

## AN IBR METHOD: PANORAMIC MOSAICS OF SLIT IMAGES WITH DEPTH\*

JIANG Chun(蒋 纯), SHI Jiao-ying(石教英)

(*State Key Laboratory of Computer Aided Design and Computer Graphics, Zhejiang University, Hangzhou 310027, China*)

Received June 21, 2000; revision accepted Sept. 18, 2000

**Abstract:** Providing a wider movement range of virtual camera is an unsolved problem for state-of-the-art image-based rendering system. In this paper, we present a new image based rendering technology called panoramic mosaics of slit images with depth that can provide large virtual camera motion region for some scenes. By limiting camera motion to a horizontal plane only, a slit image with united depth value is used as the rendering element. The panoramic mosaics of slit images with depth are easy to capture, and the data size is as small as that of panorama. We present here the capturing, construction as well as rendering process of panoramic slit images mosaic with depth. In addition, we present the join up process of multiple panoramic slit images mosaic with depth.

**Key words:** image based rendering (IBR), panorama, slit images

**Document code:** A      **CLC number:** TP391.72

### INTRODUCTION

Traditional computer graphics try to generate realistic scene view from a geometric model. Computer vision extract geometric model by analyzing images from the real scene. Combining them together gave us the idea of image-based rendering (IBR), which generates realistic scene view from some analyzed real images (photographs) along the motion of a virtual camera. The modeling of scenes photorealistically is then easier, and the rendering speed is faster.

Panorama(Chen, 1995) is one of the first image-based rendering (IBR) technologies used and is still being used widely. Using this technology, a set of photographs was taken in a viewpoint. The photographs were then registered (Brown, 1992; Szelishi et al., 1997) and stitched together to construct a panorama. The advantage of panorama is that it is easy to capture and small in data size. Its disadvantage is that the viewpoint is fixed and only the viewing directions and camera zoom can be altered.

Plenoptic Modeling (McMillan et al., 1992) is another kind of panorama-based IBR technique that supports the viewpoint motion. In Plenoptic Modeling, several panoramas are joined up, and interpolation is used to generate

the novel (i.e., synthesized) view. However, in this technique the establishing of the correspondence relation is complex.

Light Field(Levoy et al., 1996) or Lumigraph(Gortler et al., 1996) is another kind of image-based rendering technology. By sampling every ray passing through the motion range, these technologies support the motion of virtual camera. Nevertheless, their data size is large even for a small motion range.

The data size versus the virtual camera motion range has long been a trade off for most image-based rendering systems. Small data size means limited virtual camera motion, like in panorama. When virtual camera moves, like in Light Field or Lumigraph, many data will be needed. The huge data size actually limited the movement of virtual camera. Therefore providing a wider range movement of virtual camera is very hard for state-of-art image-based rendering system.

As recent research shows, geometric information can help to reduce the photos needed in new scene view synthesizing (Chai et al., 2000). In other words, with the same number of photos, more geometric information, more view can be synthesized. If we know nothing about the scene geometry, a panorama is only the pan-

\* Project(69823003) supported by NSFC

orama itself: no new view is available if we leave the original viewpoint. However, if we have some depth information, we can deduce the corresponding points in the new view of pixels in original photos. Assuming the surface is Lambertian, we can synthesize the novel view from the original photos by image warping, just like in view interpolation (Chen et al., 1993) or view morphing (Seitz et al., 1996). However, it is hard to get the depth value of every pixel in a real scene.

This paper presents panoramic mosaics of united depth slit images that support virtual camera motion. In addition, several mosaics generated by this technology may join-up to provide wider range virtual camera motion.

Slit images are 1-D images with width of only 1 pixel. We can treat them as a column of a normal image. The concept of slit image comes from computer vision (Peleg et al., 1997; Zheng et al., 1990), and it is used in several image-based rendering systems. In Multiple-Center-of-Projection (MCOP) images (Rademacher et al., 1998), a set of slit images instead of normal images is taken to achieve uniform sample. A MCOP image is a kind of range image, that is, every pixel in the image has its depth value sampled by special equipment (e. g., range finder).

Another image-based rendering technology using slit image is the concentric mosaics (Shum et al., 1999). In concentric mosaics, cameras move along planar concentric circles, and create concentric mosaics using a manifold mosaic for each circle (i. e., composing slit images taken at different locations). Concentric mosaics index all input image rays naturally in 3 parameters: radius, rotation angle and vertical elevation. Novel views are rendered by combining the appropriate captured rays in an efficient manner at rendering time. However, it still cannot provide wider range virtual camera motion.

