Automatic Disparity Control in Stereo Panoramas (OmniStereo) *

Yael Pritch Moshe Ben-Ezra Shmuel Peleg
School of Computer Science and Engineering
The Hebrew University of Jerusalem
91904 Jerusalem, ISRAEL

Abstract

An omnistereo panorama consists of a pair of panoramic images, where one panorama is for the left eye, and another panorama is for the right eye. An omnistereo pair provides a stereo sensation up to a full 360 degrees. OmniStereo panoramas can be created by mosaicing images from a rotating video camera, or by specially designed cameras.

The stereo sensation is a function of the disparity between the left and right images. This disparity is a function of the ratio of the distance between the cameras (the baseline) and the distance to the object: disparity is larger with longer baseline and close objects. Since our eyes are a fixed distance apart, we loose stereo sensation for far away objects.

It is possible to control the disparity in omnistereo panoramas which are generated by mosaicing images from a rotating camera. The baseline can be made larger for far away scenes, and smaller for nearer scenes. A method is described for the construction of omnistereo panoramas having larger baselines for faraway scenes, and smaller baseline for closer scenes. the baseline can change within the panorama from directions with closer objects to directions with further objects.

1 Introduction

The ultimate immersive visual environment should provide three elements: (i) Stereo vision, where each eye gets a different image appropriate to its location in space; (ii) complete 360 degrees view, allowing the viewer to look in any desired direction; (iii) allow free movement.

Stereo Panoramas [6, 5, 10, 14] use a new scene to image projection that enables simultaneously both (i) stereo and (ii) a complete panoramic view. No depth information or correspondences are necessary. Viewers of stereo panoramas have the ability to freely view, in stereo, all directions.

Since the scene to image projection necessary for stereo panoramic imaging can not be done with a regular camera, stereo panoramic images were generated by mosaicing images taken with rotating cameras [6, 5, 10, 14].

Short introductions are given in this section to panoramic imaging, stereo imaging, and multiple viewpoint projections. Sec. 2 discusses the multiple viewpoint projection that can be used to create stereo panoramas. Sec. 3 describes the method to create stereo panoramas by mosaicing images from a rotating camera. Sec. 4 describes the generation stereo panoramas with adaptive disparity.

1.1 Panoramic Images

A panoramic image is a wide field of view image, up to a full view of 360 degrees. Panoramas can be created on an extended planar image surface, on a cylinder, or on a sphere. Traditional panoramic images have a single viewpoint, also called the “center of projection” [8, 3, 15]. Panoramic images can be captured by panoramic cameras, by using special mirrors [9, 7], or by mosaicing a sequence of images from a rotating camera [15, 11].

1.2 Visual Stereo

A stereo pair consists of two images, from two different viewpoints, of the same scene. The disparity, which is the angular difference in viewing directions of each scene point between the two images, is interpreted by the brain as depth through a process called “stereo fusion”. Fig. 1 describes a conventional stereo setting. The disparity is a function of the point’s depth and the distance between the eyes (*baseline*). Maximum disparity change, and hence depth separation, is for scene points along the line of equal distance points between the two eyes (“principal viewing direction”). No stereo depth separation exists for points along the extended baseline.

* This research was partially funded by European ESPRIT project 26247-VIGOR, as well as by Israel Science Foundation Grant 612/97.
Contact E-Mail: {yaelpri,moshe,peleg}@cs.huji.ac.il

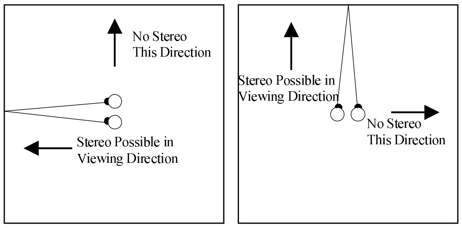

Figure 1. No arrangement of two single-viewpoint images can give stereo in all viewing directions. For upward viewing the two cameras should be separated horizontally, and for sideways viewing the two cameras should be separated vertically.

