コンピュータビジョンと
イメージメディア　119－5
（1999．11．17）

全方位画像によるレンダリング手法の提案

高橋拓二　川崎洋　池内克史　坂内正夫

東京大学生産技術研究所 電子情報系
〒106 東京都港区六本木7-22-1
Tel：03-3401-1433　E-mail：takuji@iis.u-tokyo.ac.jp

あらまし あらかじめ撮影した全方位画像を用いて、違う場所からの眺めを描画する新しい手法を提案する。この手法は3D Plenoptic Functionなるカテゴリに分類されるが、同じカテゴリの他の手法とは異なり、画像を撮影した平面(通常は地面)上のほとんど任意の点からの視界を再構築をすることができる。また、Omni-derectional cameraを用いることで計算をより簡単にすることができる。この手法の応用例として、ITS分野における運転シュミレータが挙げられる。この手法を用いれば道路をただ1度走るだけで、違う走行路からの画像をRenderingすることができる。

Rendering with Panoramic Images

Takuji Takahashi　Hiroshi Kawasaki　Katsushi Ikeuchi　Masao Sakauchi

Institute of Industrial Science
University of Tokyo
7-22-1 Roppongi, Minato-ku, Tokyo JAPAN 106
Tel : 03-3401-1433　E-mail : takuji@iis.u-tokyo.ac.jp

Abstract This paper presents a new method for rendering a novel view of prerecorded images taken from different locations. This method belongs to the family of work that employs 3D plenoptic functions; however, unlike other works of this type, this particular method allows us to render a novel view from an almost arbitrary point on the plane, on which images are taken, along a straight line. Using an omni-directional camera makes the calculation far easier. One of the applications of this method in the ITS domain is to make a driving simulator. We can generate any views on the road from images taken by running along just one. In this paper, we describe the basic theory of a 3D plenoptic function, and the applicable areas of the theory; in addition, we describe a complete working system, including the capture of images as an example. We also present rendered images of a real scene.

1 Introduction

The computer graphics community has expended much effort to create virtual environments from real scenes. We can classify such efforts into two main approaches: image based modeling (IBM) and image based rendering (IBR). An IBM approach analyzes the geometry and surface attributes of the objects in the environment and, based on those analyses, creates new views.

The IBM approach requires precise reflectance model analysis and careful image acquisitions. The IBM first captures a series of images for the analysis to be completed a later stage. This approach typically involves some combination of precise camera positions, structured lighting, and elaborate laboratory setup. By capturing the geometry and material properties of objects directly from the real world, we can create images of real objects and scenes using Computer Graphics rendering software. This method, however, is still limited in its ability to create images of relatively small objects in indoor scenes.

The IBR also captures a series of images. However, this method does not analyze those images; rather, it creates new views by reassembling or hacking those prerecorded images in a timely manner. QuickTime VR is one of the representative systems based on the IBR. A series of captured environment images, pasted on a cylindrical environment, allows a user to look around a scene from fixed points in space.

One of the key concepts developed in the IBR is the plenoptic function. Originally, a 7D plenoptic function was proposed to define the intensity of light rays passing through the camera center at every location, at every possible viewing angle, for every wavelength and at any time. It has since been shown that light source direction can also be incorporated into the plenoptic function for describing illumination environments. By ignoring time and wavelength, McMillan and Bishop generated a continuous 5D plenoptic function from a set of discrete samples. The Lumigraph and Lightfield systems presented a clever 4D parameterization of the plenoptic function.

The study of plenoptic function has focussed on maintaining the quality of rendered images while requiring fewer dimensions. Namely, the issue was how to reconstruct images from a viewpoint in a wide space with less data.

This paper proposes a new method for obtaining a 3D plenoptic function. This method first captures panoramic images through either an omni-directional camera or a combination of standard TV cameras that run along a straight line. Those captured panoramic images represent a 3D plenoptic function, brightness of a ray at any location (2D) and at any direction (1D). By selecting and combining appropriate slits in the panoramic images, we can render a novel view.

This paper is organized as follows. In Section 3, we explain the basic method for obtaining the 3D plenoptic function, and discuss some of the characteristics of the method. Section 4 presents results of experiments using a real scene, and Section 5 contains our conclusions regarding this method along with discussions of future work.

2 Reconstruct of Arbitrary View from Panoramic Images

2.1 Capturing Panoramic Images

At each location, we construct a panoramic image,which contains all rays from the capturing location; we have a total horizontal viewing field of 360 degrees at each capturing location.

