Stereo Reconstruction from Multiperspective Panoramas

Heung-Yeung Shum and **Richard Szeliski**
Vision Technology Group, Microsoft Research
One Microsoft Way, Redmond, WA 98052-6399

Abstract

This paper presents a new approach to computing depth maps from a large collection of images where the camera motion has been constrained to planar concentric circles. We resample the resulting collection of regular perspective images into a set of multiperspective panoramas, and then compute depth maps directly from these resampled images. Only a small number of multiperspective panoramas is needed to obtain a dense and accurate 3D reconstruction, since our panoramas sample uniformly in three dimensions: rotation angle, inverse radial distance, and vertical elevation. Using multiperspective panoramas avoids the limited overlap between the original input images that causes problems in conventional multi-baseline stereo. Our approach differs from stereo matching of panoramic images taken from different locations, where the epipolar constraints are sine curves. For our multiperspective panoramas, the epipolar geometry, to first order, consists of horizontal lines. Therefore, any traditional stereo algorithm can be applied to multiperspective panoramas without modification. Experimental results show that our approach generates good depth maps that can be used for image-based rendering tasks such as view interpolation and extrapolation.

1 Introduction

Traditional stereo reconstruction begins with two calibrated perspective images taken with pinhole cameras. To reconstruct the 3D position of a point in the first image, its corresponding point in the second image has to be found before applying triangulation. Perspective cameras have the property that corresponding points lie on straight lines, which are called epipolar lines. In order to simplify the search for correspondences, the two images can optionally be rectified so that epipolar lines become horizontal.

Recently, there has been a lot of work on 3D reconstruction from large collections of images. Multi-baseline stereo using several images can produce better depth maps by averaging out noise and reducing ambiguities [12]. Space sweeping approaches, which project multiple images onto a series of imaging surfaces (usually planes), also use significant data redundancy for better reconstruction [3, 18, 22, 9]. Sweeping approaches in general need to discretize the scene volume, and therefore sampling strategies are very important.

Consider the problem of building a 3D environment model from thousands of images captured on video. Many modeling approaches to date have concentrated on coarse reconstruction using structure from motion with a small number (typically hundreds) of tracked feature points. What is really desired, however, are truly photorealistic reconstructions, and these require dense 3D reconstruction. What should we do with all these images to obtain dense 3D reconstructions? What do we mean by dense reconstruction? We may not wish to compute a depth map for each input image [21] because this is computationally expensive. It seems sensible to subsample the input frames (e.g., by simply dropping neighboring frames), but this risks not having enough overlapping frames to build good correspondences for accurate 3D reconstruction.

In this paper, we present a new approach to computing dense 3D reconstruction from a large collection of images. First, we constrain our camera motion to planar concentric circles. This constraint is practical and can easily be satisfied using a number of simple camera rigs (see below). For each concentric circle, we take one or more columns of pixels from each input image (or use line scan sensors such as linear pushbroom cameras [5]) and rebin these into *multiperspective panoramas*. Rather than using the original perspective images, we perform stereo reconstruction from these resampled and rebinned multiperspective panoramas. Because our panoramas sample uniformly in radial distance and rotation angle, we can afford to use only a small amount of data (e.g., several panoramas) for reconstruction.

Next, we develop a novel multi-image cylinder sweep stereo reconstruction algorithm which generalizes the concept of plane sweep stereo. Because of the special structure of our concentric panoramas, the cylinder sweep algorithm only requires horizontal translations and vertical re-scalings of the panoramic images during matching. We show that in many cases the epipolar geometry can be well approximated by the traditional horizontal epipolar layout used by conventional stereo algorithms. This allows a wide range of pre-existing stereo matching algorithms to be applied without modification

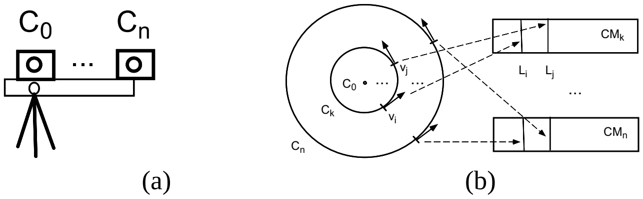


Figure 1. Concentric panoramas: (a) acquisition rig, and (b) rebinning process.

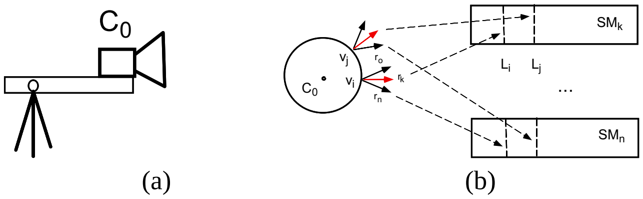


Figure 2. Swing panoramas: (a) acquisition rig, and (b) rebinning process.

to our novel representation. We demonstrate experimentally that this produces high-quality reconstructions.

