Parallel Projections for Stereo Reconstruction

Jin-Xiang Chai[1] and Heung-Yeung Shum
Microsoft Research, China

Abstract

This paper proposes a novel technique to computing geometric information from images captured under parallel projections. Parallel images are desirable for stereo reconstruction because parallel projection significantly reduces foreshortening. As a result, correlation based matching becomes more effective. Since parallel projection cameras are not commonly available, we construct parallel images by rebinning a large sequence of perspective images. Epipolar geometry, depth recovery and projective invariant for both 1D and 2D parallel stereos are studied. From the uncertainty analysis of depth reconstruction, it is shown that parallel stereo is superior to both conventional perspective stereo and the recently developed multiperspective stereo for vision reconstruction, in that uniform reconstruction error is obtained in parallel stereo. Traditional stereo reconstruction techniques, e.g. multi-baseline stereo, can still be applicable to parallel stereo without any modifications because epipolar lines in a parallel stereo are perfectly straight. Experimental results further confirm the performance of our approach.

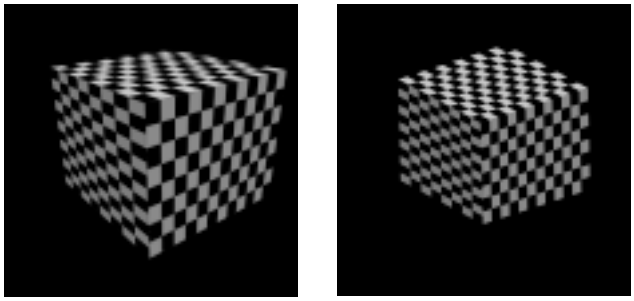
1 Introduction

Stereo reconstruction algorithms have traditionally relied on planar perspective images. A planar perspective image collects many rays at a fixed viewpoint and is easy to acquire using conventional pinhole cameras. However, 3D reconstruction based on perspective stereo has inherent shortcomings due to the *perspective foreshortening* caused by perspective projection.

Perspective foreshortening exists in typical photographic systems and the human visual system. Even though perspective foreshortening provides more realistic visual cues (see Fig.1.a), the exact shape and measurement of the objects cannot be recorded accurately. The size of the perspective projection of an object varies inversely with the distance of that object from the center of projection. Specifically, under perspective projection, parallel lines do not project to parallel lines and length ratio cannot be kept invariant. Previous work on modeling foreshortening (e.g., [4,17]) reveals that foreshortening can worsen correspondence results because 3D space is mapped nonlinearly to 2D image planes.

Foreshortening in perspective stereo also results in highly biased scene reconstruction. Even if the correspondences between two perspective images are available with high precision, the uncertainty of depth reconstruction is highly sensitive to its corresponding depth value in geometry space. An increase in depth value incurs higher range uncertainty. In practice, stereo reconstruction with perspective images is shown to generate strongly biased reconstruction because scene radiance is sampled non-uniformly.

Instead of designing algorithms to deal with foreshortening in perspective projection, we propose in this paper an alternative projection, *parallel projection*, which minimizes foreshortening effect in 3D reconstruction. Parallel projection has long been used to represent the mechanical parts in manufacturing industry using three-orthographic projections. The basic concept of parallel camera has also been applied in the area of computer graphics as an alternative image representation for synthetic scenes.



(a) (b)

Fig.1 Comparison between two projections: (a) perspective projection with significant foreshortening; (b) parallel projection.

Parallel projection has several advantages over perspective projection for image representation. First, as shown in Fig.1b, more accurate shape measurements can be obtained from parallel projection images. Second, in addition to projective invariants existing in perspective projection, parallel projection maintains parallelism and keeps length ratio invariant. Third, parallel projection samples the object space more uniformly than perspective projection. Recently, parallel projection has also been used in the area of *image-based rendering* to represent synthetic objects with multiple layered depth [5,6,8]. It has been shown that sampling uniformity in parallel images can alleviate "hole" artifact for image warping. However, vision algorithms using parallel projection images have not been studied due to the absence of parallel projection cameras in real world.

In fact, parallel projection stereo does not necessarily rely on the practical availability of parallel projection cameras in real world. In this paper, we transform a dense collection of perspective images, specifically 3D or 4D uniform *light field* dataset [3,4,9], into a set of parallel projection images. Light field data can be acquired by moving a camera uniformly along a line or on a planar grid.

[1] Currently at Robotics Institute, Carnegie Mellon University.

1.1 Related work

The work described in this paper is closely related to current research trends in multi-perspective projection for image-based representation, including *multiple-center-of-projection images* [12], *manifold mosaics* [11], and *multi-perspective panoramas for image based rendering [13], cel animation* [14], and *stereo reconstruction from multi-perspective panoramas* [14,15,19]. Similar to [14], we select the best rays for vision reconstruction. But our work differs from these previous multi-perspective approaches in that the parallel projection images are the most desirable for stereo reconstruction. As will be shown in this paper, parallel stereo enables better correspondences, allows more accurate measurement for depth estimation, and results in an epipolar geometry that is absolutely parallel between any two parallel projections. In contrast, the previous work on multiperspective stereo reconstruction in [15] has only approximately straight epipolar lines, thus limiting the use of traditional multi-baseline stereo techniques. As we will show in this paper, similar to perspective stereo, multiperspective stereo has non-uniform depth reconstruction uncertainty.

Another thread of research related to this paper is on image based rendering [3,7,13] based on light field data. However, we focus on a different aspect: how to create the depth from light field. The emergence of large and dense collection of image datasets suggests a powerful new paradigm for solving traditional graphics and vision problems. In this paper, rather than designing vision algorithms to deal with light data set, we resample light field data to obtain new images that are the most desirable for stereo reconstruction.

