representation generation unit arranged to generate at least one image representation based on said captured image. The system comprises further a game engine unit (140) arranged to generate a virtual scenario, an image processing unit (150) arranged to adapt the at least one image representation based on said generated virtual scenario so as to provide a virtual video sequence.

WO 2010/071531 A1 

(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG). **Published:**

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

Declarations under Rule 4.17:

— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*

System and method for mixing a scene with a virtual scenario

TECHNICAL FIELD

The present invention relates to a system and method for mixing a real world
5 scene with a virtual scenario in general. Specifically, it relates to perform the
mixing based on image representation generation.

BACKGROUND ART

For example, in a mixed reality presentation system, wherein a real world
10 and virtual world are merged in a three-dimensionally matched form, the
depth (front and behind) ordering of real objects and virtual objects must be
correctly recognized to render the virtual objects in a form that does not
conflict with that depth ordering. For this purpose, depth information (three-
dimensional information) of the real world must be acquired.

15

EP 0 955 606 describes a depth image measurement apparatus for acquiring
depth information of a scene. The depth image measurement apparatus
comprises means for inputting an image of a scene at a first viewpoint and
depth image generation means for generating a first depth image from the
20 scene image inputted at the first viewpoint by said image input means. The
apparatus further comprises position/posture estimation means for estimating
a position and posture information at a second viewpoint viewed from a
position and posture of the first viewpoint. The apparatus further comprises
warping means for warping the first depth image generated by said depth
25 image generation means to a second depth image at the second viewpoint
estimated by said position/posture estimation means.

SUMMARY OF THE INVENTION

30 One object of the invention is to further develop the technique known in the
art.

This has in one example been achieved by means of a system for mixing a scene with a virtual scenario. The system comprises image capturing means arranged to capture at least one image so as to cover the scene from a first viewpoint, and an image representation generation unit arranged to generate
5 at least one image representation based on said captured image. The system further comprises a game engine unit arranged to generate a virtual scenario, an image processing unit arranged to adapt the at least one image representation based on said generated virtual scenario so as to provide a virtual video sequence comprising the virtual scenario mixed with the
10 representation of the scene, and a display unit arranged to display the virtual video sequence.

The system can be used for example in military training so as to produce virtual video sequences related to a real environment or virtually created
15 environment. The produced virtual video sequences can be displayed at a display unit located at the virtual viewing point or at a remote location. As the system comprises a game engine arranged to generate virtual scenarios, and the virtual scenarios are mixed with the background in a realistic manner, goal-oriented training can be provided for soldiers and others.

20

In one example, the system comprises a warping unit for warping each image representation generated by said image representation generation unit to a representation related to a second viewpoint so as to form a representation of the scene viewed from the second viewpoint. The image processing unit is
25 then arranged to adapt the warped image representation based on said generated virtual scenario so as to provide a virtual video sequence comprising the virtual scenario mixed with the warped representation of the scene viewed from the second viewpoint. This provides flexibility, as a virtual viewing point can be chosen arbitrary within an area around a real viewing
30 point where the real images of the environment or background were taken.

The present invention also relates to a method for mixing a scene with a virtual scenario, comprising the steps of

- 5 - providing a representation of said scene related to a first viewpoint, said representation comprising depth map data and texture data,
- generating a virtual scenario by means of a game engine unit,
- adapting the depth map data based on said generated virtual scenario, and
- forming a virtual video sequence based on said adapted depth map data and based on said texture data.

10

According to an aspect of the present invention there is provided a system for providing a virtual video sequence, comprising

image capturing means arranged to capture images so as to cover a scene from a first viewpoint,

15

an image representation generation unit arranged to generate at least one image representation based on said captured images, said image representation comprising depth map data and texture data,

a game engine unit arranged to generate a virtual scenario

20

a warping unit for warping each image representation generated by said image representation generation unit to a representation related to a second viewpoint so as to form a representation of the scene viewed from the second viewpoint and comprising depth map data and texture data,

25

an image processing unit arranged to mix the scene with the virtual scenario based on the representation of the scene viewed from the second viewpoint and on the generated scenario,

wherein

the second viewpoint is arbitrarily chosen within an area around the first viewpoint, and

30

the image processing unit is arranged to adapt the warped image representation based on said generated virtual scenario so as to provide the virtual video sequence comprising the warped representation of the scene viewed from the second viewpoint mixed with the virtual scenario,

3a

the system further comprises a position/posture estimation unit for estimating position and/or posture information at said second viewpoint in relation to said first viewpoint, wherein the position/posture estimation unit is arranged to determine the field of view from the second viewpoint, wherein
5 the position/posture estimation unit is arranged to elect one or a plurality of images with associated depth maps corresponding to the determined field of view and wherein the warping unit is arranged to process the depth maps of the images elected by the position/posture estimation unit and wherein the warping unit is arranged to adapt the depth map for the second viewpoint as
10 it is moving based on updated elections by the position/posture estimation unit.

In some embodiments, the game engine unit is arranged to generate the virtual scenario in a coordinate system aligned with the coordinate system of
15 the image representations.

In some embodiments, the game engine unit is arranged to generate at least one virtual object.

20 In some embodiments, the game engine unit is arranged to generate a sequence of virtual objects in accordance with a tactical scenario.

In some embodiments, the system comprises a display unit arranged to display the virtual video sequence.
25

In some embodiments, the display unit is arranged to act as a sight.

In some embodiments, the sight is a sight of a weapon.

30 In some embodiments, the display unit is formed at a sight.

In some embodiments, the display unit is a computer screen.

3b

According to another aspect of the present invention there is provided a method for generating a virtual scenario, comprising the steps of

5 providing a representation of a scene related to a first viewpoint, said representation comprising depth map data and texture data,

10 providing a representation of the scene viewed from a second viewpoint, said representation comprising depth map data and texture data and said step comprising warping at least the depth map data of the representation of the scene so as to correspond to the second viewpoint by means of a warping unit,

generating a virtual scenario by means of a game engine unit and mixing the scene with the virtual scenario based on the representation of the scene viewed from the second viewpoint and on the generated scenario,

15 choosing the second viewpoint arbitrary within an area around the first viewpoint and

20 wherein the step of mixing the scene with the virtual scenario comprises adapting the warped image representation based on said generated virtual scenario so as to provide a virtual video sequence comprising the warped representation of the scene viewed from the second viewpoint mixed with the virtual scenario,

25 wherein a position/posture estimation unit for estimating position and/or posture information at said second viewpoint in relation to said first viewpoint, determines the field of view from the second viewpoint, elects one or a plurality of images with associated depth maps corresponding to the determined field of view and wherein the warping unit processes the depth maps of the images elected by the position/posture estimation unit and wherein the warping unit adapts the depth map for the second viewpoint as it is moving based on updated elections by the position/posture estimation unit.

