TORNADO: Omnistereo Video Imaging with Rotating Optics

Kenji Tanaka, *Member*, *IEEE*, and Susumu Tachi, *Member*, *IEEE*

Abstract—One of the key techniques for vision-based communication is omnidirectional stereo (omnistereo) imaging, in which stereoscopic images for an arbitrary horizontal direction are captured and presented according to the viewing direction of the observer. Although omnistereo models have been surveyed in several studies, few omnistereo sensors have actually been implemented. In this paper, a practical method for capturing omnistereo video sequences using rotating optics is proposed and evaluated. The rotating optics system consists of prism sheets, circular or linear polarizing films, and a hyperboloidal mirror. This system has two different modes of operation with regard to the separation of images for the left and right eyes. In the high-speed shutter mode, images are separated using postimage processing, while, in the low-speed shutter mode, the image separation is completed by optics. By capturing actual images, we confirmed the effectiveness of the methods.

Index Terms—Virtual reality, 3D/stereo scene analysis, image processing, computer vision.

1 Introduction

One of the approaches to the construction of a vision-based communication system is to control a remote vision sensor according to the motions of an observer. An example of this approach is telexistence [1], in which information, such as images or sound, is captured and displayed according to the head position, direction, or motion of a user. This approach is efficient when there is only one observer because the full bandwidth of the data channel can be used for the presented image.

Another approach is to collect more ray information than is obtainable from a single, normal perspective camera and to reconstruct images for the observer(s). For example, an omnidirectional image system can provide a wider field of view (FOV), which contributes to the sense of immersion [2]. It also reduces the sense of discomfort induced by tracking error and delay, which are inevitable in a master-slave vision system [3]. Multiviewpoint systems help multiple people view the same objects or scenes at the same time in different ways. The ideal approach is to capture and reconstruct light rays at every location (V_x, V_y, V_z), at every possible viewing angle (θ, ϕ), for every wavelength (λ), and at any time (t), as defined in the 7D plenoptic function [4]. McMillan and Bishop [5] applied the plenoptic function to image-based modeling and rendering (IBMR) by describing 5D ray information (three for the position and two for the viewing angle). Gortler et al. [6] and Levoy and Hanrahan [7] proposed a 4D description of ray information by constraining the scene to a bounding box. Since it is difficult to build a dense camera array, multiviewpoint images are often captured by moving a single camera and photographing images at certain periodic positions. Because of this substitution, many studies in this field use still scenes or objects. To capture the light field at standard video rates, NHK Science and Technical Research Laboratories developed a method for real-time, full-color integral photography with a high-definition camera and a lens array [8], while Naemura et al. implemented a digital version of an integral photography system [9]. Although the above are full-color, real-time implementations of an integral photography system, the resolution of the output image is still very low.

For this reason, there have been efforts to construct image sensors for specific applications such as omnistereo. A short introduction to omnistereo imaging is given in this section. Section 2 presents the basic concept of the proposed system TORNADO, followed by actual implementations with two different modes (Sections 3 and 4).

1.1 Omnistereo Imaging

Omnistereo imaging is characterized by capturing or reconstructing stereoscopic images for a wide FOV. Efforts are being made to construct new projection models for omnistereo imaging. Naemura et al. [10] proposed an approach to realize a multiuser projection system in which several users can experience an adequate depth sensation simultaneously. Peleg and Ben-Ezra [11] showed a theoretical basis of this type of projection by proposing circular projection as a multiple viewpoint image projection. Concentric mosaics [12], a 3D plenoptic function, constrains the camera motion to planar concentric circles.

Fig. 1 (left) shows the camera configuration corresponding to the left eye (orange circles) of an observer in omnistereo imaging. When the observer turns his or her head by θ degrees, the left eye (or the camera) moves and the light axis is rotated. In other words, the position of each camera varies according to the viewing direction. Adding this one dimension to the other two dimensions of the captured image, total 3D ray information has to be recorded

• *The authors are with the Graduate Schoool of Information Science and Technology, The University of Tokyo, 113-8656 Tokyo, Japan. E-mail: tanaken@computer.org, tachi@star.t.u-tokyo.ac.jp.*

Manuscript received 8 Dec. 2004; revised 19 Mar. 2005; accepted 12 May 2005; published online 8 September 2005.
For information on obtaining reprints of this article, please send e-mail to: tvcg@computer.org, and reference IEEECS Log Number TVCGSI-0242-1204.

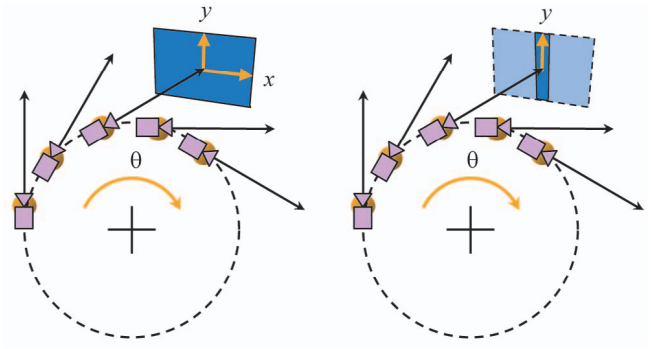

Fig. 1. Formulation of omnistereo imaging.

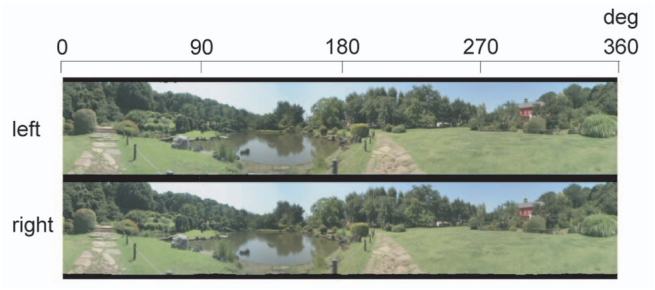

Fig. 2. A sample of omnistereo image.

at every instant. Here, we can reduce one dimension by replacing the normal camera with a line scan camera, as shown in Fig. 1 (right). With this approximation, the generated image is always accurate with regard to the observer's front direction, even if his or her head rotates, and the incoming light rays from other directions are approximated by the perpendicular incident light rays to his or her left (or right) eye when his or her head turns in another direction. Henceforth, we call this 2D light ray expression (θ and y) omnistereo imaging.

Omnistereo imaging is a subset of concentric mosaics. In omnistereo imaging, the radius parameter out of three parameters of the concentric mosaics is fixed at half the interpupilar distance (IPD). Fig. 2 is a sample of an omnistereo image constructed by panning two relatively fixed line scan cameras. Points at infinity are plotted in the same horizontal positions in the left and right images.

1.2 Omnistereo Displays

There have been efforts to build display systems for omnistereo images. Simon et al. [13] implemented a multiple-view method to generate omnistereo computer graphics and evaluated them in conventional panoramic virtual environment display systems such as CAVE using active stereo projection. Tanaka et al. [14] constructed a rotating omnistereo display called TWISTER (Telexistence Wide-angle Immersive STEReoscope), which can display panoramic and stereoscopic full-color live video sequences without the need for special eyewear or head tracking. By analyzing the distortion of the depth presentation in this rendering, it is reported that, although the estimated depth can be larger for distances larger than the radius of TWISTER, the depth distortion does not cause a serious problem and this method can also provide a sufficient sense of immersion. Omnistereo imaging is compatible with the rendering method of images for TWISTER.

