Light Field Rendering with Omni-Directional Camera

Hiroshi Todoroki and Hideo Saito †

Department of Information and Computer Science, Keio University
3-14-1 Hiyoshi Kouhoku-ku Yokohama 223-8522, Japan

ABSTRACT

This paper presents an approach to capture visual appearance of a real environment such as an interior of a room. We propose the method for generating arbitrary viewpoint images by building light field with the omni-directional camera, which can capture the wide circumferences. Omni-directional camera used in this technique is a special camera with the hyperbolic mirror in the upper part of a camera, so that we can capture luminosity in the environment in the range of 360 degree of circumferences in one image. We apply the light field method, which is one technique of Image-Based-Rendering(IBR), for generating the arbitrary viewpoint images. The light field is a kind of the database that records the luminosity information in the object space. We employ the omni-directional camera for constructing the light field, so that we can collect many view direction images in the light field. Thus our method allows the user to explore the wide scene, that can acheive realistic representation of virtual enviroment. For demonstating the proposed method, we capture image sequence in our lab's interior environment with an omni-directional camera, and succesfully generate arbitray viewpoint images for virual tour of the environment.

Keywords: mixed reality, omni-directional camera, light field rendering, Image-Based-Rendering

1. INTRODUCTION

Recently, virtualizing real environment is one of the most important subjects, in application of remote control of the robot and virtual reality. Especially, an increasing number of attempts have been reported for virtualization of large real environments.[5,6] In this paper, we propose a method for generating the arbitrary viewpoint images by building light field with omni-directional camera, which can capture the 360 circumferences.

The technique for generation of arbitrary viewpoint images is divided into two categories. One is Model-Based-Rendering (MBR), which synthesizes virtual viewpoint images from 3D models of the objects. It becomes difficult to create a 3-dimensional model if the object environment is complicated. In such a case, the arbitrary viewpoint images with exact geometrical relation can't be synthesized. The other is Image-Based-Rendering (IBR), which is the technique without creation of 3D model. The light field method used by our technique is one of the IBR.[1,4] Light field is a kind of the database that records the luminosity information in the object space. The data of luminosity information on all rays is collected from the real images taken at various viewpoint positions. Kobayashi[3] describes an approach for generating arbitrary viewpoint images which lighten the burden imposed on the camera calibration. Kawasaki[2] describes an hybrid system of light field rendering and MBR. Once the light field is built, we can synthesize arbitrary viewpoint images by taking out light information according to the given viewpoint positions. Moreover, since it is not necessary to create an exact 3D model, the light field can also be used under complicated environment for synthesizing the arbitrary viewpoint images.

Omni-directional camera used in this technique is a combination of a normal perspective camera with the hyperbolic mirror in the upper part of the camera,[8] so that we can capture the luminosity information on 360 degrees of circumferences in one image. Taking into account the parameters of the mirror shape, we

E-mail address
Hiroshi Todoroki : todoroki@ozawa.ics.keio.ac.jp
Hideo Saito : saito@ozawa.ics.keio.ac.jp
†PRESTO, Japan Science and Technology Corporation(JST)

can determine a line in 3D space that corresponds to each pixel in an input image. Thus, we can collect luminosity information of a wide area from one input image. By applying a light field method to the images of an omni-directional camera, the virtualization of wide environment is attained.

The proposed method offers a number of advantages. First, since the luminosity information on 360 degrees of circumferences is recorded, looking-around-image can be created easily. The camera position of each input image is required in light field method. In our method, we estimate the camera position by means of tracking the feature point with known 3D position, so that we do not require any additional camera position control system.

2. GEOMETRY FOR LIGHT FIELD CONSTRUCTION

2.1. Omni-directional Camera

Omni-direction camera is consists of a normal perspective camera and a hyperbolic mirror attached in the upper part of it. Since the mirror reflects the lights around, it can picturize 360 degrees of circumferences at one time.

Fig.1 shows the composition of omni-directional camera. Focal point of the mirror O_M and camera center O_C are located in two focuses of the hyperbolic mirror $(0,0,+c),(0,0,-c)$. A image plane is uv plane that is parallel to XY plane, and is separated the focal length f of a camera from O_C. A hyperboloid of the mirror is shown by the Eq.1.

