Real-Time Warps for Improved Wide-Angle Viewing

Zicheng Liu
zliu@microsoft.com

Michael Cohen
mcohen@microsoft.com

November 2002

Technical Report
MSR-TR-2002-110

Microsoft Research
Microsoft Corporation
One Microsoft Way
Redmond, WA 98052
http://www.research.microsoft.com

Abstract. Providing wide angle views is often advantageous to provide context, such as in video conferencing, or when viewing panoramic imagery. Unfortunately, wide angle images often appear distorted. The limited image size on monitors or paper together with standard viewing distances, results in the angle subtended by the image when viewed usually being much smaller than the field of view of the camera that originally captured the image. This mismatch is at the root of many perceptual problems. One problem is reflected in video conferencing systems where a wide-angle view exaggerates the depth of the room and makes the distant people look extremely small. Other artifacts include the "swimming-motion" when viewing panoramic images and distortions of objects near the corners of wide angle images. In this paper, we present a class of simple real-time image warping functions that address these perceptual problems. The key idea is to perform local scaling while preserving the global context. We show that the warping functions effectively correct the depth errors in the video conferencing applications resulting in more useful and meaningful conferencing presentations. These warping functions can also be used to remove the swimming motion artifacts when viewing panorama images. In addition, we show a specially designed five-lens camera to capture and warp wide angle images in real time without sacrificing resolution.

1 Introduction

Although wide angle perspective images are in some sense *mathematically correct*, they appear distorted under most viewing conditions. Objects near the edges of the frame appear stretched and/or sheared particularly near the corners. Relative depths between near and far objects appear exaggerated compared with our normal viewing experience of the same scene. Why does this occur? The main reason is that under normal viewing conditions, for example viewing an image on a computer monitor, the field of view subtended by the image is far less than that used in the original photographic process. As demonstrated in the Figure 1, a ray from the viewer to the image represents a very different direction ray in the original camera. A small angle from the camera towards the edge of the image represents a much larger angle for the viewer (also heading in a different direction). Just the opposite is true near the center of the image; the same angle from the camera is now much smaller from the viewer's perspective. The effect at the edges causes non-uniform scaling and shearing distortions. The effect at the center causes exaggerated foreshortening of distant objects.

These distortions are even more evident when viewing a panoramic image with most panorama viewing software. As one (virtually) turns his/her head, objects appear at one edge already stretched out, then shrink depending on distance as they pass the center, and finally are stretched out again at the other edge. This makes solid objects appear to deform and *swim* out and then in as the view is rotated.

In applications such as video conferencing wide angle views are useful to provide the meeting context and simultaneously depict all participants. This

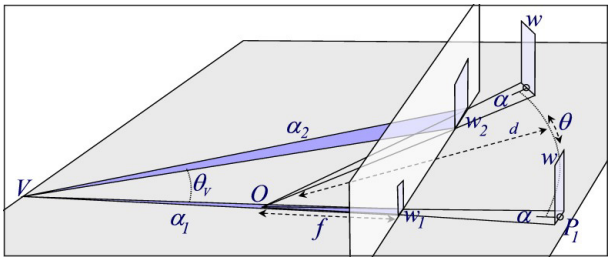

Figure 1: Two objects that subtend equal angles in a wide angle camera view do not when the viewpoint is moved back. This is the source of the distortion.

causes people in the middle of the image who are also usually further away from the camera to appear very small compared to others due to the exagerated foreshortening. For example, in our conference room, the screen is located at one end of a long conference table. Placing the camera near the screen at the end of the table results in images in which the person at the far end of the table looks much more distant than what the viewer would perceive if he/she were physically at the camera position (see Figure 2). Some researchers have proposed to place a omnidirectional video camera at the center of the table so that the people around the table have similar image sizes (when the table is close to circular) [3]. However, this does not maintain the meeting context such as who is looking where or who is talking to whom.