1.3 Omnistereo Image Sensors

There have also been efforts to capture omnistereo video sequences in real scenes instead of by rotating two cameras. Some methods use vertical parallax by placing two omnidirectional image sensors vertically. Yokoya et al. [15] constructed an omnidirectional image sensor with vertical parallax by combining two downward-pointing hyperboloidal mirrors with different parameters. The mirrors share the axes and their outer focal points are made to coincide. However, this sensor cannot provide horizontal parallax. Shimamura et al. [16] solved this problem by calculating a depth map from vertical parallax. Although this system reconstructs horizontal parallax in real time, estimating the detailed depth of objects robustly is difficult. As an approach different from those using vertical parallax, Peleg et al. [17] proposed a few methods using a curved mirror or lens to capture omnistereo video sequences. However, the proposed optical configurations were only schematic principles for the creation of circular projections using optics. Tanaka et al. [18] proposed another method using a curved mirror. Although it was a complete implementation of a wide-angle stereo imaging sensor, it could not cover a 360-degree scene. In this paper, we propose TORNADO (Tornado Omnistereo Real Natural-scene Acquisition system for Dynamic Objects) to capture omnistereo video sequences at standard video rates.

2 TORNADO: Omnistereo Capture with Rotating Optics

In this section, the basic idea of TORNADO is described. A rotator with impeller-shaped mirrors is shown in Fig. 3.1. Eight mirrors are aligned along the circumference so that their surface normals form 45-degree angles with the tangent line of the rotator. The mirrors reflect incoming light rays from the direction of the tangent line to the center of the rotator. Four of these mirrors, marked L, collect the incident light rays for the left eye of the observer, corresponding to the incident light rays in Fig. 1. The incident light rays for the right eye are collected in the same way by the four mirrors marked R. For simplicity, we assume that the width of the mirror is infinitesimal. In order to collect light rays from various vertical directions, the light rays are further reflected by two hyperboloidal mirrors (one for each eye) to two normal perspective cameras located on the upper and lower sides of the rotator. Since the inner focal points of the hyperboloidal mirrors correspond to the left and right eyes, the observer can coincide the vertical heights of the eyes by coinciding the two inner focal points of the mirrors. By turning the rotator at a sufficient speed so that a mirror for the left (or right) image reaches the next corresponding mirror while the shutter is open, the light rays for the left (or right) eye can be collected over continuous directions. We can also adjust the IPD by changing the angle formed by the surface normal of the mirrors and the tangent line (Fig. 3.2).

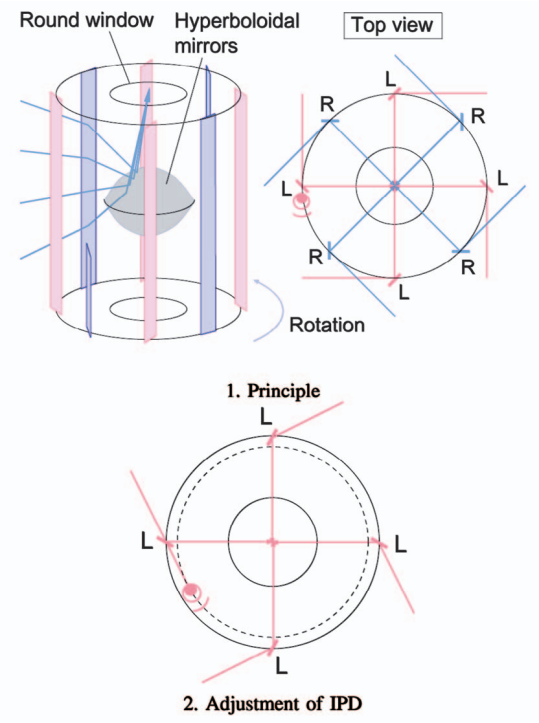

Fig. 3. Omnistereo capturing with impeller-shaped mirrors.

Note that, in the above discussion, there is no mention of a method to separate the light rays for the left and right eyes. Moreover, it is difficult to fabricate and arrange impeller-shaped mirrors of infinitesimal width. TORNADO specifically provides a solution to this problem by employing prism sheets instead of impeller-shaped mirrors in order to bend the light rays. We built and examined two versions in consideration of the separation of left-eye and right-eye images:

1. High-speed shutter mode: Images are captured at a high shutter speed and separated using image processing.
2. Low-speed shutter mode: Images are captured at a low shutter speed and separated using special optics such as polarizing film.

3 High-Speed Shutter Mode

3.1 Design and Implementation

A prism sheet is an optical element that bends incident light rays at a certain angle (Fig. 4). The incident angle for the light ray exiting perpendicularly from the prism (uneven) side of our prism sheet is 23 degrees, as calculated in Appendix A. We can replace the impeller-shaped mirrors with prism sheets. Light rays for the left and right eyes are delivered to the center of the rotator with the prism sheets for their respective eyes. In Fig. 5, the left-eye light rays and corresponding prism sheets are colored lime green and the right-eye light rays and prism sheets are colored dark green.

The left-eye and right-eye prism sheets are alternately aligned in the circumferential direction. A prism sheet is aligned so that the prism side faces inward. The distance between the extension line of an incident light ray and the

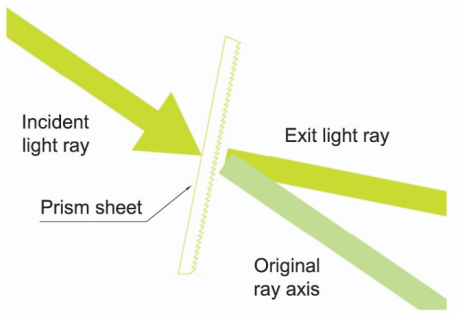

Fig. 4. Function of prism sheet.

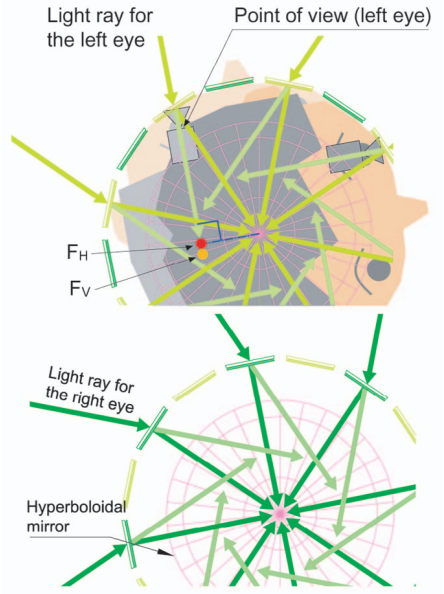

Fig. 5. Collection of the left-eye and right-eye light rays. Top view of the left-eye optics overlaid on the left eye of a human observer (upper) and top view of the right-eye optics (lower).

center of the rotator corresponds to half the IPD. Using an average figure, the IPD is set at 65 mm. The horizontal focal point F_H and the vertical focal point F_V are located on the extension line of the incident light ray. Point F_H is a foot of a perpendicular dropped from the center of the rotator. Point F_V is placed so that the distance between the prism sheet and the center of the rotator equals the distance between the prism sheet and Point F_V. Although the horizontal focal point does not exactly coincide with the vertical focal point, the difference is so subtle that there is virtually no aspect ratio problem, which occurs in the curved-mirror method [18]. Fig. 6 presents an oblique perspective figure of the optics.

Since it is difficult to fabricate a prism sheet with a curvature, the developed prototype uses oblong pieces of planar prism sheets. As shown in Fig. 7, 32 aluminum pillars and 32 prism sheets are arranged alternately in a circle to form a rotator. The height of each prism sheet is 130 mm. They are placed 105 mm from the center of the rotator. The surface of the pillar is covered with black velvet to avoid reflection. The rotator is rotated on a turntable.

