TWISTER (Telexistence Wide-angle Immersive STEReoscope) is an immersive omnidirectional full-color autostereoscopic 3D booth, designed for a face-to-face telecommunication concept called “mutual telexistence,” where people at distant locations can communicate as if they were present in the same three-dimensional space. Each human user resides in a cylindrical booth, which displays live full-color 360-degree panoramic and autostereoscopic images to the user without the need for any eyewear; at the same time, the booth captures images of the user from every angle. With multiple booths, people at remote locations can interact with each other as if they were interacting face-to-face. This concept was proposed in 1996, and prototype models TWISTER I, II, III, IV, and V were constructed based on the proposed concept. TWISTER IV displays binocular stereo images using 36 display units, each consisting of two LED arrays for the left and right eyes and a parallax barrier that rotates at a high speed (1.7 rps). This system simultaneously captures 3D images of the observer by employing cameras installed on the rotating cylinder. Thus, the observers inside two TWISTERs can view 3D images of each other in real time. In this paper, the concept and principles of TWISTER for mutual telexistence is explained, and the prototype systems developed and possible applications of TWISTER are described.

1. Introduction

Face-to-face communication is an essential and inevitable component of a person’s daily life. However, in situations where people are far apart, a face-to-face telecommunication system proves useful. One of the most essential elements of such a system is the sense of presence. Immersion being one of the essential factors for the sense of presence, it is desirable that the visual display of such a system provide an immersive stereoscopic view. Two typical conventional immersive systems include an HMD (head mounted display) and CAVE (CAVE automatic virtual environment). In such conventional systems, however, the mechanisms for binocular stereopsis, such as headgears for HMD and shutter-glasses for CAVE, not
only leave the observer with a sense of discomfort but also hinder face-to-
face communication since they obscure an observer's face, particularly
around the eyes.

Autostereoscopic 3D displays are effective for face-to-face communication
as they enable the user to view three-dimensional images without the use of
any special headgear or eyeglasses. The most common autostereoscopic
3D displays employ lenticular lenses or parallax barriers. For such
displays, if a viewer positions his/her head in a certain manner, he or she
will perceive different images with each eye, which gives rise to a stereo
image. However, when lenticular lenses or parallax barriers were used, a
system that provides a 360-degree panoramic field of view could not be
realized. Integral imaging using fly's eye lenses is a promising technology.
Nevertheless, it has not been employed in any 360-degree integral imaging
autostereoscopic 3D display.

The conventional parallax barriers, lenticular lenses, or fly's eye lenses have
another limitation: it is considerably difficult to take a picture of a user's face
while he or she is observing 3D images. This limitation stems from the
fact that the screen is too densely packed with micro lenses and LCDs to
locate cameras in between. Since recording a user's face in real time at the
eye level is essential for face-to-face telecommunication, the cameras
should be placed at the level of a human user's eyes. However, such a set-
up leads to sparse arrangement of micro lenses, which adversely affects the
performance of the display.

In order to resolve these abovementioned problems, we propose
TWISTER-Telexistence Wide-angle Immersive STEReoscope-as a booth
for face-to-face telecommunication. By adopting the moving parallax
barrier method, TWISTER can display panoramic stereoscopic images
that can be observed without the use of any special eyewear. Moreover,
we can place cameras between the display units. By rotating the cameras
simultaneously, the booth can display stereoscopic images and capture the
figure of the user simultaneously. It is also possible to place cameras
outside the booth and view the user from outside. Once these images are
obtained from various viewpoints around the user, a new view from an
arbitrary viewpoint can be constructed using an image-based rendering
technique.

With several booths, every user in a remote place can view the three-
dimensional figures of other users working in real time in the mutual virtual
environment. In such a situation, the booth works as the medium between
the real space and the virtual communication space.

In this paper, the concept, principles, developed prototypes, and possible
applications of TWISTER are described.

2. Concept, Principle and Development

The first idea of mutual telexistence using cylindrical booths that can be
used both for a panoramic display and for a real-time camera system that
takes pictures of a human user from any direction was proposed in 1996 [1].
The constructed handmade experimental system, in which the functions of
the input device and display device were separated for the sake of simplicity,
demonstrated the feasibility of the idea [1]. Fundamental studies for the
realization of the two functions-input and display- were conducted,
following which the first handmade prototype system with the display
function, dubbed TWISTER I, was constructed in 2000, and the results
were reported in 2001 [2].
The CREST project on Telexistence Communication Systems, led by the author and funded by JST (Japan Science and Technology Agency), commenced in November 2000 and ended in March 2006 [3, 4]. We have applied the concept of TWISTER to a tangible real system that can be experienced.

