NAIST-IS-TR 97015
ISSN 0919-9527

Telepresence by Real-time View-dependent Image Generation from Omnidirectional Video Streams

Yoshio Onoe, Kazumasa Yamazawa, Haruo Takemura
and Naokazu Yokoya [1]

August 1997

ABSTRACT

This paper describes a new approach to telepresence which realizes virtual tours into a visualized dynamic real world without significant time delay. We propose a novel concept of *instantour* which enables us to instantly look around a visualized space of a dynamic real world. The *instantour* is realized by the following two steps: (1) video-rate omnidirectional image acquisition and (2) real-time view-dependent perspective image generation from an omnidirectional video stream. The proposed technique is applicable to real-time telepresence in the situtation where the real world to be seen is far from an observation site, because the time delay from the change of viewing direction to the change of displayed image does not depend on the actual distance between both sites. Moreover, multiple users can look around from a single viewpoint in a visualized dynamic real world in different directions at the same time. The proposed technique is also useful for another type of telepresence which uses recorded omnidirectional video streams. We have developed a prototype of virtual *instantour* system and have successfully experimented with real dynamic scenes.

Keywords: telepresence, omnidirectional imaging, HyperOmni Vision, vision-based graphics, visualized dynamic real world, virtual reality, virtual tour, video surveillance and monitoring

---

[1] Graduate School of Information Science, Nara Institute of Science and Technology
8916-5 Takayama, Ikoma, Nara 630-01, Japan
E-mail: {yoshio-o, yamazawa, takemura, yokoya}@is.aist-nara.ac.jp

1 Introduction

Virtual reality systems can be classified into two categories: (1) 3D computer graphics-based system [7] and (2) image-based system [3, 4, 12]. The first approach fully depends upon real-time 3D computer graphics techniques for modeling and rendering virtual environments, while the second uses multiple-view, panoramic or omnidirectional images for generating virtual environments. Although the CG-based approach provides the high interactivity in general, the approach suffers from expensive cost in both modeling and rendering, which leads to low quality of visualized scene and rendering for maintaining real-time computation. The image-based approach can overcome these difficulties and moreover makes it possible to use images of real world scenes which usually have more rich details than those generated by using CG techniques [4]. However, conventional image-based systems face another difficulty of extra works in obtaining panoramic or omnidirectional images; that is, the time delay in generating such wide-angle images can not be ignored and makes it difficult to treat dynamic real worlds. For example, panoramic images are often created by camera rotation [11, 20, 28] or by integration of overlapping photographs taken with a regular camera [6]. Such processes are not suitable for acquiring image sequences of dynamic real world scenes.

The image-based approach to constructing virtual environments is closely related to telepresence [9], which aims at providing a user the natural feeling of existing in the distance. In most conventional telepresence systems, active cameras such as panning and tilting cameras are utilized to immersively present a user view-dependent images of a scene in the distance. This often suffers from the time delay from the change of viewing direction to the change of displayed image. The time delay is mainly caused by both the communication between an observation site and an observed world and the control of cameras so as to follow the user's viewing direction. Especially, the former factor depends on the actual distance between the observation and observed sites.

Our objective is to develop a methodology for realizing virtual tours into a visualized dynamic real world without significant time delay. This could achieve a novel concept of *instantour* which enables us to look around a visualized space of a dynamic real world instantly. In other words, we want to develop an instant tele-tour or telepresence system. The *instantour* requires video-rate omnidirectional imaging of dynamic real scenes and view-dependent image display which simulates camera motions such as panning, tilting, rolling and zooming [17, 18]. This paper describes a new image-based approach to virtual environment tour which meets the above objective. Our approach is based upon the integration of an omnidirectional image sensor for real-time omnidirectional imaging and an image warping technique [21] for real-time view-dependent image generation. This provides a novel approach to real-time telepresence and is applicable to the situtation where the real world to be seen is far from an observation site, because the time delay from the change of viewing direction to the change of displayed image does not depend on the actual distance between both sites. Thus this potentially enables us to realize the real-time telepresence in the long distance. For example, we could reasonably look around the Mars surface if the omnidirectional image sensor would be landed on the Mars and an image communication link would be maintained between the earth and the Mars. It should be noted that the proposed technique is also useful for another type of telepresence which uses recorded omnidirectional video streams.

This paper is structured as follows. First we briefly describe an omnidirectional image sensor called HyperOmni Vision [24, 25] which can acquire an omnidirectional image of a real scene at video-rate. Geometric transformation of an omnidirectional image is then discussed; that is, an arbitrary single-view perspective image and a cylindrical panoramic image are computed from a circular omnidirectional image taken by the HyperOmni Vision. We finally present a prototype of image-based virtual tour system which realizes the *instantour* of a visualized real scene using the HyperOmni Vision and image warping techniques, as well as some experimental results of computing looking-around image sequences from both minute-to-minute and recorded omnidirectional video streams.

2 Video-rate Omnidirectional Imaging: HyperOmni Vision

Video-rate acquisition of an omnidirectional image of a dynamic real scene requires specialized cameras. A variety of sensors have been developed to acquire omnidirectional visual information on a 3D environment [5, 15, 23]. Optics of such sensors use a fish-eye lens [1, 13], a conic mirror [22], a spherical mirror [8], a hyperboloidal mirror [24], or a paraboloidal mirror [18]. We employ the HyperOmni Vision using a hyperboloidal mirror which was originally proposed as an omnidirectional imaging sensor for robot navigation. In the following we briefly describe the geometry and characteristics of the sensor.