1.2 Overview

We address the following questions in the remainder of this paper. How are parallel images captured and generated from real perspective images? Given two parallel images, how are epipolar geometry and parallax computed? What is the reconstruction geometry involved? We provide answers to these questions using 1D parallel projection (camera moving along a line) and 2D parallel projection (camera moving on a planar grid) in Sections 2 and 3, respectively. We study the reconstruction uncertainty of parallel stereo in Section 4. Also in Section 4 we compare *parallel stereo* with two types of stereo setups for vision reconstruction, specifically, *single perspective*, and *multi-perspective*. Finally we present our experimental results on synthetic and real images in Section 5 and conclude in Section 6.

2 1D Parallel Stereo

2.1 1D parallel images from perspective images

We first describe how 1D parallel images from perspective image sequences are created. We constrain a perspective camera to move uniformly along a line, described by *s*. A large collection of single perspective images (indexed by image coordinates u and v) , denoted as $I_S(u,v)$ here, can be sampled along the *s* direction.

For a specific u, we can rebin a set of perspective images $I_s(u,v)$ into a parallel image $I_u(s,v)$, shown in Fig.2. The rebinning process is similar to constructing EPI [1].

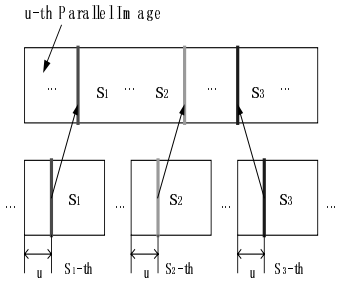


Fig.2 An image of 1D parallel projection rebinned from a sequence of perspective images.

Let *S* denote the total number of samples along the s direction, and *U* and *V* the width and height of the sampled perspective image respectively. Obviously, the width of the 1D parallel image is equal to the number of sample points along the camera trajectory. The height *V* of the original perspective image is the same as the height of the 1D parallel image. Fig. 3 shows a pair of parallel images rebinned from the u1-th and u2-th columns of original perspective images that can be used for stereo reconstruction.

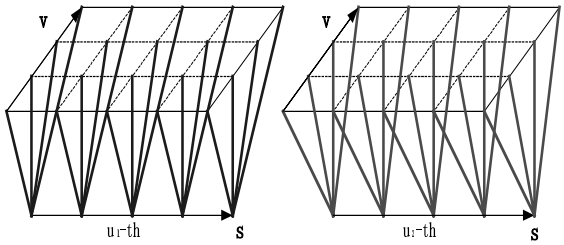


Fig.3 A pair of parallel images for stereo reconstruction rebinned from two columns, u1 and u2, of original images.

2.2 Imaging Geometry

The imaging geometry from two 1D parallel images $I_{u1}(s,v)$ and $I_{u2}(s,v)$ is shown in Fig.4. We set the origin of the world Cartesian co-ordinate system XYZ at the starting sample point on the camera path and align X and Z axes with the camera trajectory and the vertical direction of image plane. If we can establish matching between the two parallel images, the depth of corresponding points in the scene can be estimated in a similar way as the conventional perspective stereo.

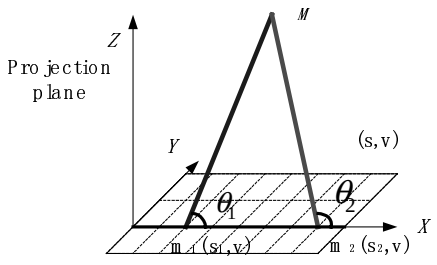


Fig.4 Imaging geometry of 1D parallel stereo.

Using the basic law of *sines* for triangles, we have

$$\frac{\Delta s \cdot (s_2 - s_1)}{\sin(\theta_2 - \theta_1)} = \frac{(x - s_1 \cdot \Delta s)/\cos\theta_1}{\sin\theta_2} = \frac{z/\sin\theta_2}{\sin\theta_1}. \quad (2.1)$$

Then,

$$x = \frac{\Delta s \cos\theta_1 \sin\theta_2 (s_2 - s_1)}{\sin(\theta_2 - \theta_1)} + s_1 \Delta s \quad (2.2)$$

$$y = \frac{v \Delta s \sin\theta_1 \sin\theta_2 (s_2 - s_1)}{f \cdot \sin(\theta_2 - \theta_1)} \quad (2.3)$$

$$z = \frac{\Delta s \sin\theta_1 \sin\theta_2 (s_2 - s_1)}{\sin(\theta_2 - \theta_1)}, \quad (2.4)$$

where Δs denotes the sample interval along the camera path, f is the camera focal length, and

$$\sin\theta_1 = \frac{u_1}{\sqrt{u_1^2 + v_1^2 + f^2}} \quad (2.5)$$

$$\sin\theta_2 = \frac{u_2}{\sqrt{u_2^2 + v_2^2 + f^2}} \quad (2.6)$$

Therefore, for a point at distance *z* seen in two parallel images I_1 and I_2, the horizontal and vertical parallax can be described as follows:

$$s_{2:1} = s_2 - s_1 = z * \frac{\sin(\theta_2 - \theta_1)}{\Delta s \sin\theta_1 \sin\theta_2} \quad (2.7)$$

$$v_{2:1} = v_2 - v_1 = 0. \quad (2.8)$$

The above horizontal epipolar line constraint (Eq.2.8) enables traditional stereo reconstruction techniques, e.g. multi-baseline stereo, to be applicable to 1D parallel stereo without any modifications. We can think of Δs as the focal length in perspective stereo. The larger the sample interval, the higher the resolution in the resulting parallel image.

