

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
10 April 2008 (10.04.2008)

PCT

(10) International Publication Number  
**WO 2008/041661 A1**

- (51) International Patent Classification:  
*H04N 13/00* (2006.01) *G06T 15/40* (2006.01)
- (21) International Application Number:  
PCT/JP2007/069121
- (22) International Filing Date:  
21 September 2007 (21.09.2007)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:  
2006-271052 2 October 2006 (02.10.2006) JP
- (71) Applicant (for all designated States except US):  
**Kabushiki Kaisha Toshiba** [JP/JP]; 1-1, Shibaura 1-chome, Minato-ku, Tokyo, 1058001 (JP).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): **YAMAUCHI, Yasunobu** [JP/JP]; c/o Intellectual Property Division, Toshiba Corporation,, 1-1, Shibaura 1-chome, Minato-ku, Tokyo, 1058001 (JP). **FUKUSHIMA, Rieko** [JP/JP]; c/o Intellectual Property Division, Toshiba Corporation,, 1-1, Shibaura 1-chome, Minato-ku, Tokyo, 1058001 (JP). **SUGITA, Kaoru** [JP/JP]; c/o Intellectual Property Division, Toshiba Corporation,, 1-1, Shibaura 1-chome, Minato-ku, Tokyo, 1058001 (JP). **HIRAYAMA, Yuzo**

[JP/JP]; c/o Intellectual Property Division, Toshiba Corporation,, 1-1, Shibaura 1-chome, Minato-ku, Tokyo, 1058001 (JP). **DOI, Miwako** [JP/JP]; c/o Intellectual Property Division, Toshiba Corporation,, 1-1, Shibaura 1-chome, Minato-ku, Tokyo, 1058001 (JP).

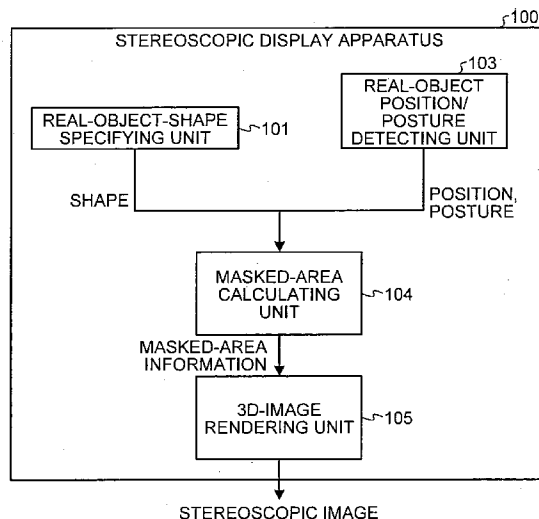
(74) Agent: **SAKAI, Hiroaki**; Sakai International Patent Office, Kasumigaseki Building, 2-5, Kasumigaseki 3-chome, Chiyoda-ku, Tokyo 1006020 (JP).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, 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, 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, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, PL,

[Continued on next page]

(54) Title: METHOD, APPARATUS, AND COMPUTER PROGRAM PRODUCT FOR GENERATING STEREOSCOPIC IMAGE



(57) Abstract: A detecting unit detects at least one of a position and a posture of a real object located on or near a three-dimensional display surface. A calculating unit calculates a masked-area where the real object masks a ray irradiated from the three-dimensional display surface, based on at least one of the position and the posture. A rendering unit renders a stereoscopic image by performing different rendering processes on the masked-area from rendering processes on other areas.

WO 2008/041661 A1



PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM,  
GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

**Published:**

— *with international search report*

## DESCRIPTION

METHOD, APPARATUS, AND COMPUTER PROGRAM PRODUCT FOR  
GENERATING STEREOSCOPIC IMAGE

5

## TECHNICAL FIELD

The present invention relates to a technology for generating a stereoscopic image linked to a real object.

## 10 BACKGROUND ART

Various methods have been used to realize a stereoscopic-image display apparatus, i.e., a so-called three-dimensional display apparatus, which displays a moving image. There is an increasing need for a flat-panel display apparatus that does not require stereoscopic glasses. There is a relatively easy method of providing a beam controller right in front of a display panel with fixed pixels such as a direct-view-type or projection-type liquid crystal display panel or plasma display panel, where the beam controller controls beams from the display panel to direct a viewer.

The beam controller is also referred to as a parallax barrier, which controls the beams so that different images are seen on a point on the beam controller depending on an angle. For example, to use only a horizontal parallax, a slit or a lenticular sheet that includes a cylindrical lens array is used as the beam controller. To use a vertical parallax at the same time, one of a pinhole array and a lens array is used as the beam controller.

30 A method that uses the parallax barrier is further classified into a bidirectional method, an omnidirectional method, a super omnidirectional method (a super omnidirectional condition of the omnidirectional method),

and an integral photography (hereinafter, "IP method"). The methods use a basic principle substantially same as what was invented about a hundred years ago and has been used for stereoscopic photography.

5           Because a visual range is generally limited, both of the IP method and the multi-lens method generate an image so that a transparent projected image can be actually seen at the visual range. For example, as disclosed in JP-A 2004-295013 (KOKAI) and JP-A 2005-86414 (KOKAI), if a  
10 horizontal pitch of the parallax barrier is an integral multiplication of a horizontal pitch of the pixels when using a one-dimensional IP method that uses only the horizontal parallax, there are parallel rays (hereinafter, "parallel-ray one-dimensional IP"). Therefore, an accurate  
15 stereoscopic image is acquired by dividing an image with respect to each pixel array and synthesizing a parallax-synthesized image to be displayed on a screen, where the image before dividing is a perspective projection at a constant visual range in the vertical direction and a  
20 parallel projection in the horizontal direction.

          In the omnidirectional method, the accurate stereoscopic image is acquired by dividing and arranging a simple perspective projection image.

          It is difficult to realize an imaging device that uses  
25 different projection methods or different distances to a projection center between the vertical direction and the horizontal direction because it requires a camera or a lens with the size equal to a subject, especially for parallel projection. To acquire parallel projection data by imaging,  
30 it is realistic to convert the image from the imaging data of the perspective projection. For example, a ray-space method based on compensation using an epipolar plane (EPI) has been known.

To display a stereoscopic image by reproducing the beams, a three-dimensional display based on integral imaging method can reproduce a high-quality stereoscopic image by increasing amount of information of the beams to be reproduced. The information is, for example, the number of points of sight in the case of the omnidirectional method, or the number of the beams in different directions from a display plane in the case of the IP method.

However, the processing load of reproducing the stereoscopic image depends on the processing load of rendering from each point of sight, i.e., rendering in computer graphics (CG), and it increases in proportion to the number of the points of sight or the beams. Specifically to reproduce a voluminous image in three dimensions, it is required to render volume data that defines medium density that forms an object from each point of sight. Rendering the volume data generally requires excessive load of calculating because tracking beams, i.e., ray casting, and calculating an attenuation rate have to be performed on all of the volume elements.

Therefore, to render the volume data on the integral-imaging three-dimensional display, the processing load further increases in proportion to the increased number of the points of sight and the beams. Moreover, when a surface-level modeling such as a polygon is employed at the same time, a fast rendering method based on the polygon cannot be fully utilized because the processing speed is controlled by a rendering process based on a ray tracing method, and the total processing load in the image generation increases.

Fusion of a real object and a stereoscopic virtual object and an interaction system use a technology such as mixed reality (MR), augmented reality (AR), or virtual

reality (VR). The technologies can be roughly classified into two groups; the MR and the AR that superposes a virtual image created by CG on a real image, and the VR that inserts a real object into a virtual world created by CG as in cave automatic virtual equipment.

By reproducing a CG virtual space using a bidirectional stereo method, a CG-reproduced virtual object can be produced in a three-dimensional position and posture as in the real world. In other words, the real object and the virtual object can be displayed in corresponding position and posture; however, the image needs to be configured every time the point of sight of a user changes. Moreover, to reproduce visual reality that depends on the point of sight of the user, a tracking system is required to detect the position and the posture of the user.

#### DISCLOSURE OF INVENTION

An apparatus for generating a stereoscopic image, according to one aspect of the present invention, includes a detecting unit that detects at least one of a position and a posture of a real object located on or near a three-dimensional display surface; a calculating unit that calculates a masked-area where the real object masks a ray irradiated from the three-dimensional display surface, based on at least one of the position and the posture; and a rendering unit that renders a stereoscopic image by performing different rendering processes on the masked-area from rendering processes on other areas.

A method of generating a stereoscopic image, according to another aspect of the present invention, includes detecting at least one of a position and a posture of a real object located on or near a three-dimensional display surface; calculating a masked-area where the real object

masks a ray irradiated from the three-dimensional display surface, based on at least one of the position and the posture; and rendering a stereoscopic image by performing different rendering processes on the masked-area from  
5 rendering processes on other areas.

A computer program product according to still another aspect of the present invention includes a computer-usable medium having computer-readable program codes embodied in the medium that when executed cause a computer to execute  
10 detecting at least one of a position and a posture of a real object located on or near a three-dimensional display surface; calculating a masked-area where the real object masks a ray irradiated from the three-dimensional display surface, based on at least one of the position and the  
15 posture; and rendering a stereoscopic image by performing different rendering processes on the masked-area from rendering processes on other areas.

#### BRIEF DESCRIPTION OF DRAWINGS

20 Fig. 1 is a block diagram of a stereoscopic display apparatus according to a first embodiment of the present invention;

Fig. 2 is an enlarged perspective view of a display panel of the stereoscopic display apparatus;

25 Fig. 3 is a schematic diagram of parallax component images and a parallax-synthesized image in an omnidirectional stereoscopic display apparatus;

Fig. 4 is a schematic diagram of the parallax component images and a parallax-synthesized image in a  
30 stereoscopic display apparatus based on one-dimensional IP method;

Figs. 5 and 6 are schematic diagrams of parallax images when a point of sight of a user changes;

Fig. 7 is a schematic diagram of a state where a transparent cup is placed on the display panel of the stereoscopic display apparatus;

Fig. 8 is a schematic diagram of hardware in a real-object position/posture detecting unit shown in Fig. 1;

Fig. 9 is a flowchart of a stereoscopic-image generating process according to the first embodiment;

Fig. 10 is an example of an image of the transparent cup with visual reality;

Fig. 11 is an example of drawing a periphery of the real object as volume data;

Fig. 12 is an example of drawing an internal concave of a cylindrical real object as volume data;

Fig. 13 is an example of drawing virtual goldfish autonomously swimming in the internal concave of the cylindrical real object;

Fig. 14 is a function block diagram of a stereoscopic display apparatus according to a second embodiment of the present invention;

Fig. 15 is a flowchart of a stereoscopic-image generating process according to the second embodiment;

Fig. 16 is a schematic diagram of a point of sight, a flat-laid stereoscopic display panel, and a real object seen from 60-degree upward;

Fig. 17 is a schematic diagram of spherical coordinate used to perform texture mapping that depends on positions of the point of sight and a light source;

Fig. 18 is a schematic diagram of a vector  $U$  and a vector  $V$  in a projected coordinate system;

Figs. 19A and 19B are schematic diagrams of a relative direction  $\theta$  in a longitudinal direction;

Fig. 20 is a schematic diagram of the visual reality when a tomato bomb hits and crashes on the real transparent



cup;

Fig. 21 is a schematic diagram of the flat-laid stereoscopic display panel and a plate;

Fig. 22 is a schematic diagram of the flat-laid stereoscopic display panel, the plate, and a cylindrical object; and

Fig. 23 is a schematic diagram of linear markers on both ends of the plate to detect a shape and a posture of the plate.

10

#### BEST MODE(S) FOR CARRYING OUT THE INVENTION

Exemplary embodiments of the present invention are explained in detail below with reference to the accompanying drawings.

15 As shown in Fig. 1, a stereoscopic display apparatus 100 includes a real-object-shape specifying unit 101, a real-object position/posture detecting unit 103, a masked-area calculating unit 104, and a 3D-image rendering unit 105. The stereoscopic display apparatus 100 further  
20 includes hardware such as a stereoscopic display panel, a memory, and a central processing unit (CPU).

The real-object position/posture detecting unit 103 detects at least one of a position, a posture, and a shape of a real object on or near the stereoscopic display panel.  
25 A configuration of the real-object position/posture detecting unit 103 will be explained later in detail.

The real-object-shape specifying unit 101 receives the shape of the real object as specified by a user.

The masked-area calculating unit 104 calculates a  
30 masked-area where the real object masks a ray irradiated from the stereoscopic display panel based on the shape received by the masked-area calculating unit 104 and at least one of the position, the posture, and the shape

detected by the real-object position/posture detecting unit 103.

The 3D-image rendering unit 105 performs rendering process on the masked-area calculated by the masked-area calculating unit 104 in a different manner from a manner used in other areas (Namely, the 3D-image rendering unit 105 performs different rendering processes on the masked-area calculated by the masked-area calculating unit 104 from rendering processes on other areas), generates a parallax-synthesized image, thereby renders a stereoscopic image, and outputs it. According to the first embodiment, the 3D-image rendering unit 105 renders the stereoscopic image on the masked-area as volume data that includes points in a three-dimensional space.

A method of generating an image on the stereoscopic display panel of the stereoscopic display apparatus 100 according to the first embodiment is explained below. The stereoscopic display apparatus 100 is designed to reproduce beams with  $n$  parallaxes. The explanation is given assuming that  $n$  is nine.

As shown in Fig. 2, the stereoscopic display apparatus 100 includes lenticular plates 203 arranged in front of a screen of a flat parallax-image display unit such as a liquid crystal panel. Each of the lenticular plates 203 includes cylindrical lenses with an optical aperture thereof vertically extending, which are used as beam controllers. Because the optical aperture extends linearly in the vertical direction and not obliquely or in a staircase pattern, pixels are easily arranged in a square array to display a stereoscopic image.

On the screen, pixels 201 with the vertical to horizontal ratio of 3:1 are arranged linearly in a lateral direction so that red (R), green (G), and blue (B) are

alternately arranged in each row and each column. A longitudinal cycle of the pixels 201 (3Pp shown in Fig. 2) is three times of a lateral cycle of the pixels 201 (Pp shown in Fig. 2).

5 In a color image display apparatus that displays a color image, three pixels 201 of R, G, and B form one effective pixel, i.e., a minimum unit to set brightness and color. Each of R, G, and B is generally referred to as a sub-pixel.

10 A display panel shown in Fig. 2 includes a single effective pixel 202 consisting of nine columns and three rows of the pixels 201 as surrounded by a black border. The cylindrical lens of the lenticular plate 203 is arranged substantially in front of the effective pixel 202.

15 Based on one-dimensional integral photography (IP method) using parallel beams, the lenticular plate 203 reproduces parallel beams from every ninth pixel in each row on the display panel. The lenticular plate 203 functions as a beam controller that includes cylindrical  
20 lenses linearly extending at a horizontal pitch (Ps shown in Fig. 2) nine times as much as the lateral cycle of the sub-pixels.

Because the point of sight is actually set at a limited distance from the screen, the number of parallax  
25 component images is nine or more. The parallax component image includes image data of a set of pixels that form the parallel beams in the same parallax direction required to form an image by the stereoscopic display apparatus 100. By the beams to be actually used being extracted from the  
30 parallax component image, the parallax-synthesized image to be displayed on the stereoscopic display apparatus 100 is generated.

A relation between the parallax component images and

the parallax-synthesized image on the screen in an omnidirectional stereoscopic display apparatus is shown in Fig. 3. Images used to display the stereoscopic image are denoted by 301, positions at which the images are acquired  
5 are denoted by 303, and segments between the center of the parallax images and exit apertures at the positions are denoted by 302.

A relation between the parallax component images and the parallax-synthesized image on the screen in a one-  
10 dimensional IP stereoscopic display apparatus is shown in Fig. 4. The images used to display the stereoscopic image are denoted by 401, the positions at which the images are acquired are denoted by 403, and the segments between the center of the parallax images and exit apertures at the  
15 positions are denoted by 402.