Figure 2: A cylindrical projection of a conference room. Compare to the results of our real-time warping function in Figure 14.

We present a family of warping functions to address these perceptual problems. The functions are designed to minimize the introduction of new distortions while overcoming most of the drawbacks of wide angle images. The main idea is to perform (almost) uniform local scaling while preserving the global context. For the video conferencing application, the warp scales up distant people without distorting each individual while preserving the context of the room. These same warping functions are also effective at removing the swimming motion artifacts when viewing panorama images.

Since the warping function do not scale all parts of the image the same, the resulting warped images have very uneven resolution. For video conferencing

applications, some regions may be zoomed up significantly resulting in poor image quality. To address this problem, we have constructed a new five-lens camera, each with a different field of view to obtain more uniform resolution after the application of our warping functions. We demonstrate real-time capture, stitching, and warping of the pixels from the five-camera array.

2 Related Work

Depictions of wide angle views using linear perspective have been created for the last six hundred years. The rules, first put in writing by Leon Battista Alberti in the fifteenth century were utilized by artists and architects who established a new human (vs. divine) centered world view. However, even as the rules were established, they were also violated by artists who recognized that objects became distorted. For example, we expect a sphere to always project as a circle (and artists continued to do this) but this is not true under linear perspective. Linear perspective always keeps straight lines straight at the expense of maintaining shape. The limitations of linear perspective projection are well known in art photography [4]. There have also been considerable studies on human perception of pictures in the psychology field [8, 5].

Wide angle and panoramic photography is almost as old as photography itself. Ingenious panoramic cameras were developed with curved film and a rotating slit lens to image very wide angle scenes. These cameras produce *cylindrical projections* that do not keep straight lines straight but are better at maintaining shape. Panoramic photographers recognized this. By assembling large groups of people to be photographed in a circular arc rather than in a straight line resulted in images in which the group appeared to be standing in a straight arrangement and each member was approximately the same size. There is a wealth of images and information on the internet (search for "panorama photography" in Google). See, for example, http://memory.loc.gov/ammem/pnhtml/pnhome.html, from the U. S. Library of Congress collection.

More recently, there has been a lot of progress on large view image sensors which can capture up to 360 degree panorama images at video frame rates [11, 10, 1, 2, 12, 6]. For video conferencing applications, one could place such an omnidirectional video camera at the center of the table to achieve a uniform resolution when the table is approximately circular or square [3]. Since the place one cuts the cylinder to flatten out the panorama is arbitrary, this can confuse the meeting context. For example, two people facing each other may appear to be facing in opposite directions in the result. Visual cues such as who is looking where and who is talking to whom, are very important for communication and we would like to preserve them.

Zorin and Barr [15] applied perceptual principles to the analysis and construction of viewing transformations in computer graphics. Our work was inspired in part by theirs. Zorin and Barr demonstrated that when 3D information is available, one can move the view point back to generate images with smaller field of views which can be perceived better by humans. They also established

two conflicting desirable aspects of images; that straight lines stay straight, and shape is maintained. The first is provided by linear perspective, the second by a spherical projection. They provided a parametric family of projections to try to balance between these. They did not, however, attempt to minimize the problems of non-rigidity in panorama viewers, nor did they address the problems of exaggerated foreshortening in scenes with large depth variations. We extend their work in these directions.