2.1 Overview of HyperOmni Vision

The HyperOmni Vision is mainly composed of a hyperboloidal mirror and a single CCD camera as illustrated in Fig. 1. The hyperboloidal mirror is constructed of a hyperboloid of two sheets of revolution, which has two focal points (O_M and O_C). Note that the axis of the camera is aligned with that of the hyperboloidal mirror and the camera center is fixed at one of the focal points (O_C) of the mirror. The hyperboloidal mirror yields a 360-degree circular omnidirectional image of a 3D environment around its axis. Given a world coordinate (X, Y, Z) and an image coordinate (x, y) as shown in Fig. 1, the shape of hyperboloidal mirror is implicitly represented as:

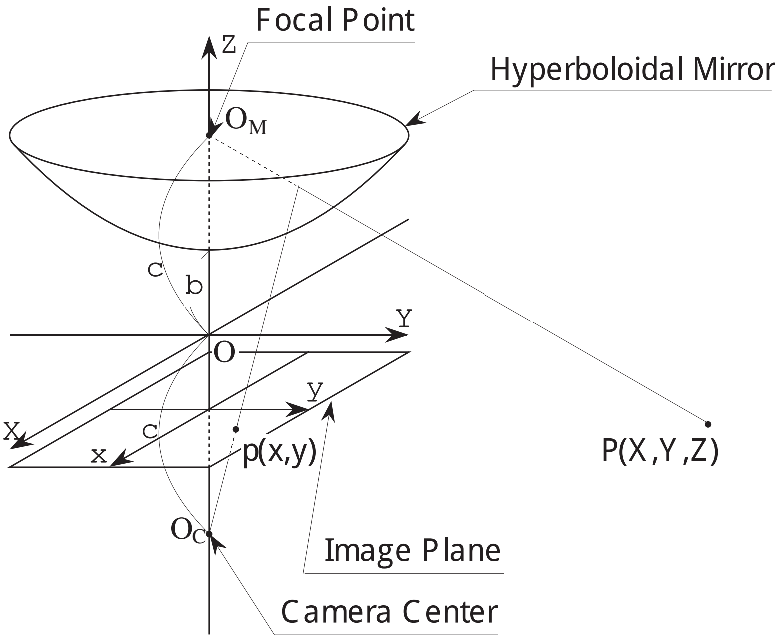


Figure 1: Geometry of HyperOmni Vision.

$$\frac{X^2+Y^2}{a^2} - \frac{Z^2}{b^2} = -1 \quad (Z > 0). \tag{1}$$

The inner focal point O_M of the mirror is at $(0,0,c)$ and the outer focal point (camera lens center) O_C is at $(0,0,-c)$ in the world coordinate, where $c = \sqrt{a^2+b^2}$.

A newly developed prototype of the HyperOmni Vision is shown in Fig. 2 and its parameters are summarized in Table 1. An example of obtained omnidirectional image of a laboratory scene is shown in Fig. 3. The CCD resolution is 768 × 494 pixels and the angular altitude range of the sensor is between +20 degrees and -90 degrees from a horizontal plane. It can be observed in Fig. 3 that the HyperOmni Vision provides a circular omnidirectional view of the surroundings including the ground (floor) and objects on it around the sensor axis.

Table 1: Parameters of HyperOmni Vision.

a	42.1mm
b	42.7mm
c	59.9mm
diameter of mirror	150mm
focal length f	5.9 ~ 47.2mm

Figure 2: Prototype of HyperOmni Vision.

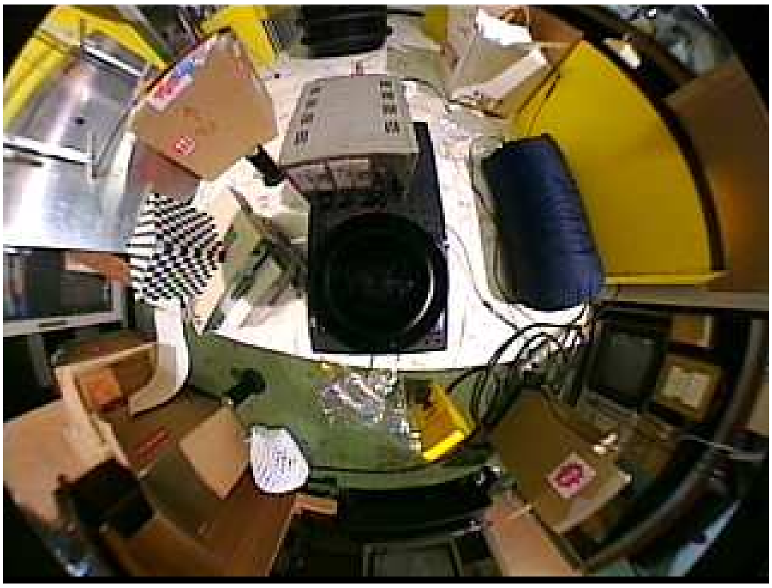


Figure 3: Omnidirectional image.

2.2 Characteristics of Omnidirectional Image

An omnidirectional image obtained by using the HyperOmni Vision is characterized by the mapping of a scene onto an image plane through a hyperboloidal mirror and a lens. The hyperboloidal mirror yields the image of a point in space as follows. As illustrated in Fig. 1, a ray going from the point $P(X,Y,Z)$ in space toward the focal point O_M of the hyperboloidal mirror is reflected by the mirror and passes through the other focal point (camera center) O_C intersecting an image plane at a point $p(x,y)$. Thus the image taken by HyperOmni Vision is equivalent to the central projection of a scene onto the hyperboloidal mirror surface and its single center of projection (viewpoint) is O_M. This projection may be called *hyperboloidal projection* and yields the following equations.

$$\begin{aligned}
Z &= \sqrt{X^2+Y^2}\tan\gamma_m + c, \\
\gamma_m &= \tan^{-1}\frac{(b^2+c^2)\sin\gamma_c - 2bc}{(b^2-c^2)\cos\gamma_c}, \\
\gamma_c &= \tan^{-1}\frac{f}{\sqrt{x^2+y^2}}, \\
\theta &= \tan^{-1}\frac{Y}{X} = \tan^{-1}\frac{y}{x},
\end{aligned} \tag{2}$$

where γ_m denotes an angle representing the direction from the focal point O_M to the point $P(X,Y,Z)$ in space, γ_c does an angle representing the direction from the camera lens center O_C to the projected point $p(x,y)$ in the image plane, f is the focal length of camera lens (the distance between the point O_C and the image plane), and θ denotes the azimuth angle of $P(X,Y,Z)$ (see Fig. 4).