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

a, b are constants which define the form of a curved surface of the mirror. All light rays towards O_M gather at lens center of a camera O_C reflected by the hyperboloid.

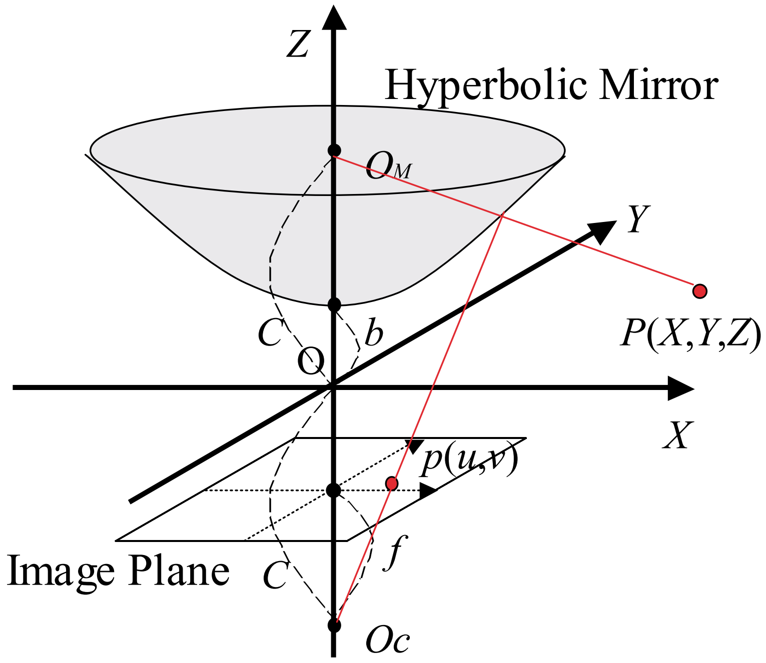


Figure 1. Composition of omni-directional camera.

With the characteristic of hyperboloid, the relationship between a point $P(X, Y, Z)$ in 3D coordinate and its projected point $p(u, v)$ on the omni-direction image can be shown in the following equation.[8]

$$\begin{cases} u = \dfrac{Xf(b^2 - c^2)}{(b^2 + c^2)(Z - c) - 2bc\sqrt{X^2 + Y^2 + (Z - c)^2}} \\ v = \dfrac{Yf(b^2 - c^2)}{(b^2 + c^2)(Z - c) - 2bc\sqrt{X^2 + Y^2 + (Z - c)^2}} \end{cases} \tag{2}$$

From Eq.2, the equation that represents the relation between X, Y, and Z for a given image point $p(u, v)$ is derived as shown in the following equation.

$$\begin{cases} Z = \dfrac{-f(b^2 + c^2) + 2bc\sqrt{u^2 + v^2 + f^2}}{(c^2 - b^2)u} X + c \\ Z = \dfrac{-f(b^2 + c^2) + 2bc\sqrt{u^2 + v^2 + f^2}}{(c^2 - b^2)v} Y + c \end{cases} \tag{3}$$

2.2. Light Field Method

Light Field Method is the technique to represent appearance at arbitrary viewpoint in a real space by all light rays that pass through the space. As shown in a Fig.2, the light rays are generally parameterized by the intersection points on the ST plane and UV plane of two paralell sheets. The four parameters that indicate the intersection points (s, t, u, v) specify a particular light ray. In this way, the light field is represented in four dimentional space. Various methods are also proposed for representing the light field information.[7]