30

In some embodiments, the generation of a virtual scenario comprises forming depth map data and/or texture data for at least one virtual object.

3c

BRIEF DESCRIPTION OF THE DRAWINGS

Fig 1 is a block scheme of a system for providing a virtual video sequence with mixed in virtual objects

5

Fig 2 is a block scheme of a system for providing image data related to a first viewpoint in the system of Fig 1.

Fig 3 is a block scheme of a system for providing image data related to a second viewpoint in the system in Fig 1.

10

Fig 4 is a block scheme of a game engine in the system in Fig 1.

Fig 5 is a schematic view of a virtual object.

15

Fig 6 is a block scheme of an image processing unit in the system in Fig 1.

Fig 7 is a flow chart over a method for generating a representation of a scene.

20

Fig 8 is a flow chart over a method for generating a virtual video sequence representing a scene with mixed-in objects.

DETAILED DESCRIPTION

In Fig 1, a system 100 for mixing a real world scene with a virtual scenario comprises a memory 110 for 3D image data related to a real viewpoint, a system 130 for providing image data related to a virtual viewpoint, a game engine 140, an image processing unit 150 for mixing in virtual objects, and a display device 160 for displaying the provided virtual video sequence with the mixed in virtual objects. The memory 110 for 3D image data related to a real viewpoint is arranged to store a representation of a scene visible from the real viewpoint. The representation comprises for example depth map data and texture data. The system 130 for providing image data related to a virtual viewpoint is arranged to warp the image data from the real viewpoint so as to form a virtual scene as viewed from the virtual viewpoint. The virtual scene is in one example formed as a representation comprising for example depth map data and texture data. A pre-assumption is that the virtual viewpoint is located in a known relation to the real viewpoint. The game engine 140 is arranged to produce data related to virtual objects to be mixed with the depth map data related to the real or virtual viewpoint. The virtual objects generated by the game engine are in one example moving objects. In an alternative example, the generated virtual objects are stationary pop-up objects. The image processing unit 150 is arranged to mix the image data from the real or virtual viewpoint with the virtual objects generated by the game engine 140. The mixed data is then fed to the display device 150. The display device is in one example associated to the sights of a weapon. Alternatively, the display device is formed as a computer screen. Alternatively, the display device is a projector arranged to display a wide screen image.

In one example, the image processing unit 150 is arranged to cut the provided data so as to relate to a predetermined or adjustable field of view. In one example, the field of view is arranged to coincide with the field of view of a given type of sight. In one example, the field of view presented by the display device is 10°. In an alternative example, the field of view is 5°.

Accordingly, the system 100 can be used for training purpose for example for training searching through a terrain with a sight having a proportionately narrow field of view. In one example, movements of the sight can be performed by marking in the image. In the example, wherein the display unit
5 is a computer screen, the marking can be performed by means of a mouse or key board inputs or touch screen inputs.

In Fig 2, a system 220 for providing 3D image data related to a real viewpoint is shown. The system 220 is for use in the system 100 of Fig 1 for mixing a
10 real world scene with a virtual scenario. The system 220 for providing 3D image data related to a real viewpoint comprises a first image capturing unit 211 positioned at said real viewpoint. The first image capturing unit 211 is in one example arranged to capture an image of a scene in front of the first image capturing unit 211. In one example, wherein the scene is wide-angle,
15 such as covering an angle over 90° , over 180° or even 360° around, the first image capturing unit 211 is arranged to scan the scene so as to capture a plurality of images, each covering a sub-part of the entire scene. In accordance with this example, the first image capturing unit 211 is mounted on a rotatable plate 212. A control unit 219 is arranged to control a servo 213
20 associated to the plate 212 so as to control the rotatable plate 212 to turn to a plurality of angles of rotation. In one example, the angles of rotation are pre-defined. The control unit 219 is further arranged to control the first image capturing unit 211 to capture one image at each of said plurality of angles of rotation. The control unit 219 is further arranged to associate each image to
25 location data related to the location of the first image capturing unit 211 at the real viewpoint. In one example, the location data comprises position and/or posture information. In one example, the position information comprises for example a coordinate for the real viewpoint given in a chosen coordinate system. In the illustrated example, a position sensor 214 is arranged in
30 association with the real viewpoint so as to provide the coordinate for the real viewpoint. In one example, the position sensor 214 is a GPS receiver arranged to provide the coordinate for the viewpoint in a global geographical

coordinate system. The posture information comprises in one example the above described angle of rotation. The control unit 219 is arranged to store each captured image along with position and/or posture information in a memory 216. This or these images forming the scene are also referred to as
5 background image(s). In the illustrated example, a direction sensor 215 such as a compass is arranged in association with the first image capturing unit 211 arranged to provide the compass bearing or the like of the direction of the first image capturing unit 211.

10 The system 220 for providing 3D image data related to the real viewpoint comprises further a second image capturing unit 221. The second image capturing unit 221 is also positioned at the real viewpoint, but a known distance apart from the first image capturing unit 211. The distance is for example 2 meters or 1 meter. Accordingly, the coordinate of the second
15 image capturing unit 221 is known. In one example, the first image capturing unit 211 and the second image capturing unit 221 are spaced apart only in a vertical direction. In one example, the second image capturing unit 221 is controlled by the control unit 219 so as to operate in synchronism with the first image capturing unit 211. In an alternative example, the second image
20 capturing unit 221 is not arranged to operate in synchronism with the first image capturing unit 211

The images captured by the second image capturing unit 221 are stored in the memory 216. In one example, wherein it can be assumed that the second
25 image capturing unit 221 is rotating in synchronism with the first image capturing unit 211, each second image is stored in the memory 216 together with the associated first image captured by the first image capturing unit 211.

In an alternative example (as illustrated in fig 2), the images from the second
30 image capturing unit 221 are stored along with associated position information and/or posture information in the memory 216. Accordingly, the control unit 219 is as described above arranged to associate each image to

location data related to the location of the second image capturing unit 221 at the real viewpoint.

5 In an alternative example, for example wherein the images are not associated to location data, image pairs are found using conventional image processing techniques for example based on recognition of objects in the images.

10 In the illustrated example, a position sensor 224 such as a GPS receiver is arranged in association with the real viewpoint so as to provide a coordinate for the second image capturing unit 221 at the real viewpoint. In the illustrated example, a directional sensor 225 such as a compass is arranged in association with the second image capturing unit 221. The direction sensor 225 is arranged to provide the compass bearing or the like, so as to relate to
15 a global coordinate system.