All points with the same azimuth in space appear on a radial line through the image center; in other words, the azimuth angle is invariant from changes in depth and height. The above equations can be rewritten as follows:

$$\begin{aligned}
x &= \frac{f(b^2-c^2)X}{(b^2+c^2)Z - 2bc\sqrt{X^2+Y^2+Z^2}}, \\
y &= \frac{f(b^2-c^2)Y}{(b^2+c^2)Z - 2bc\sqrt{X^2+Y^2+Z^2}},
\end{aligned} \tag{3}$$

which directly represents the projection of the point $P(X,Y,Z)$ in space to the point $p(x,y)$ in the image plane.

By the analysis of geometry, the hyperboloidal projection has been proven to have the advantage that a hyperboloidal projection image can be easily transformed to both a cylindrical panoramic image and a familiar perspective (distortion-free) image, where the origin of transformed coordinate is fixed at the focal point O_M [25]. This feature is important for

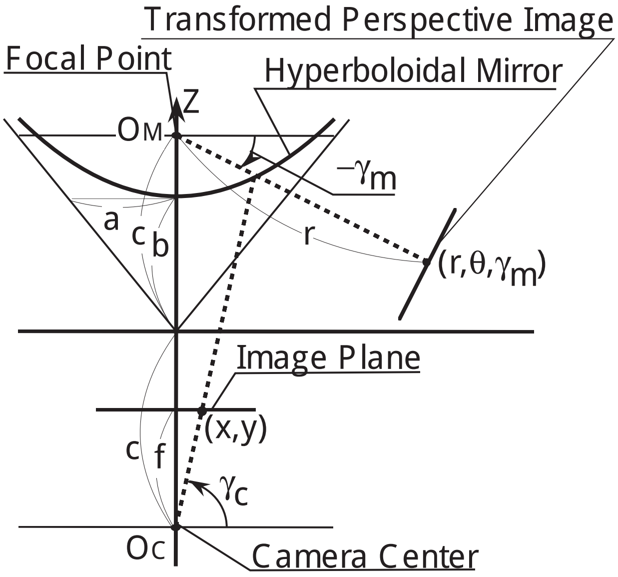


Figure 4: Inverse projection of an omnidirectional image to a virtual perspective image plane.

virtual reality and telepresence applications. Other omnidirectional sensors using conic or spherical mirrors do not have such a good feature. This is the main reason why we employ the HyperOmni Vision for the present application.

3 Geometric Transformation of Omnidirectional Images

In the following we describe the geometric transformation of an omnidirectional image to a familiar perspective image and a cylindrical panoramic image.

3.1 Generating Perspective Images

The common perspective image can be computed from the circular omnidirectional image. Figure 4 illustrates a two-dimensional geometry of the transformation of a circular omnidirectional image to a perspective image, where a pixel in a transformed image plane is represented by a polar coordinate (r, θ, γ_m) in space whose origin is located at the focal point O_M. Note that the azimuth angle θ is omitted in Fig. 4. The angle γ_c is calculated from the hyperboloidal projection as follows:

$$\gamma_c = \tan^{-1} \frac{(b^2 + c^2) \sin \gamma_m - 2bc}{(b^2 - c^2) \cos \gamma_m}, \tag{4}$$

where a, b and c are parameters of the hyperboloid (see Fig. 4). The corresponding coordinate (x, y) in the omnidirectional image is obtained by substituting (θ, γ_c) into the following equations:

Figure 5: Computed perspective image.

$$
\begin{aligned}
x &= \frac{f}{\tan \gamma_c} \cos \theta, \\
y &= \frac{f}{\tan \gamma_c} \sin \theta.
\end{aligned}
\tag{5}
$$

If each pixel location in a transformed image plane is represented by the world coordinate (X, Y, Z), the corresponding coordinate (x, y) in the source image can be obtained by applying Eq. (3). This is computationally more efficient than using Eq. (5) because Eq. (3) does not contain trigonometric functions.

By computing the above geometric transformation between (x, y) and (r, θ, γ_m), or (x, y) and (X, Y, Z), a perspective image viewing in an arbitrary direction from the fixed point O_M can be generated from an omnidirectional image. Some interpolation is needed in actual transformation. Figure 5 shows an example of computed perspective image which is obtained from a part of Fig. 3, where the bi-linear interpolation using four nearest neighbors is employed.

3.2 Generating Panoramic Images

Conventional cylindrical panoramic images can also be easily computed from the circular omnidirectional image by using equations similar to those given in the previous section, where a transformed perspective image plane should be replaced by a cylindrical image surface. Figure 6 shows an example of a cylindrical panoramic image computed from the omnidirectional image in Fig. 3.

Figure 6: Computed cylindrical panoramic image.

4 Real-time Tour into a Visualized Dynamic Real World

Now we describe a prototype of image-based interactive virtual tour system using the HyperOmni Vision. The interactivity here means the change of viewing direction and can be actually achieved by using various interactive devices such as a mouse, joystick and 3D positioning sensors. Among those, we assume to use a magnetic head-tracking facility attached on a head mounted display in the following. In our system, camera motions such as panning, tilting and rolling are simulated by generating view-dependent images from omnidirectional images.

4.1 Algorithm

A common perspective image is generated by clipping a part of circular omnidirectional image and is displayed on a head mounted display. Although such a perspective transformation can be achieved by applying the transformation described in Subsection **3.1** to all the pixels in a destination image, we actually compute the exact transformation only for sparsely spaced grid points for maintaining real-time computation.

The flow of view-dependent image generation from omnidirectional images is given in the followings.

1. At the beginning, uniformly spaced grids are simply generated on a rectangular destination image (see Fig. 7).
2. User's head position and viewing direction are measured by a magnetic head tracker attached with a head mounted display. Let the viewing direction be denoted by a rotation matrix R.