The simplest and easiest method for capturing panoramic images is to use an omni-directional camera. This type of camera has an orthographic lens that has a single effective viewpoint (see Fig.1. From the sensed omni-directional image, we can generate pure perspective images, and, thus can make panoramic images from the omni-directional image.

Using one omni-camera, we can take images with 360 degrees in a horizontal direction; those images cover the northern hemisphere of a viewing sphere. The images cover the upper directions above the image plane of the omni-directional camera. By placing two such cameras back-to-back, thereby satisfying the single viewpoint constraint, we can achieve a true omnidirectional sensor, 360 degree views in both horizontal and vertical directions. Details of the omni-directional camera with equations are shown in the Appendix.

Another method for capturing panoramic images is to arrange a few cameras cylindrically as shown in Fig.2 [7]. Those cameras' optical axes intersect at one point with rays around the center of cameras.

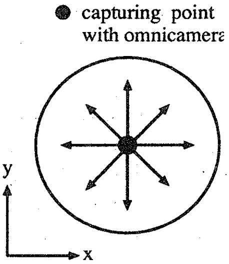

Figure 1: With omnicamera

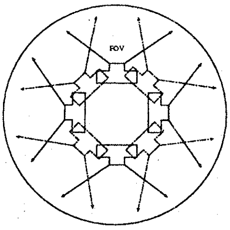

Figure 2: Configuration of some cameras

2.2 System of Capturing Images

One application of the system is to build a driving simulator. Here, we are mainly concerned with the case where a camera runs along a straight line. Namely, by running in a straight line e.g., a lane on a road, we capture omni-images with the position and direction of each panoramic image. Here the positional information is given by a GPS sensor.

We will denote the ground plane as the x-y plane. As shown in Fig.3, by moving from C_0 to C_n, we capture images and record their positions. Here we denote a panoramic image captured at (x_i, y_i) as $C_i(x_i, y_i)$. We can pick up arbitrary slits from each panoramic image for reconstruction purposes.

2.3 Reconstruct of Novel Views

Given a series of panoramic images on a straight line, we can construct a novel view from almost any arbitrary region. Consider the case of rendering the novel image at the point P, as illustrated in Fig.4. For constructing the view from P, we need rays around P from R_s to R_e as shown in the figure. By finding the slits corresponding to the rays from stored panoramic images and collecting those slits, we can synthesize a new view.

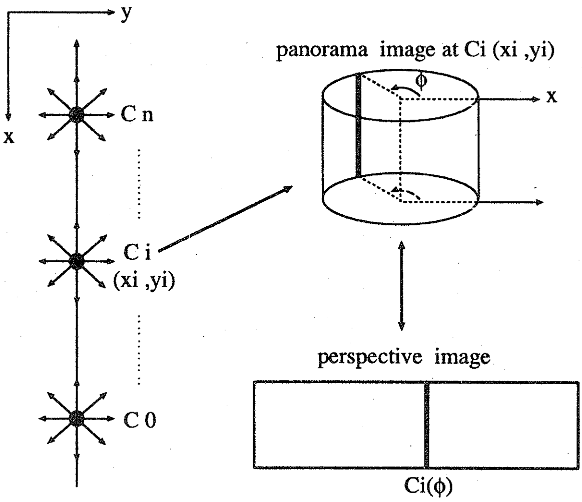

Figure 3: Capturing panoramic images

The calculation of this process are as follows. See Fig5. To reconstruct a novel view at $P(x_p, y_p)$, we need all rays from θ to θ'. Here the angle of the rays are denoted with respect to the x axis.

Now, we define vector $i = (1, 0)$, and position vectors as $P = (x_p, y_p)$ $C_i = (x_i, y_i)$. The ray, of which the angle is ϕ, is given by a panoramic image of $C_i(x_i, y_i)$; the equation is as follows.

$$\begin{aligned} \phi &= \arccos \frac{(C_i - P_i) \cdot i}{|(C_i - P_i)||i|} \\ &= \arccos \frac{(-x_i) + (-x_p)}{\sqrt{(x_i - x_p)^2 + (y_i - y_p)^2}} \end{aligned} \quad (1)$$