In an alternative example, the system comprises one single image capturing unit, which is capturing the image(s) of the scene first from the first location and then from the second location. In one example, the first and second
20 image capturing units are cameras such as a digital still camera or a video camera. In one example, the first and second image capturing units are comprised in a stereo camera. In order to provide a wide-angle scene, the stereo camera can be mounted on the plate 212 or 222.

25 As was described above, the first and second image capturing units 211, 221 are in one example arranged spaced apart in a z-direction, i.e. in a height direction. In order to improve the accuracy in the subsequent image processing, at least a third image capturing unit may be mounted in known relation to the first and second image capturing units. The third image
30 capturing unit is arranged to capture images in the same manner as the first and second image capturing units. In accordance with this example, the at least first, second and third image capturing units are arranged in a horizontal

plane. Alternatively, the at least first, second and third image capturing units are spaced apart both horizontally and vertically. In a third example, the at least first, second and third image capturing units are vertically spaced apart. Apparently, one image capturing unit can be used and moved between the
5 different locations instead of using a plurality of image capturing units.

When more than two image capturing units are provided, an image set (instead of an image pair) is formed for each direction in the scene. Even though image sets comprising more than two images can be used, the
10 following description relates to an example with two images in each image set (or pair).

When a pair of first and second images has been captured, each image pair is processed in an image representation generation unit 217. The image
15 representation generation unit 217 is arranged to calculate a depth map for each image pair. The calculations performed by the processing unit so as to provide the depth map will not be described in detail herein. However, generally, since two images of the same scene captured at a known distance from each other are provided, the distance to different objects in the images
20 can be determined by means of image processing. Objects which are very far away are located in the same position in the two images while objects which are very near the image capturing unit(s) are not located in the same position in the two images. Accordingly, the displacement of the location of an object in the two images determines the distance to the object from the image
25 capturing device(s). Thus, the image representation generation unit 217 is arranged to provide a depth map for each first and/or second image based on the displacement of the location of objects in the two images so as to provide information about the distance to the objects in said images. In one example, one depth value is associated to each point (or pixel) in the images.
30 In an alternative example, a subset of the points (pixels) distributed over the images are associated to a depth value. In an extended example (as illustrated herein), the image representation generation unit 217 is arranged

to improve the data of the depth map by using information about known distances to one or a plurality of objects identifiable in the images. In one example, the distance to the identified object(s) in the images is known based on locally stored or received information about the locations of identifiable objects and based on knowledge about the location of the real viewpoint. In the illustrated example, a laser range finder 218 is arranged at the real viewpoint. The laser range finder 218 is arranged to measure the distance to an object at which it is pointed. The image representation generation unit 217 is then arranged to correlate the distance information with the data of the depth map. The control unit 219 is arranged to feed image data representing the first and/or second images building up the scene along with associated depth map(s) to the memory 110 for image data related to the real viewpoint (as shown in Fig 1). As is understood from the above, the image data and/or depth maps are in one example each associated to a position and/or posture. The image data representing the first and/or second images building up the scene comprises for example texture data.

In Fig 3, a system 330 for providing image data related to a virtual viewpoint comprises an input unit 331 for inputting location information related to the virtual viewpoint, an interface 334 arranged to receive the data from the memory 110 for image data related to the real view point, a position/posture estimation unit 332 for estimating position and/or posture information related to said virtual viewpoint in relation to said real viewpoint, and a warping unit 333 arranged to adapt the above described depth maps so as to form a virtual scene viewed from the virtual viewpoint. The data from the memory 110, received over the interface 334 comprises, as described above, the image data representing the first and/or second images building up the scene at the real viewpoint along with associated depth map(s) and possibly also position and/or posture information related to the images building up the scene.

The input unit 331 is in the illustrated example connected to a position sensor 336 such as a GPS-receiver arranged to receive position information for the virtual viewpoint. Alternatively, position data is inputted manually by means of the input unit 331. For example, the input unit is a computer mouse or joystick arranged to mark a position in a scene displayed by the display device 150. In the illustrated example, the input unit 331 is further connected to a direction sensor 337, such as a compass, arranged to provide a compass bearing or the like so as to provide an indication of the field of view from the virtual viewpoint. The position/posture estimation unit 332 is arranged to determine a relation between the position at the virtual and the real viewpoints. In accordance therewith, the position/posture estimation unit 332 is arranged to fetch position data related to the real viewpoint from the memory 110 and to receive over the interface 334 the position data from the input unit 331 related to the position of the virtual viewpoint. The position data related to the real and virtual viewpoints is in one example given in x, y and z coordinates. The position/posture estimation unit 332 is then arranged to determine the relative distance between the real and virtual viewpoints based on the position data related to the real and virtual viewpoints. The position/posture estimation unit 332 is arranged to receive the posture information from the input unit 331 and to determine a field of view from the virtual viewpoint based on said posture information and based on the position of the virtual viewpoint. Based on the determined field of view, the position/posture estimation unit 332 is arranged to elect one or a plurality of images from the memory 110 with associated depth maps corresponding to the determined field of view.

The warping unit 333 is arranged to process the depth maps of the images elected by the position/posture estimation unit 332. The processing involves adapting the distances to different objects of the depth map(s) related to the real viewpoint to the distances to said objects seen from the virtual viewpoint. The warping unit 333 is arranged to form a depth map related to the virtual viewpoint, wherein the adapted distances are given.

In one example, the location of the virtual viewpoint is continuously moving. In accordance with this example, position/posture estimation unit is arranged to elect new images/depth maps from the memory 110 based on the
5 information from the input unit 331 related to updated positions and/or postures. The warping unit 333 is then arranged to adapt the depth map for the moving virtual viewpoint based on the updated elections. Accordingly, the depth map for the virtual viewpoint can be frequently updated so as to follow the movements of the virtual viewpoint. In one example, the updating is
10 performed in real time. The virtual viewpoint is for example associated to a person, such as a soldier or a vehicle, such as a tank.

In Fig 4, a game engine 440 is arranged to generate a scenario by means of virtual objects. The scenario can be automatically generated or generated in
15 interaction with an operator. In detail, the game engine is arranged to form virtual objects to be superimposed on images provided by either the system 220 arranged to provide image data related to the real viewpoint or the system 130 arranged to provide image data related to the virtual viewpoint. In one example, the virtual objects are formed in real time. In one example, the
20 game engine 440 is arranged to form virtual objects in the form of virtual targets in a training mode of operation of the game engine. The game engine 440 is in one example arranged to operate in an additional or alternative mode of operation, such as a combat mode of operation wherein the virtual objects comprises virtual combating forces.