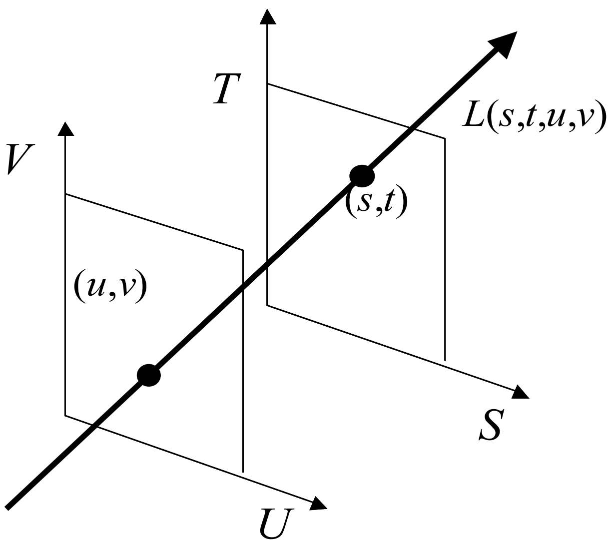


Figure 2. Parameterization of the light with two planes.

In this paper, we use a cylindrical plane whose height is infinity for representing the light field space as shown in Fig.3. It is assumed that an object does not exist in the cylinder, and the cylinder is arranged in a appropriate position, at which an assumption that each light going straight in the cylinder is constant can be satiesfied. The point on the surface of the cylinder is defined with the 2-dimensional parameter of an angle θ and height h. Then, each light can be specified using four parameters. (θ_{in}, h_{in}) is the point when it enters in a cylinder, and (θ_{out}, h_{out}) is the point when it comes out. By recording the light information which comes flying outside a cylinder, a user can feel looking over the surrounding of the cylinder virtually.

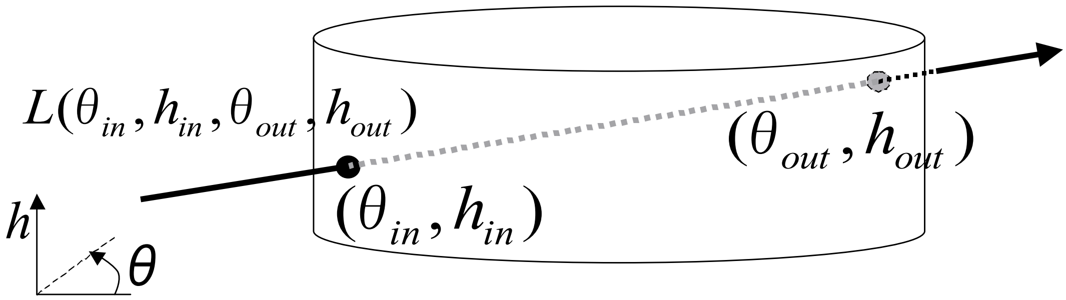


Figure 3. Specification of the light with the cylinder model.

An image is considered as a data set of the information on the lights which pass through a viewpoint position. Therefore, the number of pixels of one image is equal to the number of sampled lights that pass through the viewpoint position. Light field data is constructed by creating the database of the light that passes through the inside of a certain space using many images as input. However, since memory capacity is limited, it can record light data only on the position of the integer value in each variable $(\theta_{in}, h_{in}, \theta_{out}, h_{out})$.

3. ALGORITHM

This section describes the algorithm of the proposed method. This technique can be divided into three stages: estimation of intrinsic camera parameters, construction of omni-directional light field, and rendering of arbitrary viewpoint images. The target for capture of this technique is indoor static environment. The input is the image sequence captured with an omni-directional camera. While capturing, the omni-directional camera moves on the floor plane with keeping same pose that is vertical to the floor plane. This means that there is no rotation of the camera, but only 2D transation on the floor plane.

In the first stage, "Estimation of intrinsic parameters", the intrinsic parameter of the camera without the hyperbolic mirror is estimated. This procedure is performed only at once before the capturing of the image sequence of indoor static environment. The second stage, "Construction of omni-directional light field", is performed to every image in the captured sequence with the omni-directional camera. It is equivalent to the stage of the model creation in the technique of MBR. After building the light field, the third stage, "Rendering of arbitrary viewpoint images", is executed. This stage can be performed repeatedly according to the required viewpoint.

3.1. Estiamtion of intrinsic parameters

Intrinsic parameters of the normal perspective camera without the hyperbolic mirroris estimated in this procedure. Besides the intrinsic parameters, the parameters represents the shape of hyperbolic mirror is previously known according to the specification information provided by the mirror vendor.