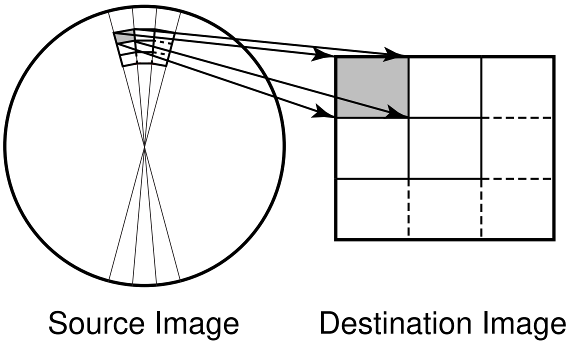


Figure 7: Transformation of omnidirectional image to perspective image.

3. According to the above measurement, the corresponding viewing area in a source omnidirectional image is determined and then clipped (see Fig. 7). Note that the user's head position (viewpoint) is assumed to locate at the inner focal point O_M of the HyperOmni Vision. Using the rotation matrix R, each grid point coordinate (u, v) in the destination image is represented in the world coordinate (X, Y, Z) as follows:

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = R \begin{pmatrix} u \\ v \\ f_h \end{pmatrix}, \tag{6}$$

   where f_h is the focal length of a virtual perspective camera which is equivalent to the distance from the viewpoint O_M to the destination image plane in space. The grid points which correspond to those in the destination image are also determined in the source image by applying the transformation described in Subsection **3.1**.

4. The destination image is computed from the clipped source image and grid points by applying an image warping technique with bi-linear interpolation [16]; that is, each quadrilateral in the source image is linearly warped to the corresponding rectangular area in the destination image. Note that the computed destination image is not exactly a perspective image but a *pseudo-perspective* image, and thus is geometrically distorted in some degree.

5. The completed destination image is finally displayed on the head mounted display via a down converter.

The above steps are iterated according to continuous measurements of user's view direction for providing an image sequence of virtual tour of a dynamic real scene. It should be noted that the computation time and the quality (the degree of geometrical distortion) of images computed by the algorithm above depend on the number of grids initially generated

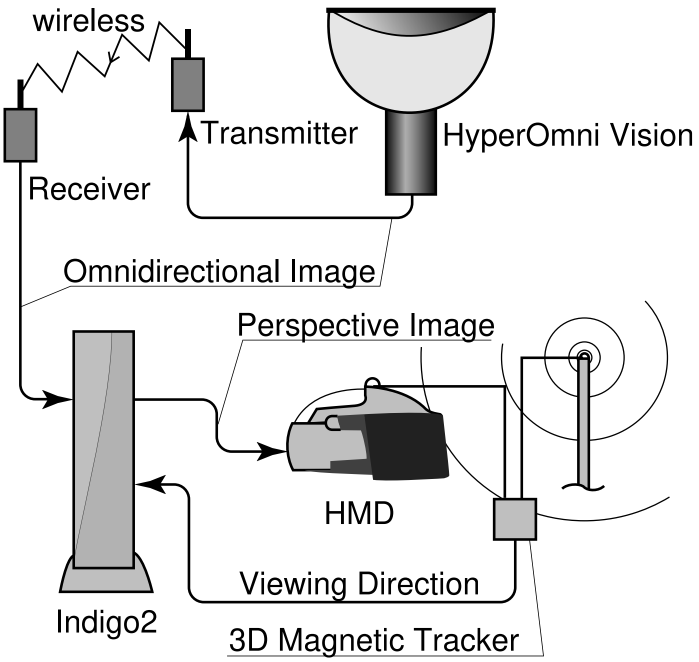


Figure 8: Hardware configuration of instantour system prototype.

in the destination image in Step 1 of the algorithm. This problem will further be discussed in later.

4.2 Implementation

A prototype of interactive virtual tour system for realizing the *instantour* in a real scene has been developed [17]. The system configuration is illustrated in Fig. 8; the system is composed of a standard workstation (Silicon Graphics Indigo2 / Maximum Impact (R4400, 250MHz)) with a video board (Impact Video), HyperOmni Vision ver.2, 3D magnetic tracker (Polhemus 3SPACE FASTRAK), wireless video communication facilities (Premier Wireless CS-200) and head mounted display (Olympus Mediamask). The 3D magnetic head tracker is attached with the head mounted display. The HyperOmni Vision can be mounted on a mobile facility such as a mobile robot and a car. Pseudo-perspective view generation described above is achieved by software with a graphics accelerator on the workstation.

5 Experiments

We have carried out two kinds of experiments: one is with indoor dynamic environments using the prototype system in Fig. 8, and the other is with recorded omnidirectional video streams of outdoor dynamic scenes. In the first experiment, dynamic omnidirectional images are transmitted from the HyperOmni Vision to the workstation via wireless video communication. In the second experiment, play-backed video streams are inputted to the workstation. In both experiments, view-dependent images are computed and are presented to a user in real-time in the same way.