TWISTER II was constructed on the basis of TWISTER I and reported in 2001 [5]. TWISTER II successfully realized the world's first omnidirectional autostereoscopic full color display. We demonstrated it worldwide at SIGGRAPH in 2002 using TWISTER III [6, 7]. TWISTER IV is a system that has both functions of display and video input [8]. It was constructed in 2004. An almost similar system called TWISTER V was constructed in 2005 and it is placed at the National Museum of Emerging Science and Innovation (Miraikan). Communication experiments are being conducted using TWISTER IV at the University of Tokyo and TWISTER V at Miraikan.

Figure 2 shows the concept of teleexistence in a mutual virtual environment using TWISTER. Each human user resides inside a booth that has a virtual cylindrical wall that can be used as a display of the virtual environment in which all users are projected as well as an input camera device to obtain the three-dimensional moving pictures of the user inside the booth. The booth is surrounded by (1) a circular mechanism that turns continuously at a relatively high speed carrying display units consisting of a pair of LED arrays and a parallax barrier and (2) video cameras.

Video cameras, which are arranged such that they cover the user completely, take pictures of the user from virtually every angle as they turn around the user inside the booth. Using this method, stereo information pertaining to the user can be obtained from virtually any angle. When the system is used by several users, each user's line of sight and relative position are measured and are used to select the appropriate picture of other users. The images of other persons thus selected are appropriately arranged three-dimensionally in a mutual environment and displayed to all the users.

Virtual transparent boards (alpha plane) are placed in a virtual environment, and the pictures taken from two positions (left and right) are texture mapped on the boards. The other parts of the scenery are not texture mapped, and they remain transparent. The corresponding points of the texture-mapped human figures on the virtual boards produce the three-dimensional figure of the user.

Since the parallax barrier obscures LED emission from the opposite side, different images are shown to the left and right eyes. At one moment, only images for the left eye and right eye are observed in areas L and R, respectively. Both images are observed in area B, while neither image is observed in area N. When the display units with barriers rotate around the user, these areas also rotate, as shown in Fig. 4. The condition of stereopsis is maintained as the display units rotate around the user. The angle of view of a stereoscopic display depends on the direction and position of the observer. If the head of the observer is fixed in one position, it exceeds 140° in an ideal condition. On the other hand, if the observer always faces the center of the region of interest of the image, the angle of view can be 360°.

With moving parallax barriers, the crosstalk between the left eye image and the right eye image is reduced to almost zero, which gives this system a powerful advantage over other stereoscopic vision systems. Moreover, the
barrier per se cannot be observed by the observer because it rotates at a speed of 600 degrees per second, whereas the human eye can only track a revolution of approximately 500 degrees per second. This assures a continuous 360-degree panoramic stereo view.

Two modes are used in rendering for TWISTER. In Mode I, we assume that we can track the direction in which the head of the user is facing and render the left and right images on a display cylinder perspective transformed from the left eye and right eye, respectively. In this mode, a precise image is obtained when the head is fixed in one position. The condition is kept if we render in real time according to the measured head direction.

In Mode II, rendering with concentric mosaics is used. In this mode, irrespective of the direction the user turns in, correct light information will be obtained at least from the front. The light information from other directions is just an approximation. However, it is not necessary to track the direction in which the head of the human user faces.

Figures 5 and 6 show the typical panoramic displayed images in TWISTER.

The stereo-video picture of a human user inside a booth is captured by 36 rotating cameras, and an image from an assigned angle is reproduced from the information captured by appropriate cameras. Figure 7 shows a schematic diagram of the system.

Fig. 7. Schematic Diagram of Human Stereo-video Picture Capturing System

Table 1 shows the specifications of TWISTER III and TWISTER IV. Incidentally, the specification for TWISTER V is the same as that of TWISTER IV. Minor modification of LED arrangements on a display unit was conducted for TWISTER V, and TWISTER V can provide LED display without a diffuser, which was used for TWISTER IV.