Our multiperspective panoramas are a special case of the more general *multiperspective panoramas for cel animation* [24], and are actually very similar to *multiple-center-of-projection images* [16], *manifold mosaics* [14], and *circular projections* [13]. They are also closely related to images obtained by *pushbroom cameras* [5] and *panning non-frontal imaging cameras* [8]. Unlike these more general approaches, however, we constrain our camera motions to be along one or more radial paths around a fixed rotation center, and always take the same column from a given camera to generate a multiperspective panorama. As we show in this paper, this enables a very simple surface sweep stereo correspondence algorithm, and results in an epipolar geometry that in most cases is well approximated by traditional horizontal epipolar lines.

The idea of resampling and rebinning has recently become popular in image-based rendering. For example, the Lumigraph and Lightfield resample all captured rays and rebin them into a 4D two-plane parameterization [4, 10]. The reason for rebinning input images is to find a representation that facilitates the desired application. For the Lumigraph, this is the re-synthesis of novel views. For our multiperspective panoramas, the application is 3D stereo reconstruction of depth maps to be associated with panoramic images, in order to support a “look around and move a little” kind of viewing scenario.

Stereo matching of multiperspective panoramas can be viewed as a method to reduce the large amount of data in the original collection of images. A similar attempt to generate smaller descriptions is the stereo matching of single-perspective panoramic images taken at several different locations [7, 11]. The sampling of corresponding pixels is, however, non-uniform in the radial and angular directions, resulting biased stereo reconstructions. For instance, points along the baseline of two panoramas cannot be reconstructed. Moreover, the epipolar geometry for panoramic images is complicated because epipolar curves are sine curves. In contrast, the epipolar geometry of multiperspective panoramas can often be well approximated by horizontal lines. Therefore, traditional stereo matching algorithms can be used without modification.

We address the following questions in the remainder of this paper. How are multiperspective panoramas captured and generated from regular images? Given two or more such multiperspective panoramic images, how do we compute the correspondences? What is the epipolar geometry, and can traditional stereo matching algorithms be used or easily modified?

2 The imaging geometry of multiperspective panoramas

To generate one of our multiperspective panoramas, the camera motion must be constrained to planar concentric circles. A multiperspective panorama is formed by selecting the same column from each of the original perspective images taken at different positions on a circle. Multiperspective panoramas differ from conventional panoramas in that parallax effects are captured because each column is taken from a different perspective [16].

Two camera rigs can easily be built to capture multiperspective panoramas [20]. (These multiperspective panoramas have been used in [20] to synthesize novel view images, but suffer from vertical distortions in the synthesized images unless per-pixel depth correction is used.) The first rig uses several slit cameras (or regular cameras where only a few columns are kept) mounted on a rotating bar (Figure 1a). It is also possible to use only a single slit camera but slide it to different locations before rotating. A multiperspective panorama is constructed by collecting all slit images at all rotation angles. We call these images *concentric panoramas*. Figure 6a shows a synthetically generated concentric panorama, while Figure 6b shows details from several such panoramas illustrating the parallax effects.

The other design is to swing a regular camera mounted on a rotating bar looking outwards. In this case, different columns are used to construct different multiperspective panoramas. We call such panoramas *swing panoramas* (Peleg and Ben-Ezra call these *circular projections* [13]). A video sequence of F frames of size $W \times H$ can be rebinned into (up to) W panoramas with size $F \times H$. Figure 7a shows some sample input images and Figure 7b shows one of the rebinned panoramic images. We can also use a more general configuration of one or more cameras angled at various orientations with respect to the swing arm with one or more columns being sampled from each camera.

The general imaging geometry and trigonometric relationships for a single concentric multiperspective panorama are shown in Figure 3. This top-down view shows the center of rotation C around which the camera rotates along a radius of

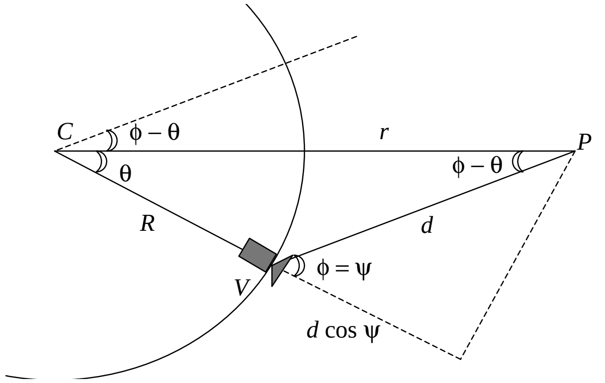


Figure 3. Angle and distance relationships in multiperspective panoramic stereo.

size R. The camera is located at V, and the selected column of pixels is looking at a scene point P, which is at a distance r from the center C, and at an in-plane distance d from the camera V. The current angle of rotation is denoted by θ, and varies between 0° and 360°. In a panoramic image, θ forms the horizontal axis. A typical rebinned panoramic image may have 1280 columns, each one taken from the same column in successive frames of a given input video. The other (vertical) axis in the panoramic image is indexed by the elevation or row number y (not shown in the top-down figure).