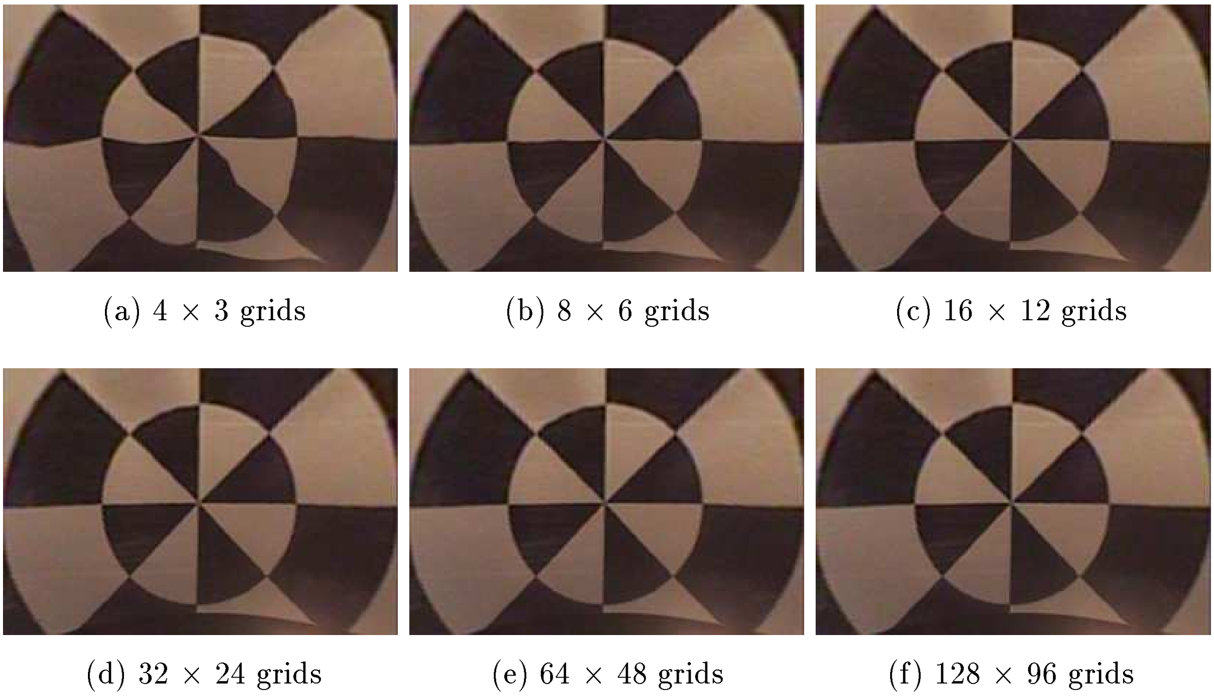
Prior to these experiments we have examined to determine the optimal number of grids for generating view-dependent images. The number of grids are experimentally analyzed in two aspects: (1) video-rate view generation and (2) unperceptible geometrical distortion of generated image.

5.1 Preliminary Experiments

As mentioned earlier, the computation time and the quality of images computed depend on the number of grids that are initially generated in the rectangular destination image. Our requirement is to reduce the computation time less than a time unit of video processing (1/30 second) and to generate high quality images so that users can not perceive geometrical distortions. In order to decide the optimal number of grids, we have preliminarily experimented with several different numbers of grid points for measuring the computation time and subjectively evaluating the degree of geometrical distortion in computed pseudo-perspective images.

Figure 9 shows examples of generated pseudo-perspective images of the size of 720 × 486 pixels with different numbers of grids, which are computed from an omnidirectional image containing a test pattern. In the subjective evaluation of image quality, straight line segments and circular arcs are mainly used as cues to examine geometrical distortions. As can be seen in Figs. 9(a) and (b), the geometrical distortion can be clearly perceived in the images computed from the small number of grids, 4 × 3 and 8 × 6, respectively. From experiments including Fig. 9, we have concluded that the number of grid points must be greater than 8 × 6 for providing images of unperceptible geometrical distortion. Figure 10 illustrates the image computation time and the image refreshing interval with respect to the number of initially generated grids. The image computation time means the measured time in determining grid point coordinates in the source image plus the theoretical time in executing the image warping to complete a pseudo-perspective destination image for each view, while the image refreshing interval implies the time interval between consecutive two views are displayed. This experimental result tells us that the number of grid points must be less than 64 × 48 for real-time image generation. Consequently the reasonable number of grids is between 16 × 12 and 32 × 24 for the present system configuration.

We note that the measurement of viewing direction, view-dependent image computation and image display are performed asynchronously with one another in the present prototype system. The viewing direction is measured every 8.3msec and the image display is refreshed



(a) 4 × 3 grids (b) 8 × 6 grids (c) 16 × 12 grids

(d) 32 × 24 grids (e) 64 × 48 grids (f) 128 × 96 grids

Figure 9: Pseudo-perspective images computed with different number of grids.

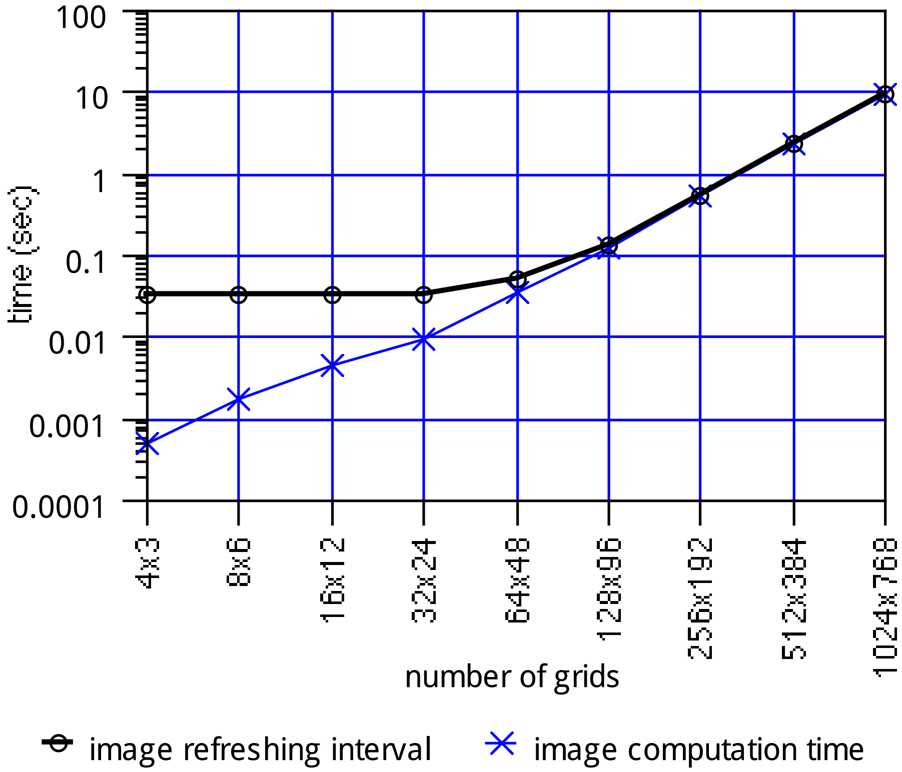


Figure 10: Image computation time and image refreshing interval with respect to number of grids.

every 33.3msec. Using 16 × 12 grids, the time delay from user's head motion to corresponding image display onto the head mounted display varies in time from 51.5msec to 93.1msec, where a computed pseudo-perspective image has the size of 720 × 486 pixels.