This paper is organized as follows. We introduce the research works we have done on depth recovery from slit images and present the concept of slit images with depth in Section 2. In Section 3 we present the sampling and constructing process of panoramic mosaics of slit images with depth as well as the rendering of novel view us-

ing these mosaics. In Section 4 we present an experiment system to verify our idea. We summarize our work with some conclusions in Section 5.

## THEORY AND CONCEPT

### 1. Slit images with depth

Most computer graphics technologies are used for simulating human motion, which is usually only in 2-D: left and right, or forward and backward. In such condition, pixel change in horizontal direction is much faster than that in vertical direction. If the motion of the virtual camera is limited only in a horizontal plane, we can make a helpful simplification: combine vertical pixels together, use the combination, i. e., slit images to be the element in image warping. A united depth value is used for the whole slit image. The scenes are assumed to be composed of a set of colored 1-D vertical lines, or pillars; the color of the lines looks the same for any angle of view.

Of course, especially in forward and backward motion, depth variance alone in slit images also cause pixel change in them. If we want this simplification to approach reality, another assumption is needed: for most pixels in a slit image, the depth variance should be small. When most objects in a scene are very far away from the viewpoint, the depth variance could also be neglected. We call a slit image with a united depth value "slit image with depth", and will use a panoramic mosaic of slit image with depth for our rendering system.

### 2. Analogical slit images

Fig. 1 shows two slit images (called analogical slit images (Jiang et al., 2000)\* captured by the same camera, from the same direction but at two different positions. They contain the projections of almost the same scene objects, especially when the camera is near the two slit images. Now consider a point  $M$  in the scene, its projection in two analogical slit images is  $m_1$  and  $m_2$  respectively.  $m_1$  and  $m_2$  are called corresponding points in terms of computer vision.

Denote the optical centers of the two cameras

\* Jiang, C., Shi J. Y., 2000. Planner slit image fields. *Chinese Journal of Software* (in Chinese, accepted)

as  $C_1$  and  $C_2$  respectively. Using a pinhole model of camera, we can deduce the following (When  $H$  is not equal to zero):

$$\begin{cases} \frac{f}{d_1} = \frac{h_1}{H} \\ \frac{f}{d_2} = \frac{h_2}{H} \end{cases} \quad (1)$$

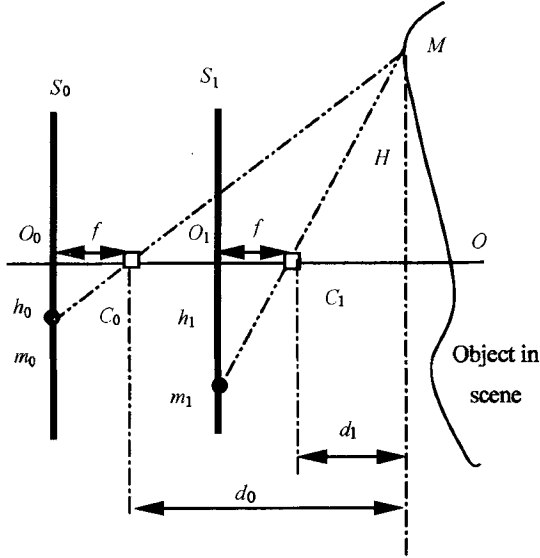


Fig. 1 Analogical slit images

Where  $f$  is the focal length of the pinhole model cameras,  $H$  is the distance from  $M$  to the optical axis of the cameras.  $d_1$  and  $d_2$  are the lengths of the projections in the optical axis of  $MC_1$  and  $MC_2$ ,  $h_1$  and  $h_2$  are distances from the optical axis to  $m_1$  and  $m_2$ , respectively.

Obviously, when  $H$  is not equal to zero, neither  $h_1$  nor  $h_2$  will equal to zero. From Equation 1 we get:

$$\frac{h_2}{h_1} = \frac{d_1}{d_2} \quad (2)$$

$$\text{or } \frac{h_1 - h_2}{h_2} = \frac{d_2 - d_1}{d_1} \quad (3)$$

These equations can be used to recover depth from looming parallax. If we knew the corresponding point of the two slit images, we can deduce the  $d$  value from the above equations. If it is assumed the slit image has a united depth value, the equation can also be used to apply to the slit image to simulate the looming parallax while the viewpoint moves:

$$\frac{l_o}{l_v} = \frac{d_v}{d_o} \quad (4)$$

Where  $d_o$  is a initial depth value of a slit image,  $l_o$  is a initial length of a slit image.  $d_v$  is depth value from a given new virtual camera, and  $l_v$  is the new length of the slit image at that viewpoint.