25

The game engine 440 comprises in Fig 4 an operator interface 441. The operator interface 441 comprises in one example input means for choosing mode of operation of the game engine 440. The operator interface 441
30 comprises in one example means for choosing for example types of virtual objects, distances to the chosen virtual objects, and orientation of the chosen virtual objects. In one example, the virtual objects are stationary. In an alternative example, the virtual objects are arranged to be moving. In

accordance therewith, the interface comprises in one example means for electing a movement pattern or for freely choosing the movements for example by indication on a map. The map is for example formed on a touch screen on the operator interface. The information about the virtual object(s)
5 inputted by means of the interface is fed to a virtual object forming unit 442. The interface is in one example also arranged to receive coordinate data. This data can be used for transferring images provided by the game engine to a coordinate system used by systems for providing image data related to the real and virtual viewpoints.

10

The virtual object forming unit 442 is arranged to form representations of the virtual objects in the form of coordinate data. In the example, wherein the virtual objects are moving, the virtual object forming unit can be arranged to provide said coordinate data in real time. The coordinate data representing
15 the virtual object or virtual objects is feed to the image processing unit 150.

20

In Fig 5, a virtual object is built up by a number of corners A, B, C, D. Lines can be drawn between the corners A, B, C, D so as to build up a contour. Each corner of the contour is associated to a coordinate. In one example, the
coordinate is given in a tree dimensional coordinate system. The coordinate system is for example a Cartesian coordinate system. The coordinate system is in one example arranged to provide latitude, longitude and height data. In an alternative example, the coordinate system is local. For example, the origin of the coordinate system can be located in the virtual viewpoint.

25

In Fig 6, an image processing unit 650 is arranged to mix the virtual objects provided by the game engine 440 with the image representation data provided by either the real viewpoint image data system 220 or the virtual viewpoint image data system 130 so as to provide a virtual video sequence.
30 The image processing unit 650 comprises an interface 653 for receiving the above mentioned information. In the illustrated example, the image processing unit 650 comprises an alignment unit 651 arranged to align the

coordinate system used by the game engine with the coordinate system used by the real viewpoint image data system 220 and the virtual viewpoint image data system 130. The aligned coordinate data related to the virtual objects provided by the game engine 440 is then mixed in a mixing unit 652 with the
5 image data from the real viewpoint image data system 220 or the virtual viewpoint image data system 130. The mixing unit 652 is arranged to superpose the virtual objects on the image(s) of the scene, wherein the virtual objects are not obscured by objects in the scene. Accordingly, the mixing unit is arranged to generate a video sequence, wherein the virtual
10 objects are mixed together with the background by means of the depth map data. In the video sequence, the virtual objects can move behind real objects in the background. The video sequence provided by the mixing unit 652 is fed to the display device 160.

15 In Fig 7, a method 770 for generating an image representation comprises a first step 771 of capturing images of a scene and a second step 772 of generating an image representation comprising depth map data and texture data or the like.

20 In Fig 8, a method 880 for mixing a scene with a virtual scenario comprises a first step of providing 883 a representation of said scene, said representation comprising depth map data and texture data or the like. The representation of the scene is based on a real world scene. The representation directly reflects a captured image. In a second step 884, the representation is warped so as
25 to reflect an image associated to a second, virtual viewpoint. In detail, the virtual viewpoint is placed somewhere in the undistorted image and the undistorted image is distorted in accordance with the location of the virtual viewpoint. This second step can be omitted if a representation related to the captured scene is desired.

30

In a third step 885, a virtual scenario is generated by means of a game engine unit. In a fourth step 886, the representation is adapted based on the

generated virtual scenario. In a fifth step 887, a virtual video sequence is formed based on said adapted representation and based on said texture data.

- 5 The generation 885 of a virtual scenario comprises in one example forming depth map data and/or texture data for at least one virtual object.

The embodiments of the invention in which an exclusive property or privilege is claimed are defined as follows:

1. A system for providing a virtual video sequence, comprising
 - image capturing means arranged to capture images so as to cover a scene from a first viewpoint,
 - an image representation generation unit arranged to generate at least one image representation based on said captured images, said image representation comprising depth map data and texture data,
 - a game engine unit arranged to generate a virtual scenario
 - a warping unit for warping each image representation generated by said image representation generation unit to a representation related to a second viewpoint so as to form a representation of the scene viewed from the second viewpoint and comprising depth map data and texture data,
 - an image processing unit arranged to mix the scene with the virtual scenario based on the representation of the scene viewed from the second viewpoint and on the generated scenario,wherein
 - the second viewpoint is arbitrarily chosen within an area around the first viewpoint, and
 - the image processing unit is arranged to adapt the warped image representation based on said generated virtual scenario so as to provide the virtual video sequence comprising the warped representation of the scene viewed from the second viewpoint mixed with the virtual scenario,
 - the system further comprises a position/posture estimation unit for estimating position and/or posture information at said second viewpoint in relation to said first viewpoint, wherein the position/posture estimation unit is arranged to determine the field of view from the second viewpoint, wherein the position/posture estimation unit is arranged to elect one or a plurality of images with associated depth maps corresponding to the determined field of view and wherein the warping unit is arranged to process the depth maps of the images elected by the position/posture estimation unit and wherein the warping unit is

arranged to adapt the depth map for the second viewpoint as it is moving based on updated elections by the position/posture estimation unit.

2. A system according to claim 1, wherein the game engine unit is arranged to generate the virtual scenario in a coordinate system aligned with the coordinate system of the image representations.

3. A system according to claim 1 or 2, wherein the game engine unit is arranged to generate at least one virtual object.

4. A system according to any one of claims 1 to 3, wherein the game engine unit is arranged to generate a sequence of virtual objects in accordance with a tactical scenario.

5. A system according to any one of claims 1 to 4, wherein it comprises a display unit arranged to display the virtual video sequence.

6. A system according to claim 5, wherein the display unit is arranged to act as a sight.

7. A system according to claim 6, wherein the sight is a sight of a weapon.

8. A system according to claim 5, wherein the display unit is formed at a sight.

9. A system according to any one of claims 1 to 7, wherein the display unit is a computer screen.

10. A method for generating a virtual scenario, comprising the steps of providing a representation of a scene related to a first viewpoint, said representation comprising depth map data and texture data,

providing a representation of the scene viewed from a second viewpoint, said representation comprising depth map data and texture data and said step comprising warping at least the depth map data of the representation of the scene so as to correspond to the second viewpoint by means of a warping unit, generating a virtual scenario by means of a game engine unit and mixing the scene with the virtual scenario based on the representation of the scene viewed from the second viewpoint and on the generated scenario, choosing the second viewpoint arbitrary within an area around the first viewpoint and

wherein the step of mixing the scene with the virtual scenario comprises adapting the warped image representation based on said generated virtual scenario so as to provide a virtual video sequence comprising the warped representation of the scene viewed from the second viewpoint mixed with the virtual scenario,

wherein a position/posture estimation unit for estimating position and/or posture information at said second viewpoint in relation to said first viewpoint, determines the field of view from the second viewpoint, elects one or a plurality of images with associated depth maps corresponding to the determined field of view and wherein the warping unit processes the depth maps of the images elected by the position/posture estimation unit and wherein the warping unit adapts the depth map for the second viewpoint as it is moving based on updated elections by the position/posture estimation unit.