The plane of pixels captured by a particular panoramic imager forms an angle ϕ with the *swing line* in the normal direction connecting V and C. When multiple columns are selected from a camera pointing outward from C (the "swing stereo" configuration), we have $\phi = \tan^{-1}((x - x_c)/f)$, where x is the column number in the input image, x_c is the center column, and f is the focal length in pixels. In the concentric panorama configuration, $\phi = 90^\circ$. Not shown (explicitly) in Figure 3 is the angle ψ between the optical axis of the camera and the column being rebinned. In swing panoramas, $\psi = \phi$, whereas in concentric panoramas, $\psi = 0$.

To summarize, each panoramic image is indexed by (θ, y), and its acquisition geometry is parameterized by (R, ϕ, ψ). In the most common acquisition setups, either $\phi_k = \tan^{-1}(kw)$, $k = -K \ldots K$ (swing panoramas), or $\phi = \pm 90^\circ$, $R_k = kR_1$, $k = 0 \ldots K$ (concentric panoramas).

3 Analysis of parallax

In this section, we derive the formulas for the horizontal and vertical parallax of a point located on a cylindrical surface or radius r (Figure 3).

Using the basic law of sines for triangles, we have

$$\frac{R}{\sin(\phi - \theta)} = \frac{r}{\sin(180^\circ - \phi)} = \frac{d}{\sin\theta}, \qquad (1)$$

or

$$\theta = \phi - \sin^{-1}\left(\frac{R}{r}\sin\phi\right). \qquad (2)$$

Therefore, the horizontal parallax $\Delta\theta_{2:1} = \theta_2 - \theta_1$ for a point at a distance r seen in two panoramic images I_1 and I_2 consists of a constant factor $\phi_2 - \phi_1$ and two terms depending on r. If we (circularly) shift each panoramic image by ϕ_i, the first factor drops out, leaving

$$\Delta\theta_{2:1} = \sin^{-1}\left(\frac{R_1}{r}\sin\phi_1\right) - \sin^{-1}\left(\frac{R_2}{r}\sin\phi_2\right). \qquad (3)$$

The vertical parallax can be derived using the following observation. Recall that according to the laws of perspective projection, the appearance (size) of an object is inversely proportional to its distance along the optical axis, e.g., $u = fx/z$, $v = fy/z$. For pixels at a constant distance r from C, and therefore a constant distance from V along the optical axis, the vertical scaling can be computed directly from this distance $z = d\cos\psi$ (Figure 3). Here, d is the in-plane distance between V and P, and $d\cos\psi$ is the distance along the optical axis (as we mentioned before, typically $\psi = \phi$ or $\psi = 0$).

We can write this scale factor as $s = f/z = f/(d\cos\psi)$. Using the law of sines (1) again, we can compute the change of scale between two panoramic images as

$$s_{2:1} = \frac{s_2}{s_1} = \frac{d_1\cos\psi_1}{d_2\cos\psi_2} = \frac{\sin\theta_1/\sin\phi_1}{\sin\theta_2/\sin\phi_2}\frac{\cos\psi_1}{\cos\psi_2}.$$

Expanding $\sin\theta$, where θ is given by (2), we obtain

$$\begin{aligned} \sin\theta &= \sin\phi\cos\left(\sin^{-1}\left(\frac{R}{r}\sin\phi\right)\right) - \cos\phi\frac{R}{r}\sin\phi \\ &= \sin\phi\left[\sqrt{1 - (\frac{R}{r}\sin\phi)^2} - \frac{R}{r}\cos\phi\right]. \end{aligned}$$

We can thus re-write the scale change as

$$s_{2:1} = \frac{\sqrt{1 - (\frac{R_1}{r}\sin\phi_1)^2} - \frac{R_1}{r}\cos\phi_1}{\sqrt{1 - (\frac{R_2}{r}\sin\phi_2)^2} - \frac{R_2}{r}\cos\phi_2}\frac{\cos\psi_1}{\cos\psi_2}. \qquad (4)$$

The first factor in this equation depends on r, and goes to 1 as $r \rightarrow \infty$. The second factor is a global scale that compensates for the off-axis angle of a given column. This is the same correction that is applied to rectilinear (perspective) images when converting them into cylindrical coordinates [23]. The vertical parallax for any pixel in row y_1 of image I_1 can be computed directly from

$$\Delta y_{2:1} = y_2 - y_1 = (s_{2:1} - 1)y_1. \qquad (5)$$

4 Cylinder sweep stereo

Plane-sweep and space coloring/carving stereo algorithms have recently become popular, because they support true multi-image matching [3], enable reasoning about occlusion relationships [18, 22, 9], and are more efficient than traditional correlation-based formulations [6]. Traditional stereo matching algorithms pick a window around each pixel in a reference image and then find corresponding windows in other images at every candidate disparity (searching along an epipolar line). Plane sweep algorithms consider each candidate disparity as defining a plane in space, and projecting all images to be matched onto that plane, using a planar perspective transform (homography) [3, 22, 1]. A per-pixel fitness metric (e.g., the variance of the corresponding collection of pixels) is first computed, and this is then aggregated spatially using an efficient convolution algorithm (e.g., a moving average box filter) or some other technique [17]. After all the cost functions have been computed, a winning disparity can be chosen. If the planes are processed in front to back order, occlusion relationships can also be included [18].