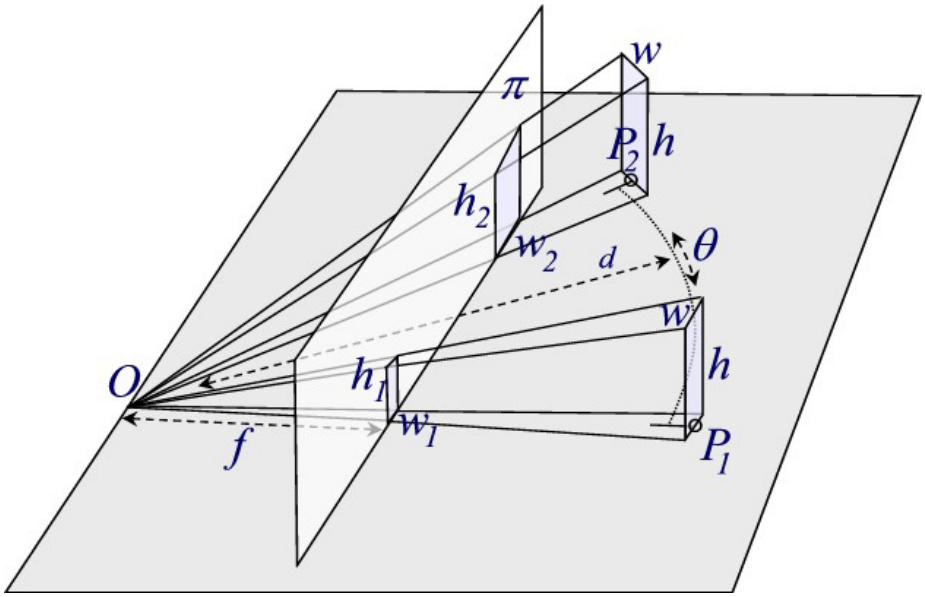

Figure 3: Perspective projection of an object rotating around the camera

3 Wide Angle Image Distortion

We first discuss in more detail the causes of apparent distortions in wide angle images. We will then address how to ameliorate these problems.

3.1 Shape Distortion and Non-rigidity Artifacts

As objects move from the center to the edges in wide-angle perspective images they appear to distort. Figure 3 shows a perspective projection of a rectangular object rotating around the camera. O is the projection center, π is the projection plan, and f is the focal length. Let w and h be the width and the height of the object. Suppose the object rotates from P_1 to P_2, where OP_1 is perpendicular to the image plane and θ is the angle between OP_2 and OP_1. Let $d = ||OP_1|| =$

$||OP_2||$. At position P_1, its width and height on the image plan are

$$w_1 = \frac{fw}{d} \text{ and } h_1 = \frac{fh}{d} \tag{1}$$

At P_2, its width and height (the lengths of the two edges adjacent to P_2) on the image plan are

$$w_2 = \frac{fw}{d \cdot cos^2\theta} \text{ and } h_2 = \frac{fh}{d \cdot cos\theta} \tag{2}$$

So we have

$$w_2 = \frac{1}{cos^2\theta} \cdot w_1 \text{ and } h_2 = \frac{1}{cos\theta} \cdot h_1 \tag{3}$$

and thus the ratio

$$\frac{w_2}{h_2} = \frac{1}{cos\theta} \frac{w_1}{h_1} \tag{4}$$

Equation 3 shows that both the width and height increases as the object rotates from the image center to the edge. The larger the field of view angle, the more the amount of increase. Equation 4 shows that the image of the object also deforms (stretches) when the object rotates from the image center to the edge. The larger the field of view angle, the more the amount of deformation. As the field of view approaches 90 degree, the deformation ratio $\frac{w_2}{h_2}$ approaches infinity.

The perspective deformations are still *correct* in that if the viewer of the image places their eye at O, the image would not appear deformed. However, when viewing the wide-angle image with a smaller field of view the viewer expects smaller increases of image sizes as well as smaller amount of deformation on the image plane as the object rotates. This is why objects appear stretched at the edges. The two objects that used to subtend equal angles (α in Figure 1) now do not (angles α_1 vs. α_2). The larger than expected changes in size and deformation on the image plane make the user feel that the scene is not rigid as if it were swimming around him/her.

The shape distortions can be ameliorated somewhat by using a cylindrical projection as in Figure 2. At least in the horizontal direction shape would remain constant, but this would come at a cost of having horizontal straight lines appear curved. Spherical projections maintain local shape but result in all straight lines appearing curved. Neither projection, however, addresses the exaggerated foreshortening (and associated depth mi-perception) evident in the teleconferencing setting.