The one-dimensional IP stereoscopic display apparatus acquires the images using a plurality of cameras disposed at a predetermined visual range from the screen, or performs rendering in computer graphics, where the number  
20 of the cameras is equal to or more than the number of the parallaxes of the stereoscopic display apparatus, and extracts beams required for the stereoscopic display apparatus from the rendered images.

The number of the beams extracted from each of the  
25 parallax component images depends on an assumed visual range in addition to the size and the resolution of the screen of the stereoscopic display apparatus. A component pixel width determined by the assumed visual range, which is slightly larger than nine pixel width, can be calculated  
30 using a method disclosed in JP-A 2004-295013 (KOKAI) or JP-A 2005-86414 (KOKAI).

As shown in Figs. 5 and 6, if the visual range changes, the parallax image seen from an observation point also

changes. The parallax images seen from the observation points are denoted by 501 and 601.

Each of the parallax component images is generally  
perspectively projected at the assumed visual range or an  
5 equivalent thereof in the vertical direction and also  
parallelly projected in the horizontal direction. However,  
it can be perspectively projected in both the vertical  
direction and the horizontal direction. In other words, to  
generate an image in the stereoscopic display apparatus  
10 based on an integral imaging method, the imaging process or  
the rendering process can be performed by a necessary  
number of the cameras as long as the image can be converted  
into information of the beams to be reproduced.

The following explanation of the stereoscopic display  
15 apparatus 100 according to the first embodiment is given  
assuming that the number and the positions of the cameras  
that acquire the beams enough and necessary to display the  
stereoscopic image has been calculated.

Details of the real-object position/posture detecting  
20 unit 103 are explained below. The explanation is given  
based on the process of generating the stereoscopic image  
linked to a transparent cup used as the real object. In  
this case, actions of virtual penguins stereoscopically  
displayed on a flat-laid stereoscopic display panel are  
25 controlled by covering them with the real transparent cup.  
The virtual penguins move autonomously on the flat-laid  
stereoscopic display panel while shooting tomato bombs.  
The user covers the penguins with the transparent cup so  
that the tomato bombs hit the transparent cup and will not  
30 fall on the screen.

As shown in Fig. 8, the real-object position/posture  
detecting unit 103 includes infrared emitting units L and R,  
recursive sheets (not shown), and area image sensors L and

R. The infrared emitting units L and R are provided at the upper-left and the upper-right of a screen 703. The recursive sheets are provided on the left and the right sides of the screen 703 and under the screen 703,

5 reflecting infrared lights. The area image sensors L and R are provided at the same positions of the infrared emitting units L and R at the upper-left and the upper-right of the screen 703, and they receive the infrared lights reflected by the recursive sheets.

10 As shown in Fig. 7, to detect the position of a transparent cup 705 on the screen 703 of a stereoscopic display panel 702, each of areas 802 and 803 where the infrared light emitted from the infrared emitting unit L or R is masked by the transparent cup 705 so as not to be  
15 reflected by the recursive sheet and to reach none of the area image sensors L and R is measured. A reference numeral 701 in Fig. 7 denotes a point of sight.

In this manner, the center position of the transparent cup 705 is calculated. The real-object position/posture  
20 detecting unit 103 can detect only a real object within a certain height from the screen 703. However, the height area in which the real object is detected can be increased by using results of detection by the infrared emitting units L and R, the area image sensors L and R, and the  
25 recursive sheets arranged in layers above the screen 703. Otherwise, by applying a frosting marker 801 on the surface of the transparent cup 705 at the same height as the infrared emitting units L and R, the area image sensors L and R, and the recursive sheets as shown in Fig. 8, the  
30 accuracy of the detection by the area image sensors L and R is increased while taking advantage of the transparency of the cup.

A stereoscopic-image generating process performed by

the stereoscopic display apparatus 100 is explained referring to Fig. 9.

The real-object position/posture detecting unit 103 detects the position and the posture of the real object in the manner described above (step S1). At the same time, the real-object-shape specifying unit 101 receives the shape of the real object as specified by a user (step S2).

For example, if the real object is the transparent cup 705, the user specifies the three-dimensional shape of the transparent cup 705, which is a hemisphere, and the real-object-shape specifying unit 101 receives the specified three-dimensional shape. By matching the three-dimensional scale of the screen 703, the transparent cup 705, and the virtual object in a virtual scene with the actual size of the screen 703, the position and the posture of the real transparent cup and those of the cup displayed as the virtual object match.

The masked-area calculating unit 104 calculates the masked-area. More specifically, the masked-area calculating unit 104 detects a two-dimensional masked-area (step S3). In other words, the two-dimensional masked-area masked by the real object when seen from the point of sight 701 of a camera is detected by rendering only the real object received by the real-object-shape specifying unit 101.

An area of the real object in a rendered image is the two-dimensional masked-area seen from the point of sight 701. Because the pixels in the masked-area correspond to the light emitted from the stereoscopic display panel 702, the detection of the two-dimensional masked-area is to distinguish the information of the beams masked by the real object from the information of those not masked among the beams emitted from the screen 703.

The masked-area calculating unit 104 calculates the masked-area in the depth direction (step S4). The masked-area in the depth direction is calculated as described below.

5        A Z-buffer corresponding to a distance from the point of sight 701 to a plane closer to the camera is considered to be the distance between the camera and the real object. The Z-buffer is stored in a buffer with the same size as a frame buffer as real-object front-depth information  
10    Zobj\_front.

Whether the real object is in front of or at the back of the camera is determined by calculating an inner product of a vector from the point of sight to a focused polygon and a polygon normal. If the inner product is positive,  
15    the polygon faces forward, and if the inner product is negative, the polygon faces backward. Similarly, a Z-buffer corresponding to a distance from the point of sight 701 to a plane in the back of the point of sight is considered to be the distance between the point of sight  
20    and the real object. The Z-buffer at the time of the rendering is stored in the memory as real-object back-depth information Zobj\_back.

The masked-area calculating unit 104 renders only objects included in a scene. A pixel value after the  
25    rendering is herein referred to as Cscene. The Z-buffer corresponding to the distance from the visual point is stored in the memory as virtual-object depth information Zscene. The masked-area calculating unit 104 renders a rectangular area that corresponds to the screen 703, and  
30    stores the result of the rendering in the memory as display depth information Zdisp. The closest Z value among Zobj\_back, Zdisp, and Zscene is considered as an edge of the masked-area Zfar. A vector Zv indicative of an area in



the depth direction finally masked by the real object and the screen 703 is calculated by

$$Z_v = Z_{\text{obj\_front}} - Z_{\text{far}} \quad (1)$$

The area in the depth direction is calculated with respect to each pixel in the two-dimensional masked-area from the point of sight.

The 3D-image rendering unit 105 determines whether the pixel is included in the masked-area (step S5). If it is included in the masked-area (YES at step S5), the 3D-image rendering unit 105 renders the pixel in the masked-area as a volume data by performing a volumetric rendering (step S6). The volumetric rendering is performed by calculating a final pixel value  $C_{\text{final}}$  to be determined taking into account the effect on the masked-area using Equation (2).

$$C_{\text{final}} = C_{\text{scene}} * \alpha * (C_v * Z_v) \quad (2)$$

The symbol "\*" indicates multiplication.  $C_v$  is color information including vectors of R, G, and B used to express the volume of the masked-area, and  $\alpha$  is a parameter, i.e., a scalar, used to normalize the Z-buffer and adjust the volume data.

If the pixel is not included in the masked-area (NO at step S5), the volumetric rendering is not performed. As a result, different rendering processes are performed on the masked-areas and other areas.

The 3D-image rendering unit 105 determines whether the process at the steps S3 to S6 has been performed on all of points of sight of the camera (step S7). If the process has not been performed on all the points of sight (NO at step S7), the stereoscopic display apparatus 100 repeats the steps S3 to S7 on the next point of sight.

If the process has been performed on all of the points of sight (YES at step S7), the 3D-image rendering unit 105 generates the stereoscopic image by converting the

rendering result into the parallax-synthesized image (step S8).

By performing the above-described process, for example, if the real object is the transparent cup 705 disposed on the screen, the internal of the cup is converted into a volume image that includes certain colors, whereby the presence of the cup and the state inside the cup are more easily recognized. When a volume effect is applied to the transparent cup, it is applied to the area masked by the transparent cup, as indicated by 1001 shown in Fig. 10.

If it is an only purpose to apply visual reality to the three-dimensional area of the transparent cup, detection of the masked-area in the depth direction does not have to be performed with respect to each pixel in the two-dimensional masked-area of the image from each point of sight. Instead, the stereoscopic display apparatus 100 can be configured to render the masked-area with the volume effect by accumulating the colors that express the volume effect after rendering the scenes that include virtual objects.

Although the 3D-image rendering unit 105 renders the area masked by the real object as the volume data to apply the volume effect in the first embodiment, the 3D-image rendering unit 105 can be configured to render the area around the real object as the volume data.

To do so, the 3D-image rendering unit 105 enlarges the shape of the real object received by the real-object-shape specifying unit 101 in three dimensions, and the enlarged shape is used as the shape of the real object. By rendering the enlarged area as the volume data, the 3D-image rendering unit 105 applies the volume effect to the periphery of the real object.

For example, to render the periphery of the

transparent cup 705 as the volume data, as shown in Fig. 11, the shape of the transparent cup is enlarged in three dimensions, and a peripheral area 1101 enlarged from the transparent cup is rendered as the volume data.

5       The 3D-image rendering unit 105 can be configured to use a cylindrical real object and render an internal concave of the real object as the volume data. In this case, the real-object-shape specifying unit 101 receives the specification of the shape as a cylinder with a closed  
10 top and closed bottom, the top being lower than the full height of the cylinder. The 3D-image rendering unit 105 renders the internal concave of the cylinder as the volume data.

To render the internal concave of the cylindrical real  
15 object as the volume data, for example, as shown in Fig. 12, the fullness of water is visualized by rendering an internal concave 1201 as the volume data. Moreover, by rendering virtual goldfish autonomously swimming in the concave internal of the cylinder as shown in Fig. 13, the  
20 user recognizes by sight that the goldfish are present in a cylindrical aquarium that contains water.

As described above, the stereoscopic display apparatus  
100 based on the integral imaging method according to the first embodiment specifies a spatial area to be focused on  
25 using the real object, and efficiently creates the visual reality independent from the point of sight of the user. Therefore, a stereoscopic image that changes depending on the position, the posture, and the shape of the real object is generated without using a tracking system that tracks  
30 actions of the user, and efficiently generates a voluminous stereoscopic image with reduced amount of process.

A stereoscopic display apparatus 1400 according to a second embodiment of the present invention further receives

an attribute of the real object and performs the rendering process on the masked-area based on the received attribute.

As shown in Fig. 14, the stereoscopic display apparatus 1400 includes the real-object-shape specifying unit 101, the real-object position/posture detecting unit 103, the masked-area calculating unit 104, a 3D-image rendering unit 1405, and a real-object-attribute specifying unit 1406. Moreover, the stereoscopic display apparatus 1400 includes hardware such as the stereoscopic display panel, the memory, and the CPU.

The functions and the configurations of the real-object-shape specifying unit 101, the real-object position/posture detecting unit 103, and the masked-area calculating unit 104 are same as those in the stereoscopic display apparatus 100 according to the first embodiment.

The real-object-attribute specifying unit 1406 receives at least one of thickness, transmittance, and color of the real object as the attribute.

The 3D-image rendering unit 1405 generates the parallax-synthesized image by applying surface effect to the masked-area based on the shape received by the real-object-shape specifying unit 101 the attribute of the real object received by the real-object-attribute specifying unit 1406.

A stereoscopic-image generating process performed by the stereoscopic display apparatus 1400 is explained referring to Fig. 15. Steps S11 to S14 are same as the steps S1 to S4 shown in Fig. 9.

According to the second embodiment, the real-object-attribute specifying unit 1406 receives the thickness, the transmittance, and/or the color of the real object specified by the user as the attribute (step S16). The 3D-image rendering unit 1405 determines whether the pixel is

included in the masked-area (step S15). If it is included in the masked-area (YES at step S15), the 3D-image rendering unit 1405 performs a rendering process that applies the surface effect to the pixel in the masked-area by referring to the attribute and the shape of the real object (step S17).

The information of the pixels masked by the real object from each point of sight is detected in the detection of the two-dimensional masked-area at the step S13. One-to-one correspondence between each pixel and the information of the beam is uniquely determined by the relation between the position of the camera and the screen. Positional relation among the point of sight 701 that looks at the flat-laid stereoscopic display panel 702 from 60 degrees upward, the screen 703, and a real object 1505 that masks the screen is shown in Fig. 16.

The rendering process on the surface effect applies an effect on an interaction with the real object with respect to each beam that corresponds to each pixel detected at the step S13. More specifically, a pixel value of the image from the point of sight finally determined taking into account the surface effect of the real object Cresult is calculated by

$$C_{result} = C_{scene} * C_{obj} * \beta * (d_{obj} * (2.0 - N_{obj} \cdot V_{cam})) \quad (3)$$

The symbol "\*" indicates the multiplication, and the symbol "." indicates the inner product. Cscene is the pixel value of the rendering result excluding the real object; Cobj is the color of the real object received by the real-object-attribute specifying unit 1406 (vectors of R, G, and B); dobj is the thickness of the real object received by the real-object-attribute specifying unit 1406; Nobj is a normalized normal vector on the surface of the real object; Vcam is a normalized normal vector directed

from the point of sight 701 of the camera to the surface of the real object; and  $\beta$  is a coefficient that determines a degree of the visual reality.

Because Vcam is equivalent to a beam vector, it can  
5 apply the visual reality taking into account the attribute of the surface of the real object, such as the thickness, to the light entering obliquely to the surface of the real object. As a result, it is more emphasized that the real object is transparent and has the thickness.

10 To render roughness of the surface of the real object, the real-object-attribute specifying unit 1406 specifies map information such as a bump map or a normal map as the attribute of the real object, and the 3D-image rendering unit 1405 efficiently controls the normalized normal vector  
15 on the surface of the real object at the time of the rendering process.

The information on the point of sight of the camera is determined by only the stereoscopic display panel 702 independently of the state of the user, and therefore the  
20 surface effect of the real object dependent on the point of sight is rendered as the stereoscopic image regardless of the point of sight of the user.

For example, the 3D-image rendering unit 1405 creates a highlight to apply the surface effect to the real object.  
25 The highlight on the surface of a metal or transparent object changes depending on the point of sight. The highlight can be realized in units of the beam by calculating Cresult based on Nobj and Vcam.

The 3D-image rendering unit 1405 defocuses the shape  
30 of the highlight by superposing the stereoscopic image on the highlight present on the real object to show the real object as if it is made of a different material. The 3D-image rendering unit 1405 visualizes a virtual light source

and an environment by superposing a highlight that is not actually present on the real object as the stereoscopic image.

Moreover, the 3D-image rendering unit 1405 synthesizes  
5 a virtual crack that is not actually present on the real object as the stereoscopic image. For example, if a real glass with a certain thickness cracks, the crack looks differently depending on the point of sight. The color information generated by the effect of the crack Ceffect is  
10 calculated using Equation (4) to apply the visual reality of the crack to the masked-area.

$$\text{Ceffect} = \gamma * \text{Ccrack} * |\text{Vcam} \times \text{Vcrack}| \quad (4)$$

The symbol "\*" indicates the multiplication, and the symbol "x" indicates exterior product. By synthesizing  
15 Ceffect with the pixel on the image from the point of sight, the final pixel information that includes the crack is generated. Ccrack is a color value used for the visual reality of the crack; Vcam is the normalized normal vector directed from the point of sight of the camera to the  
20 surface of the real object; Vcrack is a normalized crack-direction vector indicative of the direction of the crack; and  $\gamma$  is a parameter used to adjust the degree of the visual reality.