The whole structure of the system is roughly divided into an upper structure and a lower structure (Fig. 8). A rotator and a hyperboloidal mirror (SOIOS138 from Eizoh

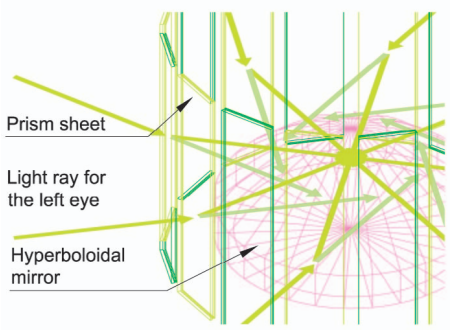

Fig. 6. Collection of the left-eye light rays. Oblique perspective view of the optics.

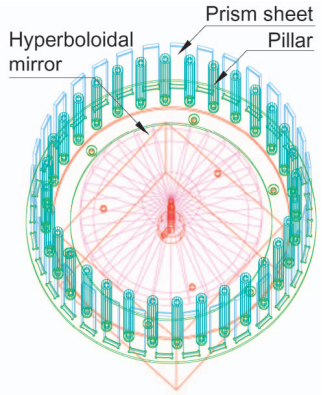

Fig. 7. Lower structure of the system. The rotator on the turntable and the hyperboloidal mirror suspended from the upper structure are shown. The upper holding ring of the rotator is detached in this figure.

Corp.) are located on the lower structure. The diameter of the mirror is 138 mm and the maximum depression angle is 16 degrees. Unlike the impeller-shaped mirror described in the previous section, a single hyperboloidal mirror is used in the actual implementation. The hyperboloidal mirror is located at the center of the rotator. The light rays that reach the inner focal point of the mirror are reflected on the outer focal point and a camera located here can capture the omnidirectional light ray information. A high-definition camera (GR-HD1 from JVC) is located on the upper structure, with the focal axis of its lens aligned to be vertical.

Since it is unnecessary to rotate the hyperboloidal mirror itself, it is suspended from the center of a clear acrylic plate placed on a round window on the floor separating the upper and lower structures. Since the upper structure is supported by a transparent acrylic tube, there is no obstacle to omnidirectional capturing (Fig. 9). The horizontal FOV is 360 degrees and the vertical FOV is 60 degrees. An overview photograph of the system is shown in Fig. 10. The shutter speed was set at 1/250 seconds to maximize the exposure and minimize the motion blur of the aperture image.

In the prototype, we used a 5-phase stepping motor (PK-545AW), a motor unit (UPK/W), and a motor controller (SG9200T), all from Oriental Motor. The motor was driven at 700 Hz, which corresponds to an angular velocity of 504 deg/sec (= 700 Hz $\times$ 0.72 deg). An accurate measurement using a video observation gave an angular velocity of 506.84 deg/sec. Under that condition, the rotator rotates at a speed of 3/64 revolutions per frame (1/29.97 sec). Since either

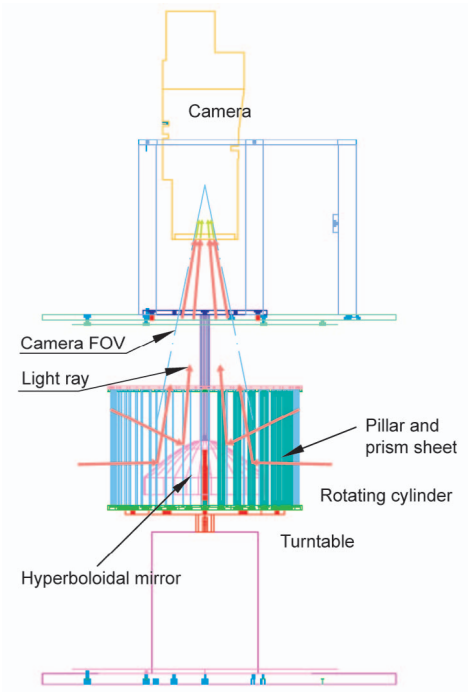

Fig. 8. The whole structure of the system in the high-speed shutter mode.

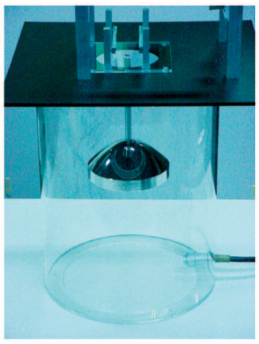

Fig. 9. A hyperboloidal mirror is suspended from the center of a clear acrylic plate placed on a round window. Thanks to this structure, the mirror is free from mechanical jitter.

Fig. 10. Overview photograph of TORNADO (high-speed shutter mode).

the left-eye or right-eye aperture is aligned with a pitch of 1/16 of the circumference and its width is 1/64 of the circumference, all omnistereo images can be reconstructed by utilizing four frames. This principle is depicted in Fig. 11.

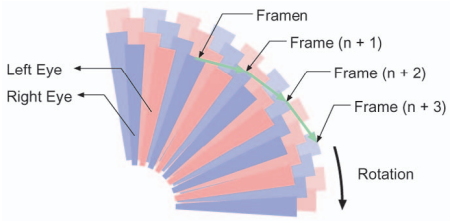

Fig. 11. Principle of the image mosaicing in the high-speed shutter mode.

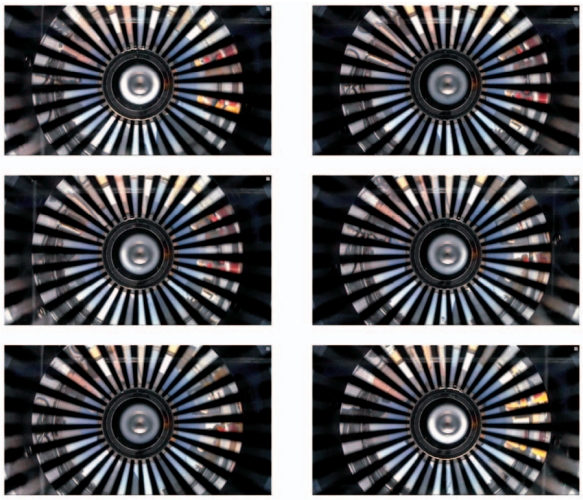

Fig. 12. Camera image sequences of a real scene in the high-speed shutter mode (left-eye images on the left column and right-eye images on the right column).

The figure shows the left-eye (red) and right-eye (blue) aperture arrangements reflected in the hyperboloidal mirror over four frames. A left-eye (right-eye) aperture rotates clockwise so that it overlaps the next left-eye (right-eye) aperture in four frames. Images are accumulated over frames to construct an omnidirectional left-eye (right-eye) image. The whole image is renewed by replacing only the earliest parts of the image with the current (latest) part of the image.

3.2 Capturing Images of a Real Scene

Fig. 12 contains the camera-image sequences of a real scene arranged for every frame. The rotation speed is 3/64 revolutions for each frame. The shutter speed is 1/250 seconds. In this example, the images are zoomed in to put the spatial resolution ahead of the capture of the whole hyperboloidal mirror. The reconstructed omnistereo sequences are shown in Fig. 13 and Fig. 14. Curved vertical patterns are due to the correction of image distortion, detailed in the next section. The image accumulation can be observed in the first three frames. We can also observe the motion of a bicycle and a figure. Comparing a stuffed toy in the right-eye images with that in the left-eye images, the parallax is confirmed. Thin dark stripes are visible because the width of the aperture was slightly less than 1/64 of the circumference. These stripes can be erased by narrowing the radius of the pillars used in the rotator.

Theoretically, the temporal resolution focused on a point is 7.5 frames per second (fps) because pixels are renewed only once every four frames. Using a GR-HD1 camera, the frame rate can be doubled at the expense of reduced spatial resolution. Since a converted image for each frame consists of images from four frames, the synthesized image contains spatial discontinuities for faster objects.