Our novel *cylinder sweep* algorithm works similarly. Instead of projecting images onto planes, however, we project our multiperspective panoramas onto cylinders of varying radii r. The transformations that map panoramas onto these cylinders, and hence onto each other, are particularly simple, as we demonstrate in the following theorem.

Theorem 1. A 3D reconstruction can be computed from a collection of concentric multiperspective panoramic images using the cylinder sweep algorithm described above. The reprojection of each panoramic image onto a surface of radius r consists of a horizontal translation and a vertical scaling.

Proof. The proof follows directly from the equations (3) and (4) for horizontal and vertical parallax. A less formal but more intuitive proof is to just observe that the image of any column of a cylinder at a fixed radius r seen by a concentric pushbroom camera is just a scaled version of the pixels lying on that cylinder. Since our representation has no preferred direction, the shift between the panoramic image and the cylinder must be the same for all pixels.

An imaging configuration of special interest is a two-frame stereo arrangement where the optical rays are symmetric with respect to the swing line $\overline{CV}$, i.e., $\phi_1 = -\phi_2$ (we also assume that $\psi_i = \phi_i$ or $\psi = 0$). This occurs when you take, for example, the left and right columns of a swing stereo sequence [13, 19], or when you fix two cameras at equal distances but at opposite ends of a rotating boom (concentric panorama).

Theorem 2. For a *symmetric pair* of concentric multiperspective panoramas (as defined above), the epipolar geometry consists of horizontal lines.

Proof. This follows directly from equation (4), i.e., $s_{2:1} = 1$. A more informal proof could be obtained by drawing another point P' in Figure 3 at an angle $-\phi$, and observing that $z = d \cos \psi$ is the same for both viewing rays.

A direct consequence of selecting such a pair of panoramic images is that *any* traditional stereo algorithm (e.g., hierarchical warp [15, 2]) can be used, instead of having to use the modified plane sweep algorithm described in Theorem 1. More general camera configurations (e.g., that do not constrain camera to a planar motion) are discussed in [19].

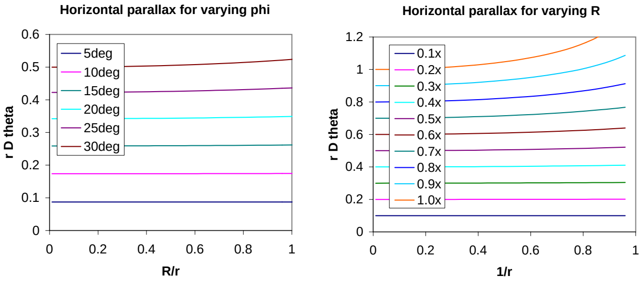


Figure 4. Plot of horizontal parallax (divided by $1/r$) for varying values of ϕ and R.

5 Small disparity and angle approximations

In practice, we would like to use more than two images, in order to obtain a more accurate and robust correspondence [12, 3], but we would like to avoid having to restrict ourselves to a modified plane sweep algorithm. In this section, we study whether the epipolar geometry is sufficiently close to a horizontal epipolar geometry so that other traditional multi-image stereo matching algorithms (such as SSSD [12]) could be used.

A requirement for a traditional multi-baseline stereo algorithm to be applicable is that the location of pixels at different depths can be explained by a collection of pinhole cameras. In our case, we further restrict our attention to the horizontal epipolar geometry case, in which case we have the requirement that horizontal parallax (3) be of the form

$$\Delta\theta_{k:0} = m_k f(r),$$

i.e., that we have a fixed linear relationship between horizontal parallax and some common function of r for all images. The horizontal parallax equation given in (3) does *not* exactly satisfy this requirement. However, if either R/r or $\sin\phi$ are small in both images, we obtain

$$\Delta\theta_{k:0} \approx [R_k \sin\phi_k - R_0 \sin\phi_0] r^{-1}. \qquad (6)$$

We therefore see that inverse r plays the same role as inverse depth (*disparity* [12]) does in multi-baseline stereo.

Figure 4 plots the exact formula (3) for horizontal parallax as a function of r^{-1} for various values of ϕ and R. The left plot shows the ratio of $\Delta\theta_{k:0}$ to $1/r$ (since it is very hard to tell the deviation from linearity by eye) for $\phi_0 = 0$ (central column) and varying ϕ_k (swing panoramas). The right plot shows the ratio of $\Delta\theta_{k:0}$ to $1/r$ for $R_0 = 0$ (no parallax) and varying R_k for concentric panoramas with $\phi_k = 90°$ and $\psi_k = 0$. As we can see from these plots, the linear approximation to parallax is quite good, as long as the nearest scene point doesn't get too close, e.g., no closer than 50% of R for moderate focal lengths. The reduced linearity in the concentric stereo rig can be mitigated by spacing the cameras more closely, i.e., not using the full baseline of the arm if scene points are too close.