## PANORAMIC MOSAICS OF SLIT IMAGES WITH DEPTH

### 1. Sampling and construction

The panoramic mosaics of slit images with depth is sampled as Fig.2(a) shows. We use a single off-centered camera that rotates along a circle (called the sample circle) to sampling the scene, and the direction of camera is always along the normal direction of the circle. At each rotation angle, a slit image is captured. (An alternative solution is to capture a normal image, and pick out the middle column of the image.) We call these captured slit images sampled slit image and a sampled slit image with its recovered united depth is called sampled slit image with depth. All sampled slit images are stitched together to construct a panoramic mosaics.

The next problem is how to get the united depth of the sampled slit images. One solution is to sample two panoramic mosaics of slit images in two concentric circles with different radius, so that every slit image in one panoramic mosaic can find its analogical slit image in the other mosaic. The depth of every slit image in the mosaic can then be recovered using the analogical slit image method introduced in Section 2.2. We have developed an algorithm that can automatically find the corresponding point between analogical slit images (Jiang et al., 2000). The corresponding relations can also be interactively specified. The depth of pixels can be derived from Equation 2. Then a small depth value range, which can best approach the depth distribution of the slit image, is selected to be the united depth of the sampled slit image. If the scene geometry is simple, the depth can even be specified directly from a predefined map. Note that for a sampled slit image with depth, the united depth of a slit image is not assigned by a single value, but by a small range. It should al-

so be noted that although the sampled slit images with depth are only one pixel wide, they still have a small horizontal field of view (FOV). These make those sampled slit images with depth look like some colored "fine pillars" with volume rather than some one-dimension lines, as shown in Fig. 2(b). Sampled slit images with depth can be treat as a set of ideal slit images with depth with the same pixel values. In the following, the term "slit image" is used to denote ideal (1-D) slit images with depth for convenience.

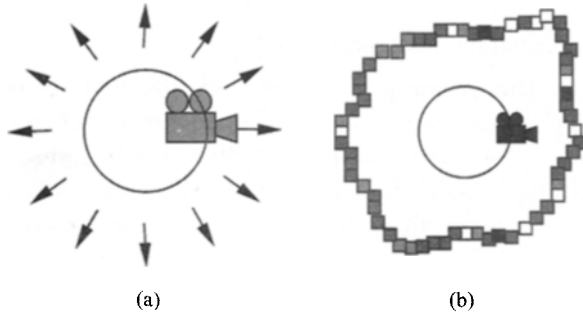


Fig. 2 The sampling of panoramic mosaics of slit images with depth

(a) sampling method; (b) sampled slit image

## 2. Rendering using single panoramic mosaic of slit images with depth

### (1) Mapping slit images (Fig.3)

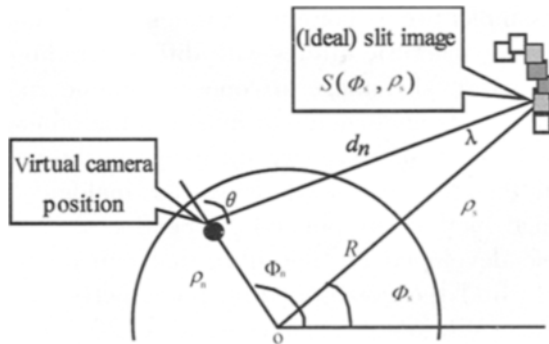


Fig.3 Mapping from slit images in the panoramic mosaic to the virtual camera

In a single panoramic mosaic of slit images with depth, every (ideal) slit image is identified by polar coordinates, written as  $S(\varphi, \rho)$ . The origin point of the polar coordinates  $O$  is located at the center of sampling circle. where  $\rho$  is the "distance" from slit image to the origin point  $O$ , and is equal to the united depth value of slit image plus the radius  $R$  of sampling circle.