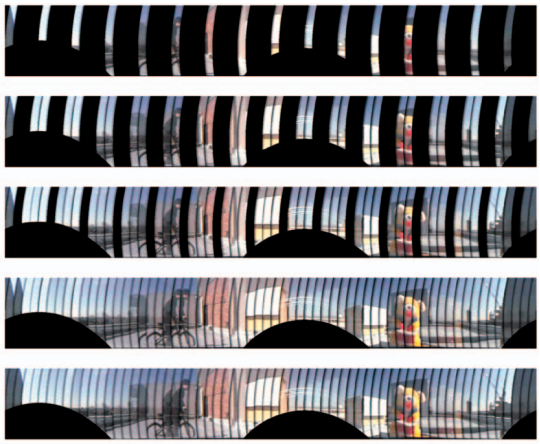

Fig. 13. Synthesized omnistereo image sequences in the high-speed shutter mode (left eye).

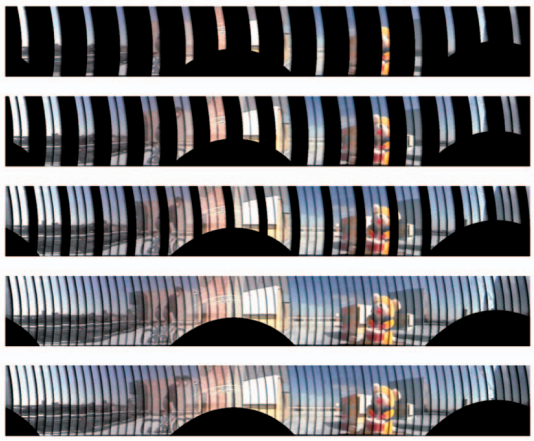

Fig. 14. Synthesized omnistereo image sequences in the high-speed shutter mode (right eye).

4 Low-Speed Shutter Mode

4.1 Principle

In the high-speed shutter mode, we digitally extract the left-eye and right-eye images from all the frames captured at high shutter speed and reconstruct omnistereo images. On the other hand, the two images are separated using special optics in the low-speed shutter mode. In this case, since there is no need for a temporal separation, we can utilize temporal blending to construct a whole omnistereo scene.

The whole structure is shown in Fig. 15 and the overview of an actual implementation is shown in Fig. 16. Circular polarizing films, a nonpolarizing beam splitter, and one extra camera were added to the high-speed shutter system. Oblong circular polarizing films were installed inside the prism sheet (Fig. 17) as well as in front of the camera lenses.

Consider how circular polarizing film works (Fig. 15). The left-eye and right-eye light rays pass through the respective pieces of circular polarizing film. The light rays that reach the inner focal point in the mirror are reflected on the outer focal point. In this reflection process, the rotation

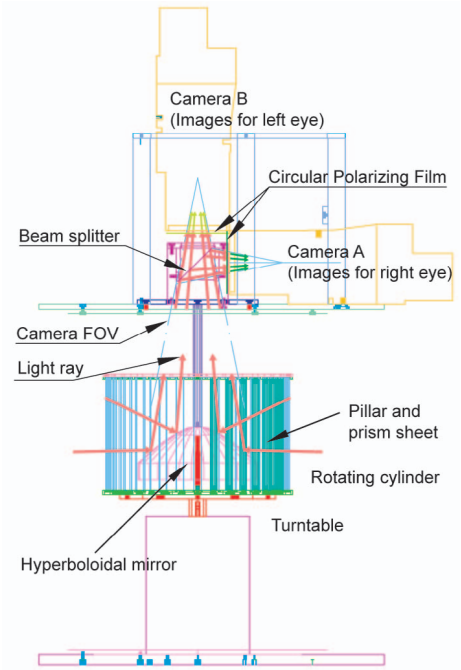

Fig. 15. The whole structure of the system in the low-speed shutter mode. The light rays reflected by the hyperboloidal mirror are split by the beam splitter into two groups of light rays with opposite-hand circular polarization. The left-eye and right-eye light rays are extracted from each group of light rays by right-handed circular polarizing films.

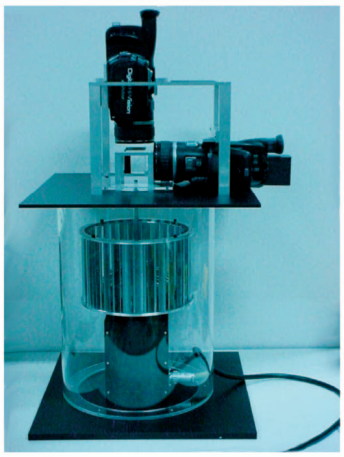

Fig. 16. Overview of an actual implementation in the low-speed shutter mode.

of the circular polarization is reversed. The light rays further enter a nonpolarizing beam splitter and two cameras, A and B. Since the rotation of circular polarization of the light rays reflected by the beam splitter is reversed, the left-eye light rays entering camera A have a left-handed circular polarization and the right-eye light rays have a right-handed one. The right-handed circular polarizing film in front of camera A extracts the right-eye light rays. On the other hand, because the rotation of circular polarization of the light rays that pass straight through the beam splitter is unchanged, the left-eye light rays entering camera B have a right-handed circular polarization and the right-eye light rays have a left-handed one. The circular polarizing film in front of camera B, which is also right-handed, extracts the

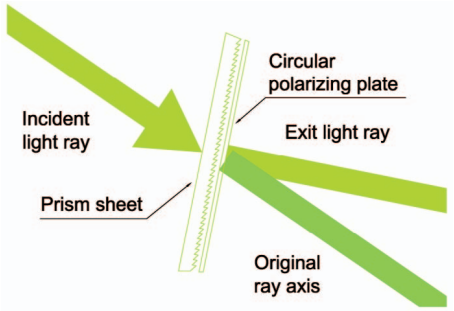

Fig. 17. Top-view configuration of prism sheets and circular polarizing films for the left-eye.

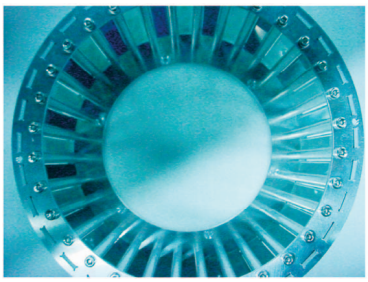

Fig. 18. Top-view photograph of the rotator.

left-eye light rays. The situation remains true during the rotation of the rotator. When the rotator is rotated at a sufficient speed for the shutter speed, both the left-eye and right-eye light rays can be captured over a continuous range of directions. By completing the above process within one frame, omnistereo video sequences can be obtained.

4.2 Design and Implementation

As mentioned above, a few devices were added to the high-speed shutter system. Two high-definition cameras (GR-HD1 from JVC), whose lens axes form a right angle, were placed on the upper structure. In addition, a nonpolarizing beam splitter was placed at the cross-point of the two axes. The right-handed circular polarizing film was placed in front of the lenses of the cameras. A hyperboloidal mirror and a rotator were placed on the lower structure. The rotator consists of oblong prism sheets and oblong circular polarizing films. The horizontal FOV is 360 degrees and the vertical FOV is 60 degrees.

4.2.1 Circular Polarizing Film

Shown in Fig. 18 is a top-view photograph of the rotator. The half in the upper left was photographed through a circular polarizing film. Apertures are observed as alternating light and dark portions. This situation is reversed with regard to the aperture image reflected by a lower aluminum ring because the rotation direction of circular polarization is reversed by the surface reflection.

4.2.2 Hyperboloidal Mirror

The hyperboloidal mirror (SOIOS138 from Eizoh Corp.) is plated with chromium and does not affect the polarization in the surface reflection. For example, a mirror with double refraction, such as an acrylic mirror, is inappropriate because it affects the polarization.