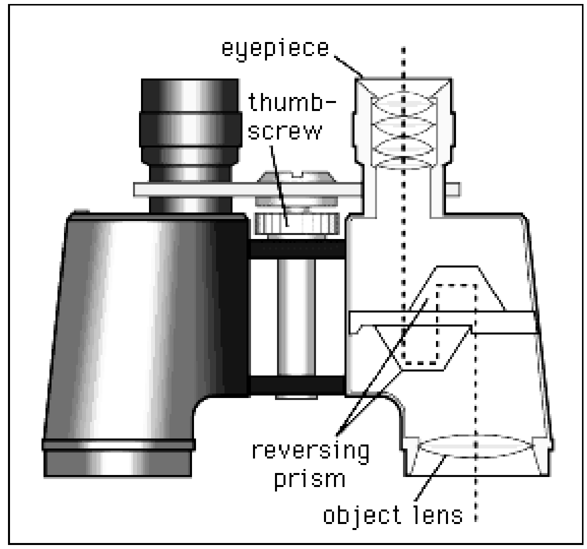

Figure 2. In order in enhance stereo perception in distant scenes, the separation between the object lens in binoculars is usually larger than the separation between the eyepieces.

People can perceive depth from stereo images if the viewpoints of the two cameras generate horizontal disparity in a specific range. Human Stereo Fusion can be obtained with disparities of up to $\pm 0.5^o$ This translates to a disparity of approximately 30 pixels on a regular screens and at a regular viewing distance.

Eye movement can change the absolute disparity by vergence. The point on which the eyes converge is called the point of fixation. Vergence can not change depth perception, as depth perception is a function of relative disparities. However vergence can make the viewing much more comfortable by setting the point of fixation close to the middle of the depth range at the viewing direction.

The maximum stereoscopic range for a human observer is approximately 670m. This limit can be considerably improved with the use of suitable optical instrument, such as binoculars. The distance between centers of the object glasses (the baseline) can reach 14cm (see Fig. 2), where in human vision the baseline is approximately 6.5 cm. This increased baseline and the lens magnification gives a maximum stereoscopic range of about 10,100 meters when using the binocular.

Depth has been computed from panoramic images having two viewpoints, one above the other [4]. However, since the disparity in this case is vertical, the images can not be used for viewing by humans having eyes which are separated horizontally.

2 Multiple Viewpoint Projections

Regular images are created by perspective projections: scene points are projected onto the image surface along projection lines passing through a single point, called the "optical center" or the "viewpoint". Multiple viewpoint projections use different viewpoints for different viewing direction, and were used mostly for special mosaicing applications. Effects that can be created with multiple viewpoint projections and mosaicing are discussed in [16, 12].

Stereo panoramic imaging uses a special type of multiple viewpoint projections, *circular projections*, where both the left-eye image and the right-eye image share the same cylindrical image surface. To enable stereo perception, the left viewpoint and the right viewpoint are located on an inner circle (the "viewing circle") inside the cylindrical image surface, as shown in Fig. 3. The viewing direction is on a line tangent to the viewing circle. The left-eye projection uses the rays on the tangent line in the clockwise direction of the circle, as in Fig. 3.b. The right-eye projection uses the rays in the counter clockwise direction as in Fig. 3.c. Every point on the viewing circle, therefore, defines both a viewpoint and a viewing direction of its own.

The applicability of circular projections to panoramic stereo is shown in Fig. 4. From this figure it is clear that the two viewpoints associated with all viewing directions, using the "left-eye" projection and the "right-eye" projection, are in optimal relative positions for stereo viewing for all directions. The vergence is also identical for all viewing directions [13], unlike regular stereo that has a preferred viewing direction.

3 Stereo Panoramas with Rotating Cameras

Representing all possible stereoscopic views with only two panoramic images presents a contradiction, as described in Fig. 1. When two ordinary panoramic images

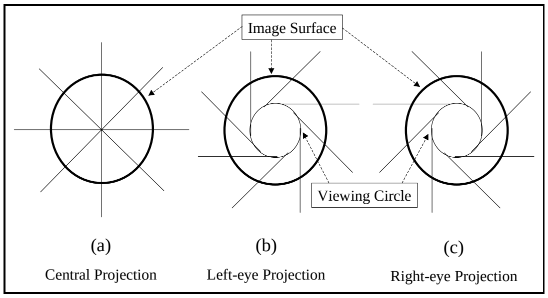

Figure 3. Circular projections. The projection from the scene to the image surface is done along the rays tangent to the viewing circle. (a) Projection lines perpendicular to the circular imaging surface create the traditional single-viewpoint panoramic image. (b-c) Families of projection lines tangent to the inner viewing circle form the multiple-viewpoint circular projections.