Using this formula, we can render a novel view

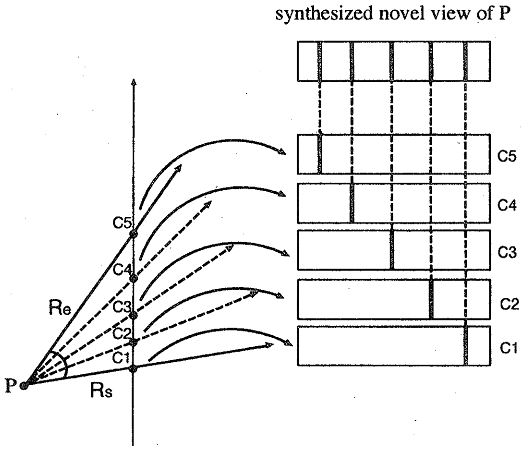

Figure 4: Reconstructing a novel view

at P.

$$\sum_{\phi=\theta}^{\theta'} \phi = \sum_{\phi=\theta}^{\theta'} \arccos \frac{(-x_i) + (-x_p)}{\sqrt{(x_i - x_p)^2 + (y_i - y_p)^2}} \quad (2)$$

2.4 Region of Reconstruct

This section will discuss the areas where it is possible to reconstruct novel views. Let us assume that the camera moves on a straight line for the interval from the position,C_0 to the position, C_n. The reconstructable angle, in this case, is shown in Fig.6. Namely, upon rendering the view from the position P, the reconstructable region around P is the shaded area.

More importantly, we can enlarge the possible region of shaded portion, as shown in Fig.6, by running for a longer distance. The boundary line moves following the arrows. By running on a straight line for a infinitely long distance, we can render a novel view at any point on the ground plane.

2.5 Singular direction

With respect to reconstruction, we can classify a novel view into two cases: with or without a singular direction. Here we define the singular direction along the direction parallel to the running direction. Namely, the first case includes the ray parallel to the singular direction, while the second case does not. When reconstructing a view toward the moving direction from the driver ' s seat, we have to consider the first case, while for a side view, we consider the second case.

First, we will discuss the case with a singular direction. In this case, as shown in Fig.7, we put together some slits from two parts of non-adjacent panoramic images. The right-hand side of the image comes from those ahead of the point A and the left-hand side of the image comes from those in back of the point B. Due to this discontinuity, there may be a distortion in the resulting image across the singular direction.

In addition, a ray parallel along the singular direction does not exist except moving for an infinite distance. We have to interpolate this ray by using a morphing technique. However, usually along the singular direction, no object or very far objects exist in the ITS applications. Thus, the distortion of any far objects is relatively small.

In the none singular case, a view to be rendered does not include the ray parallel to the singular direction. As shown Fig.8, in this case, all necessary rays are contained in the series of panoramic images. Moreover, the rays are in-

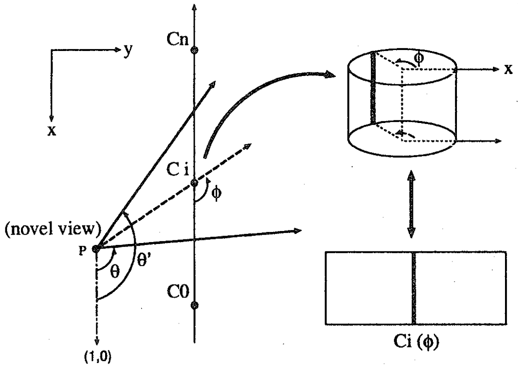

Figure 5: Calculation for rendering

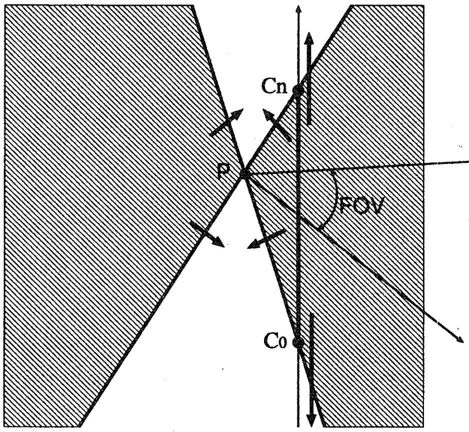

Figure 6: Reconstructable area

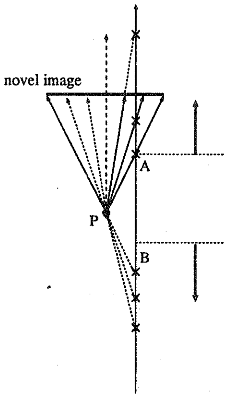

Figure 7: With a singular direction image

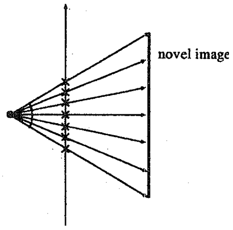

Figure 8: Without a singular direction image

cluded in a continuous part; there is no distortion in this case, as contrasted to the singular case.