We call the angle between the directions of projection related to a pair of parallel images *base-angle*. Similar to the *baseline* in perspective stereo, the *base-angle* of a parallel stereo setup is kept invariant.

2.3 1D parallel projective invariant

Projective invariants are shape descriptors that are independent of the point of view from which the shape is seen, and, therefore, are of major importance in stereo, object recognition and scene reconstruction. Projective transformation preserves type (i.e., a point remains a point and a line remains a line), incidence (i.e., whether a point lies on a line), and cross ratio. However, it does not preserve length, angle, and parallelism that are useful for measuring shapes.

Parallel projection has more measures kept invariant than perspective projection. For 1D parallel projection, it can keep the ratio of length invariant, in addition to type, incidence and *cross ratio* in perspective projection. As shown in Fig.5, let $\mathbf{M}_i=(x_i,y_i,z_i)$, i=1,2,3 be three points on the same spatial line, the length ratio defined as

$$L_r(M_1, M_2, M_3) = \frac{\overline{M_1 M_2}}{\overline{M_1 M_3}} \quad (2.9)$$

is preserved because M_1m_1, M_2m_2 and M_3m_3 are parallel to each other.

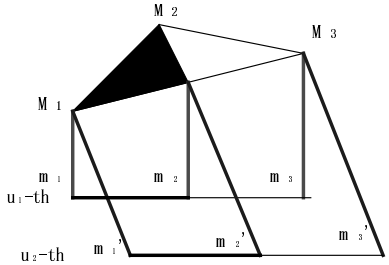
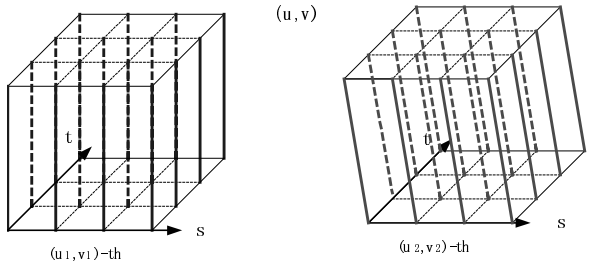


Fig.5 1D parallel projective invariant: *length ratio.*

3 2D Parallel Stereo

3.1 Parallel images from 4D light field

Similar to the formation of 1D perspective images, we can obtain 2D parallel images from perspective images. The difference is that while a 1D parallel image is implemented slit by slit from a line of images, a 2D parallel image is resampled pixel by pixel from a 2D grid of images. These 2D parallel projection images can be resampled directly from a 4D light field [7] or Lumigraph [3]. Lumigraph characterizes the flow of light through unobstructed space in a static scene with fixed illumination. Here, we follow the same definition and denotation of light field as in [7] where *s* and *t* denote the camera trajectory, and *u* and *v* represent image coordinates. Let *U* and *V* be the width and height of each single perspective image, and *S* and *T* be the number of samples along *s* and *t* directions respectively. Furthermore, let $I_{s,t}(u,v)$ and $I_{u,v}(s,t)$ represent the original perspective images and resulting 2D parallel images, respectively.



(a) (b)

Fig. 6 A pair of 2D parallel images for vision reconstruction rebinned from a grid of perspective images.

The total number of the resulting 2D parallel images is $U \times V$. The 2D parallel images are represented by the normal of projection plane *(s,t)* and the direction of projection *(u, v, f)*, where f denotes the focal length. Figure 6 shows two parallel projection images rebinned from a light field at the (u_1, v_1)-th and (u_2, v_2)-th directions.

3.2 Epipolar geometry in 2D parallel stereo

An important practical application of epipolar geometry is to aid the search for corresponding points, reducing it from the entire second image to a single epipolar line. In

multiperspective stereo [10,14,15,19], the horizontal epipolar line is only available between a symmetric pair of concentric multiperspective stereo with the same radius. In contrast, the epipolar lines among multiple 2D parallel images are absolutely horizontal. Therefore, any traditional stereo algorithms are applicable to parallel stereo without modification, *e.g.* multi-baseline stereo technique [9]. The recently developed *plane sweep algorithm* [2] can also be applied to parallel stereo where the sweeping efficiency can be further improved by encoding affine texture mapping into the warping process.

Fig. 7 shows two parallel images of a rigid scene. The 3D point *M* projects to pixel m_1 in the first image I_1, and m_2 in the second one I_2. Here, two parallel images share the same projection plane *(s,t)* but with different projection directions, denoted as $\vec{l}_1=(u_1,v_1,f)$ and $\vec{l}_2=(u_2,v_2,f)$, respectively.

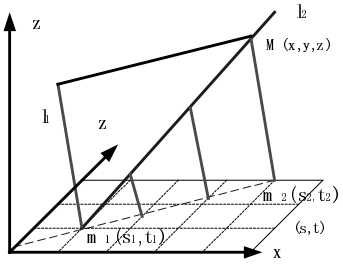


Fig.7 The epipolar line m_1m_2 between two 2D parallel images.

In a standard perspective stereo, the *epipolar plane* is defined by a 3D point *M* and the baseline that links two optical centers. The *epipolar line* is the straight line of intersection by the epipolar plane and the image plane. In contrast, the *epipolar plane* between two parallel images is defined by point *M* and two projection directions forming the *base-angle*, here m_1M and m_2M. The image point *m* in the first parallel image may correspond to an arbitrary spatial point *M* on the ray direction $\vec{l}_1=(u_1,v_1,f)$ (for example, at infinity), and the projection of m_1M in the projection plane can be determined by the intersection between the defined epipolar plane and projection plane (s,t), as shown in Fig. 6.