11. A method according to claim 10, wherein the generation of a virtual scenario comprises forming depth map data and/or texture data for at least one virtual object.

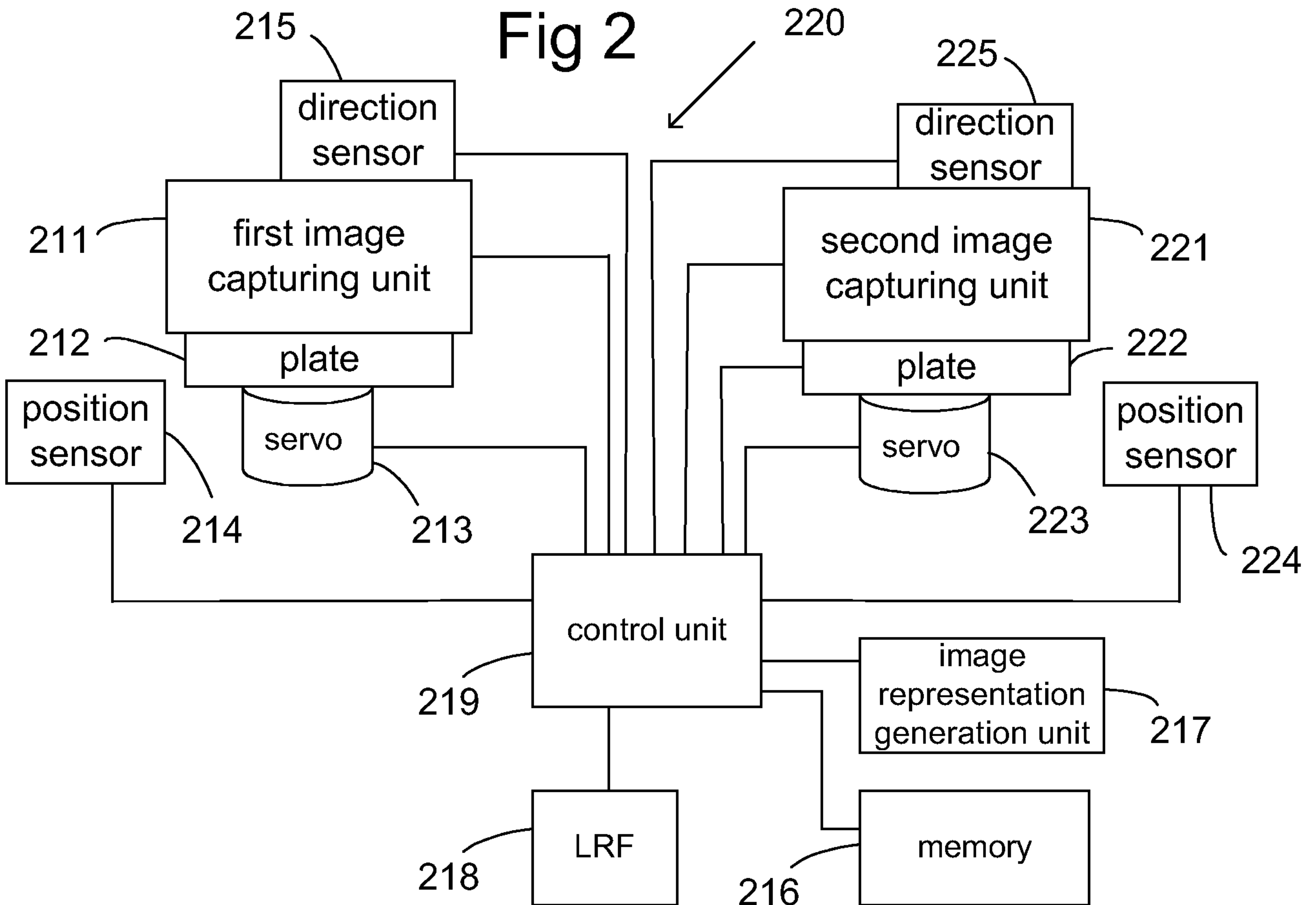
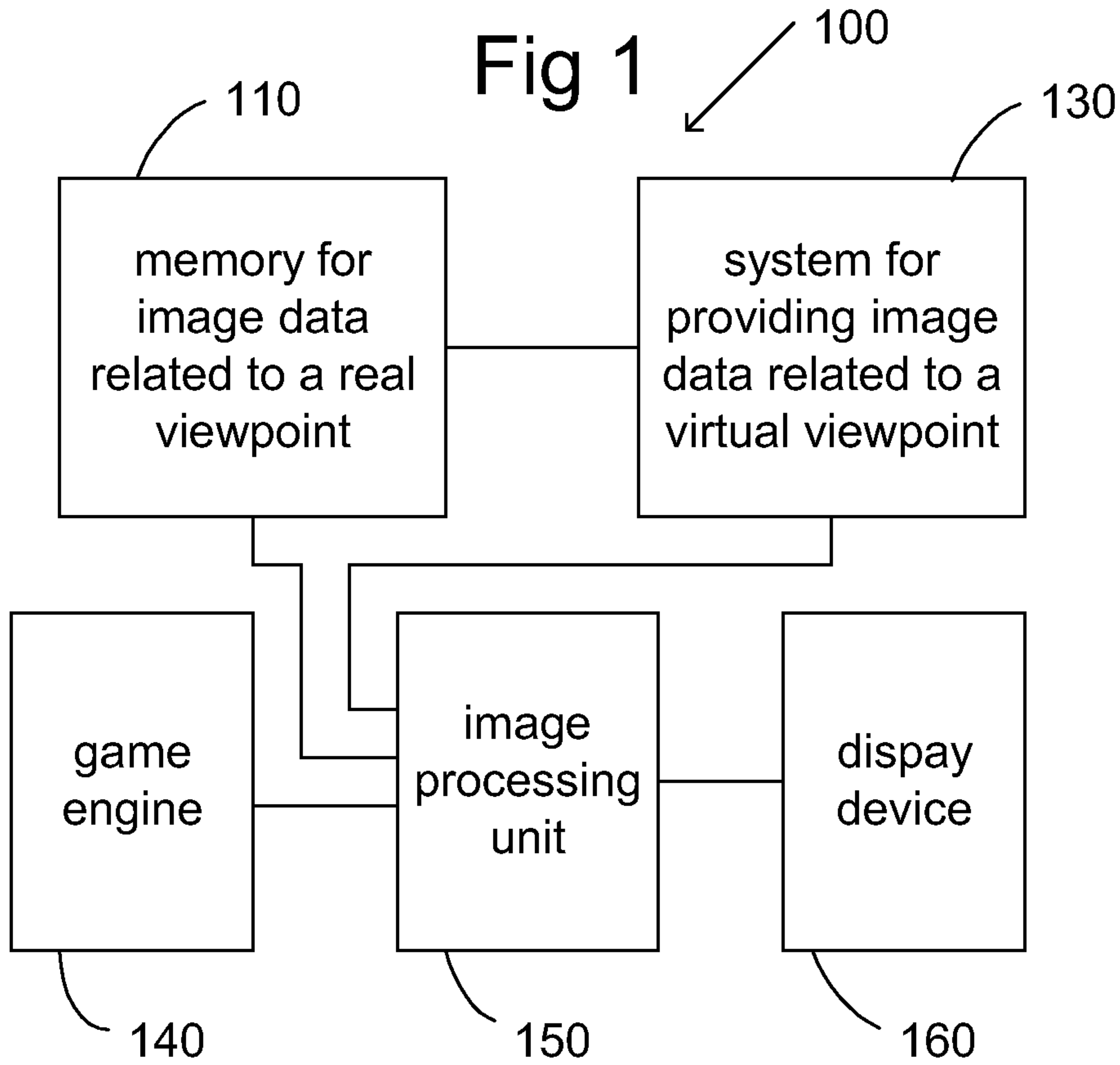