2.6 Vertical Distortion

Just as the concentric mosaics[6], this system has the effect of the vertical distortion. We can devise several methods to reduce this vertical distortion. If we know the distances between the camera optical center and the points in the scene, full perspective correction based on distances can be done. A novel view is rendered using optical flow. This method, however, requires acknowledge of geometry in the real scenes.

In many real scenes in the ITS application domain, it is reasonable to approximate that pixels along a vertical line have the same depth as walls of buildings. In this case, we need to estimate only the depth value along each vertical line. This can be done using the "dynamic EPI" analysis [8]. This method is like a video game technique in which the front object moves faster and the background moves slowly. Usually EPI analysis is done by static image analysis, but this method uses the motion vector on the EPI plane. By using both motion vectors and detected edges on the EPI plane, we can easily segment the regions with same depth (we assume that most objects consist of a plane). And by restoring the result to the video data, we can robustly retrieve the depth value. Using this estimated value, we can scale the whole line uniformly.

3 Results

We have implemented this system and created many novel views from omni- directional images. For real scenes, we have used a hyper Omni Camera and reconstructed images on an SGI workstation.

Fig.10 is an image of our laboratory scene captured by the hyper omni-directional camera. It is easy to change to perspective image, and Fig.11 is a part of perspective image which was made from Fig.10.

We have captured some panoramic images moving on a straight line. Using this method, we created some novel views without the singular direction. The positions of some of them are depicted in Fig. 12.And Fig.13, Fig.14 and Fig.15 are those views at the locations A, B and C, respectively. Their locations are not on the moving line of camera. They are further away from the line in order. Note the relation between blackboard and bookshelf behind the blackboard. In Fig.13, the blackboard almost occludes the bookshelf. In contrast, in Fig.14, which is further away from the moving line than Fig.13, a partial occlusion occurs. In Fig.15, the blackboard and bookshelf are separated.

4 Conclusion and Future Work

This paper describes a new method for reconstructing a novel view with mosaicing panoramic images. We capture panoramic images running along a straight line and record the capturing position of each image.

It is important to note that we create a novel view of an almost arbitrary point on a ground by using a relatively simple method. We need only select some suitable slits from stored panoramic images, and reassemble them to generate an image from a novel observation point. In other words, once images are run and recorded along a straight path, an arbitrary view around the path can be constructed.

Compared with the concentric mosaics, this system has more advantages. This method can generate novel views from different locations captured. This method can also correctly render occluded objects, and where they are located. By cooperating with an omni-directional camera, the method becomes very simple.

One of the most interesting extensions of this work would be to study about intervals for capturing panoramic images. Depending on the compression method, we would be able to render images with less data.

The current resolution of omni-directional image is not satisfactory for rendering novel views. It will be necessary to explore the possible use of super resolution for improving image resolution. Another method to increase the resolution would be to use several cameras so as to allow the optical axes of cameras to intersect at the center as was shown in Fig. 2. By setting up this configuration of cameras on a data collection car with a differential GPS and a speed meter, we plan to capture panoramic images with indexed time for rendering a novel view along a road for ITS purposes.

References

[1] Shree K.Nayar. Omnidirectional Video Camera. *Computer Vision and Pattern Recognition '97*, page 482 - 488, June 1997

[2] E.H.Adelson and J.Bergen. The Plenoptic function and the elements of early vision. In *Computional Models of Visual Processing*,pages 3-20,MIT Press,Cambridge,MA,1991

[3] L.McMillan and G.Bishop. Plenoptic modeling: An image-based rendering system. *Computer Graphics (SIGGRAPH'95)*,pages 39-46,August 1995

[4] M.Levoy and P.Hanrahan. Light field rendering. *Computer Graphics Proceedings, Annual Conference Series*,pages 31-42,Proc.SIGGRAPH'96(New Orleans),August 1996. ACM SIGGRAPH.

[5] S.J.Gorther,R.Grzeszczuk,R.Szeliski,and M.F.Cohon.
The lumigraph. *Computer Graphics Proceedings, Annual Conference Series*,pages 43-54, Proc.SIGGRAPH'96(New Orleans),August 1996. ACM SIGGRAPH.

[6] Heung-Yeung Shum and Li-Wei-He. *Computer Graphics Proceedings,Annual Conference Series*, pages 299-306,Proc.SIGGRAPH'99(California),August 1999. ACM SIGGRAPH.