When u axis is set as a transverse direction and v axis is set as the longitudinal direction, The intrinsic parameter is shown by Eq. 4. In this method, an intrinsic parameter is estimated by calibration method proposed by Zhang.[9]

$$\mathbf{A} = \begin{bmatrix} \alpha_u & -\alpha_u \cot\phi & u_0 \\ 0 & \alpha_v / \sin\phi & v_0 \\ 0 & 0 & 1 \end{bmatrix} \tag{4}$$

$$\left(\begin{array}{rcl} \alpha_u, \alpha_v & : & \text{scale factor of } u \text{ axis } \ v \text{ axis} \\ (u_0, v_0) & : & \text{pincipal point} \\ \phi & : & \text{angle between two axes of a image plane} \end{array} \right)$$

3.2. Construction of omni-directional light field

3.2.1. Estimation of a omni-directional camera position

In this section, we calculate the extrinsic parameters of the camera that is required for mapping of light rays. In every frame, the omni-directional camera position is estimated from the the correspondence relation of the image coordinates of feature points and the actual 3D positions.

The corresponding relation of the feature points is obtained by tracking the feature point with known 3D position. The tracking is performed by template matching using miniature bulbs as markers.

If we track the feature points with a normal perspective camera, the feature points sometimes come out of FOV with movement of the camera. On the other hand, the feature points rarely come out of FOV in the image of an omni-directional camera, even when the camera moves greatly. This is an advantage of the feature point tracking with the omni-camera.

From Eq.3, if the image coordinate of a certain feature point is given, the straight line which passes through the focal point O_M and the feature point can be computed. This means that the omni-directional camera exists on the computed line. Therefore, the intersection of two lines specified by two feature points gives the position of the omni-directional camera. Since some straight lines does not necessarily have an intersection point in 3D, the point with the minimum distance to the lines is searched as 3D position of the intersection.

3.2.2. Mapping of lights

The determination of the 4-dimensional coordinates of light ray which passes each pixel of a image and the focal of the hyperbolic mirror is called mapping. As shown Sec.2.2, in order to perform mapping, it is necessary to calculate the intersection of two points $(\theta_{in}, h_{in}), (\theta_{out}, h_{out})$ of light and a cylinder model. The 4D coordinates $L(\theta_{in}, h_{in}, \theta_{out}, h_{out})$ are calculated for each pixel of each frame by means of the parameter of the hyperbolic mirror and the estimated camera position, shown as Fig.4

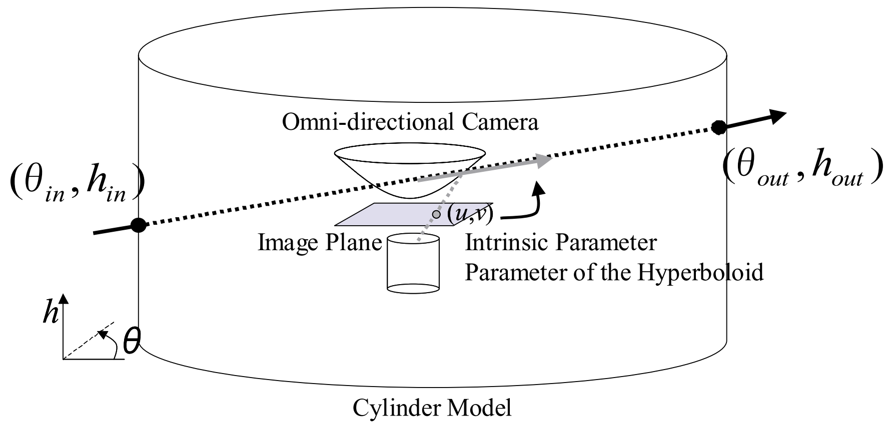


Figure 4. Mapping of lights.