Specifically, $\vec{l}_1$, $\vec{l}_2$, and epipolar line $\overline{m_1m_2}=(s_2-s_1,t_2-t_1,0)$ satisfy the coplanar constraint,

$$\overline{m_1m_2}\bullet(\vec{l}_1\times\vec{l}_2)=0\,. \tag{3.1}$$

The above equation can be simplified to

$$(v_1-v_2)s_2-(u_1-u_2)t_2=s_1(v_1-v_2)-t_1(u_1-u_2)\,. \tag{3.2}$$

We observe that the epipolar line between two parallel images indeed keeps parallel. Particularly, it becomes horizontal or vertical when either $u_2=u_1$ or $v_2=v_1$, respectively.

3.3 Imaging geometry

Fig.8 shows a 3D point M(x,y,z) and two parallel images that share the same projection plane (s,t). The directions of projection of two corresponding pixels are determined by their directional cosine, l_1 and l_2 respectively, as follows:

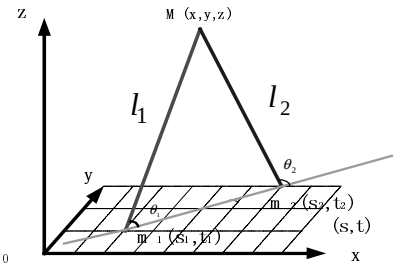


Fig.8 Imaging geometry in 2D parallel stereo.

$$l_1=(\cos\alpha_1,\cos\beta_1,\cos\gamma_1)=(u_1,v_1,f)/\sqrt{u_1^2+v_1^2+f^2} \tag{3.3}$$

$$l_2=(\cos\alpha_2,\cos\beta_2,\cos\gamma_2)=(u_2,v_2,f)/\sqrt{u_2^2+v_2^2+f^2} \tag{3.4}$$

Each point in the field of view is projected onto its projection plane (s,t) along its projection direction. Again, with the basic law of sines, we have

$$x=\frac{\cos\alpha_1\sin\theta_2\sqrt{(s_2-s_1)^2+(t_2-t_1)^2}}{\sin(\theta_2-\theta_1)}+s_1 \tag{3.5}$$

$$y=\frac{\cos\beta\sin\theta_2\sqrt{(s_2-s_1)^2+(t_2-t_1)^2}}{\sin(\theta_2-\theta_1)}+t_1 \tag{3.6}$$

$$z=\frac{\cos\gamma\sin\theta_2\sqrt{(s_2-s_1)^2+(t_2-t_1)^2}}{\sin(\theta_2-\theta_1)} \tag{3.7}$$

where θ_1 and θ_2 are determined by

$$\sin\theta_1=\frac{\sqrt{(u_2v_1+u_1v_2-2u_1v_1)^2+(u_2-u_1)^2f^2+(v_2-v_1)^2f^2}}{\sqrt{(u_2-u_1)^2+(v_2-v_1)^2}\sqrt{u_1^2+v_1^2+f^2}} \tag{3.8}$$

$$\sin\theta_2=\frac{\sqrt{(2u_2v_2-u_1v_2-u_2v_1)^2+(u_2-u_1)^2f^2+(v_2-v_1)^2f^2}}{\sqrt{(u_2-u_1)^2+(v_2-v_1)^2}\sqrt{u_2^2+v_2^2+f^2}} \tag{3.9}$$

Substituting epipolar geometry Eq. (3.2) into Eq. (3.7), we get the horizontal and vertical parallax, respectively,

$$s_{2:1}=\frac{z\sin(\theta_2-\theta_1)}{\cos\gamma\sin\theta_2\sqrt{(\frac{v_2-v_1}{u_2-u_1})^2+1}} \tag{3.10}$$

$$t_{2:1}=\frac{z\sin(\theta_2-\theta_1)}{\cos\gamma\sin\theta_2\sqrt{(\frac{u_2-u_1}{v_2-v_1})^2+1}}\,. \tag{3.11}$$

Unlike in the perspective stereo, the disparity in parallel stereo (both directions) is linear to the depth value.

3.4 Projective invariant in 2D parallel projection

Besides length ratio, parallelism is also invariant under 2D parallel projection, as illustrated in Fig.9.

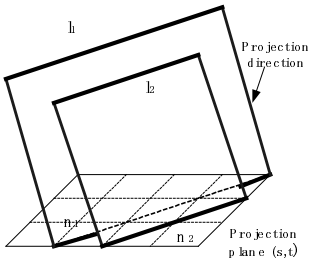


Fig.9 2D parallel projective invariant: *parallelism.*

Invariant measures such as length ratio and parallelism make correlation much easier than in perspective stereo where perspective foreshortening could be significant.

4 Depth Uncertainty Analysis

4.1 Depth Uncertainty in parallel stereo

Because of the discrete nature of the imaging system, the image coordinates of each pixel may have quantization error of up to $\pm 1/2$ pixel that, in turn, affects depth estimation.

From Equation (2.4), the depth uncertainty for 1D parallel stereo can be estimated as

$$|\Delta z| = \Delta s \cdot \Delta s_{2:1} \cdot \left| \frac{\sin\theta_1 \sin\theta_2}{\sin(\theta_2 - \theta_1)} \right| \qquad (4.1)$$

where noise $\Delta s_{2:1}$ results primarily from quantization error along the s direction. Several observations can be made from Eq.(4.1):

- There is no relationship between depth uncertainty and depth value.
- The contribution to the depth estimation error is closely related to the sampling interval Δs (and Δt in 2D parallel stereo) in light field sampling. The larger the sampling interval, the larger the depth uncertainty.