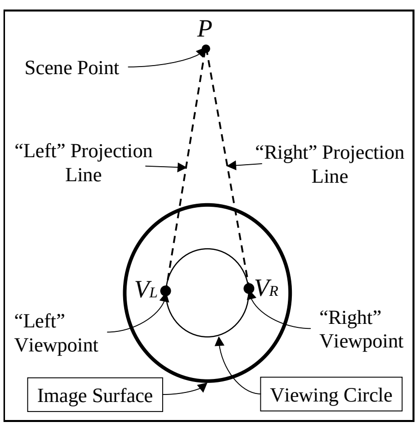

Figure 4. Viewing a scene point with "left-eye" and "right-eye" projections. The two viewpoints for these two projections are always in optimal positions for stereo viewing.

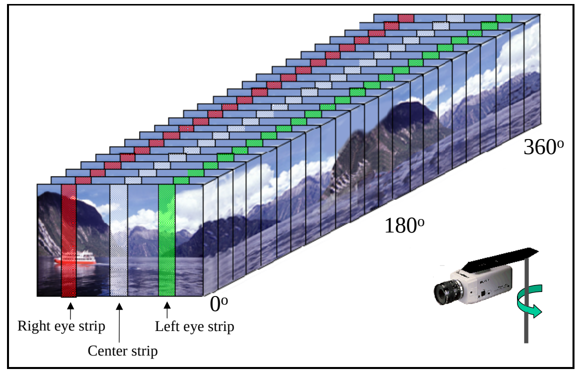

Figure 5. Stereo Panoramas can be created using images captured with a regular camera rotating about an axis behind it. Pasting together strips taken from each image approximates the panoramic image cylinder. When the strips are taken from the center of the images an ordinary panorama is obtained. When the strips are taken from the left side of each image, the viewing direction is tilted counter clockwise from the image surface, obtaining the right-eye panorama. When the strips are taken from the right side of each image, the left-eye panorama is obtained.

are captured from two different viewpoints, the disparity and the stereo perception will degrade as the viewing direction becomes closer to the baseline until no stereo will be apparent.

Generation of Image-based stereo panoramas by rotating a stereo head having two cameras was proposed in [5, 14]. A stereo head with two cameras is rotated, and two panoramic mosaics are created from the two different cameras. A Single rotating camera can also be sufficient under some conditions [10, 6, 14]. In the case of a single moving camera different sides of the same image are used to mosaic the two images for the different eyes. This can even be done in real-time [10] (See Fig. 5).

A schematic diagram of the process creating a pair of stereo panoramic images is shown in Fig 6. A camera having an optical center O and an image plane is rotated about an axis behind the camera. Strips at the left of the image are seen from viewpoints V_R, and strips at the right of the image are seen from viewpoints V_L. The distance between the two viewpoints is a function of the distance r between the rotation axis and the optical center, and the distance $2v$ between the left and right strips. Increasing the distance between the two viewpoints, and thus increasing the stereo disparity, can be obtained by either increasing r or increas-

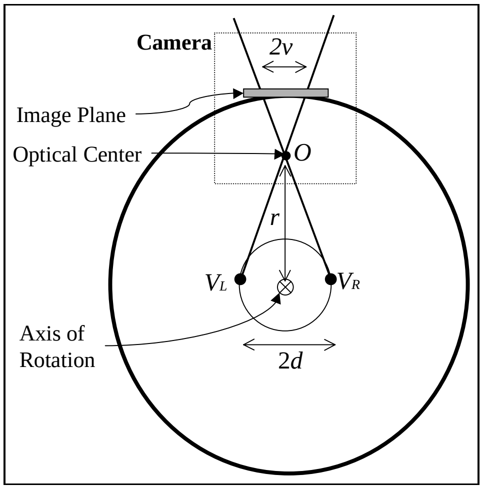

Figure 6. Schematic diagram of the system to create a pair of stereo panoramic images. A camera having an optical center "O" is rotated about an axis behind the camera. Note the "inverted" camera model, with image plane in front of the optical center.

ing v.