Furthermore, to show an image of the tomato bomb hit  
25 and crashed against the real transparent cup, the visual reality is reproduced on the stereoscopic display panel by using a texture mapping method, which uses the crashed tomato bomb as a texture.

The texture mapping method is explained below. The  
30 3D-image rendering unit 1405 performs mapping by switching texture images based on a bidirectional texture function (BTF) that indicates a texture element on the surface of

the polygon depending on the point of sight and the light source.

The BTF uses a spherical coordinate system with its origin at the image subject on the surface of the model shown in Fig. 17 to specify the positions of the point of sight and the light source. Fig. 17 is a schematic diagram of the spherical coordinate system used to perform the texture mapping that depends on positions of the point of sight and the light source.

Assuming that the point of sight is infinitely far and the light from the light source is parallel, the coordinate of the point of sight is  $(\theta_e, \phi_e)$  and the coordinate of the light source is  $(\theta_i, \phi_i)$ , where  $\theta_e$  and  $\theta_i$  indicate longitudinal angles, and  $\phi_e$  and  $\phi_i$  indicate latitudinal angles. In this case, a texture address is defined in six dimensions. For example, a texel is indicated using six variables as described below

$$T(\theta_e, \phi_e, \theta_i, \phi_i, u, v) \quad (5)$$

Each of  $u$  and  $v$  indicates an address in the texture. In fact, a plurality of texture images acquired at a specific point of sight and a specific light source is accumulated, and the texture is expressed by switching the textures and combining the addresses in the texture. Mapping of the texture in this manner is referred to as a high-dimensional texture mapping.

The 3D-image rendering unit 1405 performs the texture mapping as described below. The 3D-image rendering unit 1405 specifies model shape data and divides the model shape data into rendering primitives. In other words, the 3D-image rendering unit 1405 divides the model shape data into units of the image processing, which is generally performed in units of polygons consisting three points. The polygon



is planar information surrounded by the three points, and the 3D-image rendering unit 1405 performs the rendering process on the internal of the polygon.

The 3D-image rendering unit 1405 calculates a texture-projected coordinate of a rendering primitive. In other words, the 3D-image rendering unit 1405 calculates a vector U and a vector V on the projected coordinate when a u-axis and a v-axis in a two-dimensional coordinate system that define the texture are projected onto a plane defined by the three points indicated by a three-dimensional coordinate in the rendering primitive. The 3D-image rendering unit 1405 calculates the normal to the plane defined by the three points. A method for calculating the vector U and the vector V will be explained later referring to Fig. 18.

The 3D-image rendering unit 1405 specifies the vector U, the vector V, the normal, the position of the point of sight, and the position of the light source, and calculates the directions of the point of sight and the light source (direction parameters) to acquire relative directions of the point of sight and the light source to the rendering primitive.

More specifically, the latitudinal relative direction  $\phi$  is calculated from a normal vector N and a direction vector D by

$$\phi = \arccos(D \cdot N / (|D| * |N|)) \quad (6)$$

D·N is an inner product of the vector D and the vector N; and the symbol "\*" indicates the multiplication. A method for calculating a longitudinal relative direction  $\theta$  will be explained later referring to Figs. 19A and 19B.

The 3D-image rendering unit 1405 generates a rendering texture based on the relative directions of the point of

sight and the light source. The rendering texture to be pasted on the rendering primitive is prepared in advance. The 3D-image rendering unit 1405 acquires texel information from the texture in the memory based on the relative  
5 directions of the point of sight and the light source. Acquiring the texel information means assigning the texture element acquired under a specific condition to a texture coordinate space that corresponds to the rendering primitive. The acquisition of the relative direction and  
10 the texture element can be performed with respect to each point of sight or each light source, and they are acquired in the same manner if there is a plurality of point of sights and light sources.

The 3D-image rendering unit 1405 performs the process  
15 on all of the rendering primitives. After all of the primitives are processed, the 3D-image rendering unit 1405 maps each of the rendered textures to a corresponding point on the model.

The method for calculating the vector U and the vector  
20 V is explained referring to Fig. 18.

The three-dimensional coordinates and the texture coordinates of the three points that define the rendering primitive are described as follows.

Point P0: three-dimensional coordinate (x0, y0, z0),  
25 texture coordinate (u0, v0)

Point P1: three-dimensional coordinate (x1, y1, z1),  
texture coordinate (u1, v1)

Point P2: three-dimensional coordinate (x2, y2, z2),  
texture coordinate (u2, v2)

30 By defining the coordinates as described above, the vector  $U=(u_x, u_y, u_z)$  and the vector  $V=(v_x, v_y, v_z)$  in the projected coordinate system are calculated by

$$P2-P0=(u1-u0)*U+(v1-v0)*V$$

$$P1-P0=(u2-u0)*U+(v2-v0)*V$$

Based on the three-dimensional coordinates of P0, P1, and P2, the vector U and the vector V are acquired by solving  $u_x$ ,  $u_y$ ,  $u_z$ ,  $v_x$ ,  $v_y$ , and  $v_z$  from Equations (7)-(12)

$$5 \quad u_x = \text{idet} * (v_{20} * x_{10} - v_{10} * x_{20}) \quad (7)$$

$$u_y = \text{idet} * (v_{20} * y_{10} - v_{10} * y_{20}) \quad (8)$$

$$u_z = \text{idet} * (v_{20} * z_{10} - v_{10} * z_{20}) \quad (9)$$

$$v_x = \text{idet} * (-u_{20} * x_{10} + u_{10} * x_{20}) \quad (10)$$

$$v_y = \text{idet} * (-u_{20} * y_{10} + u_{10} * y_{20}) \quad (11)$$

$$10 \quad v_z = \text{idet} * (-u_{20} * z_{10} + u_{10} * z_{20}) \quad (12)$$

The equations are based on the following conditions:

$$v_{10} = v_1 - v_0,$$

$$v_{20} = v_2 - v_0,$$

$$x_{10} = x_1 - x_0,$$

$$15 \quad x_{20} = x_2 - x_0,$$

$$y_{10} = y_1 - y_0,$$

$$y_{20} = y_2 - y_0,$$

$$z_{10} = z_1 - z_0,$$

$$z_{20} = z_2 - z_0,$$

$$20 \quad \text{det} = u_{10} * v_{20} - u_{20} * v_{10}, \text{ and}$$

$$\text{idet} = 1/\text{det}$$

The normal is calculated simply as an exterior product of two independent vectors on a plane defined by the three points.

25 The method for calculating a longitudinal relative direction  $\theta$  is explained referring to Figs. 19A and 19B. A vector B of the direction vector indicative of the point of sight or the light source projected on the model plane is acquired. A direction vector of the point of sight or the light source  $D=(dx, dy, dz)$ , a normal vector of the model plane  $N=(nx, ny, nz)$ , and the vector of the direction vector D projected on the model plane  $B=(bx, by, bz)$  are calculated by:

30

$$B = D - (D \cdot N) * N \quad (13)$$

The equation (13) is represented by elements as shown below.

$$b_x = d_x - \alpha n_x,$$

$$5 \quad b_y = d_y - \alpha n_y,$$

$$b_z = d_z - \alpha n_z, \text{ and}$$

$\alpha$  is equal to  $d_x * n_x + d_y * n_y + d_z * n_z$ , and the normal vector  $N$  is a unit vector.

The relative directions of the point of sight and the light source are acquired from the vector  $B$ , the vector  $U$ ,  
10 and the vector  $V$  as described below.

An angle between the vector  $U$  and the vector  $V$   $\lambda$  and an angle between the vector  $U$  and the vector  $B$   $\theta$  are calculated by

$$15 \quad \lambda = \arccos(U \cdot V / (|U| * |V|)) \quad (14)$$

$$\theta = \arccos(U \cdot B / (|U| * |B|)) \quad (15)$$

If there is no distortion in the projected coordinate system,  $U$  and  $V$  are orthogonal, i.e.,  $\lambda$  is  $\pi/2$  (90 degrees). If there is a distortion,  $\lambda$  is not  $\pi/2$ . However, if there  
20 is the distortion in the projected coordinate system, a correction is required because the texture is acquired using the directions of the point of sight and the light source relative to the orthogonal coordinate system. The angles of the relative directions of the point of sight and  
25 the light source need to be properly corrected according to the projected UV coordinate system. The corrected relative direction  $\theta'$  is calculated using one of the following Equations (16)-(19):

$$30 \quad \begin{aligned} &\text{Where } \theta \text{ is smaller than } \pi \text{ and } \theta \text{ is smaller than } \lambda; \\ &\theta' = (\theta / \lambda) * \pi / 2. \end{aligned} \quad (16)$$

Where  $\theta$  is smaller than  $\pi$  and  $\theta$  is larger than  $\lambda$ ;

$$\theta' = \pi - ((\pi - \theta) / (\pi - \lambda)) * \pi / 2. \quad (17)$$

Where  $\theta$  is larger than  $\pi$  and  $\theta$  is smaller than  $\pi + \lambda$ ;

$$\theta' = (\theta - \pi) / \lambda * \pi / 2 + \pi. \quad (18)$$

Where  $\theta$  is larger than  $\pi$  and  $\theta$  is larger than  $\pi + \lambda$ ;

$$\theta' = 2\pi - ((2\pi - \theta) / (\pi - \lambda)) * \pi / 2. \quad (19)$$

The longitudinal relative directions of the point of sight and the light source to the rendering primitive are acquired as described above.

The 3D-image rendering unit 1405 renders the texture mapping in the masked-area by performing the process described above. An example of the image of the tomato bomb crashed against the real transparent cup with the visual reality created by the process is shown in Fig. 20. The masked-area is denoted by 2001.

Moreover, the 3D-image rendering unit 1405 renders a lens effect and a zoom effect to the masked-area. For example, the real-object-attribute specifying unit 1406 specifies the refractive index, the magnification, or the color of a plate used as the real object.

The 3D-image rendering unit 1405 scales the rendered image of only the virtual object centered on the center of the masked-area detected at the step S13 in Fig. 15, and extracts the masked-area as a mask, whereby scaling the scene through the real object.

By scaling the rendered image of the virtual scene centered on a pixel on which a straight line that runs through the three-dimensional zoom center on the real object and the visual point intersects with the screen 703, a digital zoom effect that uses the real object to resemble a magnifying glass is realized.

To explain the positional relation between the flat-laid stereoscopic display panel and the plate, as shown in

Fig. 21, a virtual object of the magnifying glass can be superposed in the space that contains a real plate 2105, whereby increasing the reality of the stereoscopic image.

The 3D-image rendering unit 1405 can be configured to  
5 render the virtual object based on a ray tracing method by simulating refraction of a beam defined by the position of each pixel. This is realized by the real-object-shape specifying unit 101 specifying the accurate shape of the three-dimensional lens for the real object, such as a  
10 concave lens or a convex lens, and the real-object-attribute specifying unit 1406 specifying the refractive index as the attribute of the real object.

The 3D-image rendering unit 1405 can be configured to render the virtual object, so that a cross-section thereof  
15 is visually recognized, by arranging the real object. An example that uses a transparent plate as the real object is explained below. The positional relation among the flat-laid stereoscopic display panel 702, a plate 2205, and a cylindrical object 2206 that is the virtual object is shown  
20 in Fig. 22.

More specifically, as shown in Fig. 23, markers 2301a and 2301b for detection are applied to both ends of a plate 2305, which are frosting lines. The real-object position/posture detecting unit 103 is formed by arranging  
25 at least two each of the infrared emitting units L and R and the area image sensors L and R in layers in the height direction of the screen. In this manner, the position, the posture, and the shape of the real plate 2305 can be detected.

30 In other words, the real-object position/posture detecting unit 103 configured as above detects the positions of the markers 2301a and 2301b as explained in the first embodiment. By acquiring the positions of the

corresponding marker from the results detected by the infrared emitting units L and R and the area image sensors L and R, the real-object position/posture detecting unit 103 identifies the three-dimensional shape and the three-dimensional posture of the plate 2305, i.e., the posture and the shape of the plate 2305 are identified as indicated by a dotted line 2302 from two results 2303 and 2304. If the number of the markers is increased, the shape of the plate 2305 is calculated more accurately.

10       The masked-area calculating unit 104 is configured to determine an area of the virtual object sectioned by the real object in the computation of the masked-area in the depth direction at the step S14. In other words, the masked-area calculating unit 104 refers to the relation  
15       among depth information of the real object Zobj, front-depth information of the virtual object from the point of sight Zscene\_near, and back-depth information of the virtual object from the point of sight Zscene\_far, and determines whether Zobj is located between Zscene\_near and  
20       Zscene\_far. The Z-buffer generated by rendering is used to calculate the masked-area in the depth direction from the point of sight as explained in the first embodiment.

      The 3D-image rendering unit 1405 performs the rendering by rendering the pixels in the sectioned area as  
25       the volume data. Because the three-dimensional information of the sectional plane has been acquired by calculating the two-dimensional position seen from each point of sight, i.e., the information of the beam and the depth from the point of sight, as the information of the sectioned area,  
30       the volume data is available at this time point. The 3D-image rendering unit 1405 can be configured to set the pixels in the sectioned area brighter so that they can be easily distinguished from other pixels.

Tensor data that uses vector values instead of scalar values is used to, for example, visualize blood stream in a brain. When the tensor data is used, an anisotropic rendering method can be employed to render the vector information as the volume element of the sectional plane. For example, anisotropic reflective brightness distribution used to render hair is used as a material, and a direction-dependant rendering is performed based on the vector information, which is volume information, and point of sight information from the camera. The user senses the direction of the vector by the change of the brightness and the color in addition to the shape of the sectional plane of the volume data by moving his/her head. If the real-object-shape specifying unit 101 specifies a real object with thickness, the shape of the sectional plane is not flat but stereoscopic, and the tensor data can be visualized more efficiently.

Because the scene that includes the virtual object seen through the real object changes depending on the point of sight, the point of sight of the user needs to be tracked to realize the visual reality according to the conventional technology. However, the stereoscopic display apparatus 1400 according to the second embodiment receives the specified attribute of the real object and applies various surface effects to the masked-area based on the specified attribute, the shape, and the posture to generate the parallax-synthesized image. As a result, the stereoscopic display apparatus 1400 generates the stereoscopic image that changes depending on the position, the posture, and the shape of the real object without using a tracking system for the motion of the user, and efficiently generates the stereoscopic image with more real surface effect with reduced amount of the process.



In other words, according to the second embodiment, the area masked by the real object and the virtual scene through the real object are specified and rendered in advance with respect to each point of sight of the camera  
5 required to generate the stereoscopic image. Therefore, the stereoscopic image is generated independent of the tracked point of sight of the user, and it is accurately reproduced on the stereoscopic display panel.

A stereoscopic-image generating program executed in  
10 the stereoscopic display apparatuses according to the first embodiment and the second embodiment is preinstalled in a read only memory (ROM) or the like.

The stereoscopic-image generating program can be recorded in the form of an installable or executable file  
15 recorded in a computer-readable recording medium such as a compact disk read only memory (CD-ROM), a flexible disk (FD), a compact disk recordable (CD-R), or a digital versatile disk (DVD) to be provided.

The stereoscopic-image generating program can be  
20 stored in a computer connected to a network such as the Internet and provided by downloading it through the network. The stereoscopic-image generating program can be otherwise provided or distributed through the network.

The stereoscopic-image generating program includes  
25 each of the real-object position/posture detecting unit, the real-object-shape specifying unit, the masked-area calculating unit, and the 3D-image rendering unit, and the real-object-attribute specifying unit as a module. When the CPU reads and executes the stereoscopic-image  
30 generating program from the ROM, the units are loaded into a main memory device, and each of the units is generated in the main memory.

Additional advantages and modifications will readily

occur to those skilled in the art. Therefore, the invention in its broader aspects is not limited to the specific details and representative embodiments shown and described herein. Accordingly, various modifications may  
5 be made without departing from the spirit or scope of the general inventive concept as defined by the appended claims and their equivalents.