Fig 3

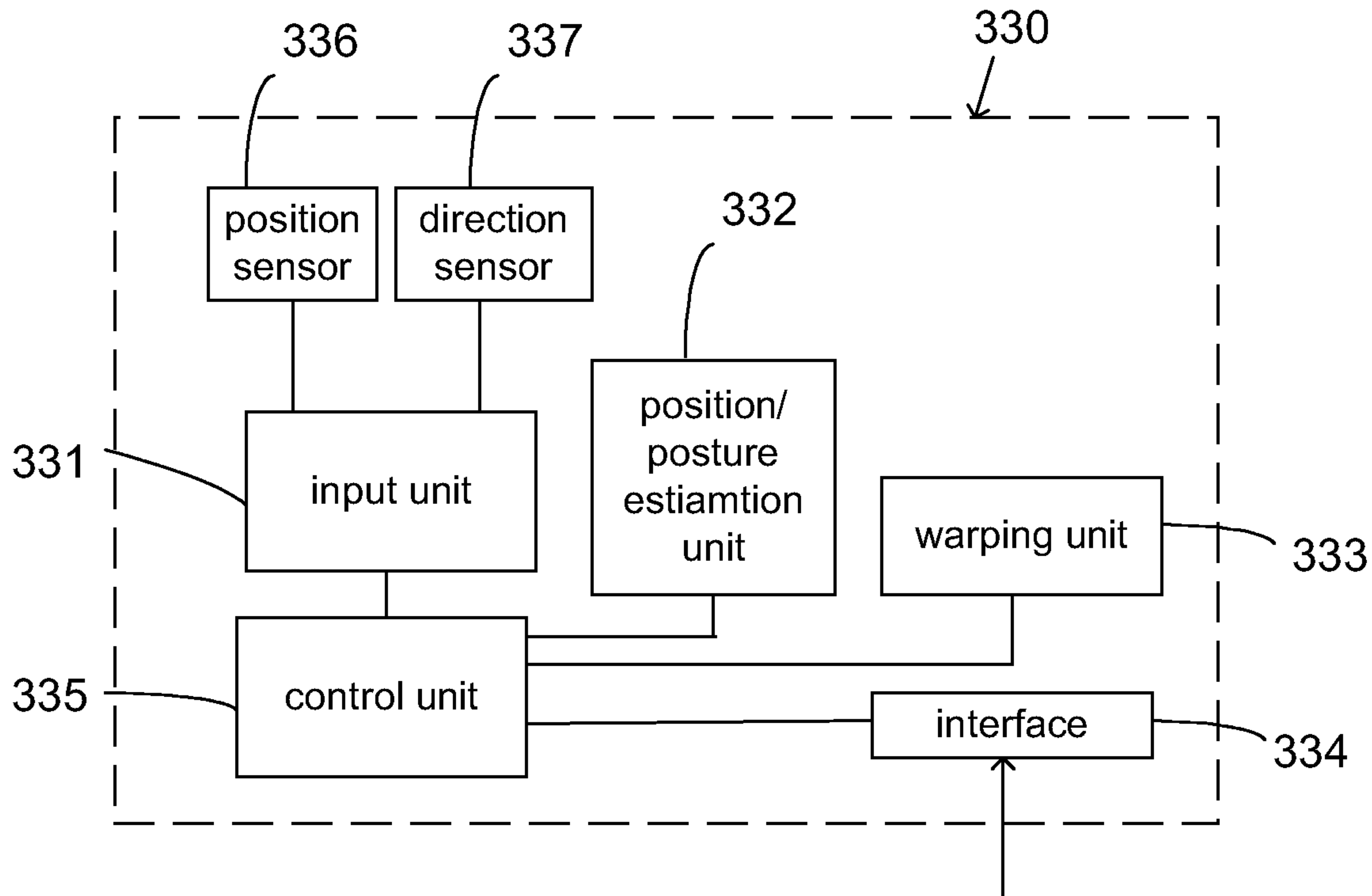


Fig 4

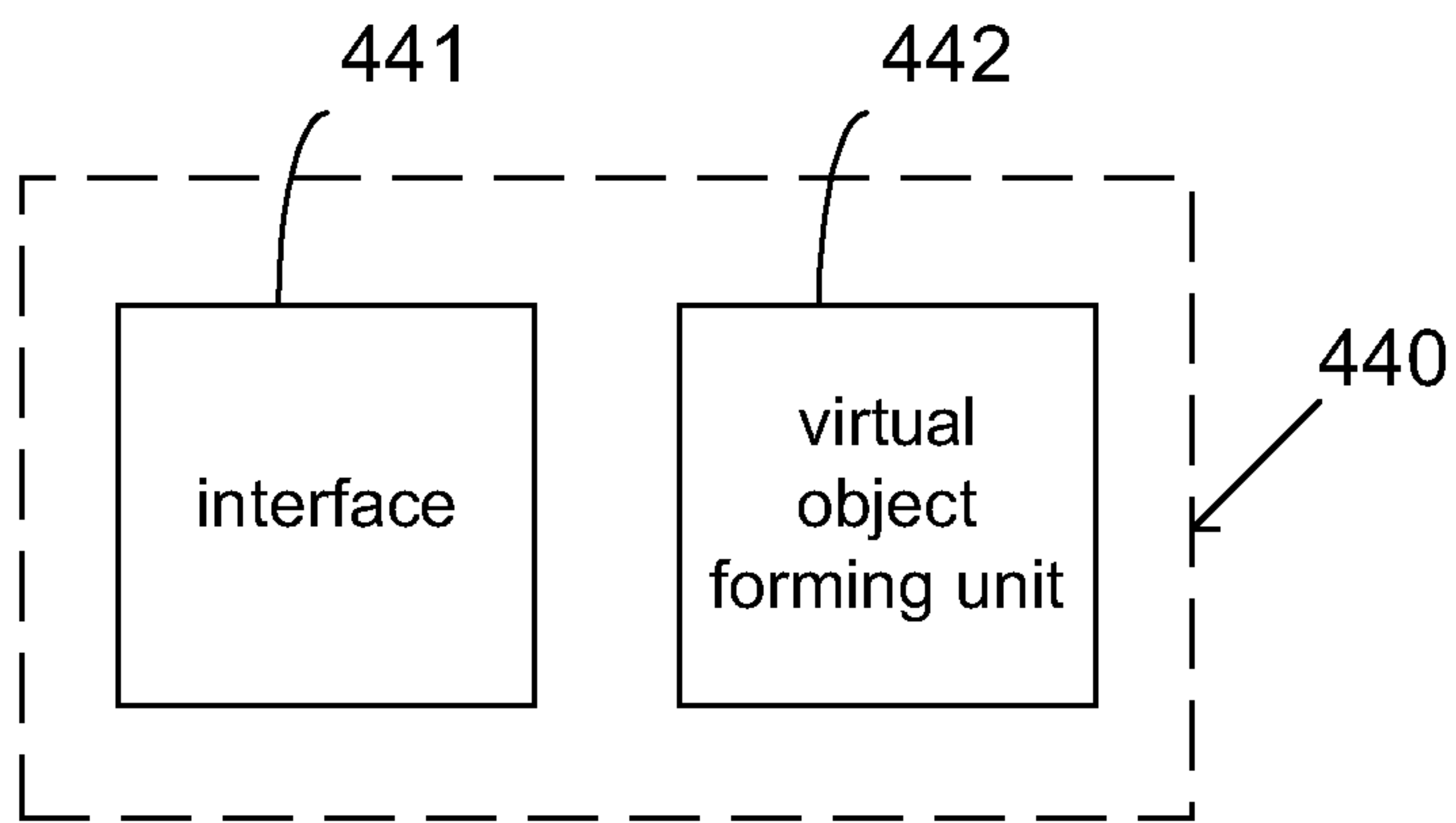