4 Automatic Disparity Control

As shown in Fig. 7, the disparity in omnistereo panoramas generated with a rotating camera is a function of the the distance r between the rotation axis and the optical center (the "arm length"), the focal length f and the separation $2v$ between the left and right strips. This relation is expressed as follows:

$$\begin{aligned} d &= R\sin(\beta) \\ \beta &= \tan^{-1}(\tfrac{v}{f}) \\ \alpha &= \sin^{-1}(\tfrac{d}{z}) = \sin^{-1}(\tfrac{R\sin(\tan^{-1}(\frac{v}{f}))}{z}) \end{aligned} \quad (1)$$

The length r of the arm between the rotation axis and the camera can be changed during the image capture time. Far away scenes will need a longer arm than closer scenes. The distance $2v$ between the left and right strips can be changed during mosaic construction. Far away scenes will need larger distance between the strips.

In this section we will describe how to measure the disparity and adjust it to make for best stereo perception. This will be done by adjusting the separation between the strips.

A schematic description of this process appears in Fig. 8, where a set of images from a rotating camera is represented as an x-y-t volume. A regular panorama can be described as a collection of strips making a planar cut by an y-t plane in the center of the x-y-t volume. In omnistereo mosaicing the left and right panoramas are each a planar cut by an y-t plane, but at the sides of the x-y-t volume and at fixed distance from the center. With the automatic disparity control each panorama is also a cut in the x-y-t volume, but the distance of the strip from the center is changing according to the desired baseline. For far away regions, where the baseline should be larger to increase disparity, the strips will be further away from the center. In closer area, where the baseline should be shortened to decrease disparity, the strips will be closer to the center.

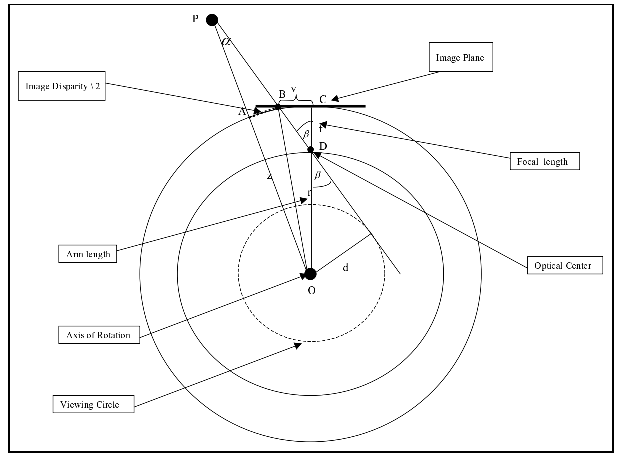

Figure 7. Computing the disparities in an omnistereo mosaicing system.

4.1 Measuring Image Disparity

The image disparity is measured in pixels. The process consists of the following steps:

1. Two panoramas are constructed, a left eye panorama and a right eye panorama, using a fixed separation between the strips.

2. The two panoramas are aligned for vergence at infinity. Objects at infinity will have no disparity.

3. The horizontal image disparity between the two panoramas is computed. To increase efficiency, rather than computing the disparity for every pixel, we compute only three disparities for each column: for its top, central, and bottom parts. These disparities are computed using correlation for a window around each column.

4. The maximum disparity for each column in the mosaic is selected from the three computed disparities (for the top, center, and bottom parts of the column). This gives a disparity value for each column.

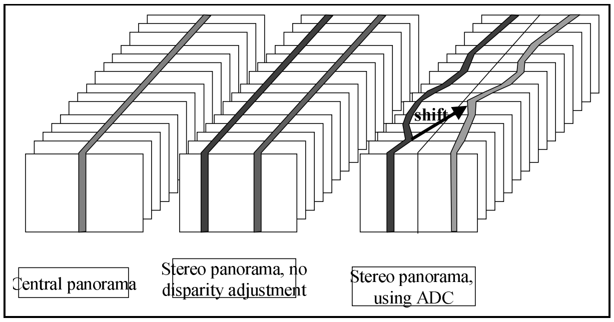

Figure 8.

Schematics of adjusting disparity by adjusting the strip separation. A set of images from a rotating camera is represented as an x-y-t volume. A regular panorama is a planar cut by an y-t plane in the center of the x-y-t volume. In omnistereo each panorama is a planar cut by an y-t plane, but at the sides of the x-y-t volume and at fixed distance from the center. With the automatic disparity control the distance of the strip from the center is changing according to the desired baseline.