Now if the virtual camera located at position  $(\varphi_n, \rho_n)$ , the relations between the virtual camera and the slit images are:

$$\begin{cases} \frac{\rho_s}{\sin(\pi + \theta)} = \frac{d_n}{\sin(\varphi_n - \varphi_s)} = \frac{\rho_n}{\sin\lambda} \\ \pi + \theta + \varphi_n - \varphi_s + \lambda = \pi \end{cases} \quad (5)$$

Where  $\theta$  is the angle of view where the slit image  $S(\varphi_s, \rho_s)$  will appear in the novel view,  $d_n$  is depth value of a slit image in the virtual camera, and  $\lambda$  is the angle of incidence to the slit image. Using Equation 5, we can define a forward mapping from slit images in the panoramic mosaic to the virtual camera, and construct the novel view.

### (2) Synthesizing images

Normally, the novel view is not a panoramic one, it is specified by a horizontal field of view (FOV)  $\omega$  and the virtual camera direction  $\theta_n$ . That is, after the mapping, only slit images map to the region  $\pi - \omega/2 \leq \theta \leq \pi + \omega/2$  is needed. The mapped slit images are then scaled up/down to simulate the looming parallax as introduced in Section 2.1.

Holes and overlaps are key problems with forward mapping. Hole here means on one angle of view  $\theta$  in the novel view, no slit image is mapped to. Overlap here means more than one slit image is mapped to one angle of view  $\theta$  in the novel view. To solve the overlap problem (note that  $d_n$  in Equation 5 is depth value of a slit image in the virtual camera) we can pick the slit image with minimum  $d_n$  for this angle of view. For the hole problem, we can use a method similar to (Chen et al., 1993), fill the hole by interpolating nearby slit images.

Holes are caused by insufficient sampling. The interpolation is just an inaccurate method. To solve this problem, we need more scene information.

Although there is no explicit limit to the motion of the virtual camera in the 2-D plane, a single panoramic mosaic of slit images can offer just a finite sample of scene. The farther the virtual camera leaves the sampling circle center, the fewer are the slit images that can be used for the novel view. This will cause the degradation of image quality and potentially limit the motion range of the virtual camera.

### 3. Rendering using multiple panoramic mosaics of slit images with depth

One advantage of panoramic mosaic of slit images with depth is it supports the joining up of multiple panoramic mosaics of slit images with depth (Fig.4). Holes may be filled by new information, slit images from different mosaics can prevent the degradation of image quality and expand the potential motion range of the virtual camera.

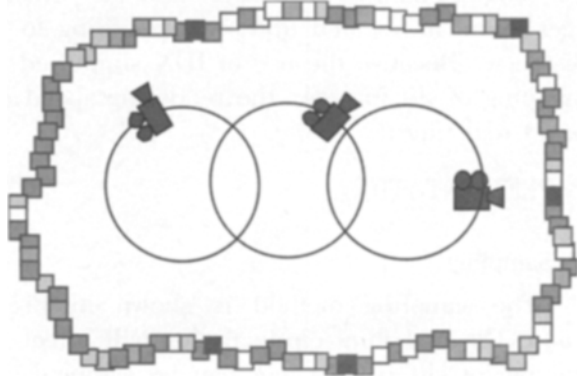


Fig.4 Multiple panoramic mosaic of slit images with depth

However, the forward mapping will consume more time as the number of slit images increase. This will slow down the rendering process. To solve this problem, we define a two-stage mapping. First, in the pre-mapping stage, we do forward map of slit images to some points called reference points. Then in the image synthesizing stage, we use an inverse mapping to find the slit images needed for a novel view.

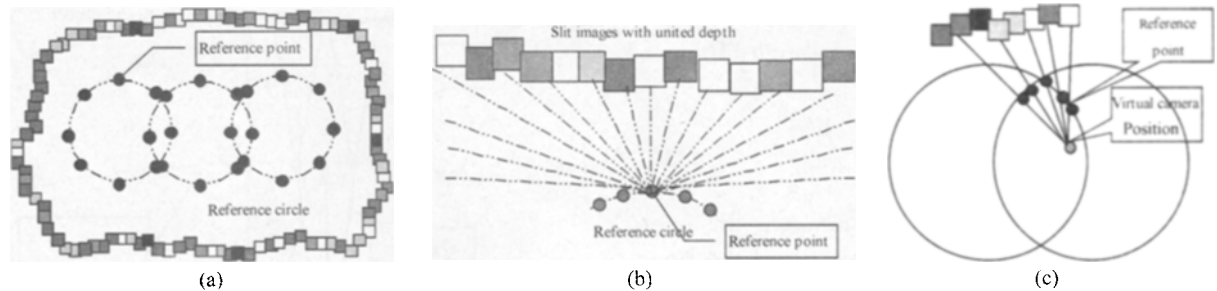


Fig.5 A two-stage mapping process in multiple panoramic mosaic of slit images with depth