Fig. 19. Capture of omnistereo video sequences (low-speed shutter mode).

4.2.3 Nonpolarizing Beam Splitter

After a careful survey, we selected a 40-mm angle nonpolarizing beam splitter. The reflectance and transmittance for s-polarization and p-polarization are almost constant over a wide range of wavelengths within visible light. This condition is true as long as the angle of reflection is 45 degrees. Because the light rays that pass through the periphery of the beam splitter have an angle of reflection less than or greater than 45 degrees, the reflectance and transmittance are not always constant, which may lead to cross talk between the left-eye and right-eye images. However, the effect was negligible.

4.3 Capturing and Reconstructing Omnistereo Video Sequences

Omnistereo video sequences of an outside scene were captured using the system (Fig. 19).

To increase the disparity, the stuffed toy was placed a short distance, 20 cm, from TORNADO. Sample camera images are shown in Fig. 20. The shutter speed is 1/30 seconds. In the left column are the left-eye images and in the right column are the right-eye images. The rotation of the apertures is observed counterclockwise for the left-eye images and clockwise for the right-eye images. Shown in the first row is the rotator before rotation. The effect of the circular polarizing film can be confirmed since the apertures are alternately light and dark. When the rotator rotates, the light areas remain light and the dark areas remain dark. As the rotator rotates at sufficient speed, the dark (light-blocked) areas are temporally blended and all the areas can be clearly seen. Thus, images that are exclusively for the left or the right eye can be observed.

Here, we observe some cross talk between the two images due to inadequate light blocking. The images in the first row show that the opacity of the apertures, which should block light, varies according to the direction. The inadequate light blocking is caused by the following factors:

- If there is a difference in the transmittance (reflectance) of the beam splitter for s-polarized and p-polarized components, the light ray exiting the beam splitter has an ellipsoidal polarization (not a circular one) for the left (right) image. For the left image, the axis ratio of the ellipsoid is defined by T_s/T_p, while, for the right image, the axis ratio is defined by R_p/R_s, where T and R stand for the transmittance and the reflectance, respectively, and suffixes s and p stand for s and p-polarized components, respectively. According to the data sheet of the beam splitter, $T_s/T_p = 0.40/0.47 = 0.85$ and $R_p/R_s = 0.53/0.60 = 0.88$. From this calculation, it is shown that the light blocking is more inadequate and the cross talk is greater in the left-eye image than in the right-eye image.
- There are two major light-blocking situations with circular polarization: 1) The slow (fast) axis of one waveplate matches the fast (slow) axis of the other waveplate and 2) the slow (fast) axis of one waveplate matches the slow (fast) axis of the other waveplate. In the former situation, since the phase difference between the two axes is zero, the light blocking is sufficient and does not depend on the wavelength of the light ray. On the other hand, in the latter situation, the phase difference between the two axes is 2π. Therefore, the light blocking is inadequate and depends on the wavelength. Fig. 20 (first row) shows this situation. In both the left and right images, the light blocking is inadequate in the directions of 3 and 9 o'clock. In these directions, the image is colored because of a dependency on the wavelength, which is why the cross talk depends on the horizontal position.

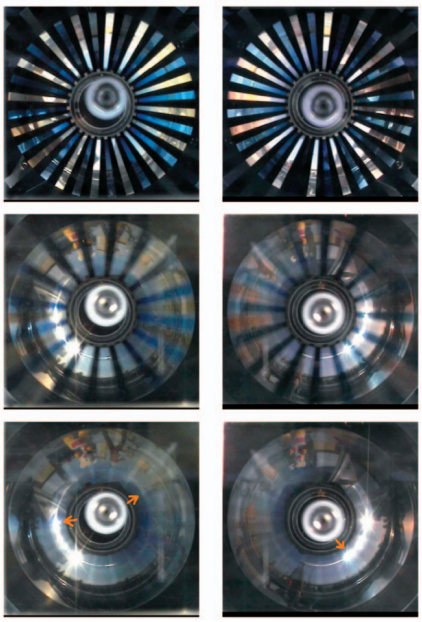

Fig. 20. Sample camera images in the low-speed shutter mode (left-eye images on the left column and right-eye images on the right column). The shutter speed is 1/30 seconds. The rotation of the apertures is observed counterclockwise for the left-eye images and clockwise for the right-eye images. Shown on the first row is the rotator before rotation. When the rotator rotates (the second row), the light areas remain light and the dark areas remain dark. As the rotator rotates at a sufficient speed (the third row), the dark (light-blocked) areas are temporally blended and all the areas can be clearly seen. Thus, images that are exclusively for the left or the right eye can be observed. We can observe some cross talk (orange arrow) in the mirrored position of the original object in the opposite side image.

Images in the second row show the rotator rotating below its stable speed and radial nonuniformities of luminance can be observed because the angular velocity of the rotation is not a multiple of 1/16 revolutions per frame, which corresponds to the angle between adjacent light apertures.

Fig. 21. Correction of image distortion. Straight vertical lines of the buildings actually appear curved in the upper figure and are corrected in the lower figure.

Fig. 22. Images before (upper) and after (lower) the brightness correction.

The nonuniformities are reduced at a stable rotating speed of two revolutions per second or 1/15 revolutions per frame (the images on the third row). As the angular velocity of the rotation is slightly more than 1/16 revolutions per frame, narrow radial regions become lighter.

4.3.1 Correction of Image Distortion

In fact, the incident angle, which is projected on the top view, for the light ray exiting perpendicularly from the prism side of the prism sheet varies slightly depending on the elevation angle of the light ray (Appendix B). The actual images were accurately restored by taking this effect into consideration (Fig. 21).

4.3.2 Brightness Correction

Because of the cross talk between the two images, the black offset reaches a high level and the dynamic range of the image intensity is decreased. The brightness was corrected by widening the dynamic range. After the correction, the black offset was adequately reduced and the contrast increased (Fig. 22).

4.3.3 Image Reconstruction

Shown in Fig. 23 are the video sequences reconstructed with the correction of image distortion and brightness. The shutter speed is 1/30 seconds. Shown in the first row are the images before rotation. In the second row are the images before the stable rotation. Other images are arranged for every frame.

Disparities can be observed in the video sequences and the performance as an omnistereo image sensor can be confirmed. A ghost-like cross talk on the left (right) of an object in the right-eye (left-eye) image can also be observed. The intensity of the cross talk depends on the horizontal position (therefore, direction) of the object and it is greater in the left-eye image than in the right-eye image.

Fig. 23. Reconstructed images (left-eye images on the left column and right-eye images on the right column).

4.4 Image Separation by Linear Polarizing Films

In order to overcome the ghosting problem observed in the image separation using circular polarization, we employed linear polarizing films instead of circular polarizing films. Oblong linear polarizing films were installed inside the prism sheet. The left-eye and right-eye light rays pass through horizontal and vertical linear polarizing films, respectively. In this case, circularly arranged linear polarizing films were installed in front of the camera lenses. The films in front of camera A (vertical extractor) extract vertically polarized light rays, while the films in front of camera B (horizontal extractor) extract horizontally polarized light rays (Fig. 24). Linear polarizing film was precisely cut into 16 triangle-shaped pieces and pieced together. Fig. 25 shows a horizontal extractor photographed with linear polarizing films. Fig. 26 shows the rotator photographed with and without linear polarizing films and the rotator photographed with a horizontal extractor in front of a camera. Fig. 27 shows the result of image capture when this method is used. In these images, cross talk is considerably reduced. We also confirmed that, by opening the apertures of the cameras, seams of the polarizing film pieces are completely invisible.