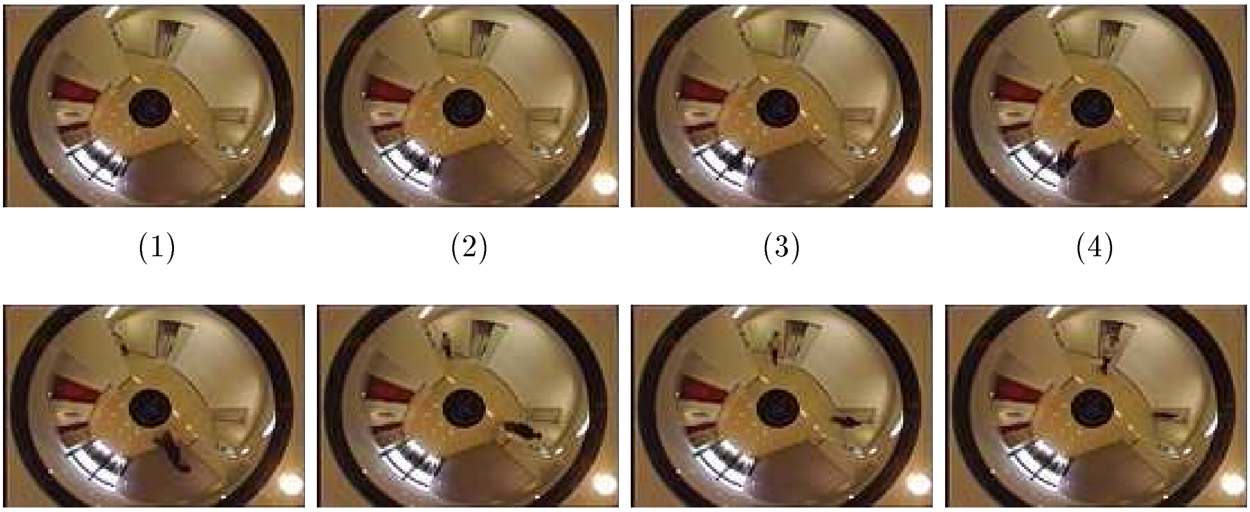
5.2 Indoor Scene

Now we present an application of the proposed technique to real-time telepresence. Figure 11 shows a sampled *instantour* sequence of virtually looking around a real indoor scene which contains a walking person, as well as the corresponding sequence of input omnidirectional images taken in the period of 6.5sec. In this experiment, view-dependent images are computed using 32 × 24 grids. It can be seen that two persons appear in the scene in the period and the view of a user of the system is tracking one of two walkers. This provides a user the natural feeling of looking around a real dynamic 3D environment at a fixed view point. It should again be noted that the time delay does not depend on the actual distance from the user (viewer) to the world to be seen, while conventional telepresence systems based on active regular cameras suffer from the significant time delay caused by the communication between both two sites. The experiment above exhibits the potential of applying the proposed technique to instant tour, tele-operation, video surveillance and monitoring (VSAM) [10, 15] and so on.

5.3 Outdoor Scene

Figure 12 shows experimental results of interactively viewing an outdoor scene from a recorded omnidirectional video stream. In this experiment, omnidirectional video is obtained in advance from the HyperOmni Vision mounted on the roof of a car (see Fig. 13) which is manually driven on a public road in the Keihanna Science City area located in Nara and Kyoto. A video stream is recorded by a standard Hi-8 video recorder. Note that the circular omnidirectional images contain 360-degree views around the moving car as well as the camera and the car roof in the center of each image. The computed pseudo-perspective video stream in Fig. 12 shows an image sequence of looking around the scene at a crossing on the road. This provides a user the feeling of looking around a real dynamic world from moving viewpoints.

The experiment here exhibits another possibility of the proposed technique of telepresence using recorded omnidirectional video streams. This could lead to promising applications such as virtual sightseeing and interactive video. We are planning to acquire omnidirectional video streams of traveling along a row of stores and houses on a street in Nara City, an ancient capital of Japan. This will also contribute to visual archiving and appreciation of historic scenes.



(1) (2) (3) (4)

(5) (6) (7) (8)

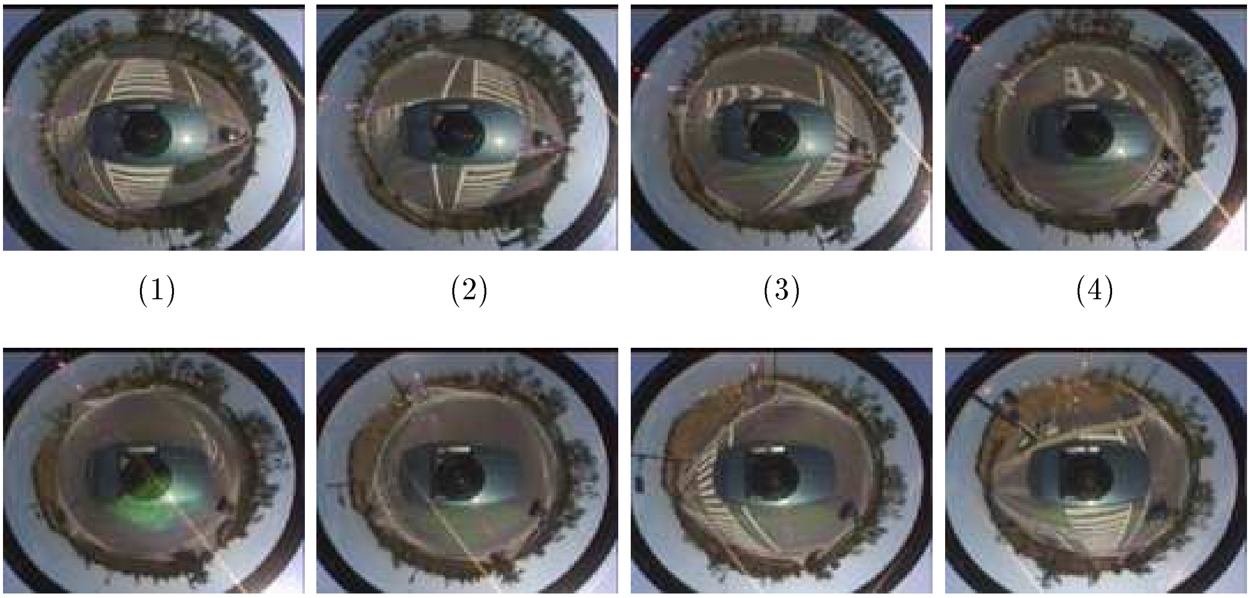
(a) Omnidirectional images in time sequence.