Fig 5

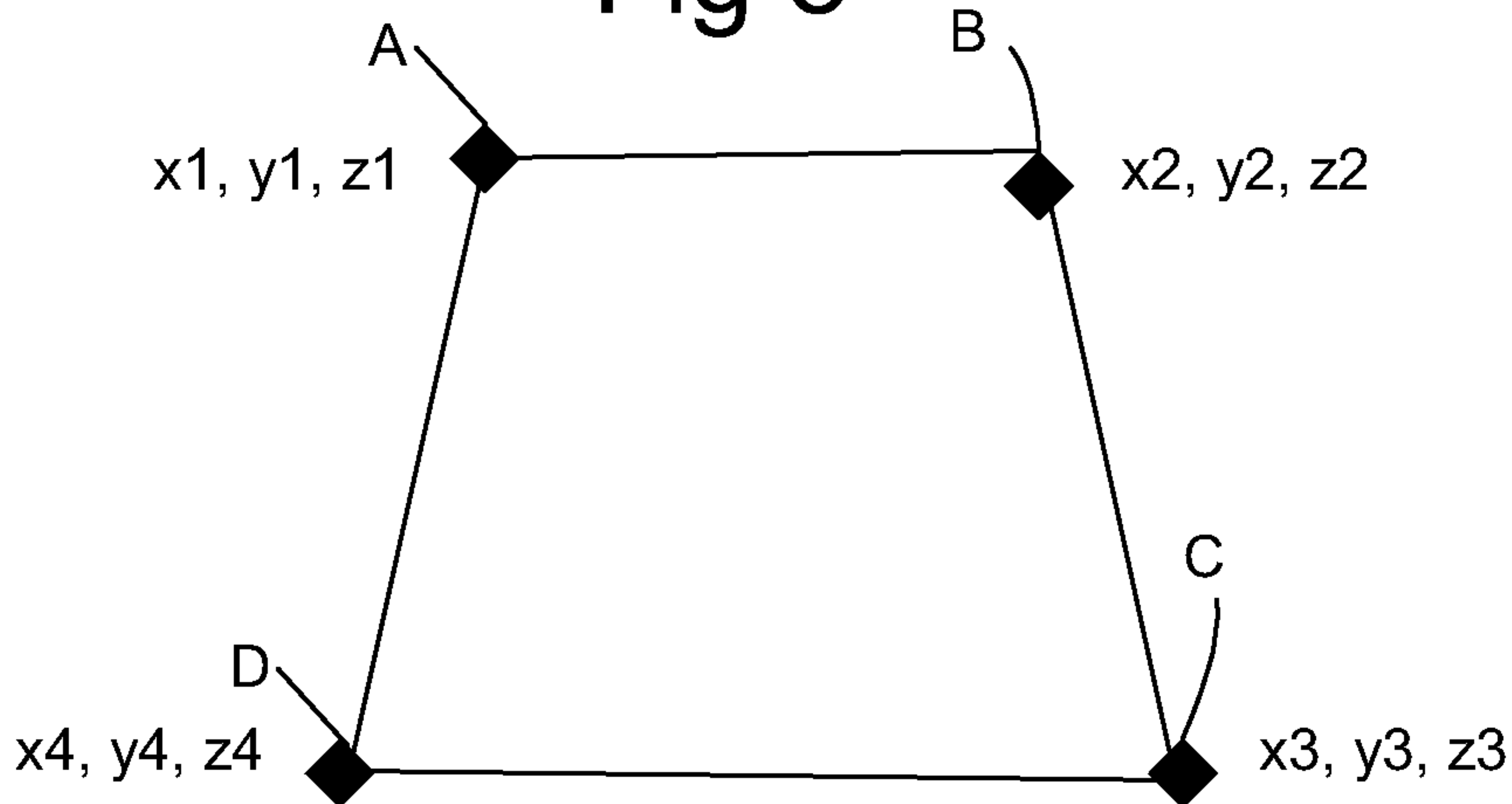


Fig 6