4.5 Real-Time Image Restoration

We developed a real-time image restoration system that carries out the geometry conversion and generates whole omnistereo images in real time. The MPEG2 output stream of the camera (GR-HD1) was transformed to an uncompressed video stream using an MPEG2 transport stream

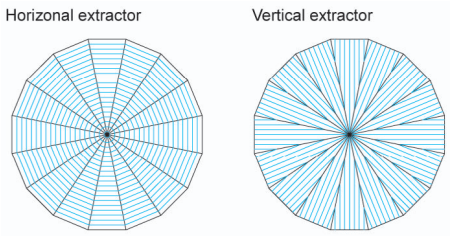

Fig. 24. Light ray extraction using circularly arranged linear polarizing films.

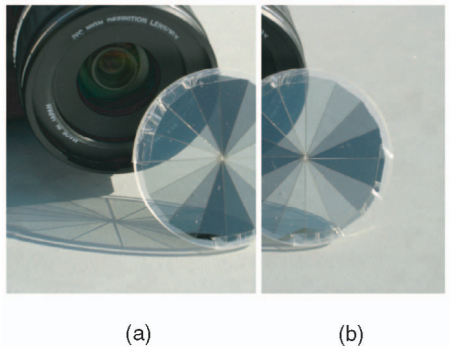

Fig. 25. Horizontal extractor shot with (a) a vertically polarizing film and (b) a horizontally polarizing film.

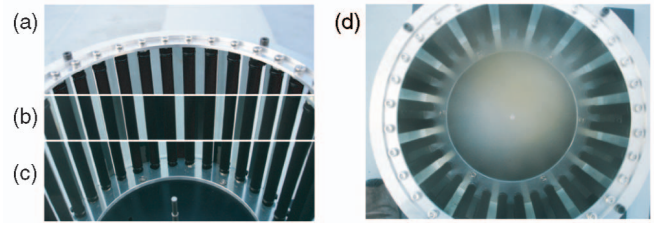

Fig. 26. (a) Rotator photographed without a polarizing film. Rotator photographed with (b) a vertically polarizing film and (c) a horizontally polarizing film. (d) Rotator photographed with a horizontal extractor.

(TS) splitter and MPEG2 decoder. The real-time geometry conversion was implemented with texture mapping techniques in a Microsoft DirectX platform, with which we can warp arbitrary points to the desired points in the converted image.

The resolution of the converted image is $1{,}920 \times 256$ pixels. We assigned one PC for each eye. Using a Pentium 4 machine, the frame rate was 7 fps for one monitor and 3 fps for two monitors. There was a delay of about 0.5 seconds due to MPEG2 decoding. With higher-performance PCs, we can expect a real-time conversion of 30 fps video sequences.

5 Discussion

5.1 Comparison of This Method with Previous Works

Unlike the curved-mirror method [18], TORNADO achieves a 360-degree omnistereo capture with a relatively compact body and is also free from the aspect ratio problem.

Since there are moving parts in this method, special attention must be given to the precision of the mechanism. In fact, a slight jitter can be observed in the generated video sequences. This jitter is likely due to the turntable directly attached to the motor shaft. By using ball bearings to support the turntable, we expect to reduce the jitter.

On the other hand, TORNADO is fairly robust compared with a model-based approach, such as that described by Shimamura et al. [16] because there is basically no recognition or estimation process.

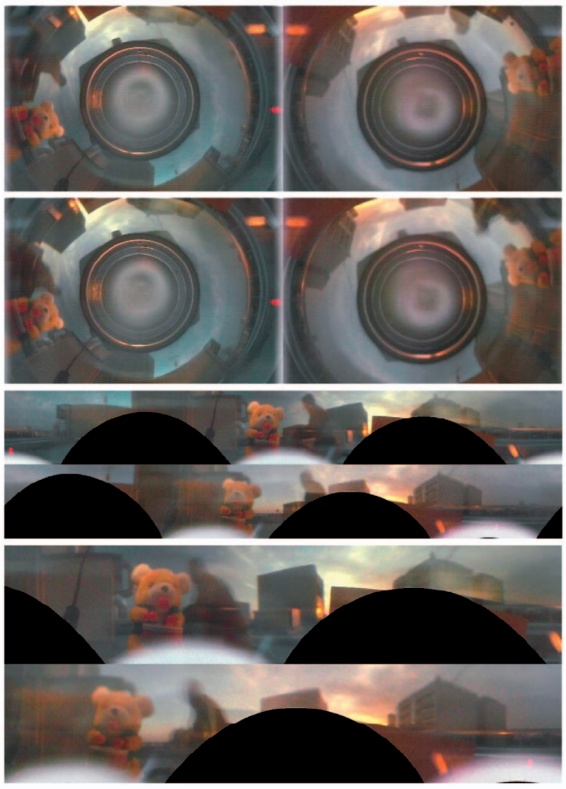

Fig. 27. Image capture with linear polarization separation in the low-speed shutter mode. Shown in the first and second rows are the original images for frame n and frame $n+1$, respectively (left-eye images on the left column and right-eye images on the right column). Shown in the next two rows are the converted images (left-eye images on the upper side and right-eye images on the lower side).

5.2 Comparison of High-Speed and Low-Speed Shutter Modes

In the high-speed shutter mode, we can capture a bright and clear image without cross talk because we use simple optics with only a prism sheet. Furthermore, there is no need for frame synchronization or color adjustment. On the other hand, both the quantity of information of the images and the frame rate are roughly one-fourth that of the low-speed shutter mode, considering that we use only a single camera and there is no information in the images of pillars. Other shortcomings are discontinuities in the synthesized image and a need for postimage processing. This mode is suitable for scenes with slow-moving objects.

In the low-speed shutter mode, there is no spatial discontinuity in the converted images and there is no need for accurate speed control of the rotator. The quantity of information of the images and the frame rate are roughly four times those of the high-speed shutter mode. The shortcomings are darker images and cross talk between the left-eye and right-eye images due to the use of circular polarizing films and a beam splitter. The cross-talk problem

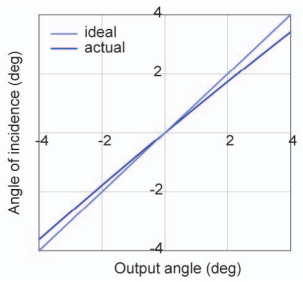

Fig. 28. Relation between the output angle and the incident angle of the prism sheet.

is almost solved by using linear polarizing films. We can conclude that this mode is suitable for more general scenes with fast-moving objects.

5.3 Refraction with a Prism Sheet and Spatial Resolution of the Converted Image

Here, we evaluate the error in the converted image due to the use of planar prism sheets without curvature. The prism sheet is made of acryl and has a refractive index of 1.5. As mentioned above, the incident angle for the light ray exiting perpendicularly from the prism side of the prism sheet is 23 degrees, as calculated in Appendix A. Since a prism sheet has a certain width, a light ray passing through at a point other than the horizontal center of the piece does not exit perpendicularly from the prism sheet. Fig. 28 shows the relation between the output angle and the incident angle, measured for the counterclockwise direction.

If the difference in the incident angle is the same as that of the output angle, light rays incoming from directions that have shifted by 23 degrees reach the center of the rotator. However, the incident angle actually increases by 0.88 degrees when the output angle is 1 degree; hence, there is an error angle of 0.12 degrees. The actual range of the output angle is -3.3 degrees to 3.3 degrees and the corresponding error of the incident angle is -0.32 degrees to 0.45 degrees. This calculation implies that the light rays incoming from a certain direction include a -0.32 to 0.45-degree error in this direction.