5. A median filter is applied to the disparity values to reduce the effects of noise.

This process is illustrated in Fig 9

4.2 Disparity Adjustment

Once the disparity has been measured, we would like to adjust it for better stereo perception:

1. Keep the stereoscopic disparity within the fusion limits of the human brain. This capability is approximately $\pm 0.5^o$, which is about 30 pixels of image disparity.
2. Stretch the disparity in areas where all objects have small disparity range and hence increase the stereo depth sensation.

Under the assumption that the strip separation is considerably smaller than the radius of rotation and the scene depth, then the relation between the strip separation and image disparity is nearly linear.

For enhanced stereo, the maximal disparity which was computed in the previous section for each column should be modified to be approximately 30 pixels. The separation between the strips is increased or decreased proportionately to the desired increase or decrease in disparity. To avoid sharp disparity changes a median filter is applied to the obtained separation values to discard isolated peaks.

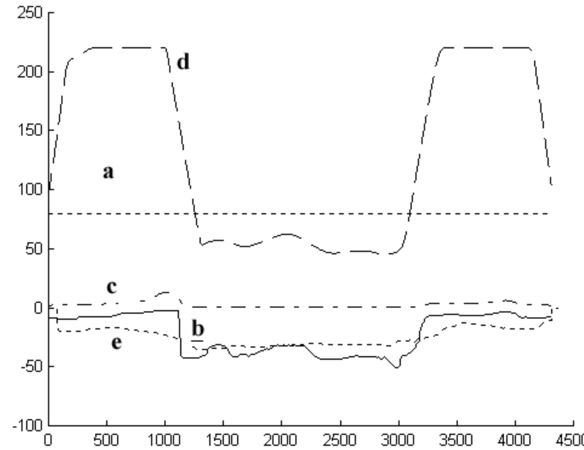

Figure 10. Adjusting separation between strips.
(a) Initial strip gap of 80 pixels.
(b) Disparities measured on the mosaics using (a).
(c) Correcting factor (see text).
(d) Adjusted strip separation, larger for distant objects and smaller for close objects.
(e) Disparities after adjustment of separation between strips.

Once the new separation between the strips is computed, both panoramas are computed using the modified strip separation values. Note that since the left panorama and the right panorama view the same location in different frames, the modification of the separation between the strip in the right panorama and the left panorama for a certain column occurs at different frames as seen in Fig 8.

Fig 10 shows the process of the disparity adaptation. The resulting panoramas are shown in Fig 11 Zooming into a detail of the panorama is done in Fig. 12.

4.3 Dynamic Vergence Control

The vergence between two panoramas greatly affects stereo viewing. Knowing the disparity range for the top, middle, and bottom part of each column allows the setting of the vergence in a convenient location usually the middle of the disparity range.

This can be implemented when the panorama is viewed by a dynamic device, such as an HMD, which enables the modification of the vergence according to the current viewing direction. Changing the vergence is accomplished simply by selecting the appropriate starting location of the right image and left image.