## CLAIMS

1. An apparatus for generating a stereoscopic image, comprising:

a detecting unit that detects at least one of a position and a posture of a real object located on or near a three-dimensional display surface;

a calculating unit that calculates a masked-area where the real object masks a ray irradiated from the three-dimensional display surface, based on at least one of the position and the posture; and

a rendering unit that renders a stereoscopic image by performing different rendering processes on the masked-area from rendering processes on other areas.

2. The apparatus according to claim 1, further comprising a first specifying unit that receives a specification of a shape of the real object, wherein

the calculating unit calculates the masked-area further based on specified shape.

3. The apparatus according to claim 2, wherein the rendering unit renders the masked-area with volume data in a three-dimensional space.

4. The apparatus according to claim 2, wherein the rendering unit renders an area around the real object in the masked-area with volume data in a three-dimensional space.

5. The apparatus according to claim 2, wherein the rendering unit renders an area of a concave portion of the real object in the masked-area with volume data in a three-dimensional space.

6. The apparatus according to claim 2, further comprising a second specifying unit that receives a specification of an attribute of the real object, wherein

5 the rendering unit performs different rendering processes on the masked-area from rendering processes the other areas, based on specified attribute.

7. The apparatus according to claim 6, wherein the  
10 attribute is at least one of thickness, transparency, and color of the real object.

8. The apparatus according to claim 6, wherein the rendering unit performs the rendering process on the  
15 masked-area based on the specified shape.

9. The apparatus according to claim 7, wherein the rendering unit performs a rendering process that applies a surface effect on the masked-area based on the specified  
20 attribute.

10. The apparatus according to claim 7, wherein the rendering unit performs a rendering process that applies a highlight effect on the masked-area based on the specified  
25 attribute.

11. The apparatus according to claim 7, wherein the rendering unit performs a rendering process that applies a crack on the masked-area based on the specified attribute.

30

12. The apparatus according to claim 7, wherein the rendering unit performs a rendering process that maps texture to the masked-area based on the specified attribute.

13. The apparatus according to claim 7, wherein the rendering unit performs a rendering process that scales the masked-area based on the specified attribute.

5

14. The apparatus according to claim 7, wherein the rendering unit performs a rendering process of displaying a cross section of the real object with respect to the masked-area based on the specified attribute.

10

15. A method of generating a stereoscopic image, comprising:

detecting at least one of a position and a posture of a real object located on or near a three-dimensional

15 display surface;

calculating a masked-area where the real object masks a ray irradiated from the three-dimensional display surface, based on at least one of the position and the posture; and

rendering a stereoscopic image by performing different rendering processes on the masked-area from rendering processes on other areas.

20

16. A computer program product comprising a computer-usable medium having computer-readable program codes embodied in the medium that when executed cause a computer to execute:

25

detecting at least one of a position and a posture of a real object located on or near a three-dimensional display surface;

30

calculating a masked-area where the real object masks a ray irradiated from the three-dimensional display surface, based on at least one of the position and the posture; and rendering a stereoscopic image by performing different

rendering processes on the masked-area from rendering processes on other areas.

FIG.1

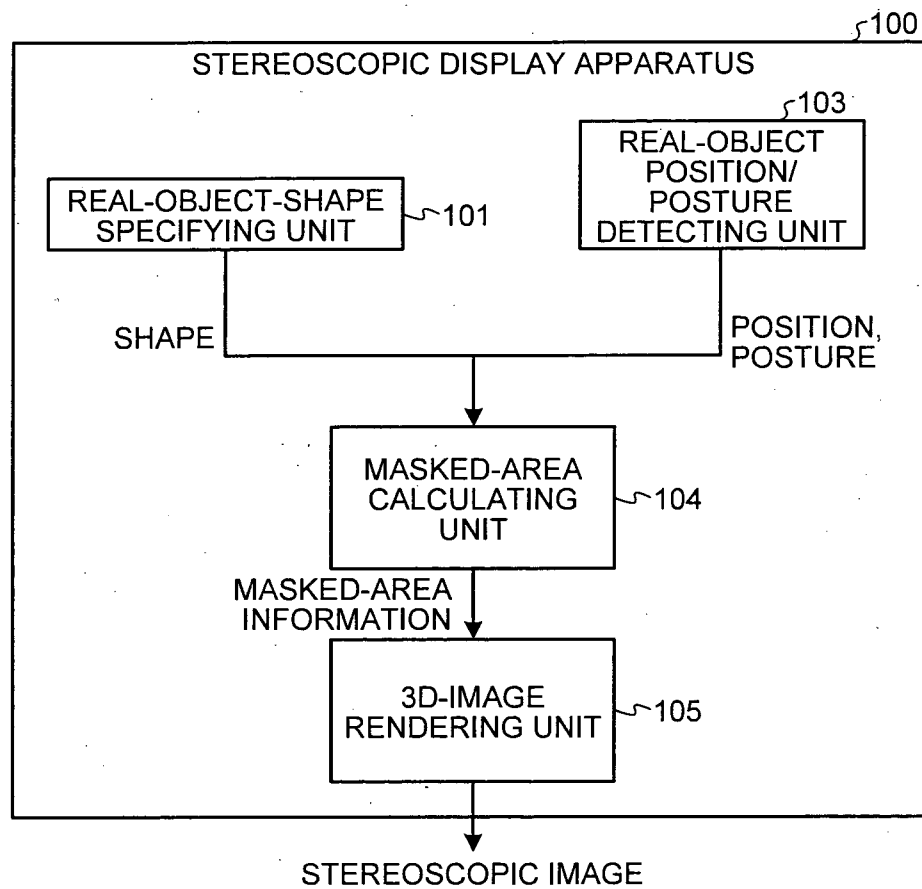
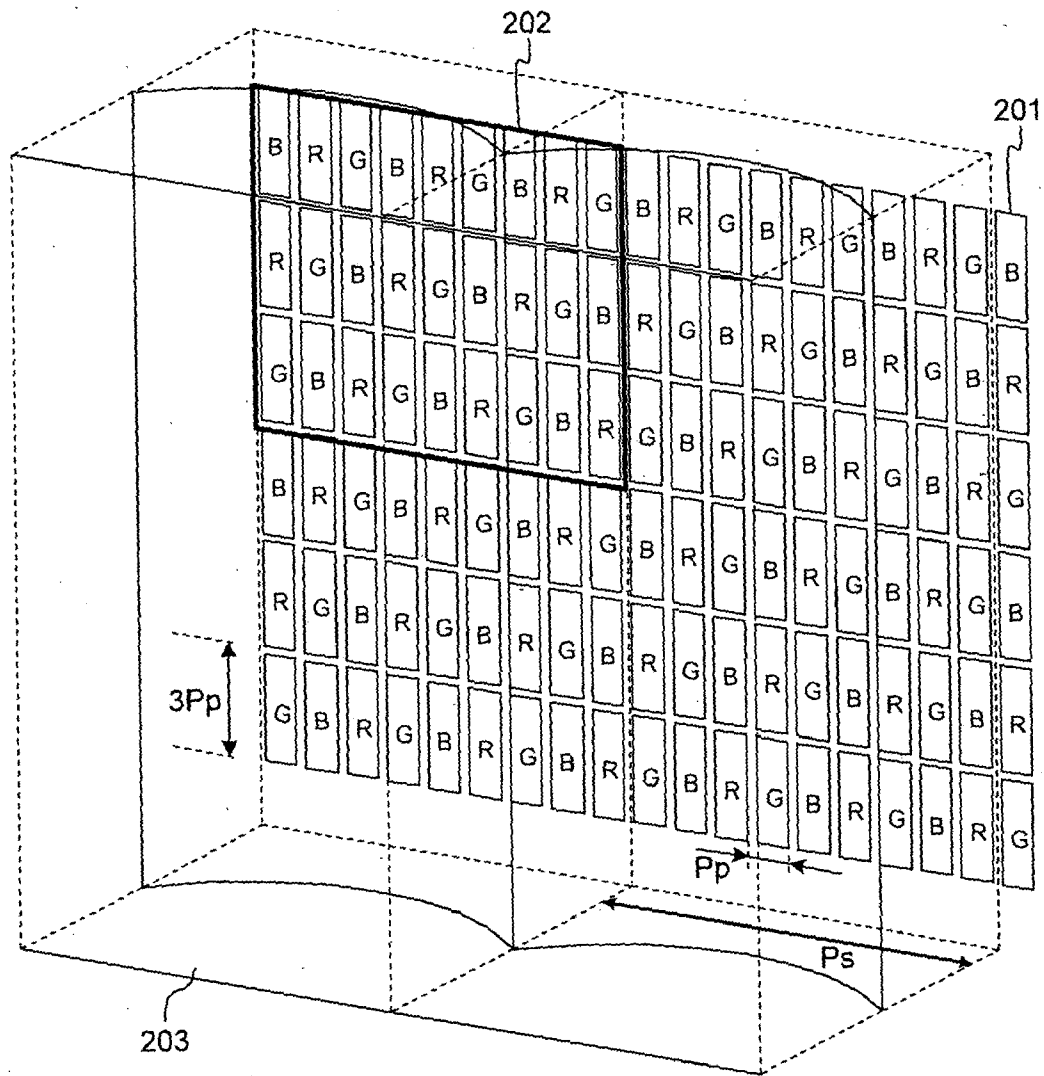