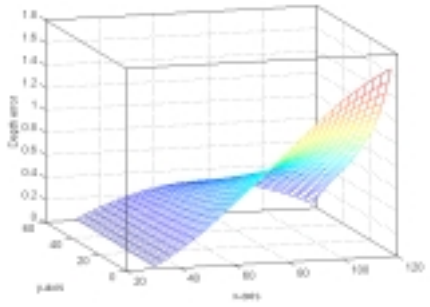


Fig.10 Depth error for different 1D parallel stereo. Here X axis denotes θ_1, Y axis denotes $\theta_2 - \theta_1$, and Z axis denotes the depth error involved.

- The uncertainty also depends on the *base-angle* between two parallel projections for vision reconstruction. Fig.10 illustrates the error associated with different values of $(\theta_1, \theta_2 - \theta_1)$. The error is well under control when a large *base-angle* is used. In particular, a minimum depth error occurs when $(\theta_1 = \theta_2 - \pi/2)$. However, with a large base-angle, only objects close to the cameras can be reconstructed.

Similarly, we can derive the depth uncertainty for 2D parallel stereo from Equation (3.7),

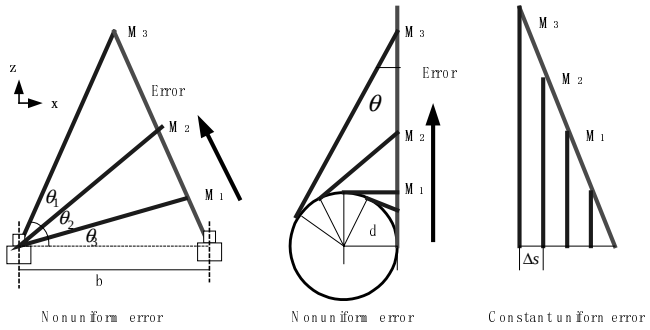
$$|\Delta z| = \left(\frac{|s_{2:1}|\Delta s_{2:1}\Delta s + |t_{2:1}|\Delta t_{2:1}\Delta t}{\sqrt{s_{2:1}^2 + t_{2:1}^2}} \right) \left| \frac{\cos\gamma_3 \sin\theta_6}{\sin(\theta_6 - \theta_5)} \right| \qquad (4.2)$$

where noise $\Delta s_{2:1}$ and $\Delta t_{2:1}$ are due mainly to quantization error along s and t directions, respectively. Substituting Eq. (3.8) and Eq.(3.9) into Eq.(4.2), we observe that the depth error does not depend on the actual depth value either.

4.2 Comparison with other stereo setups

In this section we show that parallel projection stereo is superior to other stereo techniques available today in its robust depth uncertainty. Depth estimation error in parallel stereo does not vary with the changes in depth value, but rather remains constant under different depth values. In other words, scene depth is reconstructed uniformly in all depth ranges, yielding uniform depth resolution and depth reconstruction error. In contrast, stereo reconstruction obtained from perspective or multi-perspective images are shown to generate strongly biased scene reconstruction (see fig.11).

In the following, we analyze in detail the depth uncertainty to noise for traditional perspective stereo and recently developed multi-perspective stereo.



(a) (b) (c)

Fig.11 Three typical stereo setups and their related depth reconstruction error: (a) traditional perspective stereo with non-uniform depth reconstruction error; (b) multiperspective stereo with non-uniform depth reconstruction; (c) parallel stereo with uniformly controlled depth reconstruction.

4.2.1 Depth error in perspective stereo

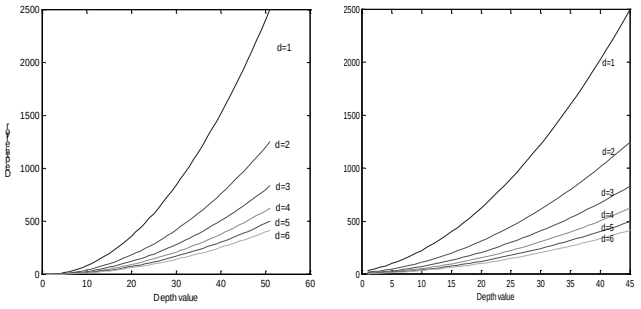
In a standard perspective stereo setup (see Fig.11a),

$$z = \frac{bf}{u_{2:1}} \quad . \tag{4.3}$$

Thus the derived depth uncertainty is proportional to the square of its related depth value,

$$|\Delta z| = \frac{\Delta |u_{2:1}| bf}{u_{2:1}^2} = |\Delta u_{2:1} * (\frac{z^2}{bf})| \tag{4.4}$$

where b denotes the baseline between two perspective images, $\Delta u_{2:1}$ is the error associated with quantization error of $u_{2:1}$. Fig.12a illustrates the theoretical error bounds versus depth value for several different baseline values. We observe that the larger the baseline, the smaller the depth error. Furthermore, the larger the depth, the larger the depth uncertainty.



(a) (b)

Fig12. Depth uncertainty for: (a) standard perspective stereo; (b) multiperspective stereo.

4.2.2 Depth error in multiperspective stereo

Recently developed multiperspective stereo [14] can yield isotropic depth reconstruction by sampling the scene radiance in all directions. However, the reconstruction accuracy of every point in the scene is not truly uniform. Uniform reconstruction of 3D points is limited to 2D manifold with the same radius, but not along the radius direction.