(1) (2) (3) (4)

(5) (6) (7) (8)

(b) Generated looking-around sequence of images.

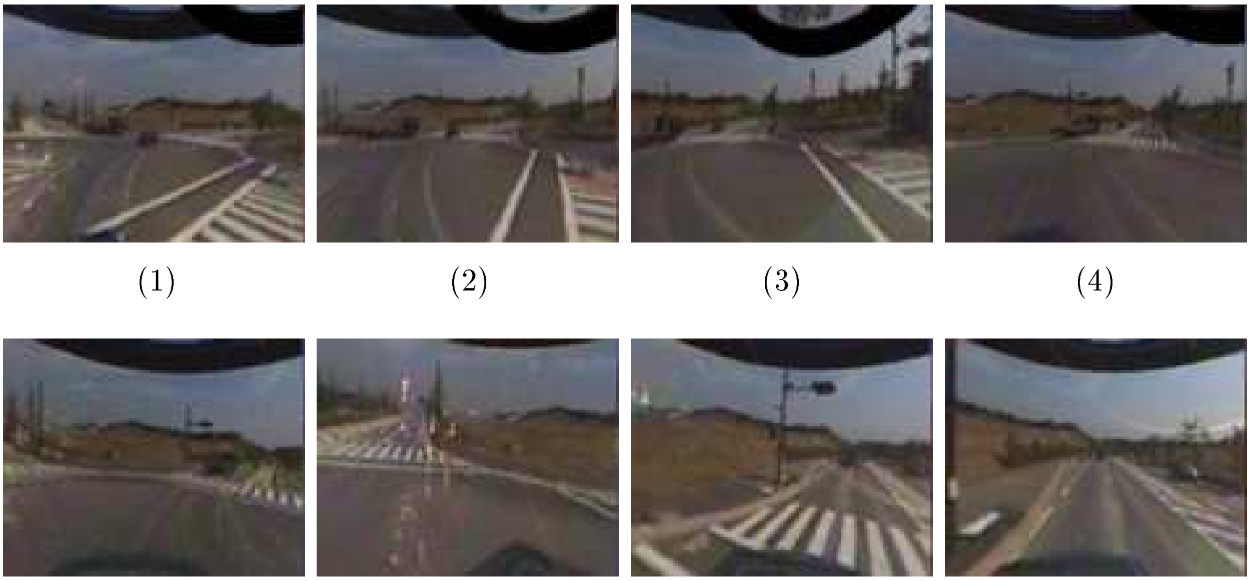
Figure 11: Experimental result with a real indoor scene. The input omnidirectional video stream was obtained from HyperOmni Vision located at a fixed position.



(1) (2) (3) (4)

(5) (6) (7) (8)

(a) Omnidirectional images in time sequence.



(1) (2) (3) (4)

(5) (6) (7) (8)

(b) Generated looking-around sequence of images.

Figure 12: Experimental result with a real outdoor scene. The input omnidirectional video stream was obtained from HyperOmni Vision mounted on a moving car.

Figure 13: HyperOmni Vision mounted on a car.

5.4 Discussions

Experimental results including those in Figs. 11 and 12 have exhibited the satisfactory performance of the proposed virtual tour system. The real-time telepresence system enables us to tour into a visualized space of a dynamic real world in real-time. In other words, we can virtually observe a "mirrored" world which varies in concurrence with the corresponding real 3D environment. Telepresence with recorded omnidirectional video streams leads to virtual sightseeing and interactive video applications.

It has been experimentally proven that the prototype system provides us the natural feeling of looking around a real dynamic world. However, there still remain some problems; for example, the quality of images displayed on a head mounted display is low, especially in an area of scene which is projected around the center of the hyperboloidal mirror of omnidirectional image sensor. This comes from the intrinsic property of the HyperOmni Vision. This could be resolved by the development of a high-resolution CCD camera or a high-resolution omnidirectional image sensor based on different optics [14, 26]. Another problem is the blurring in some parts of the image, which is caused by the varying depth of field on the curved mirror surface of HyperOmni Vision. This could be relaxed to some extent by adjusting the lens aperture. One possibility of improving the sensor is stereoscopic omnidirectional imaging [26, 27], which will provides the capability of stereoscopic viewing of a visualized world. These should further be investigated.

Although the experiments above are carried out in a single user mode, the system can be easily extended to a multiple user mode by independently computing multiple different views from a single omnidirectional video stream. The multiple user mode system could enable multiple users to look around a scene in different directions at the same time. This feature may further be exploited in some telepresence applications.

6 Conclusions

We have proposed a new virtual tour system by which we can instantly look around a visualized world of a real dynamic scene. This system is characterized by real-time omnidirectional image acquisition and real-time generation of view-dependent images from dynamic omnidirectional images. Experiments have shown the feasibility of the prototype system. The proposed approach has the following advantages over conventional telepresence systems based on active regular cameras: (1) the time delay from viewers' head motion to corresponding image display does not depend on the actual distance between viewers and the real world to be seen; (2) multiple users can look around in different directions at the same time.

The future work includes: (1) extension of omnidirectional imaging for acquiring stereoscopic information on a 3D environment [2, 19, 26, 27]; (2) development of a new high-resolution omnidirectional image sensor using pyramidal mirrors and multiple cameras [14, 26]; (3) presentation of (stereophonic) auditory information synchronized with the visual scenes. These could drastically improve the reality of telepresence in the future.