3.2 Depth Misperception

A related problem caused by the smaller field of view at viewing time is the misperception of depth. Wide-angle images exaggerate the depth disparity between near and far objects. One important visual cue of the depth of an object

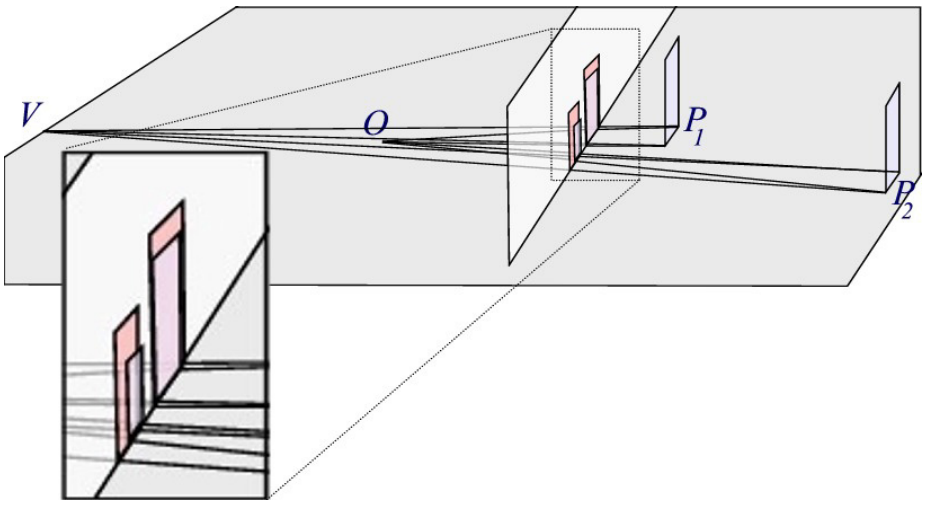

Figure 4: The change in size between near and far objects is much greater in the wide-angle image (in blue) than one would expect from the new viewing distance (in pink). This causes exaggerated depth disparities.

in a scene is the ratio (called *depth foreshortening ratio*) between the image sizes of similar objects placed at near and far locations. A smaller field of view results in a smaller foreshortening ratio as can be seen in Figure 4. Figure 5 (a) shows a photograph taken with a regular field of view, while (b) shows the same scene taken with a wide-angle lens (from a nearer vantage point). Notice the exaggerated difference between the sizes of the two people in (b), and thus the distance between the two people appears much larger in the wide angle view. When we view the image of a deep scene such as the image shown in Figure 2 on a computer monitor, our field of view is, in general, much smaller than the field of view of the actual images. Therefore, the depth we perceive is much larger than the actual depth. In such a teleconferencing application, the wide-angle view results in distant participants being extremely small in the scene (see perspective view in Figure 11).

4 Spatially Varying Uniform Scaling Function

In this section, we describe a parametric class of image warping functions that attempt to minimize the image perception problems described above caused by the viewer having a smaller field of view than the imaging apparatus. We try to balance the introduction of new distortions (e.g., straight lines becoming curved, loss of aspect ratio, unnatural rotations) with the benefits of limiting wide-angle distortion and depth misperception. We will use teleconferencing as the driving application and then apply results to panoramic images in general.

We call the class of warping functions we define *Spatially Varying Uniform*

(a) (b)

Figure 5: (a): Regular field of view. (b): Wide angle field of view.

Scaling functions, or SVU scaling for short. These functions locally resemble a uniform scaling function to preserve aspect ratios, however, the scale factor varies over the image to create the warp. The class of *conformal* projections can provide local uniform scaling, however, they introduce rotations which are visually disturbing. This led us to the SVU scaling functions that avoid rotations at some costs in terms of introducing shear.