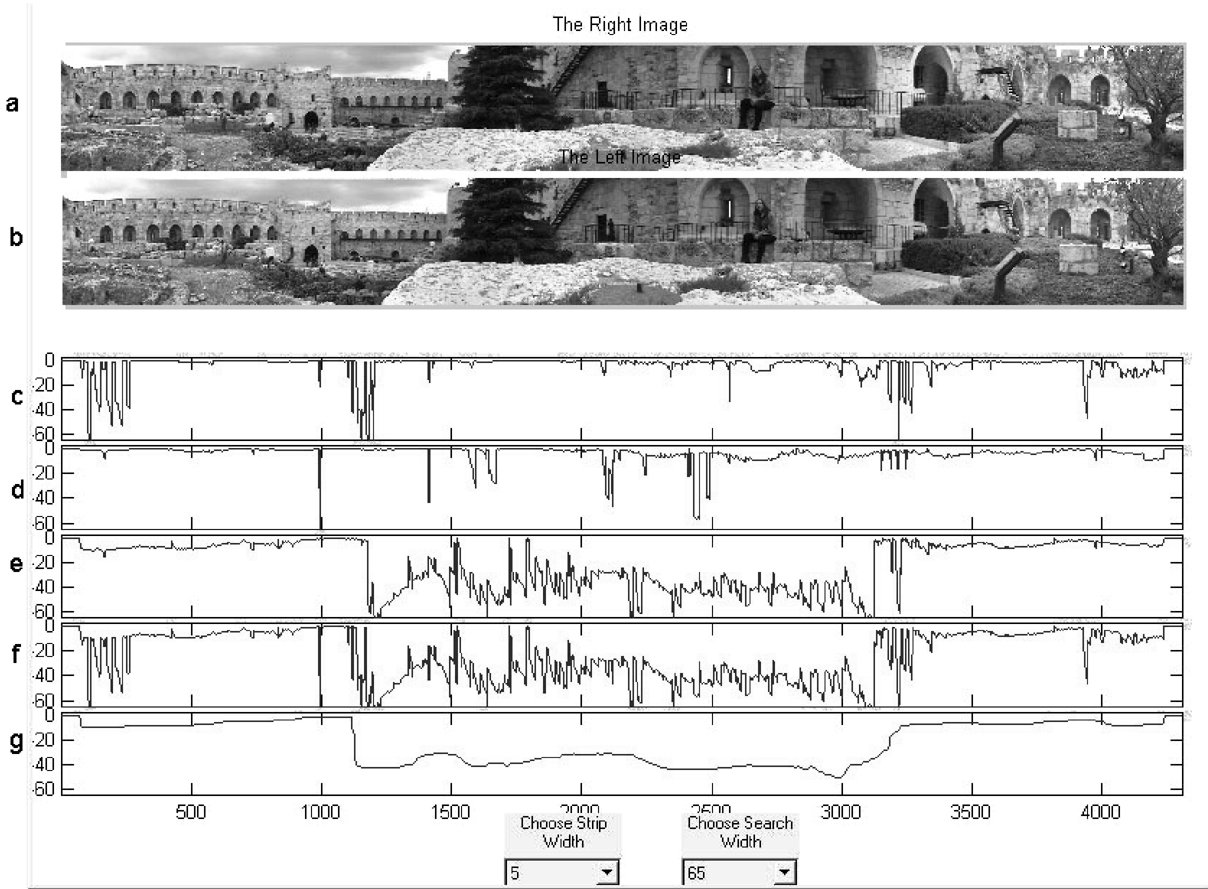

Figure 9. Measuring disparities.
(a) Original left eye panorama. (b) Original right eye panorama.
(c-e) A graph showing image disparities in pixels for the top, center, and bottom parts of the panorama.
(f) Maximum disparity for each column. (g) Median filter of (f) to reduce the effect of noise.

5 Concluding Remarks

In omnistereo panoramic imaging, one fixed baseline is not appropriate for all directions. This paper presents a method that enables the use of different stereo baselines for different directions. A smaller baseline will be used for closer areas, and a larger baseline will be used for far away areas. The baseline determination is done by an automatic selection of the separation between the strips used to build the omnistereo images.

Bibliography

[1] *IEEE Conference on Computer Vision and Pattern Recognition*, San Juan, Puerto Rico, June 1997.

[2] *Seventh International Conference on Computer Vision*, Kerkyra, Greece, September 1999. IEEE-CS.

[3] S. Chen. Quicktime VR - an image-based approach to virtual environment navigation. In *SIGGRAPH'95*, pages 29–38, Los Angeles, California, August 1995. ACM.

[4] J. Gluckman, S. Nayar, and K. Thoresz. Real-time omnidirectional and panoramic stereo. In *DARPA IUW'98*, pages 299–303, Monterey, California, November 1998. Morgan Kaufmann.

[5] H.-C. Huang and Y.-P. Hung. Panoramic stereo imaging system with automatic disparity warping and seaming. *Graphical Models and Image Processing*, 60(3):196–208, May 1998.

[6] H. Ishiguro, M. Yamamoto, and S. Tsuji. Omni-directional stereo. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 1992.