(a) define reference circles and reference points; (b) slit images that can be seen outside the circle from a reference point; (c) at each viewpoint, find the corresponding reference points

For every reference point on each reference circle, every angle of view (note the angle of view is also discrete) defines a data structure to record the slit images:

#### (1) Pre-mapping

As shown in Fig.5(a), we first define some reference circles in the scene. The virtual camera motion will be limited to inside those circles. Some points along every circle are uniformly picked out to be the reference points. For clarity, only a few reference points are drawn in figure 5 (a), but in practice a large number of reference points in each reference circle will be needed. At each reference point, all slit images that can be seen "outward" the circle (Figure 5b) are recorded. The process can also be considered as putting a virtual camera on the reference point. The virtual camera direction is along the normal direction of the reference circle and horizontal FOV is  $\pi$ . Then find slit images for the novel scene are found just like in Section 2. Instead of generating a novel image, slit images and their distance (i.e.,  $d_n$  in Equation 5) are recorded.

calculate the distance current dis from  $s$  image to  $r$  point;

if current dis is less than the distance value in current IDX ( $r$  circle,  $r$  point,  $r$  angle)

```
{
  record  $s$  image to IDX ( $r$  circle,  $r$  point,  $r$  angle), update distance;
  IDX ( $r$  circle,  $r$  point,  $r$  angle), ( $s$  image; current dis);
}
```

```
}
```

```
}
```

```
}
```

```
}
```

```
}
```

```
}
```

```
}
```

slit-image-ID;

```
}
```

distance;

The IDX structure is indexed by the three parameters reference-circle, reference-point, view-angle as described above. Its content is a global ID of a slit image and its distance to the reference point. Now if a slit image whose global ID is  $s$ -ID, the angle of view relative to a reference point  $r$ -point is  $v$ -angle, and the distance from the slit image to the slit image is  $dis$ , record these as follows:

IDX ( $r$ -circle,  $r$ -point,  $v$ -angle): ( $s$ -ID;  $dis$ )  
 $r$ -circle is the reference circle to which  $r$ -point belongs to. The angle of view is defined as the angle between the normal direction of the reference circle at the reference point and the line from the reference point to the slit image. The counter-clockwise direction is used as the positive direction, so that a valid ("outward") angle of view should be between  $-\pi/2$  and  $\pi/2$ . Note that the IDX recorded slit image will appear at the given reference point, from the given viewing direction.

If more than one slit image appears at the same view angle relative to a reference point, only the slit image with the minimum distance is recorded.

An algorithm for this step is presented as follows, using a C-like style:

```
init all IDX, set distance to be maximum distance value, set slit-image-ID to be null;
for every slit image  $s$ -image in scene
{
```

```
    for every reference point  $r$ -point in every reference circle  $r$ -circle
```

```
    {
        calculate the viewing angle  $v$ -angle from  $s$ -image to  $r$ -point;
        if  $v$ -angle  $\in (-\pi/2, \pi/2)$ 
```

## (2) Images Synthesizing

While the virtual camera moves, for every angle of view in the field of view, find the nearest reference point (Fig. 5c) and calculate the angle between the view line and the normal direction of the reference circle at that reference point. Now if the nearest reference point is  $n$ - $r$ -point, belong to the reference circle  $n$ - $r$ -circle the angle between the view line and the normal direction of the reference circle at that reference point (also called the angle of incidence for that

reference point) is  $i$ -angle, the slit image can be found from the IDX structure IDX( $n$ - $r$ -circle,  $n$ - $r$ -point,  $i$ -angle).

Note the viewpoint, reference point and the slit image here are in one line. The distance between the current virtual camera and the slit image is equal to the distance from virtual camera to the reference point plus the distance field of the IDX( $n$ - $r$ -circle,  $n$ - $r$ -point,  $i$ -angle).

The novel view is synthesized using almost the same method as that in Section (2) Slit images found are scaled up/down according to the distance. Because the use of IDX simplified the mapping of slit images, the rendering speed can be in real time.