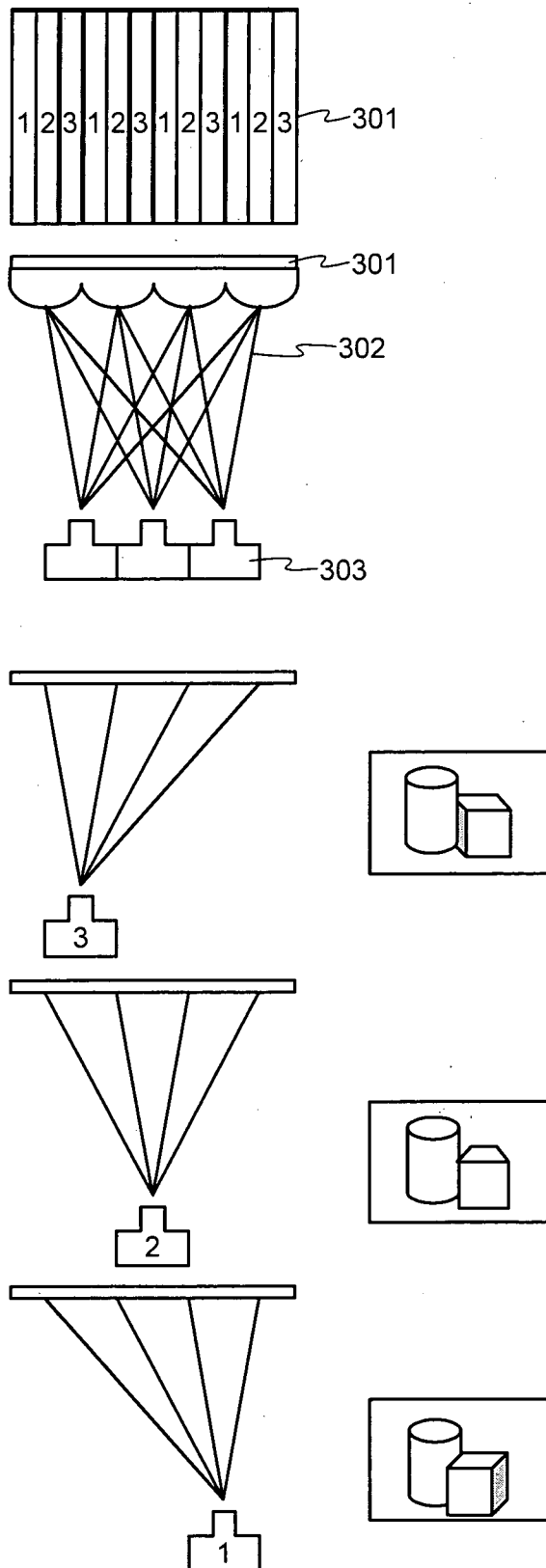
FIG.2





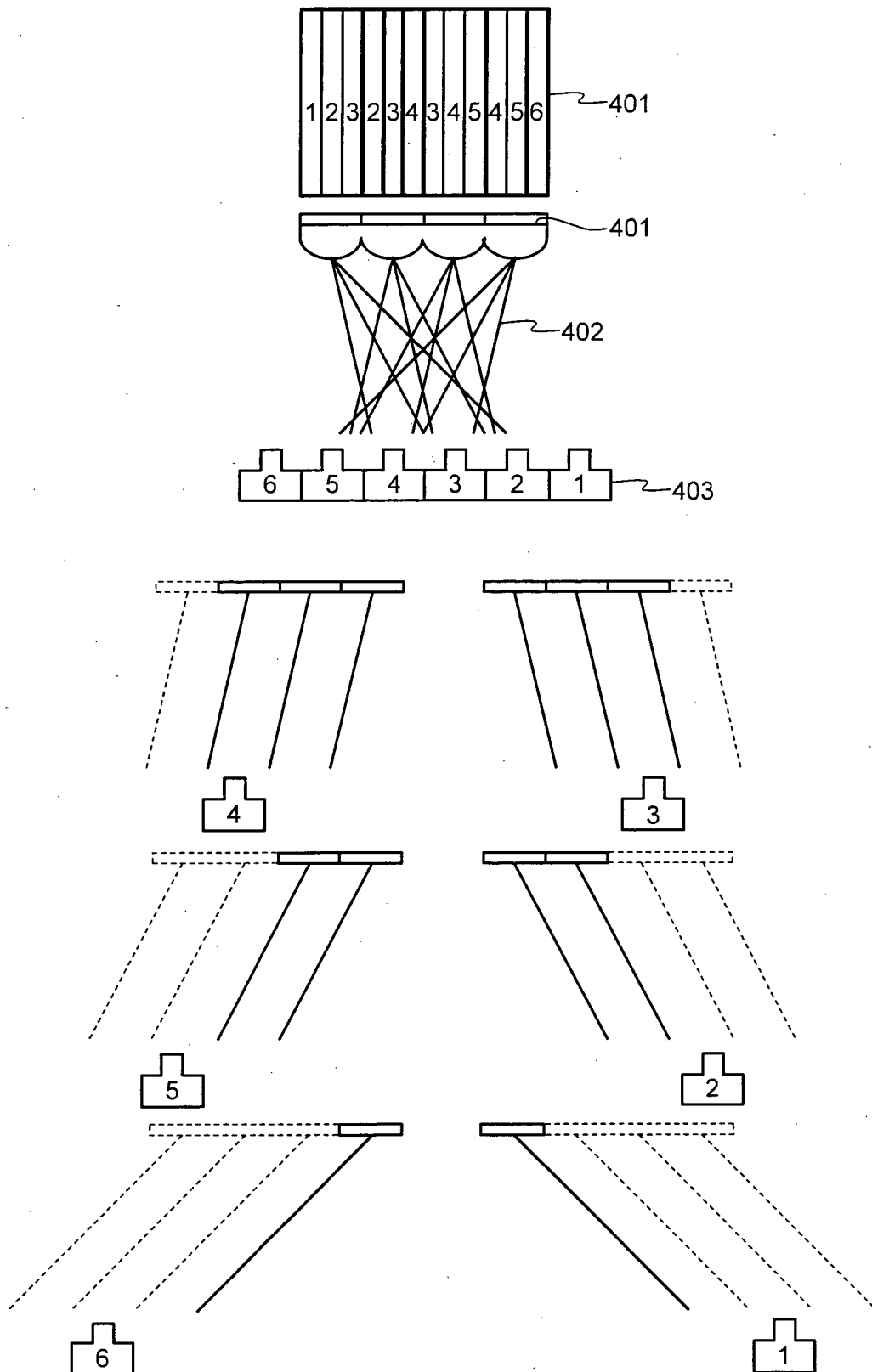
3/22

FIG.3



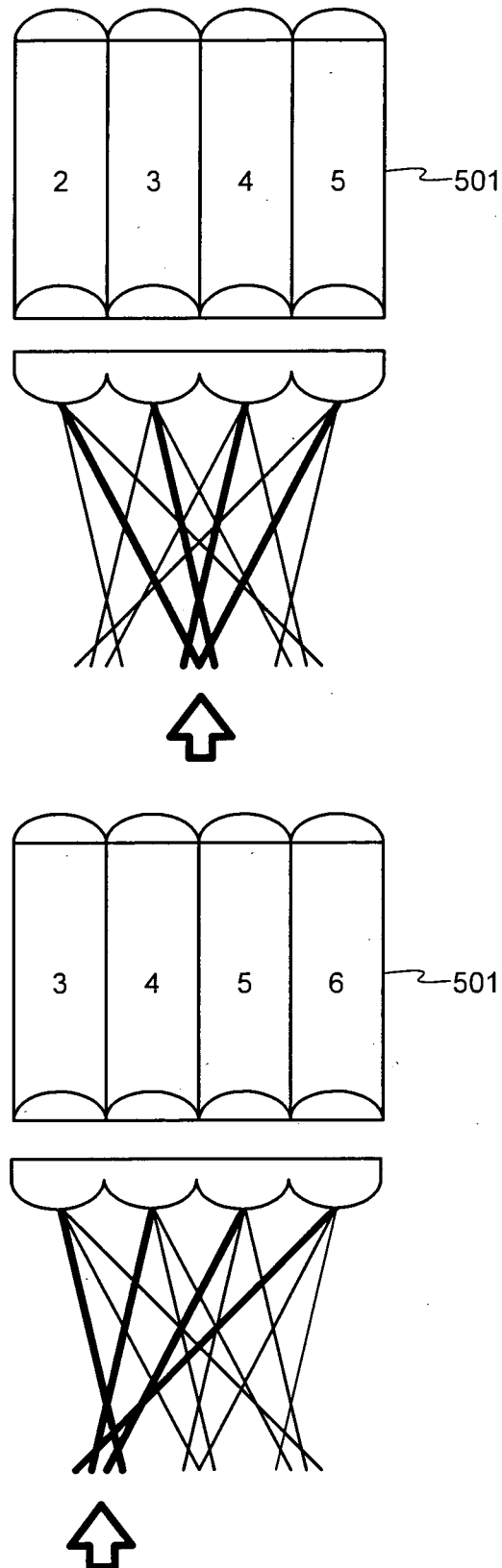
4/22

FIG.4



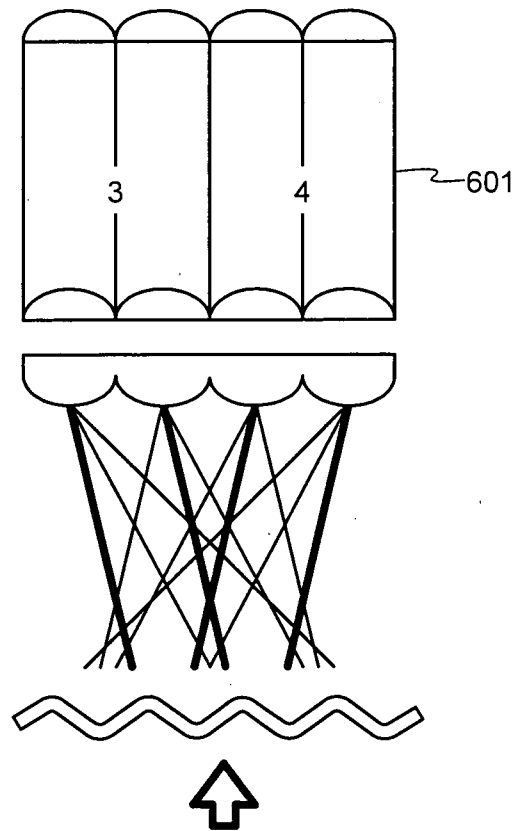
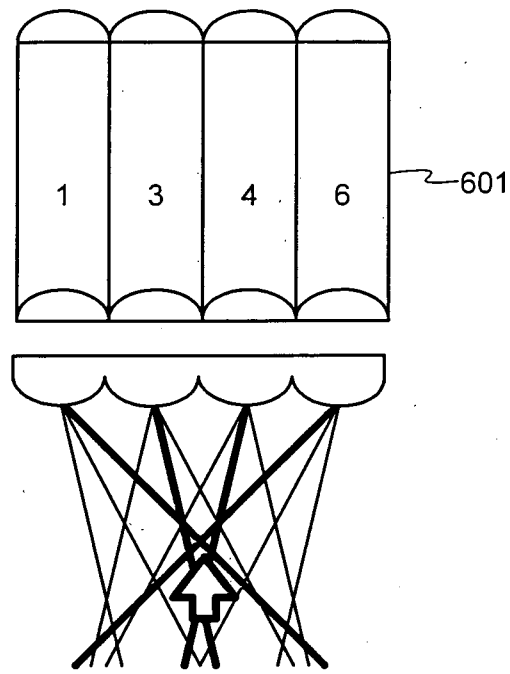
5/22