[7] T. Kawanishi, K. Yamazawa, H. Iwasa, H. Takemura, and N. Yokoya. Generation of high-resolution stereo panoramic images by omnidirectional sensor using hexagonal pyra-

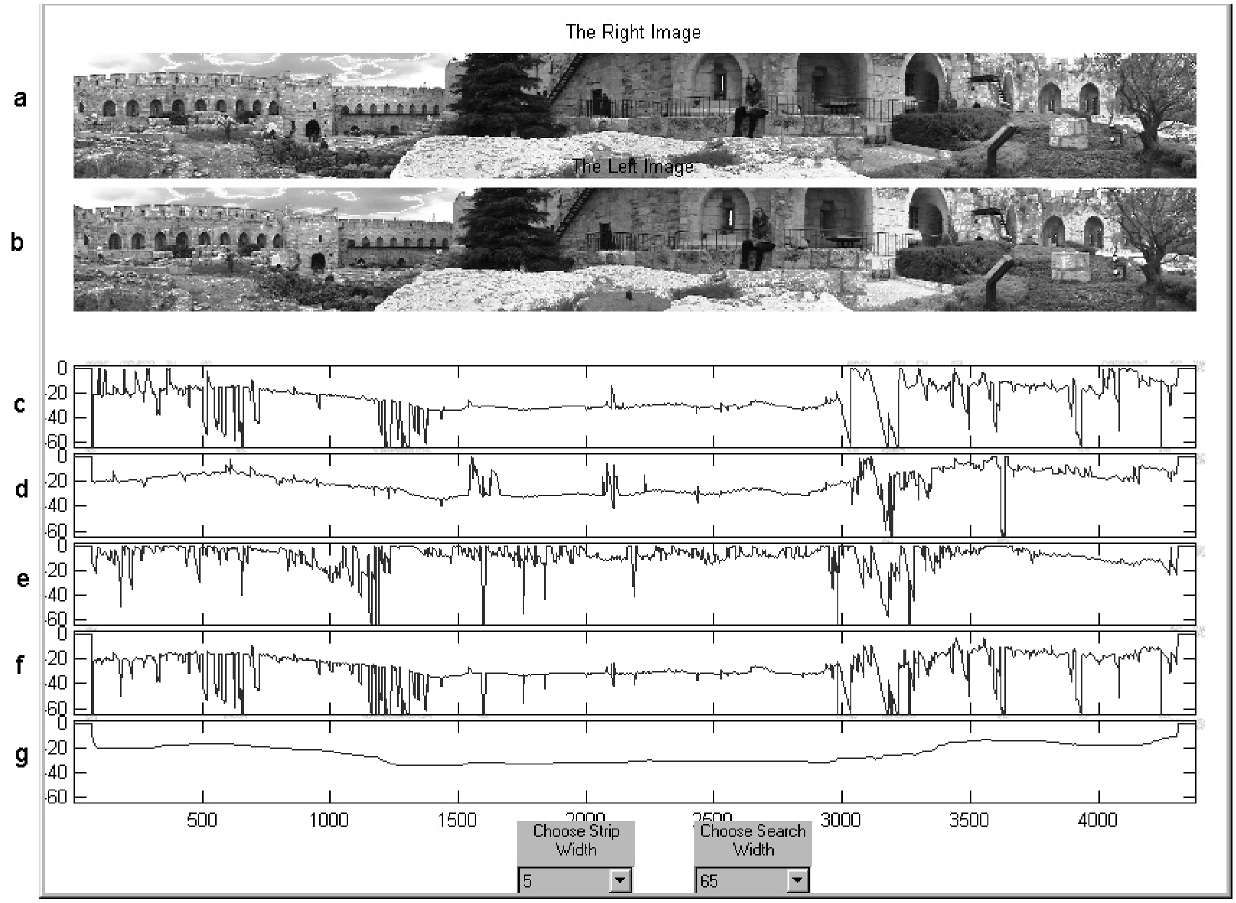

Figure 11. Measuring disparities for the enhanced panoramas.
(a) Original left eye panorama. (b) Original right eye panorama.
(c-e) A graph showing image disparities in pixels for the top, center, and bottom parts of the panorama.
(f) Maximum disparity for each column. (g) Median filter of (f) to reduce the effect of noise.

midal mirrors. In *14th International Conference on Pattern Recognition*, pages 485–489, Brisbane, Australia, August 1998. IEEE-CS.

[8] S. Mann and R. Picard. Virtual bellows: Constructing high quality stills from video. In *First IEEE International Conference on Image Processing*, volume I, pages 363–367, Austin, Texas, November 1994.

[9] S. Nayar. Catadioptric omnidirectional cameras. In *IEEE Conference on Computer Vision and Pattern Recognition* [1], pages 482–488.