3.2.3. Resampling of light ray

In this technique, the luminosity information of light ray $L(\theta_{in}, h_{in}, \theta_{out}, h_{out})$ can be recorded only on the position of the integer value in each variable $(\theta_{in}, h_{in}, \theta_{out}, h_{out})$. If a motion of a camera is controlled mechanically and images are captured at even intervals, the light information on the ideal sampling interval can be acquired easily. However, since the input of this technique is the image sequence captured with moving omni-directional camera at an variable speed, the sample of light is disordered and non-uniform. Thus the 4D coordinates of every light ray are not restricted with an integer value, and the obtained sample has very intense roughness and fineness by the position. Therefore, a resampling is performed in order to obtain data at equal intervals shown as Fig. 5. Here, we adopted the resampling technique of Gortler.[1]

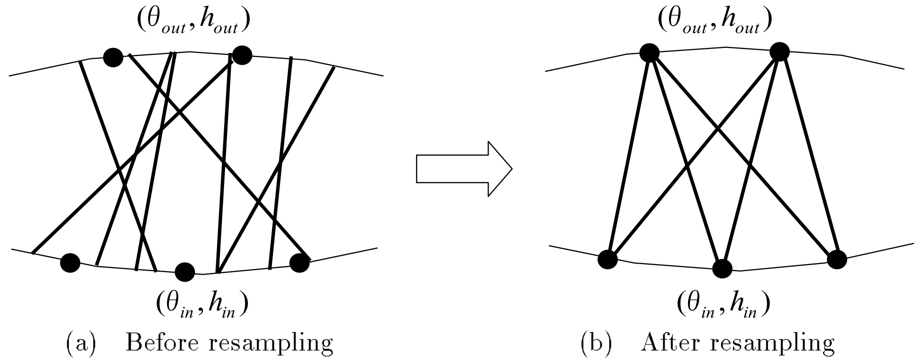


Figure 5. Resampling of light date.

3.3. Rendering of arbitrary viewpoint images

3.3.1. Specification of a virtual viewpoint

The 3D coordinates of a virtual viewpoint and the image plane are specified. This is equivalent to setting up the direction and FOV of a camera which perform a rendering. By the position of the 3D coordinates of a virtual viewpoint and a cylinder model, corresponding lights are mapped about each pixel of a synthetic image, and its two intersections with a cylinder model $(\theta_{in}, h_{in}), (\theta_{out}, h_{out})$ are calculated.

3.3.2. Extraction and interpolation of light ray

In the resampled light field, the light ray with the 4D parameters required for synthesizing the virtal viewpoint image hardly exists. Therefore, the luminosity of required light ray must be computed by interpolating with surrounding data of the required light ray.

First, four lattice points around required coordinates (θ_{in}, h_{in}), and four lattice points around required coordinates (θ_{out}, h_{out}) are chosen. Second, 16 lights around required light is extracted from these eight points, and interpolation is performed in consideration of the weighted distance, shown as Fig.6. In this way, the luminosity of each pixel is determined, and then the arbitrary viewpoint image is synthesized.

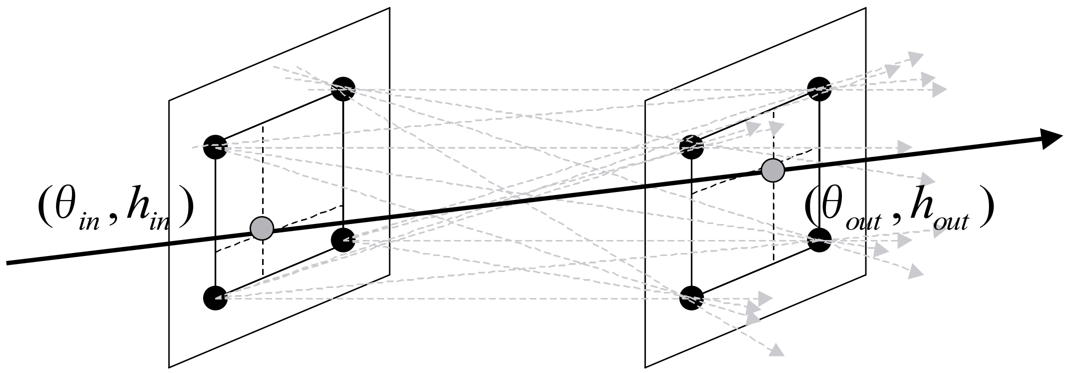


Figure 6. Interpolation of light ray.