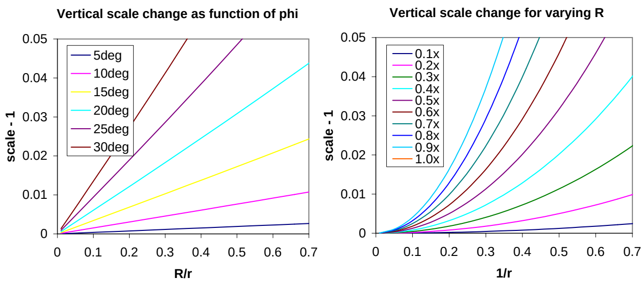


Figure 5. Plot of vertical parallax (scale change - 1) for varying values of ϕ and R.

A second requirement of assuming a horizontal epipolar geometry is that the vertical parallax needs to be negligible (preferably under one pixel). For images of about 240 lines (a single field from NTSC video), we would like $|\Delta y| \leq 120|s_{2:1} - 1| < 1$ (120 is the half-height of the image), i.e., $|s_{2:1} - 1| < 0.008$.

We can approximate the vertical parallax equation (4) under two different conditions. For swing stereo ($R_1 = R_2, \psi = \phi$), assume that $\phi_2 = 0$ (central column) and ϕ_1 is small. We can expand (4) to obtain

$$\begin{aligned} s_{2:1} &\approx \cos\psi_1[1 - \frac{R^2}{2r^2}\sin^2\phi_1 - \frac{R}{r} + \frac{R}{2r}\sin^2\phi_1][1 + \frac{R}{r}] \\ &\approx \cos\psi_1[1 + \frac{R}{2r}\sin^2\phi_1 - \frac{R^2}{r^2}] \qquad (7) \end{aligned}$$

Thus, once the global scale change by $\cos\psi_1$ (which is independent of depth) is compensated for, we have a vertical parallax component that is linear in R/r and quadratic in $\sin\phi$.

For concentric panoramas, $\phi = 90°$ and $\psi = 0$, which gives us

$$s_{2:1} = \frac{\sqrt{1 - R_1^2/r^2}}{\sqrt{1 - R_2^2/r^2}} \approx 1 + \frac{1}{r^2}(R_2^2 - R_1^2).$$

The vertical parallax is inversely proportional to squared distance r, and proportional to the difference in squared radii R. Figure 5 plots the exact formula (4) for vertical parallax as a function of r^{-1} for various values of ϕ and R. The left plot shows scale change $s_{k:0} - 1$ for $\phi_0 = 0$ (central column) and varying ϕ_k (swing panoramas). The right plot shows $s_{k:0} - 1$ for $R_0 = 0$ (no parallax) and varying R_k (concentric panoramas with $\phi_k = 90°$ and $\psi_k = 0$). As we can see from these plots, the amounts of vertical parallax are quite small ($< 1\%$) if the field of view is moderately small (say $30°$) or the ratio of the nearest scene point to the variation in radial camera positions is large.

To wrap up this section, we can now state the following theorem, whose proof follows from the analyses in this section.

Theorem 3 When the disparity R/r and/or off-axis angle ϕ are small, we obtain a nearly horizontal epipolar geometry (classic multi-baseline stereo geometry [12]), after compensating once for the vertical scaling of each image.

6 Experiments

We have applied our stereo algorithm to both synthetic data and real data. Figure 6 shows the stereo reconstruction from 7 concentric panoramas that were synthesized with a slit camera rotating around circles of different radii $(0.4, 0.5, \ldots, 1.0)$. The right and the center panoramas are shown in (a) and (b), respectively. Horizontal parallax can be observed from closeup of regions of three original panoramas shown in Figure 6(f–h). Because a small field of view ($24°$) camera was used, these concentric panoramas have negligible vertical parallax. Using the estimated depth map shown in Figure 6(c), we synthesize a panorama shown in (d) with the same camera parameters as in (a). The new panorama is almost indistinguishable from (a) except in the regions where significant occlusion occurs as shown in the closeup (h). Notice that the spotlight is synthesized well even though its depth estimation is clearly wrong.