Table 1. Specifications of TWISTER III and TWISTER IV

<table>
<thead>
<tr>
<th>Specifications</th>
<th>TWISTER III</th>
<th>TWISTER IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Horizontal resolution [pixel/360°]</td>
<td>1920</td>
<td>3168</td>
</tr>
<tr>
<td>Vertical resolution [pixel/74°]</td>
<td>250</td>
<td>600</td>
</tr>
<tr>
<td>Radius of the display [mm]</td>
<td>800</td>
<td>1000</td>
</tr>
<tr>
<td>Height of the display [mm]</td>
<td>960</td>
<td>1200</td>
</tr>
<tr>
<td>Pixel pitch [mm]</td>
<td>3.75</td>
<td>2.00</td>
</tr>
<tr>
<td>Visual acuity</td>
<td>0.09</td>
<td>0.15</td>
</tr>
<tr>
<td>Frame rate [fps]</td>
<td>30</td>
<td>60</td>
</tr>
<tr>
<td>Luminance gradient RGB 8 bit x 3</td>
<td>RGB 10 bit x 3</td>
<td></td>
</tr>
<tr>
<td>Input format</td>
<td>NTSC(240 x 256) x 8ch</td>
<td>DVR(1600 x 1200) x 2ch</td>
</tr>
<tr>
<td>Rotation speed [fps]</td>
<td>1.2</td>
<td>1.2</td>
</tr>
<tr>
<td>Number of display unit [units]</td>
<td>30</td>
<td>36</td>
</tr>
<tr>
<td>Camera resolution</td>
<td>-</td>
<td>VGA</td>
</tr>
<tr>
<td>Radius of camera rotation circle [mm]</td>
<td>-</td>
<td>950</td>
</tr>
<tr>
<td>Camera angle of view [deg]</td>
<td>-</td>
<td>45</td>
</tr>
<tr>
<td>Camera frame rate [fps]</td>
<td>-</td>
<td>30</td>
</tr>
<tr>
<td>Camera output format</td>
<td>-</td>
<td>IEEE1394 x 2ch</td>
</tr>
</tbody>
</table>

Although it was proved that the proposed method can capture a human user's stereo-video picture from an arbitrary direction employing rotating cameras inside a booth [8] using TWISTER IV and V, it became apparent that the use of this method is impractical at present because it requires (1) an expensive high-speed large-capacity rotary joint and (2) precise alignment and calibration of video cameras.

Thus, we propose an alternative method of image capturing from an arbitrary direction. As clearly seen from Figures 8 and 9, one of the salient characteristics of TWISTER is that the human user inside a booth can be seen from outside of the booth.
This implies that the human user’s picture can be taken from any direction from outside of the booth. Figure 10 shows a schematic diagram of the proposed method. A concentric rail is arranged outside of the booth, on which a movable stereo camera is set. The camera is controlled according to the angle from which the human picture should be taken. Stereo cameras can be used when several people observe the user inside. When it is necessary to cross two cameras, they can virtually be crossed by changing the roll of both cameras.

Consider the situation shown in Figure 11. Each of the three human users is inside a TWISTER booth. Each booth has two stereo cameras. Although the human users are far apart, they can virtually meet together in a virtual environment using the TWISTER booths.

First, 360° 3D virtual environment is displayed on each of the booths. Next, the two stereo cameras of each booth capture the pictures of the user inside TWISTER from the angles calculated based on the relative position and orientation of the three users in the virtual environment.

Then, the captured images are transmitted using communication links to the other TWISTERs. At each booth, the human figures of other users are placed in the virtual environment according to the relative position and orientation. Thus, the users have the feeling as if they were near at hand in the virtual environment.

Mutual telexistence is one of the most important technologies for the realization of networked telexistence, because users “telexisting” in a robot must know whom they are working with over the network. Figure 12 shows how TWISTER is used as a cockpit for mutual telexistence.

Human user A can observe remote environment B using omnistereo camera [9] on a surrogate robot A’. This provides A the panoramic stereo view of the remote environment displayed on TWISTER booth. A’ is controlled by A by using the telexistence master-slave control method [3, 4].
Cameras B' and C' mounted on the booth are controlled by the position and orientation of Humans B and C relative to A', respectively. B and C can observe A projected on A' by using HMP (Head Mounted Projector). A' is covered with a retro-reflective material, which makes it possible to project both images on the same robot and still both the images are separately observed by B and C.

This method of mutual telexistence using RPT (Retro-reflective Projection Technology) was proposed and proved to be quite effective [11].

Figure 13 shows an example of the projection of a human figure on a robot.

TWISTER will be widely used for several applications in 3C’s and 3E’s, i.e., Communication, Control, Creation, Education, Entertainment and Elucidation.

3. Conclusion
Technology based on the concept of networked telexistence enables users to meet and talk as if they share the same space and time, even if they are far apart from each other. This is the goal of the development of mutual telexistence communication systems, as a natural progression from telephone to telexistence-videophone. TWISTER (Telexistence Wide-angle Immersive STERoscope) was developed and its efficacy was demonstrated and proved. It will be widely used as a telexistence booth for communication and control with a sensation of presence.

4. Acknowledgement
This study on TWISTER has been conducted in collaboration with many of the former and present students and staff members of Tachi Laboratory of the University of Tokyo, including Yutaka Kunita, Kenji Tanaka, Junya Hayashi, Masaaki Koyanagi, Hiroki Yokoyama, Naoko Ogawa, Atsushi Sakuma, Taro Maeda, Yasuyuki Yanagida, Masahiko Inami, Naoki Kawakami, Dairoku Sekiguchi, Ichiro Kawabuchi and Hideaki Nii.

5. Reference