4. EXPERIMENTAL RESULTS

The input is the image sequence captured while the omni-directional camera moves parallel with a stage as shown Fig.7. After performing camera position estimation, omni-directional light field is constructed. Then, the arbitrary viewpoint images are synthesized continuously, so that the virtual tour movie can be generated, which is walk-through and looking around in the scene virtually.

The object of capture is the inside of our lab's room. The examples of an input image are shown in Fig. 8. Resolution of image is 512×384, and the number of input images is 190. The 4D-parameter of light field is $(\theta_{in}, h_{in}, \theta_{out}, h_{out}) = (50, 25, 800, 320)$. The radius of the cylinder for representing light rays in the light field is set to 2.5m in accordance with the size of the room. Any object does not exist inside of the cylinder. We use four markers for camera position estimation, and they are attached on the floor in the cylinder.

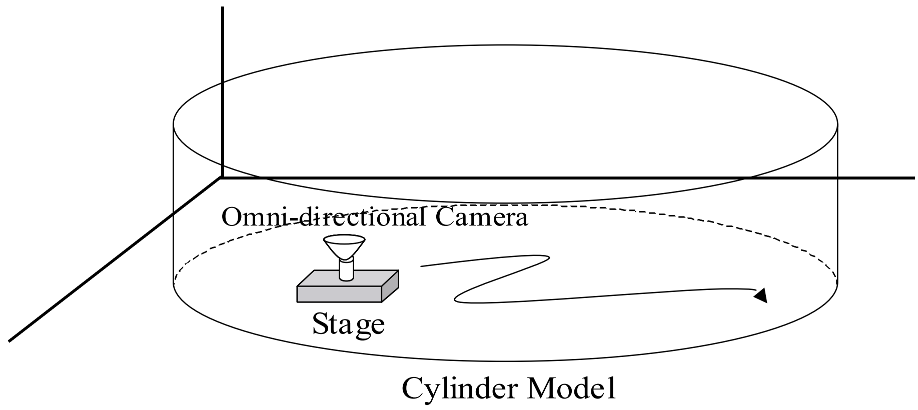
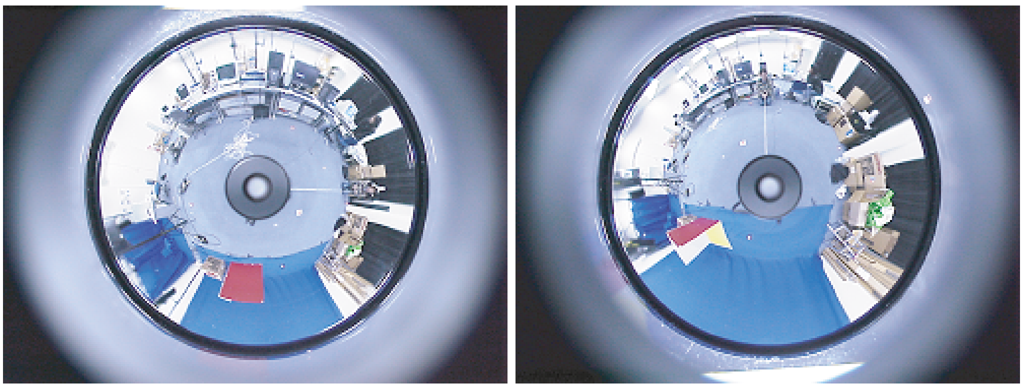


Figure 7. Capture of input image sequence.



(a) frame 50 (b) frame 120

Figure 8. Example of input images.