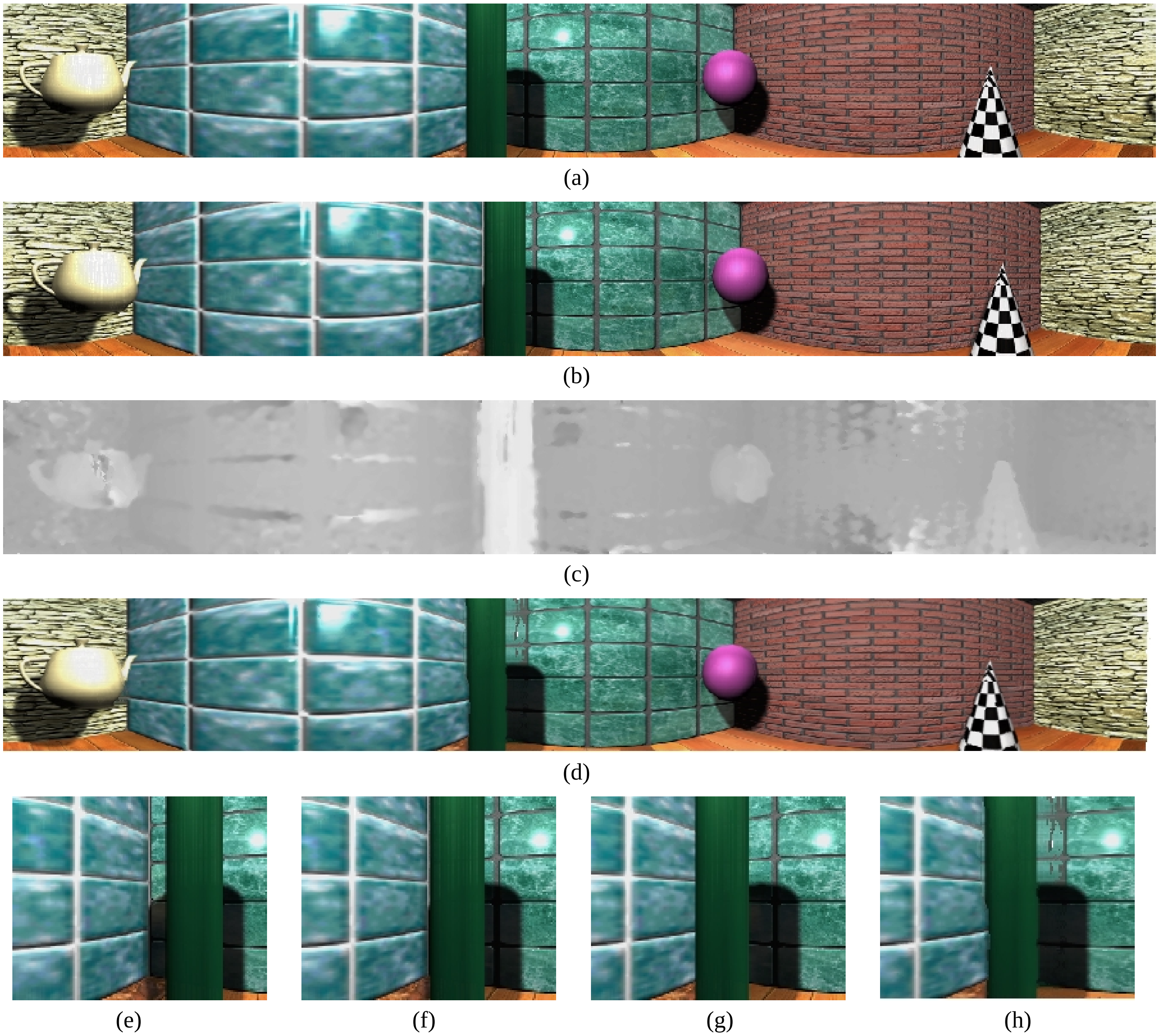
Two rebinned panoramas (left column and center column) from a real swing sequence of a lab scene are shown in Figures 7(a) and (b). We used a digital video camera in portrait mode with a field of view of around $27°$ by $36°$ and a digitized image size of 240 by 320 pixels. The camera was mounted off-centered on a plate rotated with a stepper motor which provided accurate rotation parameters.

After scaling the images vertically by $\cos\psi$, we found that there was a small 0.5 pixel drift remaining between the panoramas. This was probably caused by a slight rotation of the camera around its optical axis. In practice, compensating for such registration errors is not difficult: a simple point tracker followed by linear regression can be used to obtain the best possible horizontal epipolar geometry. The panoramas after vertical scaling and drift compensation are shown in the closeup regions in Figure 6(f–h). As you can see, very little vertical parallax remains.

The reconstructed depth map is shown in Figure 6(c), and a synthesized *novel* panorama (from an *extrapolated* viewpoint) and its closeup are shown in Figures 6(d) and 6(f), respectively.

7 Discussion and Conclusions

In this paper, we have introduced a novel representation, concentric multiperspective panoramas, that efficiently captures the parallax available in a scene or environment. Concentric multiperspective panoramas are constructed by resampling and rebinning perspective images from one or more off-center rotating cameras. While in this paper we have assumed that these cameras are all moving in the same plane, there is actually nothing preventing the cameras from being located in different (parallel) planes, so long as their rotation axes are all the same (i.e., the camera motions are *co-axial*, rather than concentric). The only difference in this case is the addition of some extra depth-dependent *vertical* parallax, which can easily be accomodated in both traditional stereo algorithms and in our novel cylinder sweep algorithm.