FIG.5



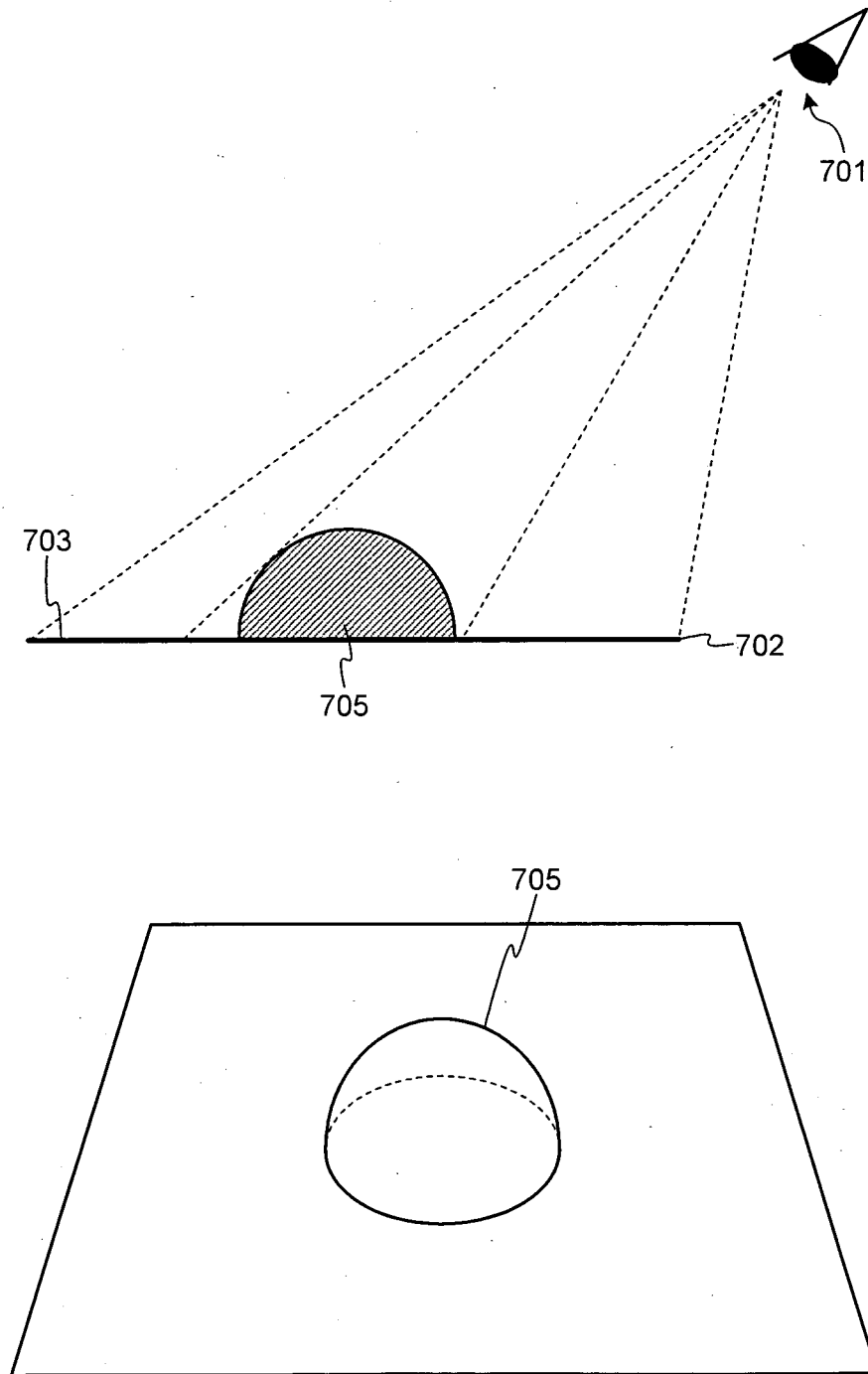
6/22

FIG.6



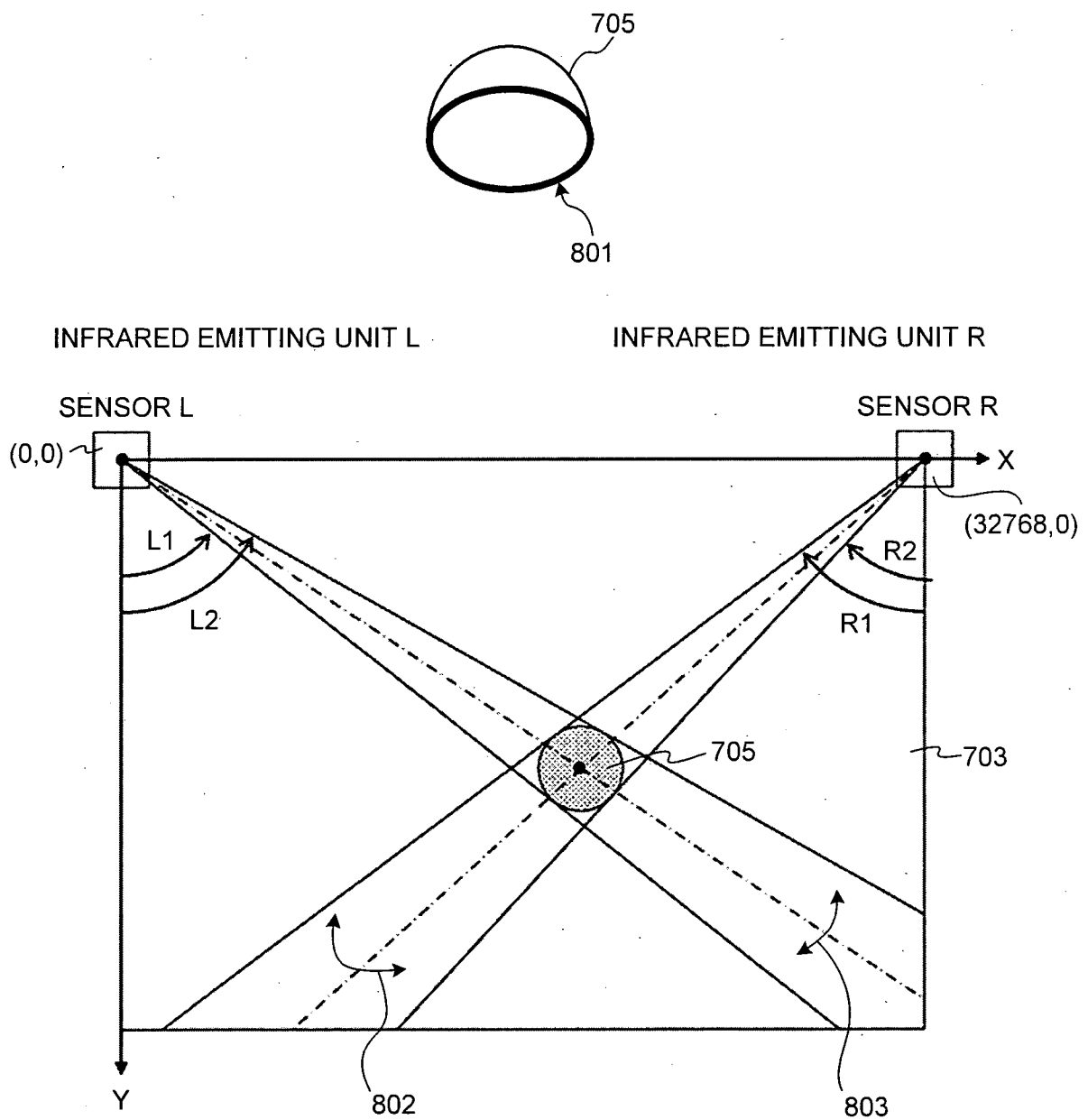
7/22

FIG.7



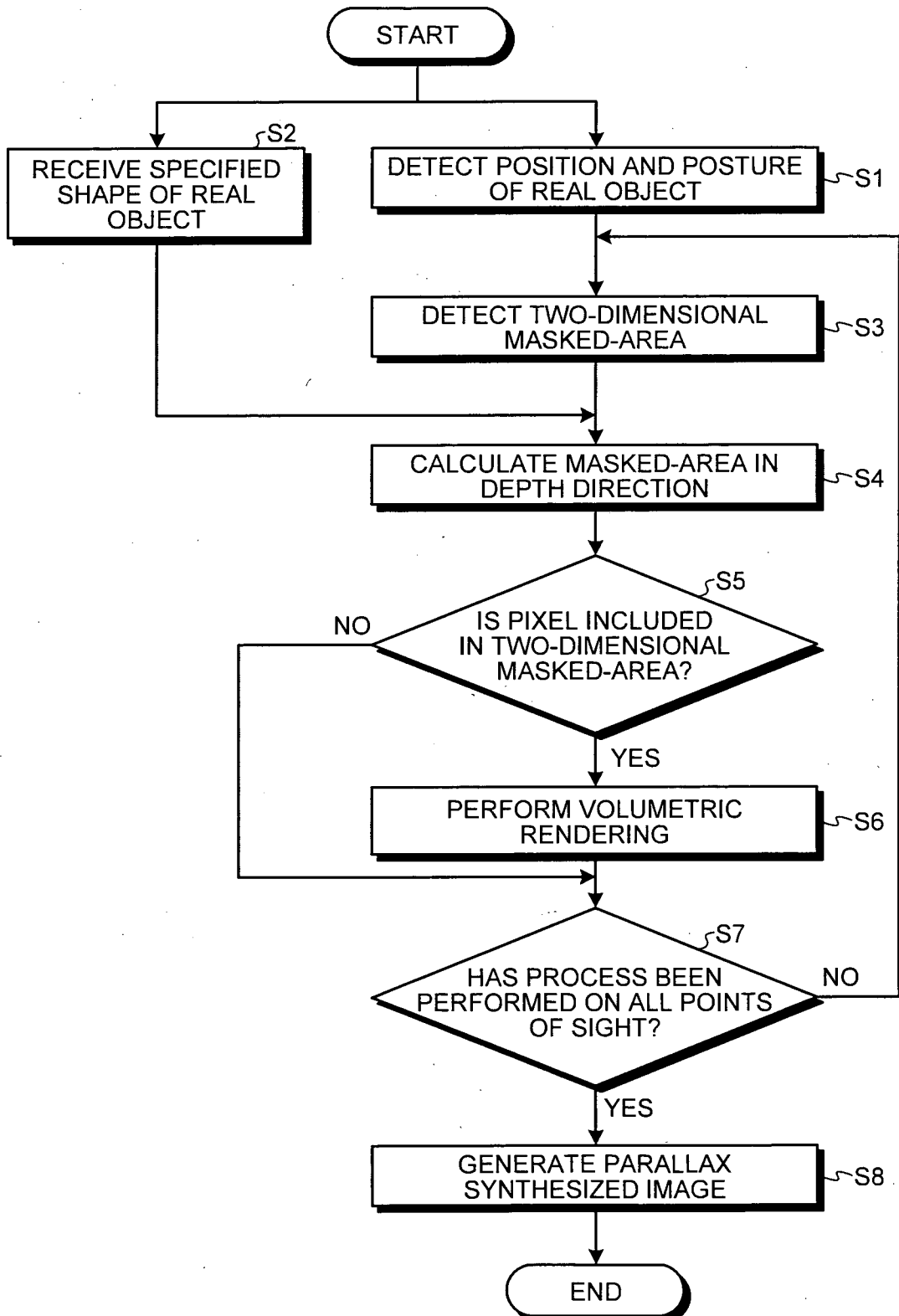
8/22

FIG.8



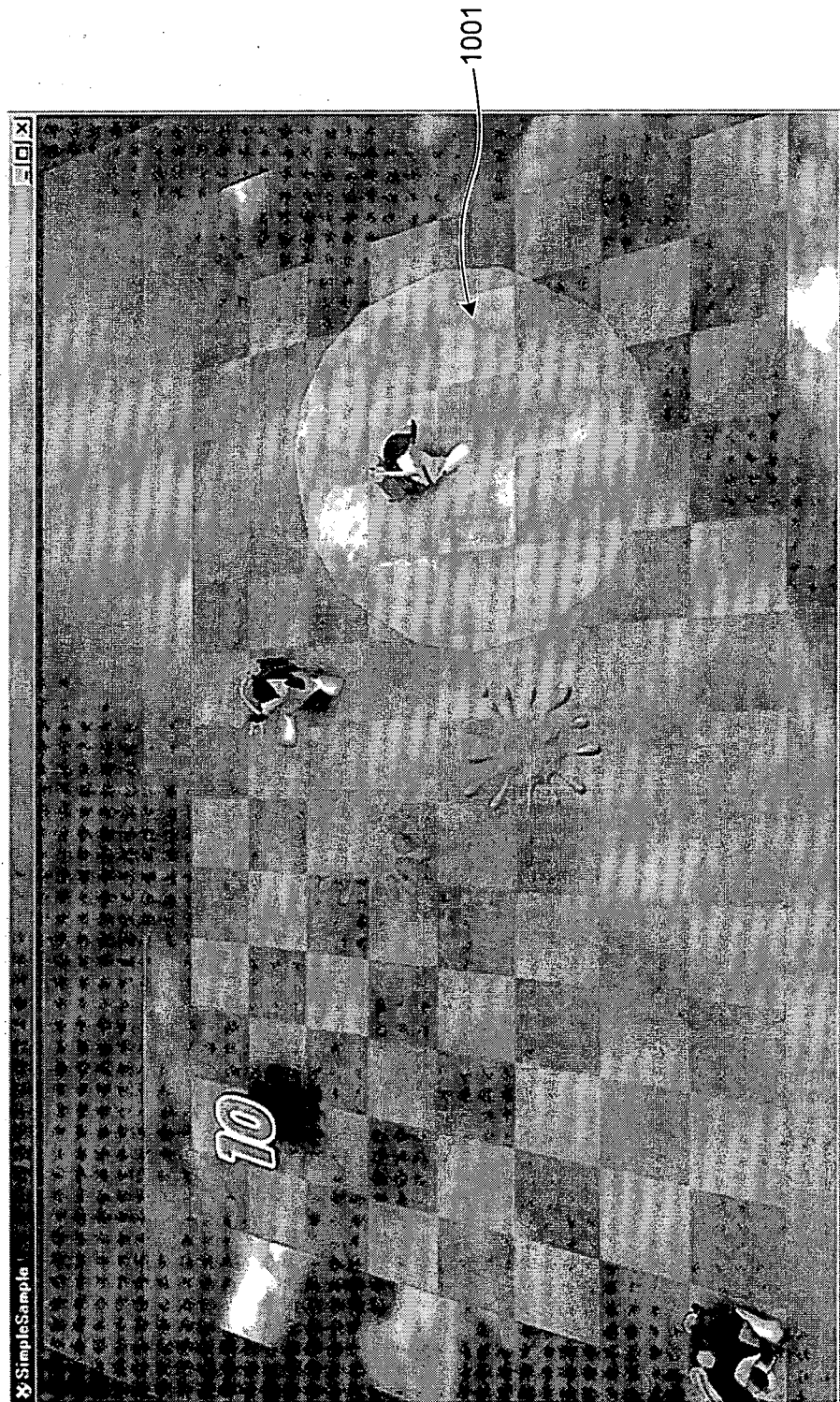
9/22

FIG.9



10/22

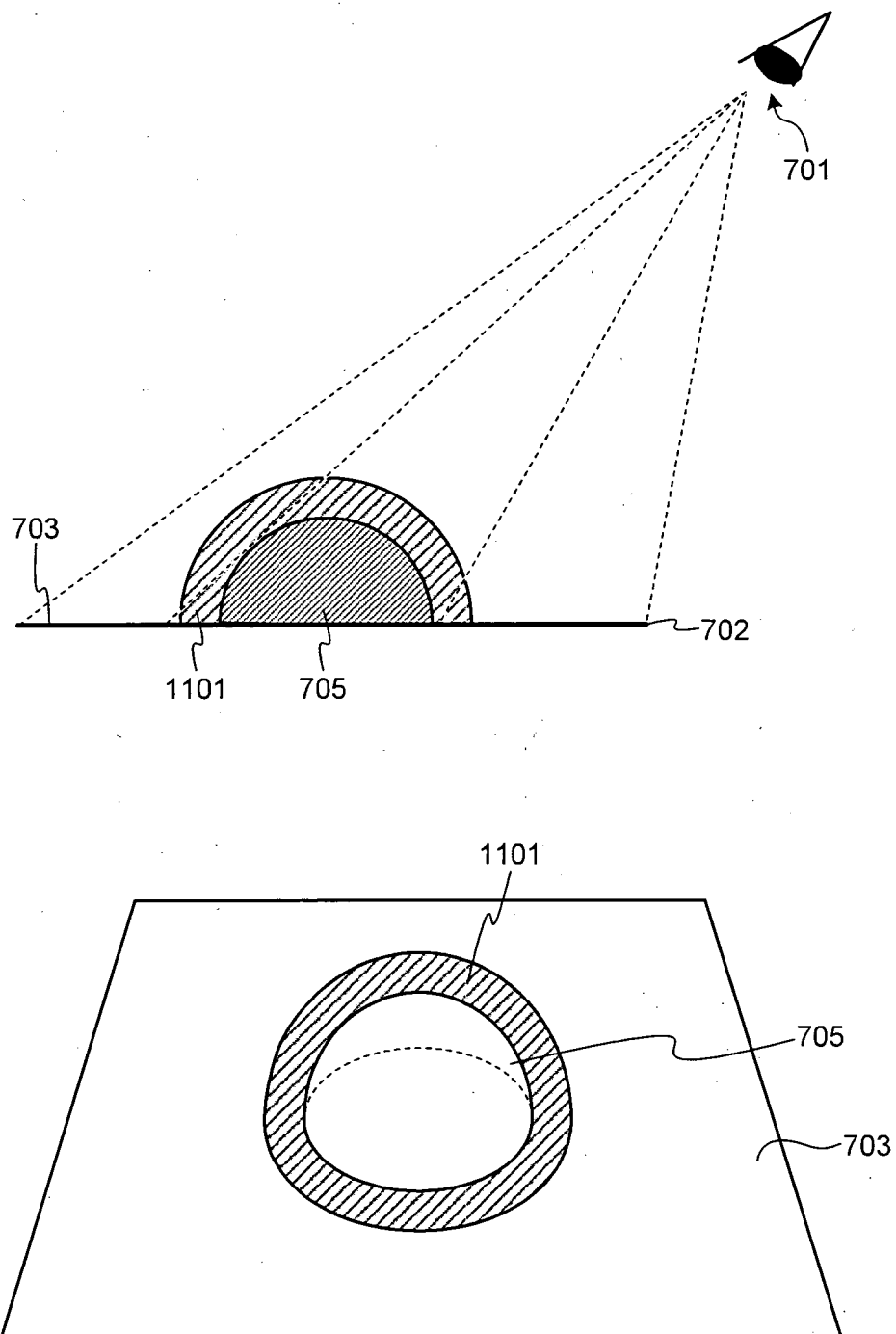
FIG.10





11/22

FIG.11