On the other hand, we can calculate the maximum resolution in visual acuity of the converted image, which is limited by the camera resolution. The vertical resolution of the camera is 720 pixels and the image of the hyperboloidal mirror has to be within a 720 $\times$ 720 square. Under this condition, the circumference of the mirror has a length of 2,260 pixels. This figure roughly corresponds to an angular resolution of 2,260/360 [pixels/deg] and a visual acuity of 2/20 at the bottom of the converted image. When measured at medium height, since the radius is half that of the mirror, the resolution becomes 1,130/360 [pixels/deg] and has a visual acuity of 1/20. Now, one pixel becomes 0.3 degrees and the level of error approaches the error of the planar prism sheet approximation.

6 Conclusion

We designed and implemented an omnidirectional stereo (omnistereo) imaging sensor with rotating optics consisting of prism sheets and a hyperboloidal mirror. We proposed two modes: the high-speed shutter mode with a single camera and the low-speed shutter mode with two cameras and circular or linear polarizing films. We confirmed the following performances for both modes:

- Omnistereo image capture, in which the generated image is always accurate with regard to the observer's front direction even if he or she rotates his or her head.
- Horizontal FOV of 360 degrees and vertical FOV of 60 degrees.
- 7.5 fps capture for the high-speed shutter mode and 30 fps capture for the low-speed shutter mode.

The advantages of the high-speed shutter mode are as follows:

- There is, in principle, no cross talk between the left-eye and right-eye images.
- Because only one camera is used, neither frame synchronization nor color adjustment is necessary.
- Bright and clear images can be obtained.

The advantages of the low-speed shutter mode are as follows:

- There are no spatial discontinuities in the converted image.
- There is no need for extra image processing, such as image mosaicing.
- There is no need for accurate speed control of the rotator.
- Since two cameras are used, more information can be obtained.

The horizontal resolution of the converted image corresponds to a visual acuity of 1/20 at medium height. The degree of image blur due to the use of a planar prism sheet is roughly the same as that induced by the limitation of the camera resolution. Using linear polarizing films, the cross talk between the left and right-eye image is considerably reduced. Supplemental videos can be found on the Computer Society Digital Library at http://computer.org/tvcg/archives.htm.

6.1 Future Work

We are now working to improve the image quality by removing the spatial discontinuities in the high-speed shutter mode, reducing the cross talk between the left-eye and right-eye images in the low-speed shutter mode, and improving spatial resolution. Some of the means to improve the image quality are as follows:

- We can use cameras of higher resolution. By utilizing cameras of full high definition (1,920 $\times$ 1,080 pixels, e.g., HDR-FX1 from Sony Corp.), the horizontal and vertical resolutions are expected to be 1.5 times higher than those in the experiments mentioned above.
- The chromatic aberration due to the use of lenses can be cancelled with postimage processing.
- By utilizing a prism sheet with curvature or by bending a thin prism sheet, the image blur mentioned in Section 5.3 can be avoided and we can

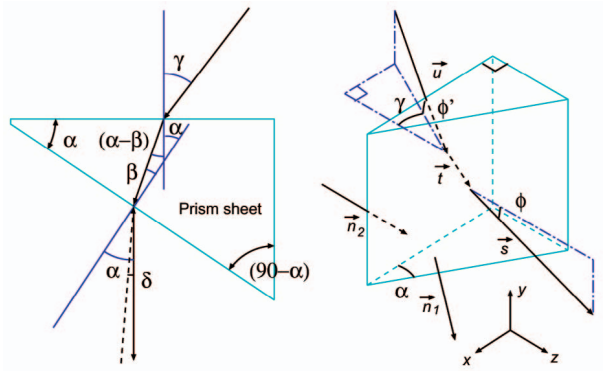

Fig. 29. Refraction by a prism sheet. Top view (left) and isometric view (right).

improve the spatial resolution of the generated image in the low-speed shutter mode.

- By increasing the number of prism sheets and by narrowing the width of them, we can expect the same effect as above. Furthermore, this time we can use a higher shutter speed than 1/30 seconds to avoid motion blur.

Moreover, with a flexible mechanism using a thin prism sheet with variable curvature, it is possible to construct an omnistereo camera with a variable parallax. Using a simple similar triangles argument, it is clear that a double-sized TORNADO has a doubled IPD, assuming a constant refraction angle of the prism sheet. Using omnistereo image sensors and omnistereo displays like TWISTER, one-to-many visual communications can be achieved in which stereoscopic images for arbitrary horizontal directions are reconstructed from images captured with a single omnistereo image sensor in a remote place and presented to multiple users according to the directions of their heads. Moreover, mutual communication can be achieved by utilizing a viewer-image extraction technique such as [19] and a sensor camouflage technique such as [20].

Appendix A

Calculation of the Angle of Refraction with a Prism Sheet

Let α represent the complementary angle of the vertex angle of each prism of the prism sheet. The light rays exit the prism (uneven) side, as shown in Fig. 29 (left). Since the image captured with TORNADO is a collection of light rays that reach a focus of the hyperboloidal mirror and these light rays exit perpendicularly from the prism side, the prism sheets are arranged with their prism sides facing inward.

The incident angle of the second refraction (β) is calculated as:

$$\beta = \arcsin\frac{1}{n\sin\alpha}, \tag{1}$$

where $n \simeq 1.5$ is the index of refraction of the prism sheet. Since the output angle of the first refraction is represented as $(\alpha - \beta)$, the incident angle (γ) is calculated as follows:

$$\gamma = \arcsin(n\sin\alpha - \beta). \tag{2}$$

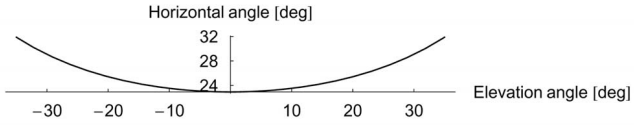

Fig. 30. The dependency of the horizontal component of the incident angle on the elevation angle.

In actual dimensions, $\alpha = 41$ degrees and, therefore, $\gamma = 22.9$ degrees. When the output angle is represented by $(\alpha - \delta)$ instead of α, β' is represented as follows:

$$\beta' = \arcsin\frac{1}{n\sin\alpha - \delta}, \tag{3}$$

therefore,

$$\gamma = \arcsin(n\sin\alpha - \beta'). \tag{4}$$

In actual dimensions, the rate of change in γ to the change in δ is approximately 0.88. From this calculation, we can estimate the error in the direction of incident light rays to the hyperboloidal mirror due to the use of a planar prism sheet. When δ varies ± 1 degree, the error in the direction of the light rays is estimated as 0.12 degrees at maximum.

Appendix B

The Dependency of the Horizontal Component of the Incident Angle on the Elevation Angle

Let $\vec{n}$ represent the normal of the interface between two media (M_1 and M_2), $\vec{v_1}$ the direction vector (whose norm is 1) of a light ray in M_1, $\vec{v_2}$ the direction vector of a light ray in M_2, and r the ratio of the index of refraction of M_1 to M_2. Snell's law gives the following relationship:

$$\vec{v_2} = \left[\sqrt{1 - r^2\{1 - (\vec{n}\cdot\vec{v_1})^2\}} - r(\vec{n}\cdot\vec{v_1})\right]\vec{n} + r\vec{v_1}. \tag{5}$$

In Fig. 29 (right), $\vec{s}$, $\vec{t}$, and $\vec{u}$ are the direction vectors of a light ray with an elevation angle ϕ, the light ray in the acryl, and the incident light ray, respectively. They are described as follows:

$$\vec{s} = (0, -\sin\phi, \cos\phi)^T, \tag{6}$$

$$\vec{t} = \left[\sqrt{1 - \frac{1 - (\vec{n_1}\cdot\vec{s})^2}{i^2}} - \frac{\vec{n_1}\cdot\vec{s}}{i}\right]\vec{n_1} + \frac{\vec{s}}{i}, \tag{7}$$

$$\vec{u} = \left[\sqrt{1 - i^2\{1 - (\vec{n_2}\cdot\vec{t})^2\}} - i(\vec{n_2}\cdot\vec{t})\right]\vec{n_2} + i\vec{t}, \tag{8}$$

where $\vec{n_1} = (\sin\alpha, 0, \cos\alpha)^T$ and $\vec{n_2} = (0, 0, 1)^T$. By substituting (7) for (8) and by using $\alpha = 41\text{deg}$, $i = 1.5$, $(\vec{u})$ is represented as follows:

$$\vec{u} = \begin{pmatrix} -0.495\cos\phi + 0.656\sigma \\ -\sin\phi \\ \sqrt{-0.538 + 0.510\cos^2\phi + 0.650\sigma\cos\phi} \end{pmatrix}, \tag{9}$$

where $\sigma = \sqrt{1.25 + 0.570\cos^2\phi}$. It is notable that the elevation angle ϕ' equals the elevation angle ϕ. Fig. 30

shows the dependency of the horizontal component of the incident angle (γ) on the elevation angle (ϕ'). The curve is used to correct the image geometry.