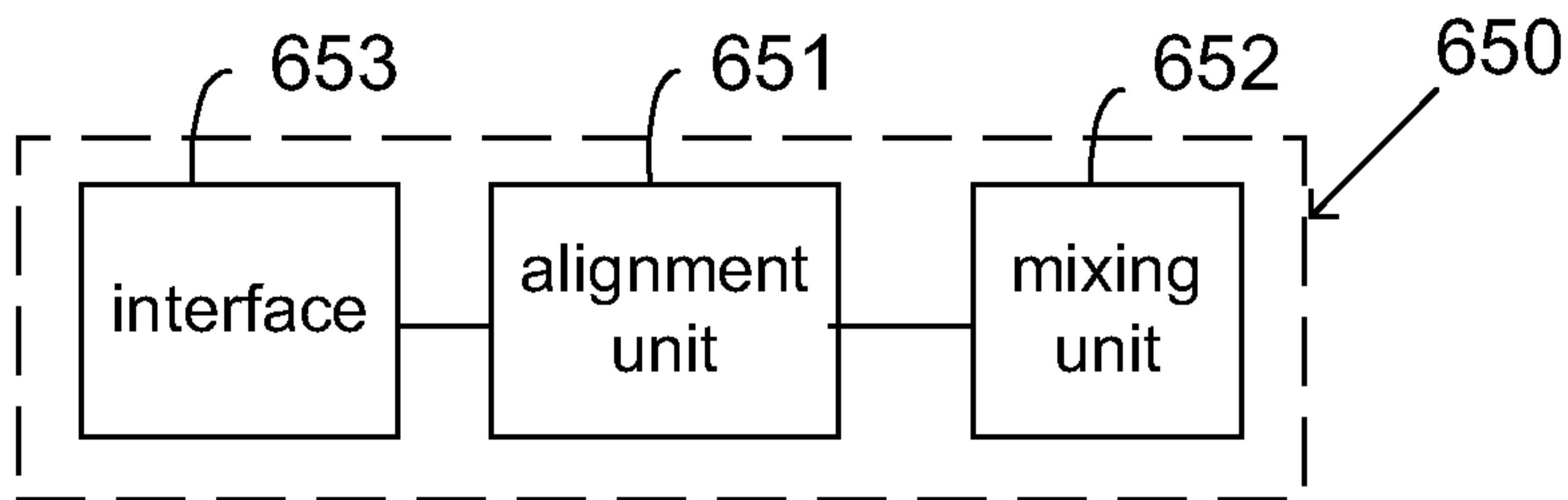


Fig 7

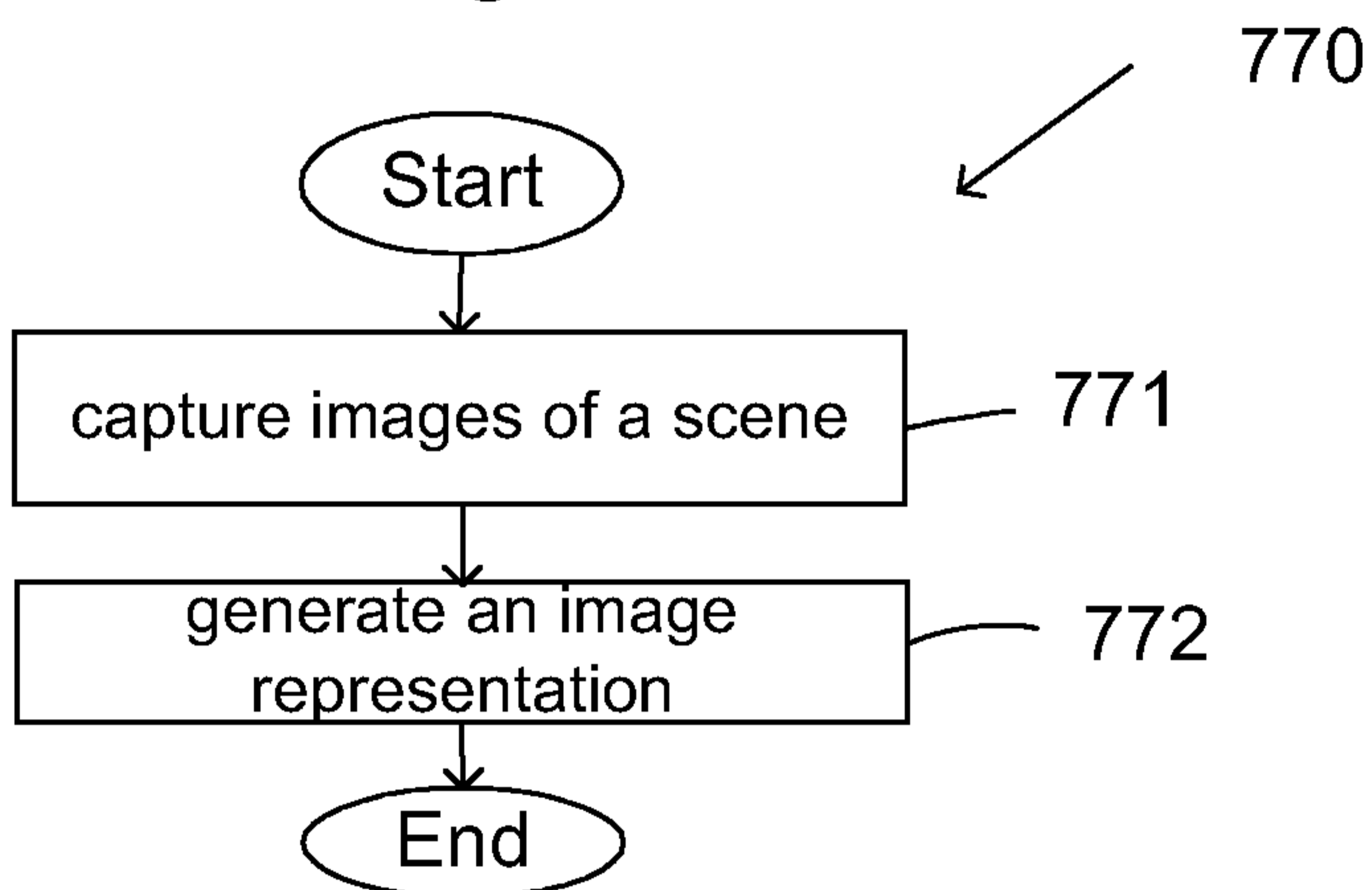


Fig 8