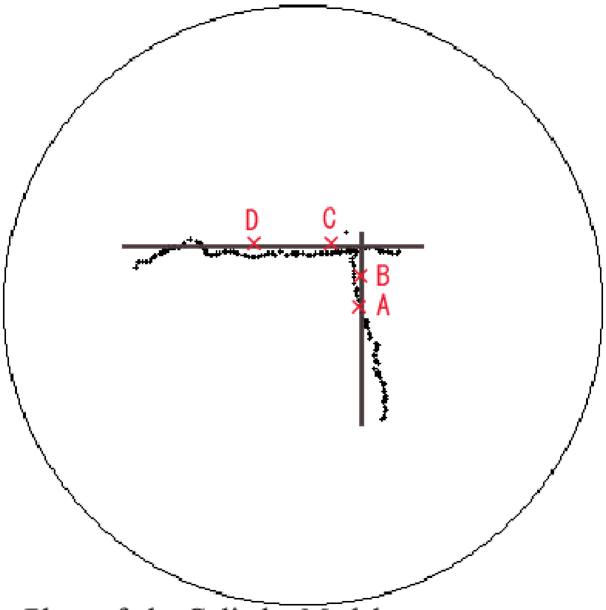


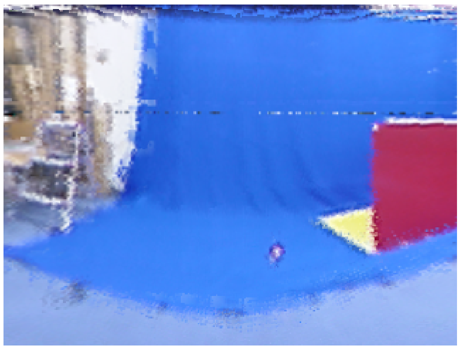
Figure 9. Presumed camera position(top view).

Fig.9 shows the camera position presumed, and illustrates the top view image of the room. The circle indicates the place of the cylinder, the continuous line indicates the actual locus of omni-directional camera, and the point indicates the place of estimated camera position.

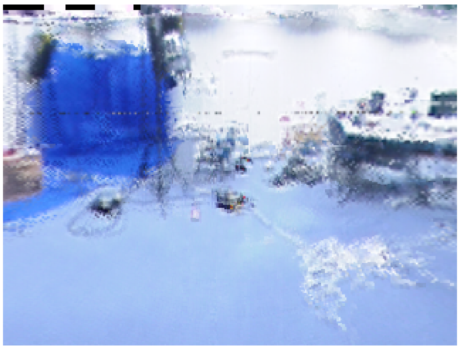
Fig. 10 shows the example of the synthesized images in a virtual tour movie. The capital alphabet shown under the image indicates the 2D position of the viewpoint in Fig.9, and the parameter indicates the 3D-coordinates of a viewpoint position, and direction of a camera. In the coordinate system here, the center of the cylinder on the floor is the origin. Fig.11 shows the example of the synthesized panorama images, which illustrates the one of the advantage in this proposal method.

The omni-directional cameras can capture 360 degrees of circumferences at the same time by attaching a hyperbolic mirror in the upper part of the camera. On the other hand, the image resolution in the vertical angle is low. Especially, since a photograph is taken circularly, resolution of angle for depression is low. In this technique, although the angular resolution with an omni-directional camera is low, capturing at various positions can compensate the low resolution by integrating all light rays in the light field.

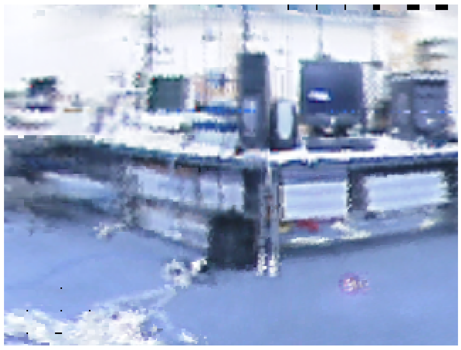
(a) A:(500,0,1400),(0,10)



(b) B:(500,200,1400),(90,10)



(c) C:(300,550,1400),(180,10)



(d) D:(-500,550,1400),(260,10)

Figure 10. Example of the synthesized images.
(The 2D position of the viewpoint is indicated as the captal in figure 9.)