ACKNOWLEDGMENTS

Part of the research presented here was conducted under the CREST Telexistence Communication Systems Project, supported by JST (Japan Science and Technology Corporation).

REFERENCES

[1] S. Tachi, K. Tanie, K. Komoriya, and M. Kaneko, "Tele-Existence(I)-Design and Evaluation of a Visual Display with Sensation of Presence," *Proc. RoManSy '84 Fifth CISM-IFToMM Symp.,* pp. 245-254, 1984.
[2] S.E. Chen, "Quicktime VR—An Image-Based Approach to Virtual Environment Navigation," *Proc. SIGGRAPH '95,* pp. 29-38, 1995.
[3] C. Cruz-Neira, D.J. Sandin, and T.A. DeFanti, "Surround-Screen Projection-Based Virtual Reality: The Design and Implementation of the CAVE," *Proc. SIGGRAPH '93,* pp. 135-142, 1993.
[4] E.H. Adelson and J.R. Bergen, "The Plenoptic Function and the Elements of Early Vision," *Computation Models of Visual Processing,* M. Landy and J.A. Movshon eds. pp. 3-20, MIT Press, 1991.
[5] L. McMillan and G. Bishop, "Plenoptic Modeling: An Image-Based Rendering System," *Proc. SIGGRAPH '95,* pp. 39-46, 1995.
[6] S.J. Gortler, R. Grzeszczuk, R. Szeliski, and M.F. Cohen, "The Lumigraph," *Proc. SIGGRAPH '96,* pp. 43-54, 1996.
[7] M. Levoy and P. Hanrahan, "Light Field Rendering," *Proc. SIGGRAPH '96,* pp. 31-41, 1996.
[8] F. Okano, H. Hoshino, J. Arai, and I. Yuyama, "Real-Time Pickup Method for a Three-Dimensional Image Based on Integral Photography," *Applied Optics,* vol. 36, no. 7, pp. 1598-1603, 1997.
[9] T. Naemura, T. Yoshida, and H. Harashima, "3-D Computer Graphics Based on Integral Photography," *Optics Express,* vol. 8, no. 2, pp. 255-262, 2001.
[10] T. Naemura, M. Kaneko, and H. Harashima, "Multi-User Immersive Stereo," *Proc. IEEE Int'l Conf. Image Processing,* vol. I, pp. 903-907, 1998.
[11] S. Peleg and M. Ben-Ezra, "Stereo Panorama with a Single Camera," *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* vol. I, pp. 1395-1401, 1999.
[12] H.-Y. Shum and L.-W. He, "Rendering with Concentric Mosaics," *Proc. SIGGRAPH '99,* pp. 299-306, 1999.
[13] A. Simon, R.C. Smith, and R.R. Pawlicki, "Omnistereo for Panoramic Virtual Environment Display Systems," *Proc. IEEE Virtual Reality 2004,* pp. 67-279, 2004.
[14] K. Tanaka, J. Hayashi, M. Inami, and S. Tachi, "TWISTER: An Immersive Autostereoscopic Display," *Proc. IEEE Virtual Reality 2004,* pp. 59-278, 2004.
[15] N. Yokoya, K. Yamazawa, and H. Takemura, "Omnidirectional Stereo Image Sensor Using Compound Hyperboloidal Mirror," *Proc. Nat'l Conf. IEICE,* D-12-145, 1997.
[16] J. Shimamura, N. Yokoya, H. Takemura, and K. Yamazawa, "Construction of an Immersive Mixed Environment Using an Omnidirectional Stereo Image Sensor," *Proc. IEEE Workshop Omnidirectional Vision,* pp. 62-69, 2000.
[17] S. Peleg, M. Ben-Ezra, and Y. Pritch, "Omnistereo: Panoramic Stereo Imaging," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 23, pp. 279-290, 2001.
[18] K. Tanaka, J. Hayashi, T. Endo, and S. Tachi, "A Method for Panoramic Stereo Image Acquisition," *Proc. Int'l Conf. Artificial Reality and Telexistence (ICAT 2003),* pp. 257-262, 2003.
[19] Y. Kunita, M. Inami, T. Maeda, and S. Tachi, "Real-Time Rendering System of Moving Objects," *Proc. IEEE Workshop Multi-View Modeling and Analysis of Visual Scenes,* pp. 81-88, 1999.
[20] M. Inami, N. Kawakami, D. Sekiguchi, Y. Yanagida, T. Maeda, and S. Tachi, "Visuo-Haptic Display Using Head-Mounted Projector," *Proc. IEEE Virtual Reality 2000,* pp. 233-240, 2000.

Kenji Tanaka received the BE degree in mathematical engineering, the MS degree in information physics, and the PhD degree in advanced interdisciplinary studies from the University of Tokyo in 1995, 1997, and 2004, respectively. His research interests are in image processing, virtual reality devices, and panoramic stereoscopic sensors and displays. He is currently with Sony Corporation. He is a member of the IEEE, the IEEE Computer Society, the International Society for Optical Engineering (SPIE), the Society of Instrument and Control Engineers (SICE), and the Virtual Reality Society of Japan (VRSJ). Further information about his publications can be found at: http://www.star.t.u-tokyo.ac.jp/~tanaken.

Susumu Tachi received the BE, MS, and PhD degrees in mathematical engineering and information physics from the University of Tokyo in 1968, 1970, and 1973, respectively. He joined the Faculty of Engineering at the University of Tokyo in 1973 and, in 1975, he moved to the Mechanical Engineering Laboratory, Ministry of International Trade and Industry, where he served as the director of the Biorobotics Division. In 1989, he rejoined the University of Tokyo, where he is currently a professor in the Department of Information Physics and Computing. From 1979 to 1980, Dr. Tachi was a Japanese Government Award Senior Visiting Scientist at the Massachusetts Institute of Technology, Cambridge. Dr. Tachi currently leads the CREST Project on Telexistence Communication Systems (2000-). His present research covers robotics, augmented reality, and virtual reality, with special focus on mutual telexistence (TELESAR and TWISTER), real-time remote robotics (R-Cubed), augmented haptics, and retro-reflective projection technology (RPT), including optical camouflage. Professor Tachi is a founding director and a fellow of the Robotics Society of Japan (RSJ), a fellow of the Society of Instrument and Control Engineers (SICE), a fellow of the Japan Society of Mechanical Engineers (JSME), the founding president of the Virtual Reality Society of Japan (VRSJ), and a member of the IEEE. Since 1988, he has served as the chairman of the IMEKO (International Measurement Confederation) Technical Committee 17 on Measurement in Robotics. He also served as a cogeneral chair of ICAT2004.