For a multiperspective stereo (see Fig.11b),

$$r = \frac{d}{\sin(\theta_{2:1}/2)} \tag{4.5}$$

$$|\Delta r| = \frac{\Delta\theta_{2:1} \cdot d \cdot \cos(\theta/2)}{\sin(\theta/2)^2} = \Delta\theta_{2:1} * (\frac{r\sqrt{r^2 - d^2}}{2d}), \quad r \geq d \tag{4.6}$$

where $\Delta\theta_{2:1}$ represents the error related to image resolution along the angle direction, and d is the radius for multi-perspective images. Fig.12.b shows the depth uncertainty associated with multi-perspective stereo.

We can observe that the reconstruction error increases as the radius of spatial point becomes larger. This is somewhat similar to traditional perspective stereo, whose reconstruction uniformity is also limited to 2D surface related to the same depth value. But parallel stereo can provide uniform reconstruction in 3D space because the reconstruction error from parallel stereo is unrelated to specific locations of spatial points.

5. Experiments

In our experiments, the standard correlation based technique with fixed window sizes is used for stereo reconstruction.

5.1 Synthetic data

We have rendered 512 perspective images (with size of 640×480) by moving a camera uniformly along a line. These images are then put together to form two parallel images (512×480), shown in fig.13.a and 13.b, respectively. These two 1D parallel images are created by stacking the 320-th and the 350-th column, respectively, from the original image sequences. The choice of base-angle in a parallel stereo depends on the tradeoff between the reconstruction accuracy and the range of scene coverage. The smaller the *base-angle*, the larger the depth uncertainty, and the larger the field of view reconstructed. Additionally, the smaller the *base-angle*, the less the occlusion.

The reconstructed depth map is shown in Fig.13.c, and its associated uncertainty map, defined as the correlation value in the resulting depth map, is given in Fig.13.d. A higher intensity value in the correlation map corresponds to a larger correlation value for the resulting disparity map. Note that a smaller disparity means a smaller depth in parallel stereo, in contrast to any previous stereo reconstruction methods where a small disparity means a larger depth. We can observe that the reconstruction result is fairly good.

5.2 Real data

Experiments were also performed on the MPEG2 *garden sequence.* The total number of images in this sequence is 150. Our tracking program shows the camera motion for the flower garden sequence is approximately linear and equally spaced. Thus, the image sequence can be rebinned to a set of approximate 1D parallel images. Fig.14(a)(b) show a pair of such parallel images rebinned from the 90-th and the 95-th column in the original image sequence. The disparity maps with window sizes 11 and 15 are shown in Fig.14(c)(d). The recovered disparity map is good even though the camera path is only approximately along a straight line and the sampling interval is unknown. These results could be further improved if denser image sequences were available.

5.3 Comparison with traditional perspective stereo

Finally, we compare parallel stereo with perspective stereo in terms of their matching performance using correlation technique. Fig15.(a-b) and Fig15.(c-d) give a pair of images of a chess board cube under parallel and perspective projections respectively. The estimated disparity maps and correlation maps with square template sizes 11 and 23 are shown in Fig.15(e)(f)(i)(j) for parallel stereo and Fig.15(g)(h)(k)(l) for perspective stereo, respectively. The following observations can be made:

1). Parallel stereo has better matching performance than conventional perspective stereo, as supported by "hole" artifacts appeared in disparity map under perspective stereo.

2). In perspective stereo, matching becomes worse when the angle between the surface normal and the ray direction is larger because of foreshortening.

3). For both stereo setups, a larger window size results in a larger local distortion (see the deterioration of their related correlation map).

4). Foreshortening effect in parallel stereo is uniformly distributed, as shown in Fig.16.(d-f). Foreshortening in perspective projection is non-uniform (Fig.16.(j-l)), thus, deteriorating the disparity results.

6. Conclusion

Based on the belief that traditional perspective images are not the most appropriate for solving many computer vision tasks, this paper introduces a new image representation under *parallel projections*. A key idea behind this approach is that the input images or rays for computer vision algorithms can be carefully chosen in order to produce optimal results. Rather than physically constructing parallel projection cameras, we describe in this paper how to create parallel projection images by processing a large sequence of perspective input images from light field.

We classify the parallel projections into two types: 1D and 2D parallel projections. In a 1D parallel projection image, the image is under parallel projection only along one dimension (along which the camera is moving) whereas the other is under perspective projection. In 2D parallel projection images, the images are under parallel projection in both directions. Parallel projection images represent the appearance of a scene from many perspective viewpoints, as opposed to traditional single perspective viewpoints. They differ from recently developed multiperspective images in that their pixel rays are parallel to each other.

The epipolar geometry, parallax, range reconstruction and projective invariant for both 1D and 2D parallel projection stereos have then been studied. Our theoretical analysis on depth reconstruction uncertainty shows that the performance of the new stereo is superior to that of the traditional perspective stereo and the more recently developed multiperspective stereo. More importantly, our experimental result shows that parallel stereo produces accurate 3D scene models with the smallest reconstruction error compared to any stereo reconstruction available today.

The essential idea of parallel projection image representation described here opens up many other applications in image based modeling and rendering, including layered depth image, high resolution images composition, and view-dependent texture mapping, etc. These are very exciting topics for our future work.

References

[1] R.Bolles, H.H.Baker, and D.H.Marimont. Epipolar Plane Image Analysis: An Approach to Determing Structure From Motion. In *International Journal of Computer Vision, 1:7-55. Boston, 1987.*

[2] R.T.Collins. A space-sweep approach to true multi-image matching. In *IEEE Comp. Soc. Conf. On Computer Vision and Pattern Recognition(CVPR'96*), pages 358-363, San Francisco, June 1996.