12/22

FIG.12

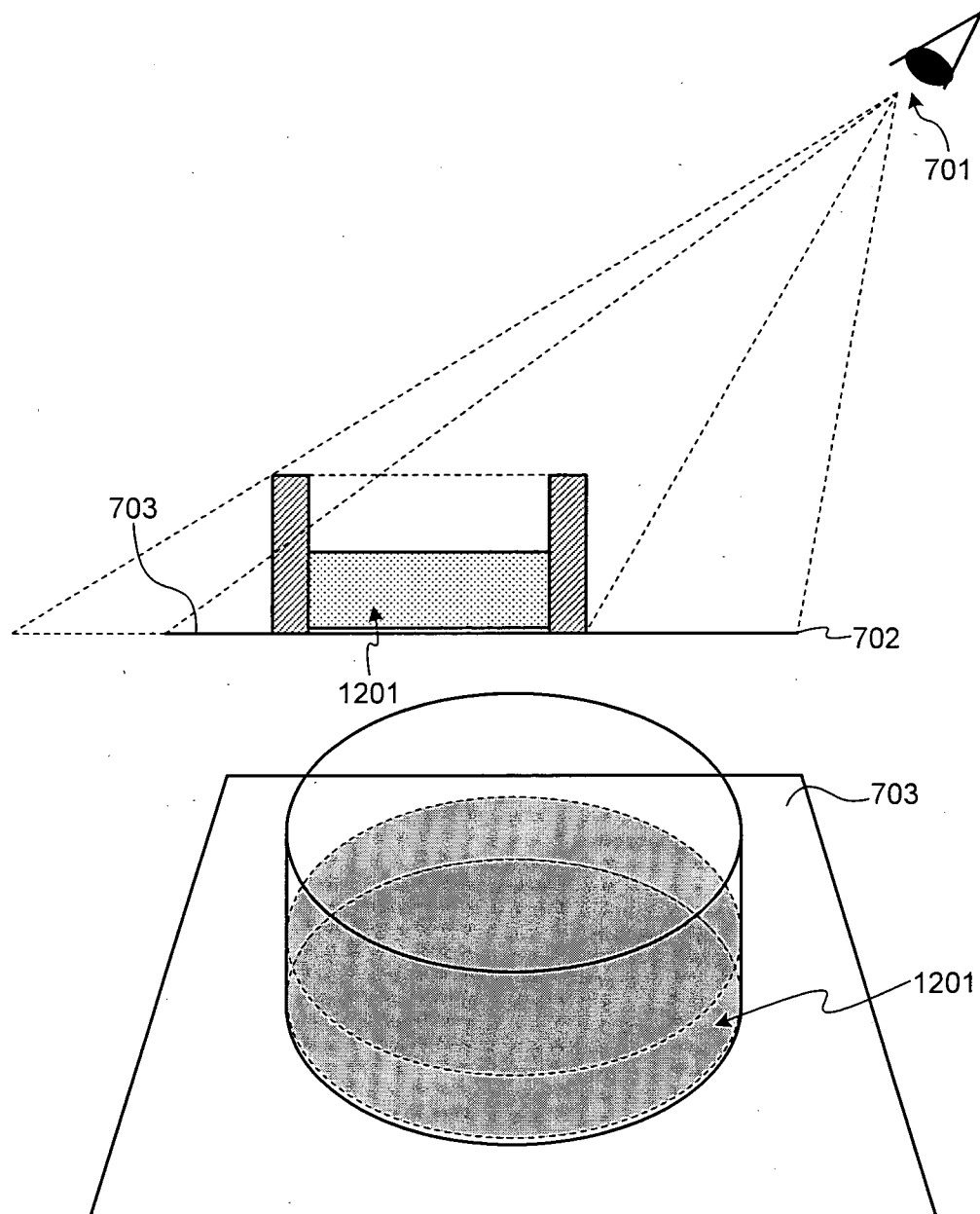


FIG.13

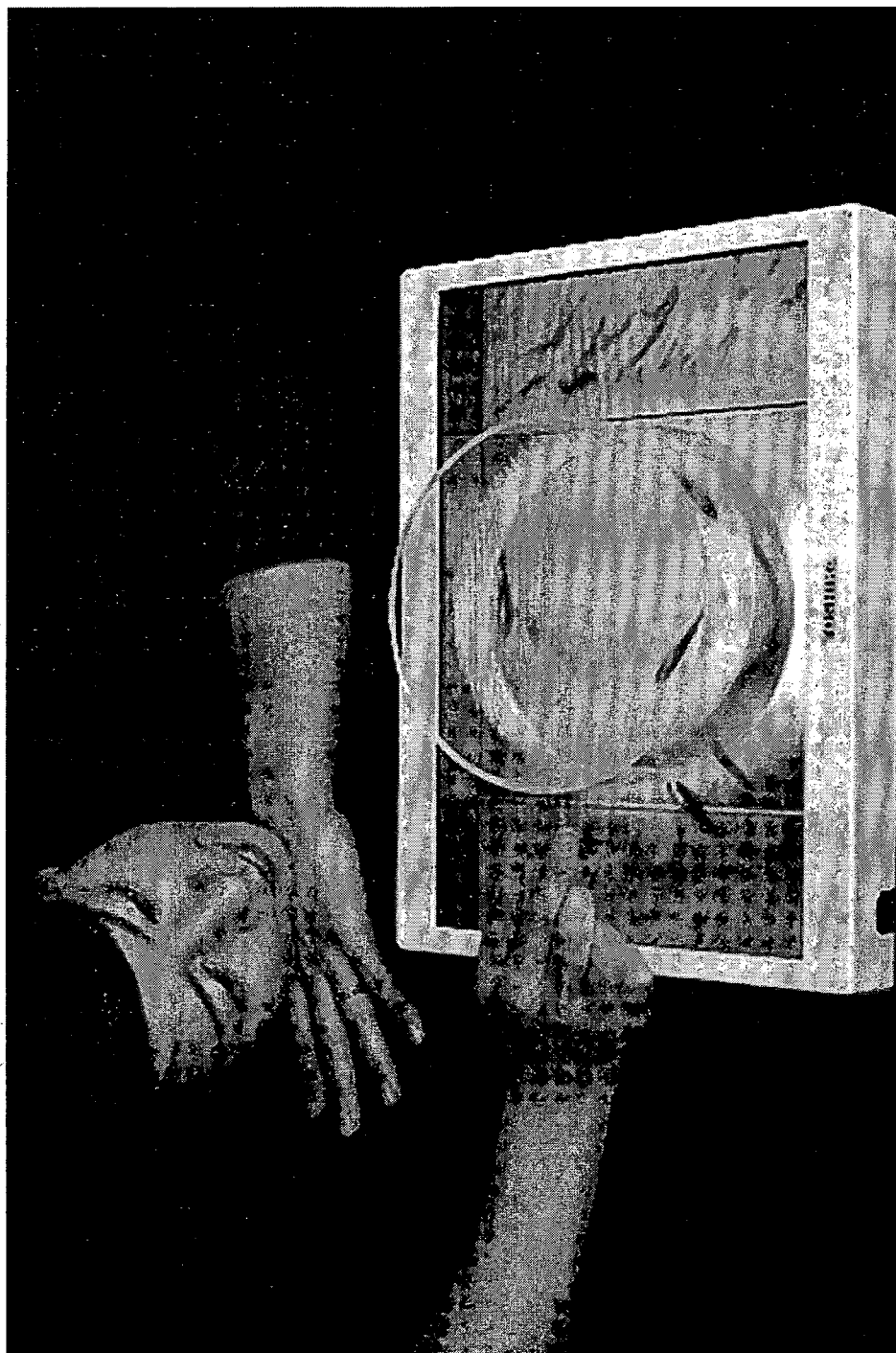
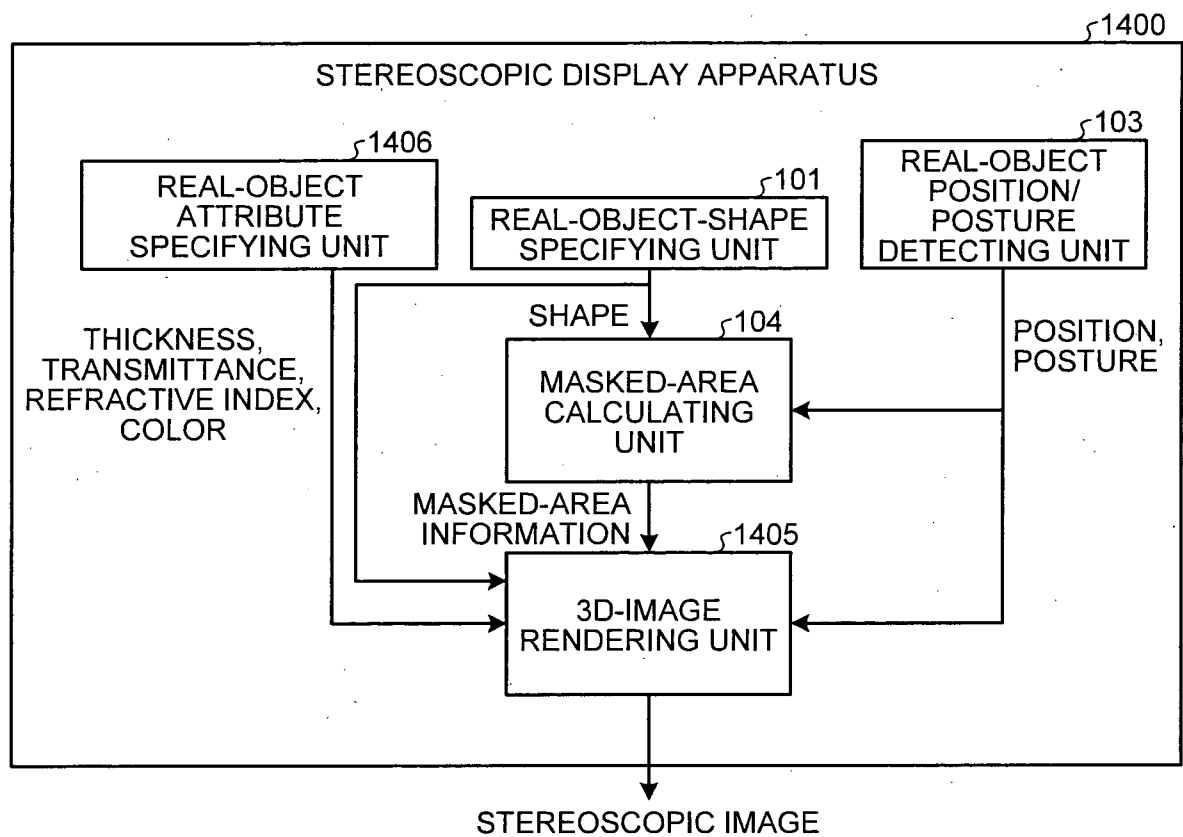
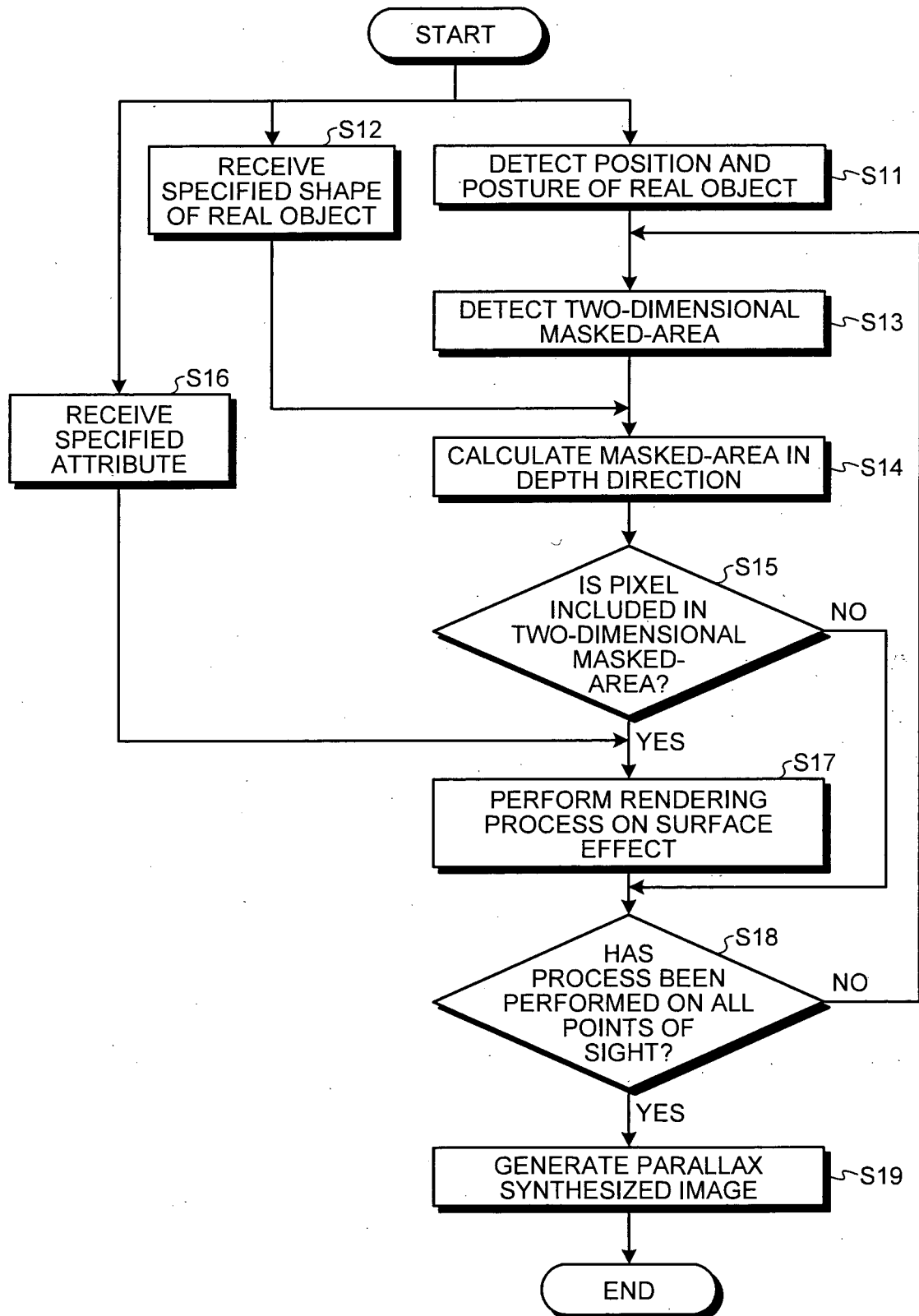


FIG.14



15/22

FIG.15



16/22

FIG.16

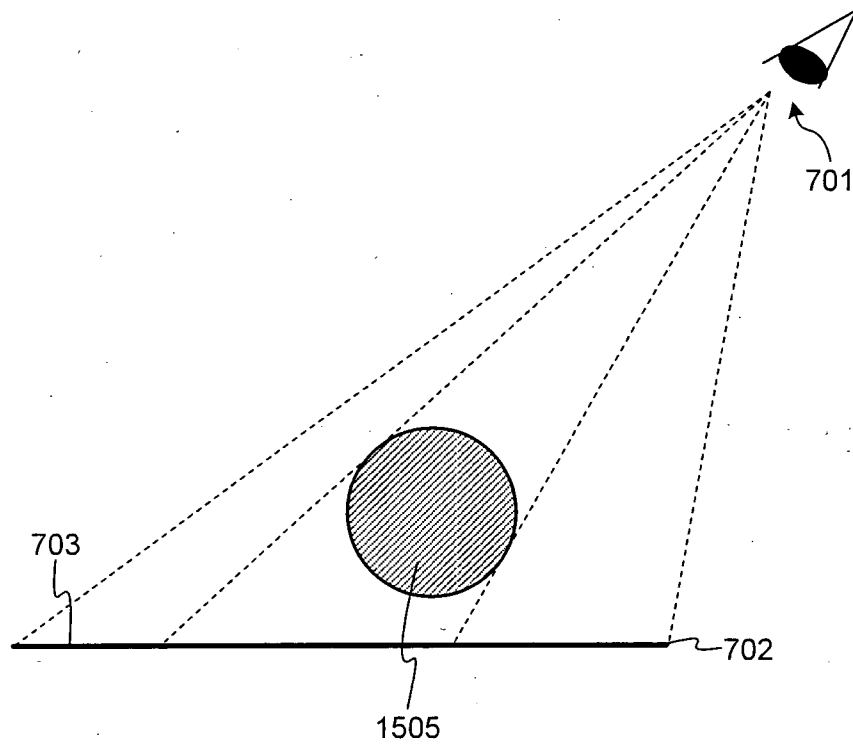
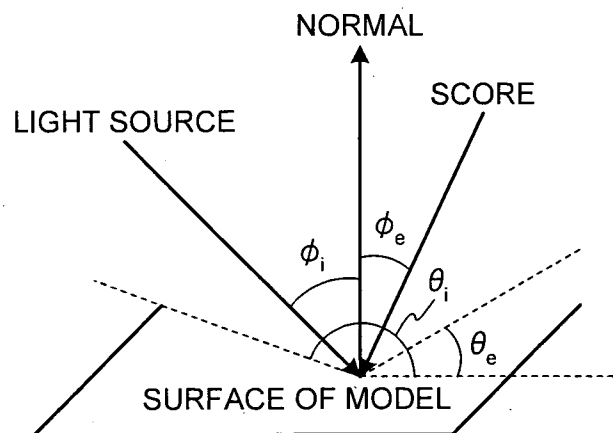