[7] Hirose M. and Takaaki E.,Building a Virtual World from the Real World. *Proceeding of International Symposium on Mixed Reality.* (1999-3). 183-197

[8] H. Kawasaki, T. Yatabe, K. Ikeuchi and M. Sakauchi Construction of a 3d city map using epi analysis and dp matching , Asian Conference on Computer Vision2000 (2000).

5 Appendix

5.1 Equations for HyperOmni Vision

The following is equation for HyperOmni Vision. Fig.9 shows the structure of HyperOmni Vision; there are two rotation hyperbolical surfaces and two focusses (in Fig.9,Om, Oc). These focusses of the camera are set at Oc.

When we assume a 3D world coordinate, as shown in Fig.9, the equation of the hyperbolical mirror face is (3), and Om(focus of hyperbolical mirror) is $(0,0,+c)$ $(c=\sqrt{a^2+b^2})$ and Oc (center of camera lens) is $(0,0,-c)$.

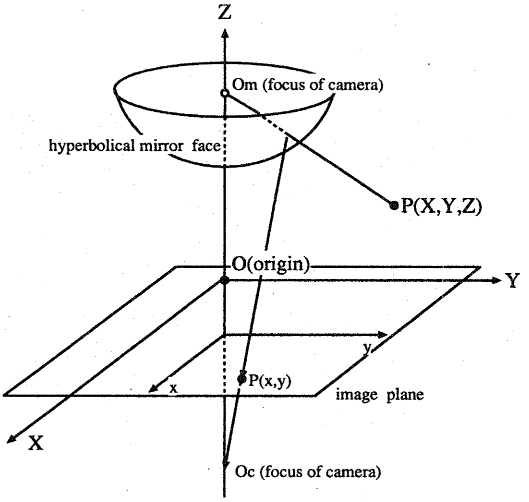

Figure 9: Hyperomni

$$\frac{X^2+Y^2}{a^2} - \frac{Z^2}{b^2} = -1 \qquad (3)$$

$$(Z > 0)$$

$$(0,0,+c) \qquad (c=\sqrt{a^2+b^2}) \qquad (4)$$

$$(0,0,-c) \qquad (5)$$

$$(a,b,c\ are\ paramerter\quad of\quad hyperbolical\ mirrorface)$$

All rays focused to Om come together at Oc reflected by hyperbolical mirror surface.

$$Z = \sqrt{X^2+Y^2}\tan\alpha + c \qquad (6)$$

$$\alpha = \arctan\frac{(b^2+c^2)\sin\gamma - 2bc}{(b^2-c^2\cos\gamma)} \qquad (7)$$

$$\gamma = \arctan\frac{f}{\sqrt{x^2+y^2}} \qquad (8)$$

$$\theta = \arctan\frac{Y}{X} = \arctan\frac{y}{x} \qquad (9)$$

α : *an zenith between* Om *and* P

γ : *an zenith between* Os *and* p

θ : *an azimuth angle*

Above all, the relation between point P(X,Y,Z) and p(x,y,z) are follows (p on the omnidirectional image is projected to P in the 3D world coordinates).

$$x = X \times f \times \frac{b^2-c^2}{(b^2+c^2)Z - 2bc\sqrt{X^2+Y^2+Z^2}} \quad (10)$$

$$y = Y \times f \times \frac{b^2-c^2}{(b^2+c^2)Z - 2bc\sqrt{X^2+Y^2+Z^2}} \quad (11)$$

Namely, eq.10 and 11 do not included any trigonometry. So, the cooedinates of point p can be calculated from the coordinates of P, simply and speedy.

6 Images

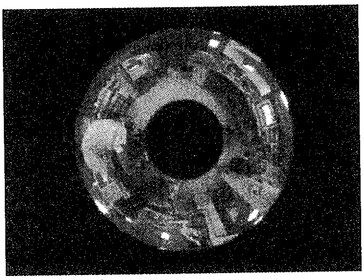

Figure 10: omni-directical image

Figure 11: perspective image made from omnidorectical image

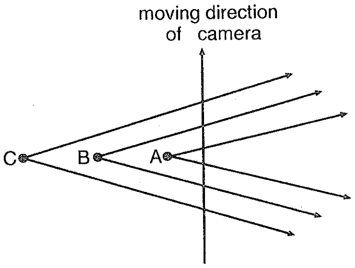

Figure 12: configuration of rendering image

Figure 13: novel view at A

Figure 14: novel view at B

Figure 15: novel view at C