## IMPLEMENTATION

### 1. Sampling

The sampling method is shown in Fig. 2 (a). The sampling circle radius will affect the number of slit images that can be sampled. As shown in figure 6, if the maximum number of slit images that can be sampled in a circle is  $N$ , there will be the following relation:

$$\text{tg}(\pi/N) \approx 2\text{tg}(\alpha/2)(z - R)/z \quad (6)$$

Where  $R$  is the sampling circle radius,  $z$  is the distance between the scene object and sampling circle center, and  $\alpha$  is the horizontal field of view of a slit image. If the number of slit images in a sampling circle is more than  $N$ , the scene view of neighboring slit images will overlap and no new information can be used.

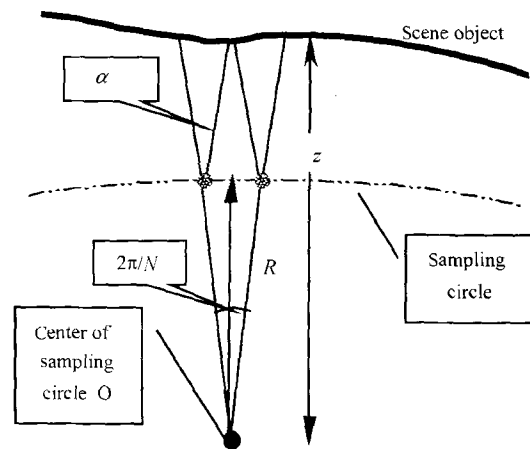


Fig. 6 The maximum number of slit images that can be sampled in a circle

A bigger radius of sampling circle can gather more information in the horizontal direction-but as shown in Section 2.2, less information in the vertical direction will be gathered. While sampling, this trade off should be considered.

The depth can be deduced from corresponding points between analogical slit images or imported from a predefined scene map. If the latter is used, depths must be calculated using plane geometry.

## 2. Preprocessing

The preprocessing stage is used when using multiple panoramic mosaics for rendering. The process should:

(1) Identify every sampled slit image with depth by a global ID composed of the sampling circle ID and the camera azimuth when the slit image is captured.

(2) Position all slit images captured at different sampling circles to one united coordinate.

(3) Specify the reference circle and generate the reference point.

(4) Using algorithm described in Section 3.3.1, generate the IDX structures.

Note that while generating IDX structure, some IDX may find no slit image mapped to this reference at this angle of view (i.e., after executing the algorithm in Section (1), the resulting Slit-image-ID field in an IDX is still null), as the "hole" described in Section (2). If this

occurs, the IDX will be tagged as "hole".

## 3. Interactive Rendering

For a single panoramic mosaic, a forward mapping process is executed in the interactive rendering stage. Because the number of slit images is small, the forward mapping process given below will not take much time.

(1) Decide slit images in the novel view using forward mapping. Calculate the slit images' depth value (distance) relative to the virtual camera. Compare the depth value to solve the overlap problem. Tag holes.

(2) Scale slit images up/down according the depth value to simulate the looming parallax. Copy them to the novel image.

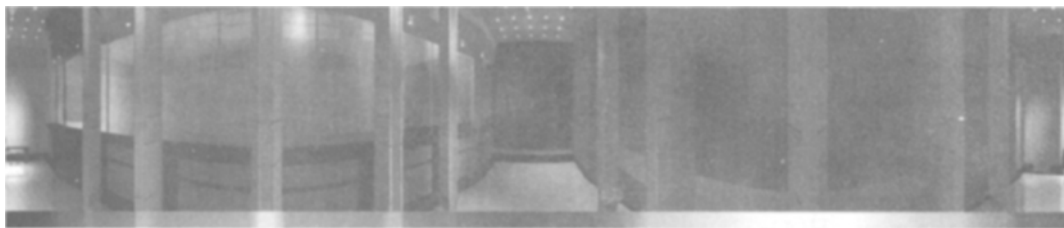
(3) Fill holes by interpolating the nearby slit images.

For multiple panoramic mosaics, only the first step is different. That is:

(1) For every angle of view of the novel view, find the nearest reference point and calculate the angle of incidence. Find corresponding IDX. From the IDX retrieve the slit image ID. Calculate the depth for every slit image will be used.