FIG.17



17/22

FIG.18

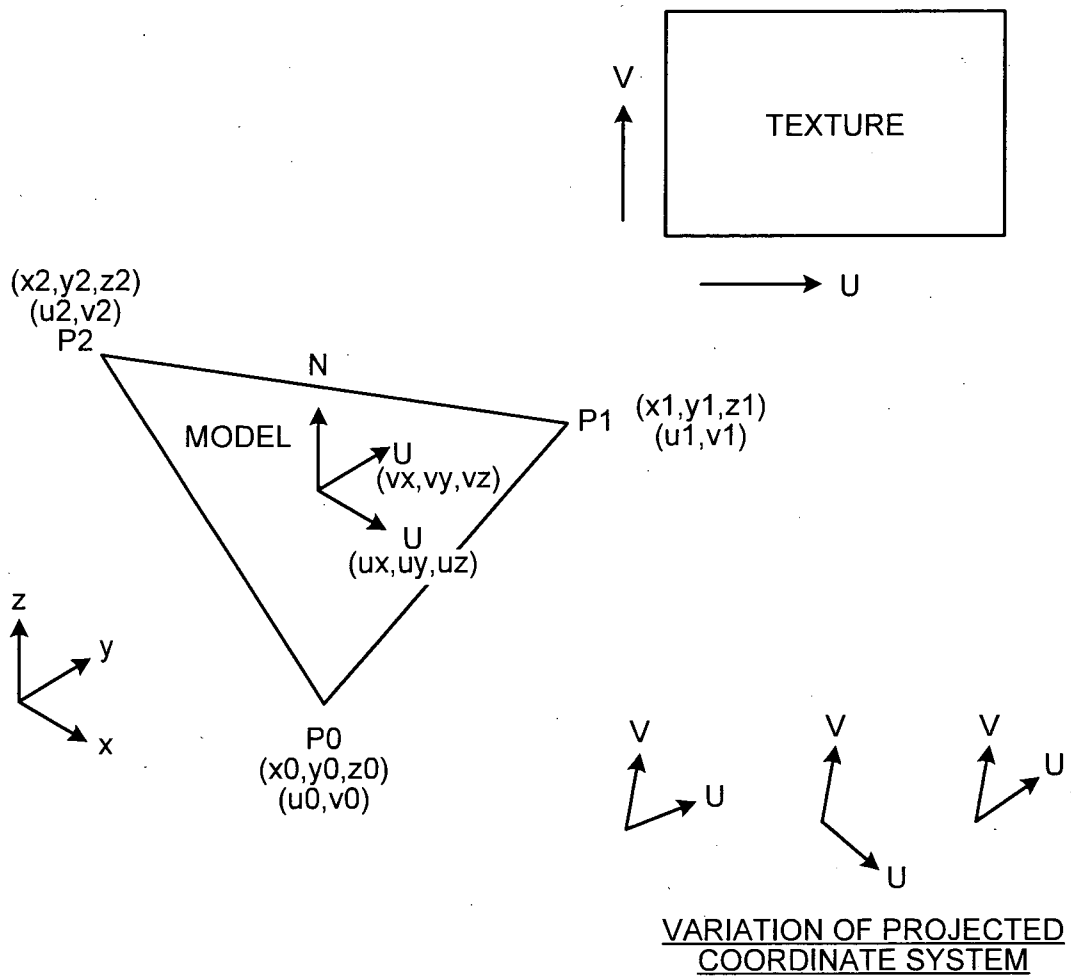


FIG.19A

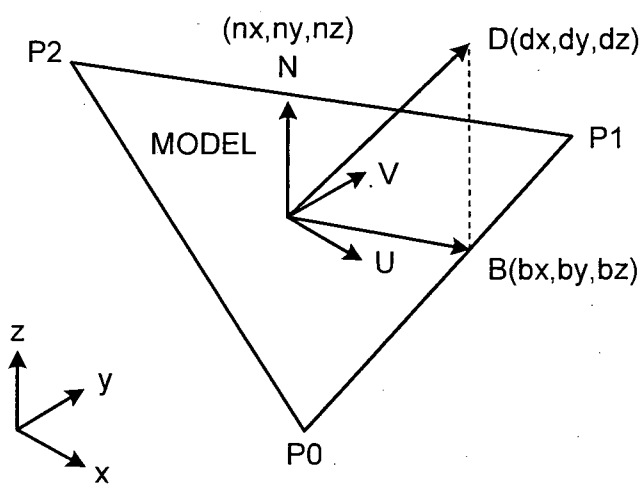


FIG.19B

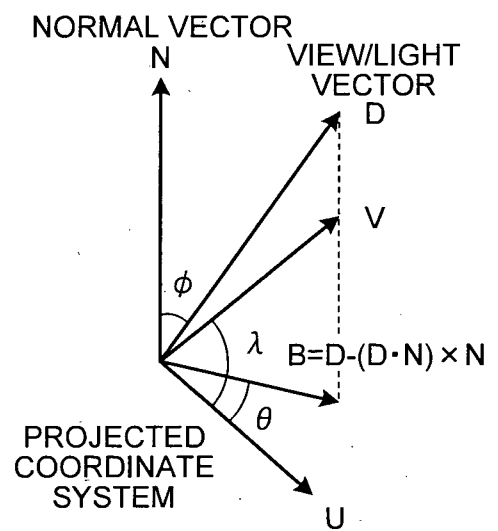
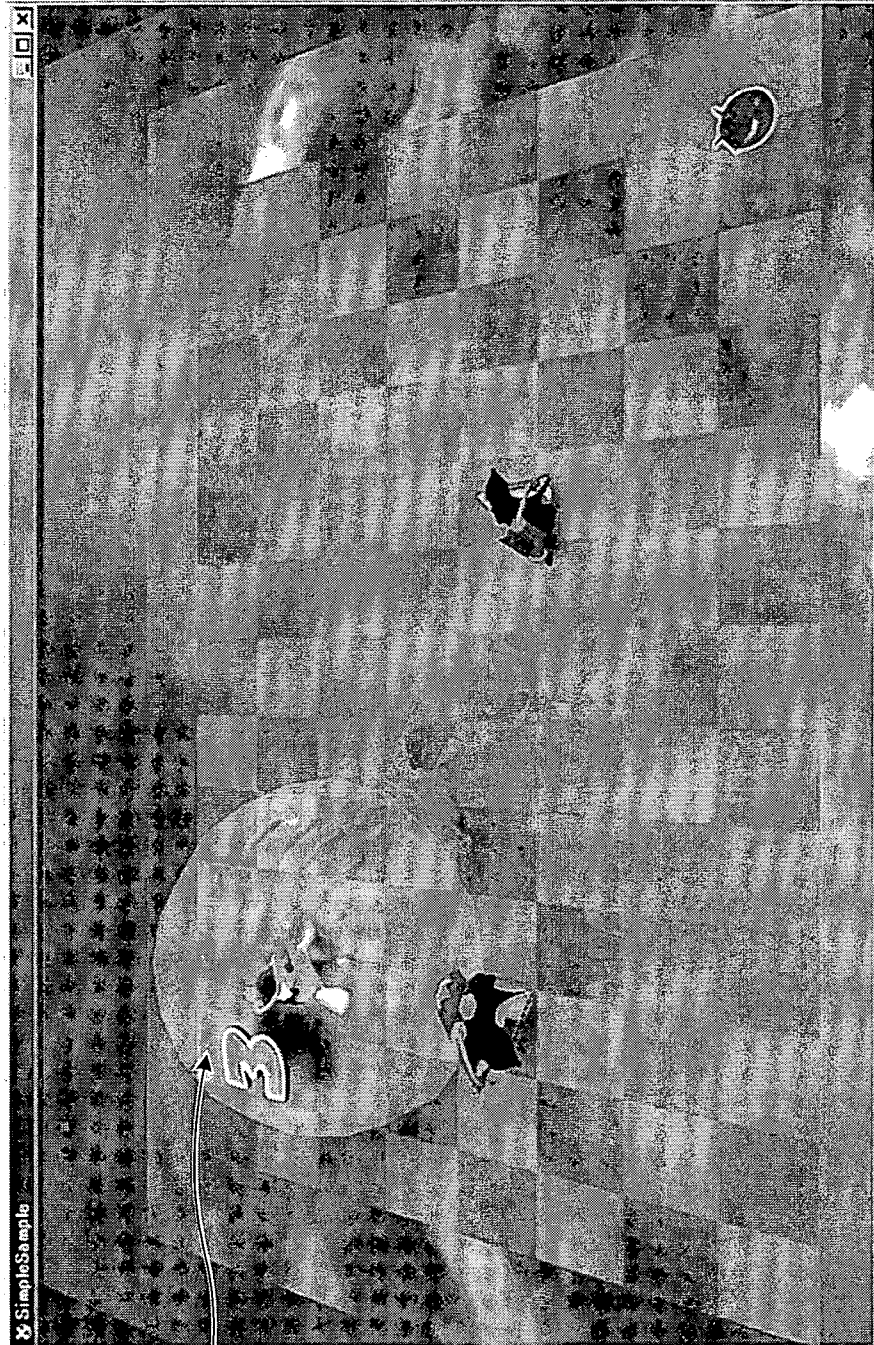




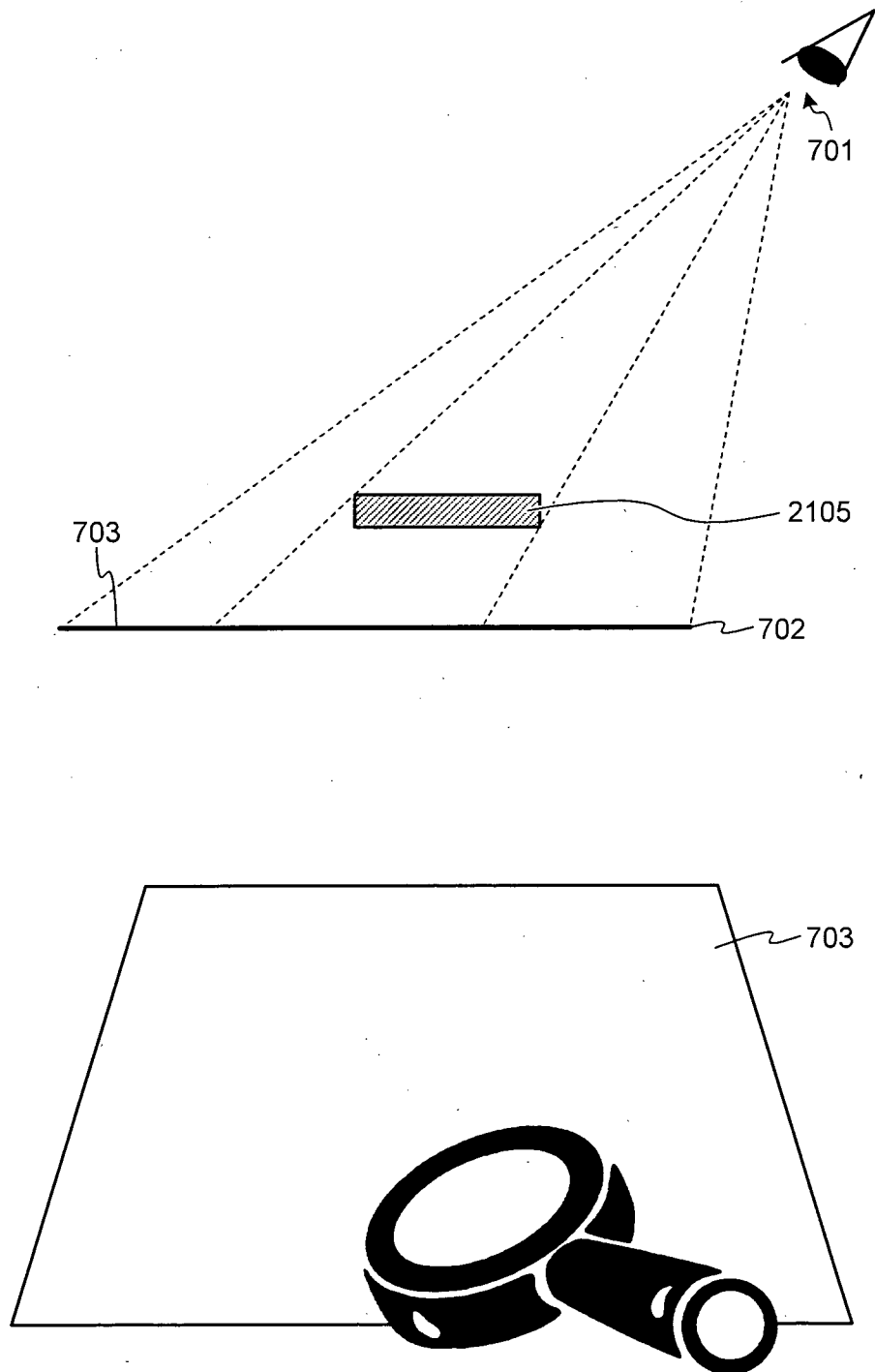
FIG.20



2001

20/22

FIG.21



21/22

FIG.22

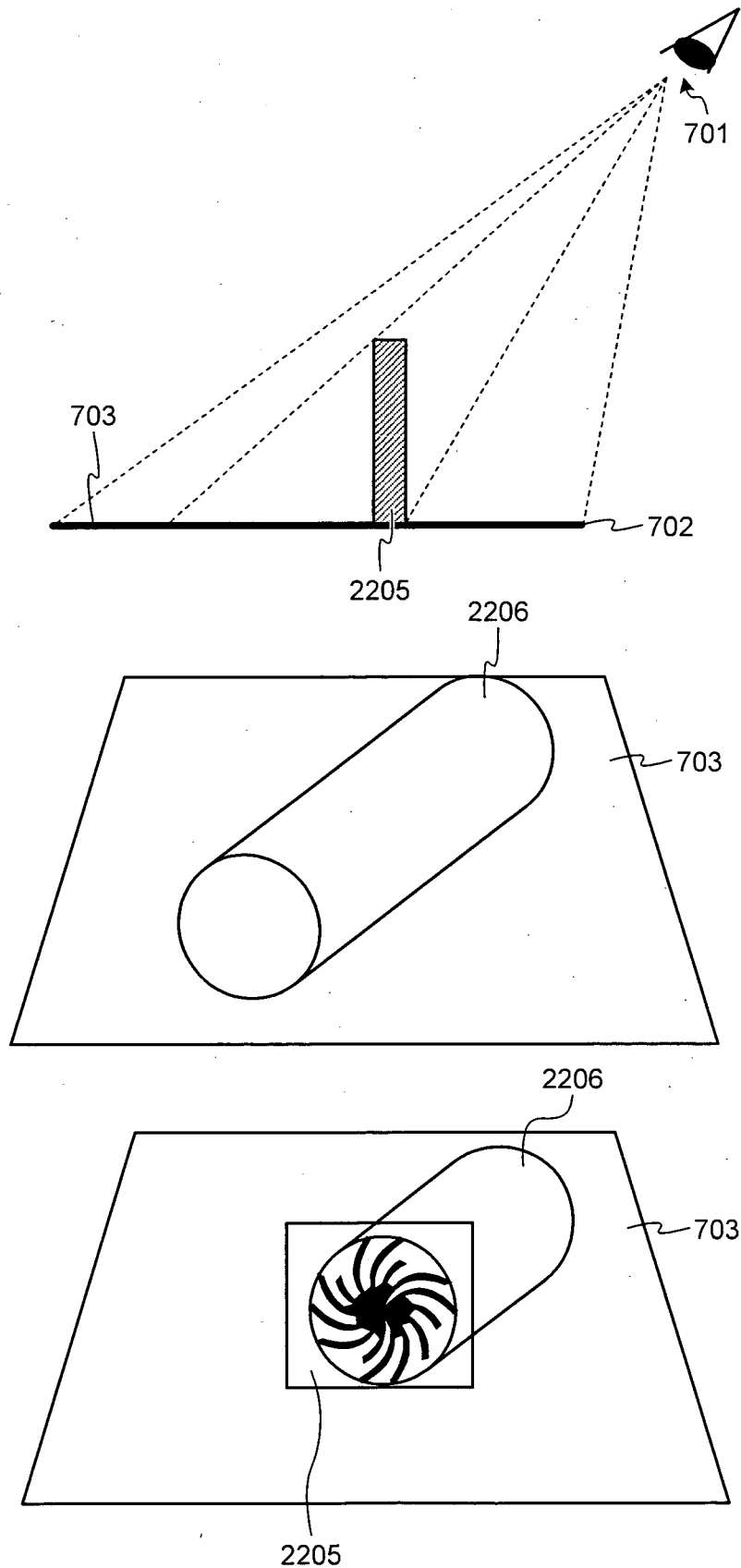
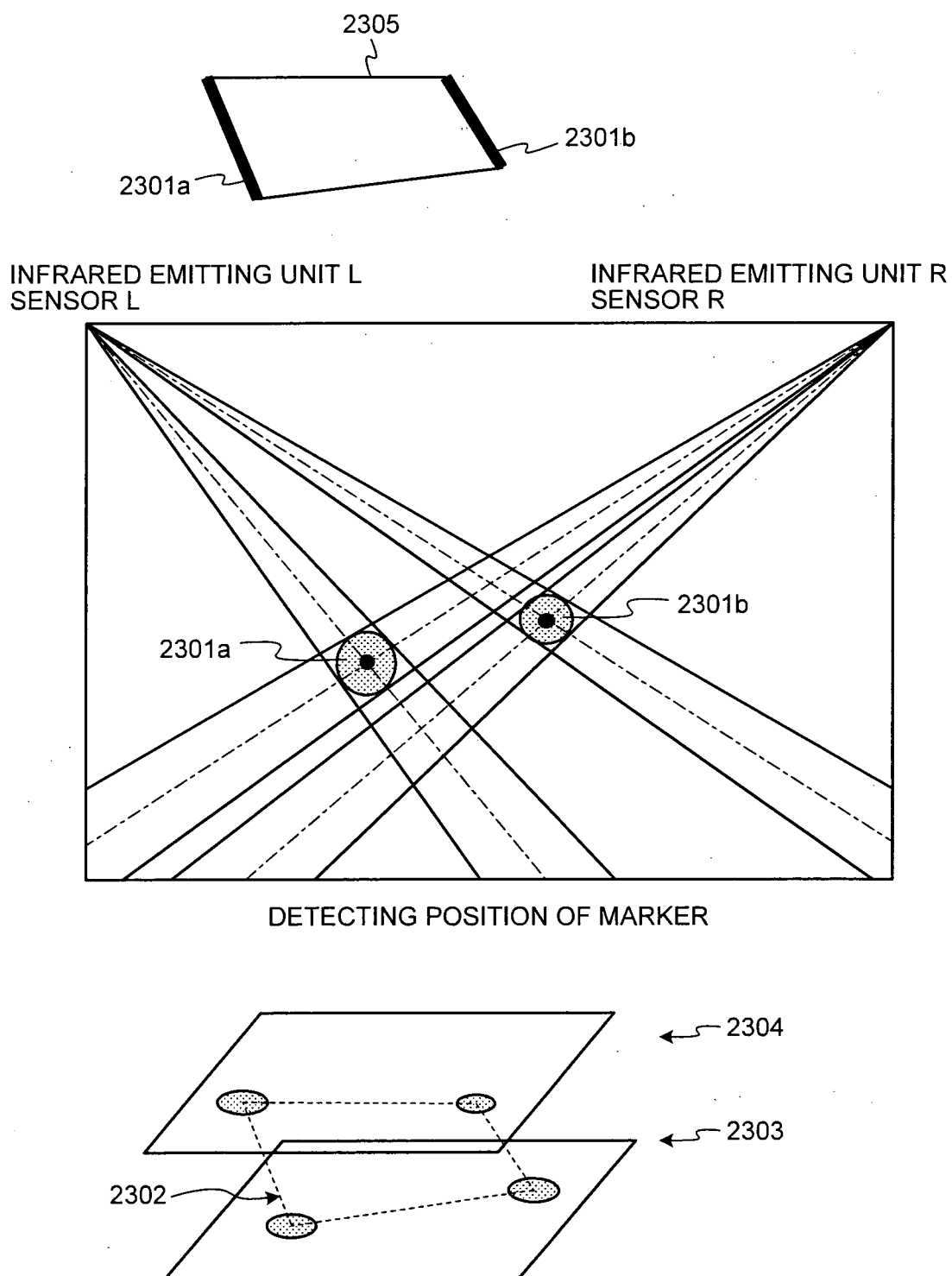


FIG.23



# INTERNATIONAL SEARCH REPORT

International application No

PCT/JP2007/069121

**A. CLASSIFICATION OF SUBJECT MATTER**  
INV. H04N13/00 G06T15/40

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06F H04N G06T

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 6 417 969 B1 (DELUCA MICHAEL [US] ET AL) 9 July 2002 (2002-07-09)	1,15,16
Y	column 4, line 21 - column 7, line 20 column 7, line 59 - column 8, line 65	2-14
Y	EP 0 607 000 A (SUN MICROSYSTEMS INC [US]) 20 July 1994 (1994-07-20) page 4, line 48 - page 5, line 35 page 7, line 52 - page 9, line 41	2-14

☐ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

\* Special categories of cited documents:

- \*A\* document defining the general state of the art which is not considered to be of particular relevance
- \*E\* earlier document but published on or after the international filing date
- \*L\* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- \*O\* document referring to an oral disclosure, use, exhibition or other means
- \*P\* document published prior to the international filing date but later than the priority date claimed

- \*T\* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- \*X\* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- \*Y\* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- \*G\* document member of the same patent family

Date of the actual completion of the international search

20 November 2007

Date of mailing of the international search report

05/12/2007

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Penchev, Petyo

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/JP2007/069121

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 6417969	B1	09-07-2002	NONE
EP 0607000	A	20-07-1994	DE 69425913 D1 26-10-2000
			DE 69425913 T2 26-04-2001
			DE 69432994 D1 04-09-2003
			DE 69432994 T2 22-04-2004
			JP 3575622 B2 13-10-2004
			JP 6317762 A 15-11-1994
			US 5394202 A 28-02-1995