[3] S.J.Gortler, R.Grzeszczuk, R.Szeliski, and M.F Cohen. The Lumigraph. In *Computer Graphics Proceedings*, pages 43-54, Proc.SIGGRAPH'96 (Orlando), August 1996.

[4]. Mark W. Maimone and Steven A. Shafer, Modeling the Foreshortening in Stereo Vision using Local Spatial Frequency, *CMU Computer Science Technical Report TR95*. Carnegie Mellon University, January 1995 (32 pages)

[5]. N.Max. Hierarchical rendering of trees from precomputed multi-layer Z-buffer, *Rendering Techniques'96 (Proc. 7th Eurographics Workshop on Rendering*, pages 165-174, Springer-Verlag, 1996.

[6]. N.Max, C.Mobley, B. Keating, and E-H Wu,Plane-parallel radiance transport for global illumination in vegetation. *Rendering Techniques'97 (Proc. 8th Eurographics Workshop on Rendering)*, pages 239-250, Springer-Verlag, 1997.

[7] M.Levoy and P.Hanarahan. Ligh Field Rendering. In *Computer Graphics Proceedings*, pages 31-42, Proc.SIGGRAPH'96 (Orlando), August 1996.

[8]. D.Lischinski and A. Rappoport, Image Base Rendering for non-diffuse Synthetic Scenes. *Rendering Techniques'98* (*Proc. 9th Eurographics Workshop on Rendering*), Springer-Verlag, 1998

[9] M.Okutomi and T.Kanade. A Multiple-baseline Stereo. IEEE Trans. On Pattern Analysis and Machine Intelligence, 15(4): 353-363, 1985.

[10] R.Peleg and M.Ben-Ezra. Stereo panorama with a single camera. In CVPR'99, pages395-401, FortCollins, June 1999.

[11] R.Peleg and J,Herman. Panoramic mosaics by manifold projection. In *Conf. On Computer Vision and Pattern Recognition*(CVPR'97), pages 338-343, San Juan, June 1997.

[12] P.Rademacher and G.Bishop. Multiple-center-of-Projection Images. In *Computer Graphics Proceedings*, pages 199-206, Proc.SIGGRAPH'98 (Orlando), July 1998.

[13] H.-Y.Shum and L.-W.He. Rendering with Concentric Mosaics. In *Computer Graphics Proceedings*, pages 299-306, Proc.SIGGRAPH'99 (Los Angeles), August 1999.

[14] H.-Y.Shum, A.Kalai, and S.M.Seitz. Omnivergent stereo. In *Seventh International Conference on Computer Vision* (ICCV'99), Greece, September 1999.

[15] H.-Y.Shum and R.Szeliski. Stereo.Reconstruction from Multiperspective Panorama. In *Seventh International Conference on Computer Vision* (ICCV'99), Greece, September 1999.

[16] D.N.Wood et al. Multiperspective panorama for cel animation. In *Computer Graphics Proceedings*, pages 243-250, Proc.SIGGRAPH'97 (Los Angeles), August 1997.

[17]. Y.-L. Xiong and L. Matthies, Error Analysis of a Real-Time Stereo System**,** In *IEEE Conference on Computer Vision and Patter Recognition (CVPR)*, 1997, pp. 1087-1093

[18] G.Xu and Z.Zhang, Epipolar Geotry in Stereo, Motion and Object Recognition: A Unified Approach. 1996. Kluwer Academic Publishers.

[19] J.-Y.Zheng and S.Tsuji. Panorama Representation for Route Recognition by a Mobile Robot. In *International Journal of Computer Vision.* 9(1):55-76. Netherlands, 1992. Kluwer.

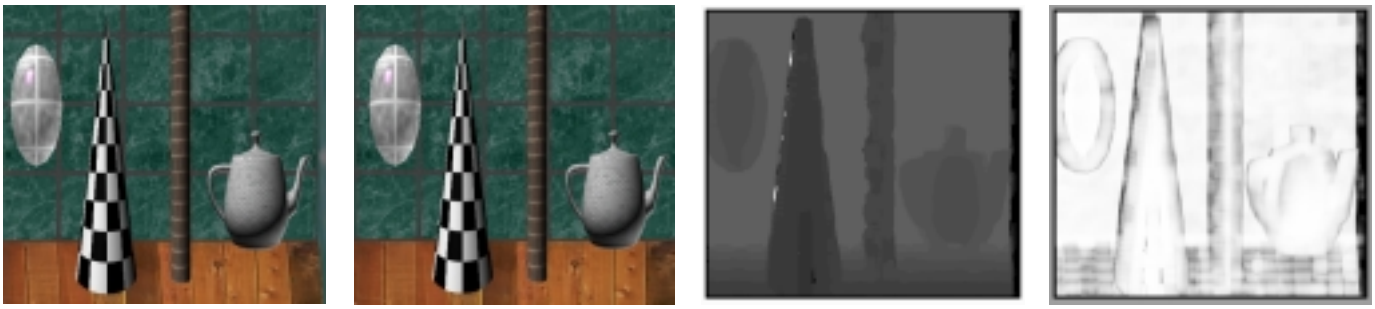


Fig.13 Parallel stereo with synthetic images: (a)(b) image pair; (c) result; (d) correlation map.