(a)

(b)

(c)

(d)

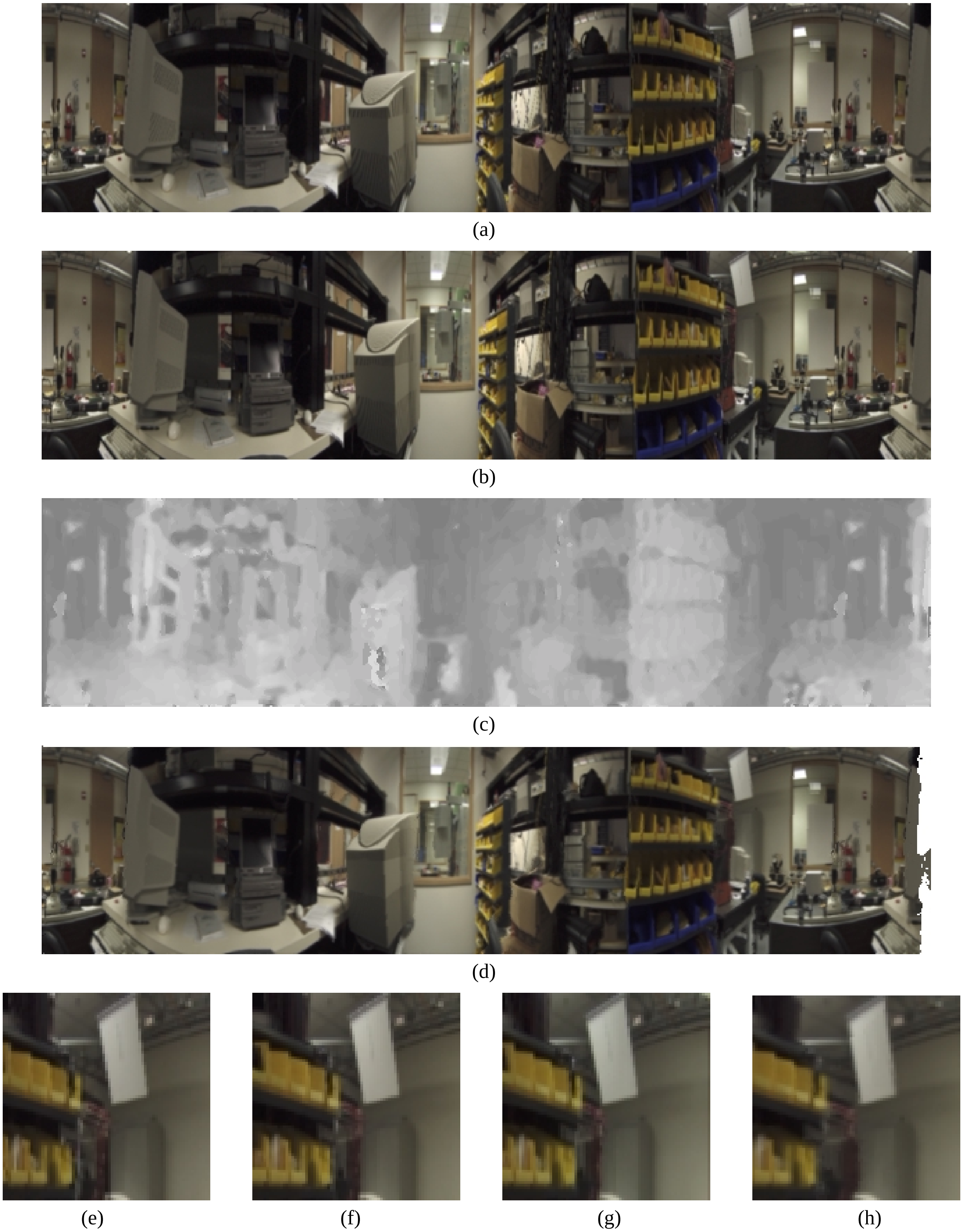
(e)

(f)

(g)

(h)

Figure 6. Concentric stereo results: (a) an input panorama (rightmost camera, $R = 1.0$), (b) another input panorama (center camera, $R = 0.7$), (c) estimated depth map for the center panorama, (d) panorama re-synthesized from center panorama and depth, (e–g) closeup of input panoramas (note the horizontal parallax), (h) closeup of re-synthesized panorama.



(a)

(b)

(c)

(d)

(e)

(f)

(g)

(h)

Figure 7. Swing stereo results: (a) an input panorama from left column, (b) input panorama from center column, (c) estimated depth map for the center panorama, (d) novel panorama extrapolated from center panorama and depth, (e-g) closeup of input panoramas, (h) closeup of extrapolated panorama.

Concentric multiperspective panoramas are ideally suited for stereo reconstruction of 3D scenes. Instead of using many original images, only several rebinned multiperspective panoramas need to be used. Unlike a collection of single-perspective panoramas taken from different locations, there are no preferred directions or areas where the matching fails because the disparity vanishes. We have developed a novel reconstruction algorithm, cylinder sweep stereo, which only requires us to translate and scale the panoramas during the matching phase.