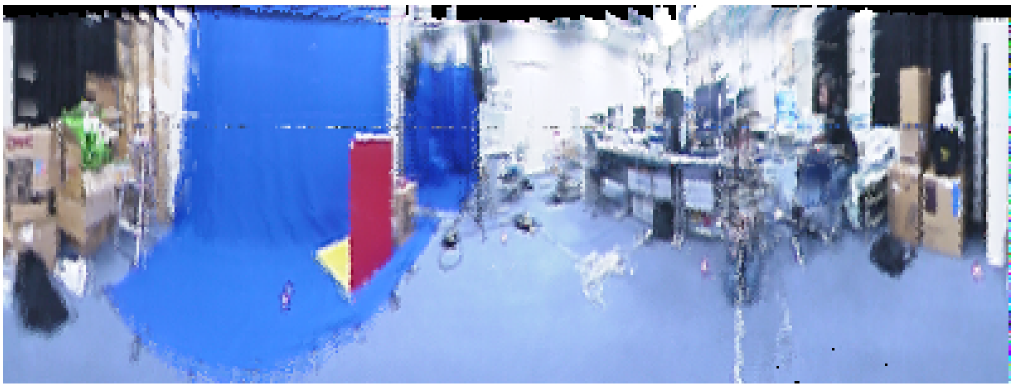


Figure 11. The synthesized panorama images at A:(500,0,1400).

5. CONCLUSION

This paper presents an effective scheme for virtualizing wide environment. This representation does not require deriving any geometric shape of the scene such as depth. Therefore, this method is an applicable technique in various environment or complicated one.

In the experiments shown in this paper, we input the image sequence of omni-directional camera, and synthesize the arbitrary viewpoint images. Then we generate the virtual tour movie, which make us feel walking through all around virtually. Since the 3D position of a camera is estimated from each image, we can generate arbitrary viewpoint images from image sequence that is captured without any the machine control of the camera.

There are some future works to be done on this topic. The crucial problem is that huge memory space is required if wide area is captured by this method. It will be important to consider the new structure of light field data and to develop compression method. Increase of the limit of the 4D parameter on the cylinder model will also improve the resolution of synthesized image.

REFERENCES

1. Steven J. Gortler, Radek Grzeszczuk, Richard, Szeliski, Michael F. Cohen, “The Lumigraph”. SIGGRAPH96, Conputer Graphics Proceeding,1996
2. Hiroshi Kawasaki, Katsushi Ikeuchi, Masao Sakauchi, “Light Field Rendering for Large-Scale Scenes” IEEE Computer Vision and Pattern Recognition(CVPR 2001), Hawaii,USA,2001
3. Kenkichi Kobayashi, Hideo Saito, “Arbitrary Viewpoint Image Synthesis Based on Creating Light Field from Uncalibrated Image Sequence”. Transaction on The Institute of Electronics, Information and Communication Engineers D-II, VOL.J86-D-II,2003 (in Japanese)
4. Marc Levoy, Pat Hanrahan, “Light Field Rendering”. SIGGRAPH96, Conputer Graphics Proceeding,1996
5. Shinji Morita, Kazumasa Yamazawa, and Naokazu Yokoya, “Internet telepresence by real-Time view-dependent image generation with omnidirectional video camera”. Proc. SPIE Electronic Imaging, Vol. 5018, pp. 51-60,2003
6. Camillo J Taylor, “VideoPlus”. IEEE Workshop on Omnidirectional Vision,2000
7. Daniel N. Wood, Daniel I. Azuma, Ken Aldinger, Brian Curless, Tom Duchamp, David H. Salesin, Werner Stuetzle, “Surface Light Fields for 3D Photography”. SIGGRAPH2000,2000.
8. Kazumasa Yamazawa, Yasushi Yagi, and Masahiko Yachida, “ HyperOmni Vision: Visual Navigation with an Omnidirectional Image Sensor”. Systems and Computers in Japan, Vol.28, No.4, pp.36-47,1997
9. Zhengyou Zhang, “Flexible Camera Calibration By Viewing a Plane From Unknown Orientations ”. International Conference on Computer Vision,1999