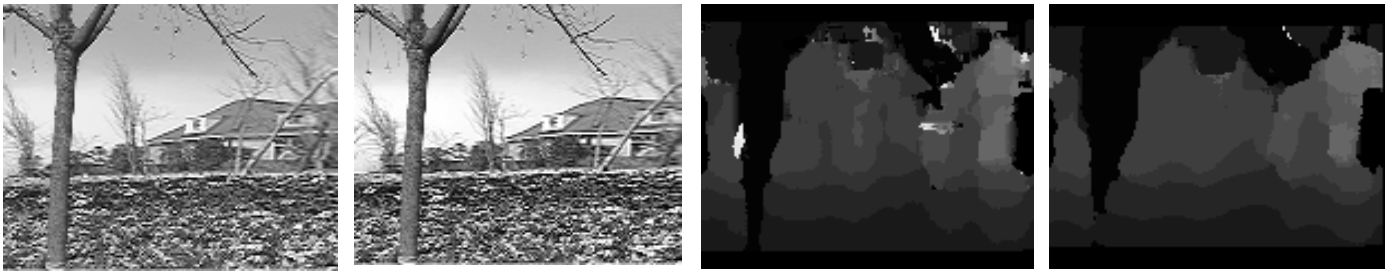


Fig.14 Parallel stereo with real images: (a)(b) rebinned image pair; results with window (c) size 11; (d) size 15.

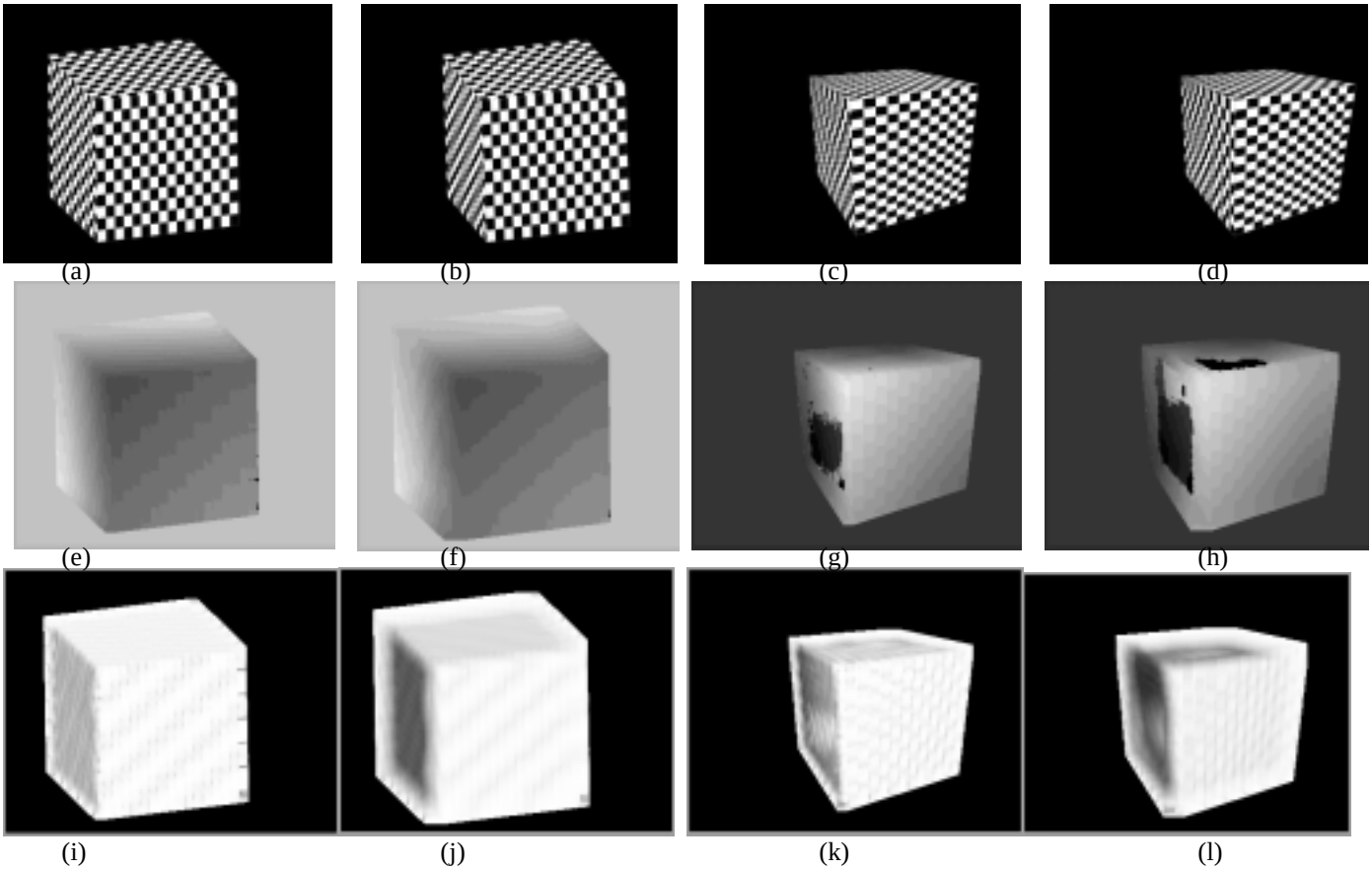


Fig. 15 Matching performance comparison between perspective stereo and parallel stereo: (a)(b) a pair of parallel images; (c)(d) a pair of pespective images; (e)(f) reconstruction of parallel stereo with window sizes 11 and 23; (i)(j) associated correlation maps of (e)(f); (g)(h) result of perspective stereo with window sizes 11 and 23; (k)(l) associated correlation maps of (g)(h).