(2) Scale slit images up/down according the depth value to simulate the looming parallax. Copy them to the novel image.

(3) Fill holes by interpolating the nearby slit images.



(a)



(b)

Fig. 7 Two panoramic mosaics of slit images with depth, the strip below indicates the depth, darker is farther

An example of panoramic mosaics of slit images with depth is shown in Fig. 7. It was captured from a hall. The novel view alone with virtual camera move is shown in Fig. 8. We gener-

ate the novel views using the algorithm described above, Fig. 7 shows the panoramic mosaics of slit images with depth.

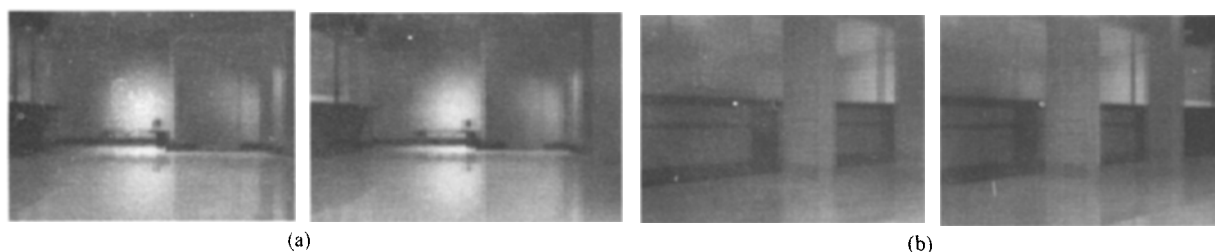


Fig. 8 Novel view from the panoramic mosaics of slit images with depth  
(a) moving forward and backward; (b) moving left and right

## CONCLUSIONS

In this paper, we present a new image based rendering technology called panoramic mosaics of slit images with depth. Compared with panorama, this method can support the movement of the virtual camera. Multiple panoramic mosaics of slit images with depth can be joined up to provide a wider range of virtual camera movement. The panoramic mosaics of slit images with depth are also easy to capture, and the data size is as small as panorama.

By assuming most pixels in slit images have small depth variance, we use a united depth for the slit images. This assumption can not be used in some complex scenes. This method is most suitable for large scenes.

## References

- Brown, L. G., 1992. A survey of image registration techniques. *Computing Surveys*, **24**(4):325 - 376.
- Chai, J. X., Tong, X., Chan, S. C. et al., 2000. Plenoptic sampling, *In: Proc SIGGRAPH*.
- Chen, S. E., 1995. QuickTime VR-an image-based approach to virtual environment navigation. *Computer Graphics (SIGGRAPH'95)*, p. 29 - 38.
- Chen, S. and Williams, L., 1993. View interpolation for image synthesis. *Computer Graphics (SIGGRAPH'93)*, p. 279 - 288.
- Gortler, S. J., Grzeszczuk, R., Szeliski, R. et al., 1996. The lumigraph. *In: Computer Graphics (SIGGRAPH'96)* p.43 - 54.
- Levoy, M., Hanrahan, P., 1996. Light field rendering. *In Computer Graphics (SIGGRAPH'96)* p.31 - 42.
- McMillan, L., Bishop, G., 1995. Plenoptic Modeling: An Image-Based Rendering System, *Computer Graphics (SIGGRAPH'95 Conference Proceedings)*, p. 39 - 46.
- Peleg, S., Herman, J., 1997. Panoramic mosaics by manifold projection. *In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'97)*, p. 338 - 343.
- Rademacher, Paul. Bishop, Gary, 1998. Multiple-Center-of-projection Images. *In: Computer Graphics (SIGGRAPH'98)*, p.199 - 206.
- Seitz, S. M., Dyer, C., 1998. View morphing: Synthesizing 3D metamorphoses using image transforms. *In: Computer Graphics (SIGGRAPH'96)*, p. 21 - 30.
- Shum, H. Y., He, L. W., 1999. Rendering with concentric mosaic, *In: Computer Graphics (SIGGRAPH'99)* p.299 - 306.
- Szeliski, R., Shum, H. Y. 1999. Creating full view panoramas image mosaics and environment maps, *In: Computer Graphics (SIGGRAPH'97)*, p. 251 - 258.
- Zheng, J. Y., Tsuji, S., 1990. Panoramic representation of scenes for route understanding. *In: Proc. of the 10th Int. Conf. Pattern Recognition*, p. 161 - 167.