We will use the example shown in Figure 2 to describe the SVU scaling. The images are captured in real-time using a five-lens device we describe later. After stitching, this provides us with a full 180 degree cylindrical projection panoramic image. A perspective projection would have infinite width so only a smaller field of view projection is possible, for example, the 110 degree view in Figure 11. Both the cylindrical and (especially) the perspective projections depict very exaggerated depths. We will work from the cylindrical image since we want to preserve the widest possible view for the teleconferencing application.

To correct for exaggerated depth, we need to enlarge the image of the distant participants relative to the people who are close to the camera. In other words,

we would like the warping function to be such that it zooms up the center more than the sides while locally mimicking a uniform scaling. We would like to avoid rotations (as might appear in conformal projections), particularly keeping vertical lines vertical. The warp we initially describe induces some vertical shear, thus slanting horizontal lines. We describe at the end of this section a modification that corrects for much of this at some cost to aspect ratio near the top and bottom boundaries.

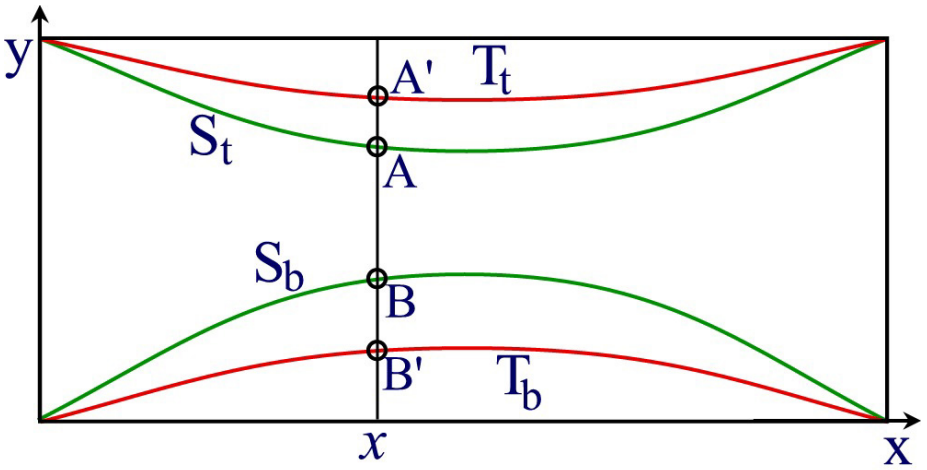

Figure 6: The warping function is determined by two sets of curves: source (green) and target (red) curves.

The user, via a simple user interface, provides the parameters for the SVU scaling function. The user is asked to define two cubic curves (see Figure 6). These two *source* curves define common (real world) horizontal features such as the tops of people's heads, and the edge of the table. A factor, α also chosen by the user determines how much the image is warped.

Let $y = S_t(x)$ and $y = S_b(x)$ be the equations of the top and bottom source curves respectively. Two *target* curves (where points on the source curves will move to) are determined by the source curves and α. If we denote the equation of the line between the end points of $S_t(x)$ as $y = y_t(x)$, and the equation of line connecting the bottom source ends as $y = y_b(x)$, then the top target curve is $T_t(x) = (1-\alpha)S_t(x)+\alpha y_t(x)$, and $T_b(x) = (1-\alpha)S_b(x)+\alpha y_b(x)$. An $\alpha = 0$ will leave the image untouched. An $\alpha = 1$ will pull pixels on source curves to the lines between the end points. For example, the four curves shown in Figure 6 consist of two green source curves and two red target curves.

Given any vertical scanline x as shown in Figure 6, let A, B denote its intersections with the source curves, and A', B' the intersections with the target

curves. The SVU scaling function will scale AB to $A'B'$. Let

$$\begin{aligned} r(x) &= \frac{||A'B'||}{||AB||} \\ &= \frac{T_t(x) - T_b(x)}{S_t(x) - S_b(x)} \end{aligned} \tag{5}$$

We scale the line vertically by $r(x)$, and to preserve aspect ratio we also scale the scanline horizontally by $r(x)$. Therefore, the total width of the new image, w', becomes

$$w' = \int_0^w r(x)dx \tag{6}$$

where w is the width of the source image.