Acknowledgments

This work was supported in part by the Telecommunications Advancement Organization of Japan and the Telecommunications Advancement Foundation, and also by Grant-in-Aid for Scientific Research under Grant No. 09555127 from the Ministry of Education, Science, Sports and Culture. Kazumasa Yamazawa would like to thank Yasushi Yagi and Masahiko Yachida of Osaka University for their guidance in the work of developing omnidirectional image sensors.

References

[1] Z.L. Cao, S.J. Oh and E.L. Hall: “Dynamic omnidirectional vision for mobile robots”, *Journal of Robotic Systems*, Vol.3, No.1, pp.5-17, 1986.

[2] A. Chaen, K. Yamazawa, N. Yokoya and H. Takemura: “Omnidirectional stereo vision using HyperOmni Vision”, IEICE Tech. Report, IE96-122, February 1997 (in Japanese).

[3] S.E. Chen and L. Williams: "View interpolation for image synthesis", *Proc. SIGGRAPH'93*, pp.279-288, Anaheim, CA, August 1993.

[4] S.E. Chen: "QuickTime VR — An image-based approach to virtual environment navigation", *Proc. SIGGRAPH'95*, pp.29-38, Los Angeles, CA, August 1995.

[5] K. Der, A. Basu and J. Reyda: "Interactive viewing of panoramic images", *Proc. Vision Interface '97*, pp.162-169, Kelowna, British Columbia, May 1997.

[6] N. Greene: "Creating raster omnimax images from multiple perspective views using the elliptical weighted average filter", *IEEE Computer Graphics and Applications*, Vol.6, No.6, pp.21-27, 1986.

[7] L.F. Hodges, B.A. Watson, G.D. Kessler, B.O. Rothbaum and D. Opdyke: "Virtually conquering fear of flying", *IEEE Computer Graphics and Applications*, Vol.16, No.6, pp.42-49, November 1996.

[8] J. Hong, X. Tan, B. Pinette, R. Weiss and E.M. Riseman: "Image-based homing", *Proc. 1991 IEEE Int. Conf. on Robotics and Automation*, pp.620-625, Sacramento, CA, April 1991.

[9] Special Issue on Immersive Telepresence, *IEEE MultiMedia*, Vol.4, No.1, pp.17-56, January 1997.

[10] T. Kanade and K. Ikeuchi: "Image understanding research at CMU", *Proc. DARPA Image Understanding Workshop*, pp.95-109, New Orleans, May 1997.

[11] A. Krishnan and N. Ahuja: "Panoramic image acquisition", *Proc. 1996 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, pp.379-384, San Francisco, CA, June 1996.

[12] G. Miller, E. Hoffert, S.E. Chen, E. Patterson, D. Blackketter, S. Rubin, S.A. Applin, D. Yim and J. Hanan: "The virtual museum: Interactive 3D navigation of a multimedia database", *Journal of Visualization and Computer Animation*, Vol.3, pp.183-197, 1992.

[13] K. Miyamoto: "Fish eye lens", *Journal of Optical Society of America*, Vol.54, No.8, pp.1060-1061, August 1964.

[14] V. Nalwa: "A true omnidirectional viewer", Tech. Report, Bell Laboratories, Holmdel, NJ, February 1996.

[15] S.K. Nayar and T.E. Boult: "Omnidirectional VSAM systems: PI report", *Proc. DARPA Image Understanding Workshop*, pp.55-61, New Orleans, May 1997.

[16] J. Neider, T. Davis and M. Woo: *OpenGL Programming Guide*, Addison Wesley, 1993.

[17] Y. Onoe, K. Yamazawa, H. Takemura and N. Yokoya: "Telepresence by real-time view-dependent image generation from omnidirectional images", IEICE Tech. Report, PRMU 97-20, May 1997 (in Japanese).

[18] V. Peri and S.K. Nayar: "Generation of perspective and panoramic video from omnidirectional video", *Proc. DARPA Image Understanding Workshop*, pp.243-245, New Orleans, May 1997.

[19] D. Southwell, A. Basu, M. Fiala and J. Reyda: "Panoramic stereo", *Proc. 13th Int. Conf. on Pattern Recognition*, Vol.I, pp.378-382, Vienna, Austria, August 1996.

[20] R. Szeliski: "Video mosaics for virtual environments", *IEEE Computer Graphics and Applications*, Vol.16, No.2, pp.22-30, March 1996.

[21] G. Wolberg: *Digital Image Warping*, IEEE Computer Society Press, Los Alamitos, CA, 1990.

[22] Y. Yagi, S. Kawato and S. Tsuji: "Real-time omnidirectional image sensor (COPIS) for vision-guided navigation", *IEEE Trans. on Robotics and Automation*, Vol.10, No.1, pp.11-22, February 1994.

[23] Y. Yagi: "Omnidirectional imaging", IEICE Tech. Report, PRMU 96-140, January 1997 (in Japanese).

[24] K. Yamazawa, Y. Yagi and M. Yachida: "New real-time omnidirectional image sensor with hyperboloidal mirror", *Proc. 8th Scandinavian Conf. on Image Analysis*, Vol.2, pp.1381-1387, Tromso, Norway, May 1993.

[25] K. Yamazawa, Y. Yagi and M. Yachida: "Obstacle detection with omnidirectional image sensor HyperOmni Vision", *Proc. 1995 IEEE Int. Conf. on Robotics and Automation*, pp.1062-1067, Nagoya, Japan, May 1995.

[26] K. Yamazawa, H. Iwasa, H. Takemura and N. Yokoya: "High-resolution omnidirectional stereo imaging using pyramidal mirrors", *Proc. 1997 IEICE General Conf.*, No.D-12-146, March 1997 (in Japanese).

[27] N. Yokoya, K. Yamazawa and H. Takemura: "Omnidirectional stereo imaging sensor using a composite hyperboloidal mirror", *Proc. 1997 IEICE General Conf.*, No.D-12-145, March 1997 (in Japanese).

[28] J.Y. Zheng and S. Tsuji: "Panoramic representation of scenes for route understanding", *Proc. 10th Int. Conf. on Pattern Recognition*, Vol.I, pp.161-167, Atlantic City, NJ, June 1990.