We have also shown, both analytically and experimentally, that the epipolar geometry of concentric multiperspective panoramas is often well approximated by a traditional horizontal geometry. This property allows us to apply traditional multi-baseline or multiview stereo algorithms without any modification. We have shown experimentally that good stereo reconstructions can be obtained from such panoramas, and that the original parallax in the scene can be recreated from just one panorama and one panoramic depth map. It is also possible to extrapolate novel views from original panoramas and the recovered depth map.

The novel representation of concentric multiperspective panoramas and good quality stereo reconstruction from such panoramas suggest a powerful new way of modeling and rendering a large environment. Instead of using a single global model for the whole environment, we envision using a collection of local models for overlapping sub-regions of the environment. Each sub-region is represented by a small set of multipespective panoramas and their associated depth maps. At each sub-region, the user is free to "look around and move around" inside the circular region using the local panoramas and depth maps. As the user moves from sub-region to another, a different local model is activated. We are developing novel rendering algorithms based on this representation that will bring the "third dimension" back into panoramic photography and the viewing and exploration of virtual environments.

References

[1] S. Baker, R. Szeliski, and P. Anandan. A layered approach to stereo reconstruction. In *IEEE Comp. Soc. Conf. on Computer Vision and Pattern Recognition (CVPR'98)*, pages 434–441, Santa Barbara, June 1998.

[2] J. R. Bergen, P. Anandan, K. J. Hanna, and R. Hingorani. Hierarchical model-based motion estimation. In *Second European Conference on Computer Vision (ECCV'92)*, pages 237–252, Santa Margherita Liguere, Italy, May 1992. Springer-Verlag.

[3] 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.

[4] S. J. Gortler, R. Grzeszczuk, R. Szeliski, and M. F. Cohen. The lumigraph. In *Computer Graphics Proceedings*, pages 43–54, Proc. SIGGRAPH'96 (New Orleans), August 1996.

[5] R. Gupta and R.I. Hartley. Linear pushbroom cameras. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(9):963–975, September 1997.

[6] T. Kanade et al. A stereo machine for video-rate dense depth mapping and its new applications. In *IEEE Comp. Soc. Conf. on Computer Vision and Pattern Recognition (CVPR'96)*, pages 196–202, San Francisco, June 1996.

[7] S. B. Kang and R Szeliski. 3-D scene data recovery using omnidirectional multibaseline stereo. *International Journal of Computer Vision*, 25(2):167–183, November 1997.

[8] A. Krishnan and N. Ahuja. Range estimation from focus using a non-frontal imaging camera. *International Journal of Computer Vision*, 20(3):169–185, March/April 1996.

[9] K. N. Kutulakos and S. M. Seitz. A theory of shape by space carving. In *Seventh International Conference on Computer Vision (ICCV'99)*, Greece, September 1999.

[10] M. Levoy and P. Hanrahan. Light field rendering. In *Computer Graphics Proceedings*, pages 31–42, Proc. SIGGRAPH'96 (New Orleans), August 1996.

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

[12] M. Okutomi and T. Kanade. A multiple baseline stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(4):353–363, April 1993.

[13] R. Peleg and M. Ben-Ezra. Stereo panorama with a single camera. In *CVPR'99*, pages 395–401, Fort Collins, June 1999.

[14] S. 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.

[15] L. H. Quam. Hierarchical warp stereo. In *Image Understanding Workshop*, pages 149–155, New Orleans, Louisiana, December 1984. Science Applications International Corporation.

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

[17] D. Scharstein and R. Szeliski. Stereo matching with non-linear diffusion. *International Journal of Computer Vision*, 28(2):155–174, July 1998.

[18] S. M. Seitz and C. M. Dyer. Photorealistic scene reconstrcution by space coloring. In *IEEE Comp. Soc. Conf. on Computer Vision and Pattern Recognition (CVPR'97)*, pages 1067–1073, San Juan, June 1997.

[19] H.-Y. Shum et al. Omnivergenet stereo. In *Seventh International Conference on Computer Vision (ICCV'99)*, Greece, September 1999.

[20] H.-Y. Shum and L.-W. He. Rendering with concentric mosaics. In *SIGGRAPH'99*, Los Angeles, August 1997.

[21] R. Szeliski. A multi-view approach to motion and stereo. In *IEEE Comp. Soc. Conf. on Computer Vision and Pattern Recognition (CVPR'99)*, pages 157–163, Fort Collins, June 1999.

[22] R. Szeliski and P. Golland. Stereo matching with transparency and matting. In *Sixth International Conference on Computer Vision (ICCV'98)*, pages 517–524, Bombay, January 1998.

[23] R. Szeliski and H.-Y. Shum. Creating full view panoramic image mosaics and texture-mapped models. *Computer Graphics (SIGGRAPH'97)*, pages 251–258, August 1997.

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