[10] S. Peleg and M. Ben-Ezra. Stereo panorama with a single camera. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 395–401, Ft. Collins, Colorado, June 1999.

[11] S. Peleg and J. Herman. Panoramic mosaics by manifold projection. In *IEEE Conference on Computer Vision and Pattern Recognition* [1], pages 338–343.

[12] P. Rademacher and G. Bishop. Multiple-center-of-projection images. In *SIGGRAPH'98*, pages 199–206, Orlando, Florida, July 1998. ACM.

[13] H. Shum, A. Kalai, and S. Seitz. Omnivergent stereo. In *Seventh International Conference on Computer Vision* [2], pages 22–29.

[14] H. Shum and R. Szeliski. Stereo reconstruction from multiperspective panoramas. In *Seventh International Conference on Computer Vision* [2], pages 14–21.

[15] R. Szeliski. Video mosaics for virtual environments. *IEEE Computer Graphics and Applications*, 16(2):22–30, 1996.

[16] D. Wood, A. Finkelstein, J. Hughes, C. Thayer, and D. Salesin. Multiperspective panoramas for cel animation. In *SIGGRAPH'97*, pages 243–250, Los Angeles, California, August 1997. ACM.

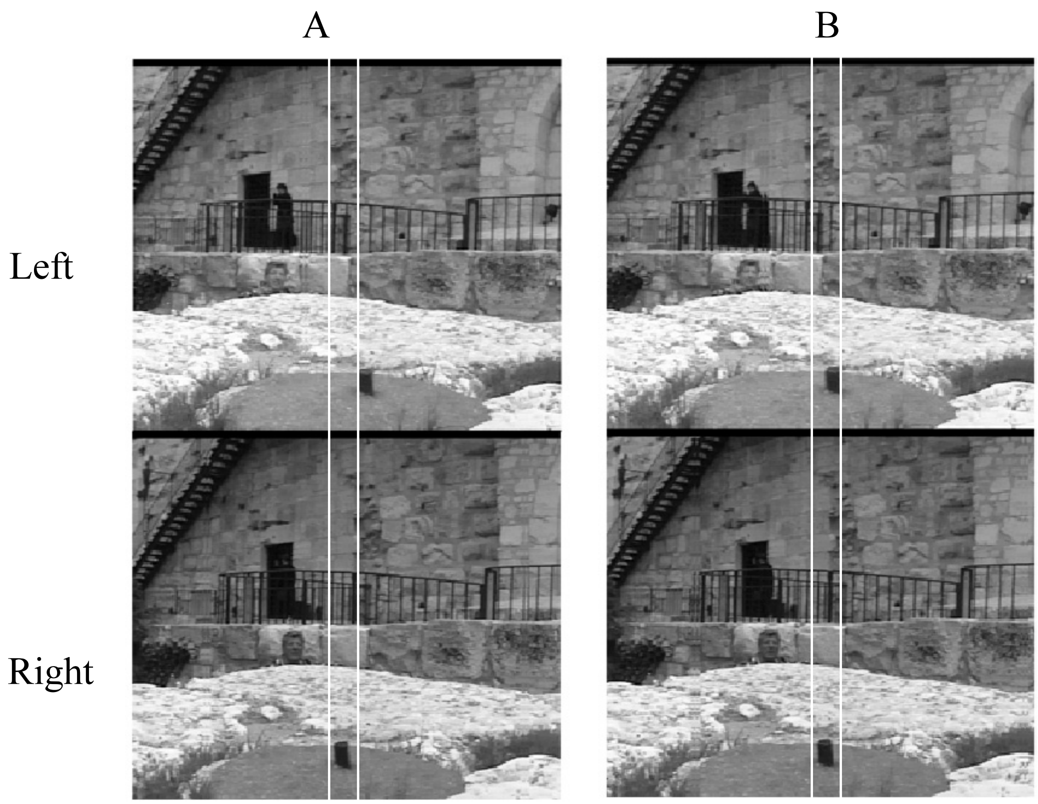

Figure 12. A detail in the disparity adjustment. In the original stereo panorama on the left (A), the disparity of the black pole is too big for stereo fusion, since it is very close to the camera (left eye image is shown above the right eye image). After automatic disparity adjustment (B), the disparity is within the fusion range, enhancing stereo.