For any pixel (x, y) in the source image, let (x', y') denote its new position in the warped image. We have

$$\begin{aligned} x' &= \int_0^x r(x)dx \\ y' &= T_t(x) + r(x) * (y - S_t(x)) \end{aligned} \tag{7}$$

This is the forward mapping equation for the SVU scaling function. The SVU scaling function is not a perfect uniform scaling everywhere. It is easy to prove that the only function that is a perfect uniform scaling everywhere is a uniform global scaling function.

The SVU scaling function is similar to a projection onto a generalized cylindrical surface. However, as shown in the appendix, such a simple projection does not produce locally uniform scaling.

4.1 Horizontal Distortion Correction

While the SVU-scaling function maintains vertical lines as vertical, it distorts horizontal lines. The distortions are smallest between the source curves and largest near the top and bottom. Scenes often contain horizontal surfaces near the top or bottom, such as a table and the ceiling on a room for which the distortions may be noticeable (see Figure 2). To minimize this problem we relax the uniformity of the scaling and nonlinearly scale each vertical scanline. The portion of the image between the source curves is scaled by $r(x)$ as described above. The portions outside the source curves are scaled less in the vertical direction. The horizontal scaling remains the same (i.e., $r(x)$) to maintain the straightness of vertical lines. To maintain continuity, the vertical scaling function smoothly transitions as it crosses the source curves.

Consider the vertical line in Figure 6. Denote $g(y)$ to be the vertical scale factor at any point y on this vertical line (see Figure 7). Note that $g(y)$ is dependent on x. $g(y)$ is controlled by two parameters s and ω. The portion of the vertical scanline more than $\omega/2$ distance from the source curves is scaled by

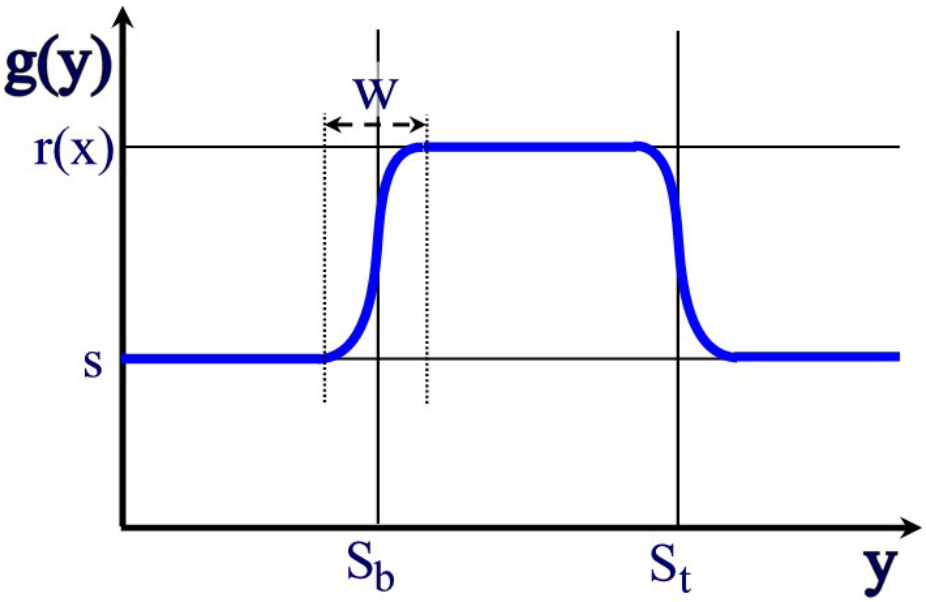

Figure 7: The vertical scale function.

$r(x)$ between the source curves and by s outside the source curves. The three constant segments are glued together by two cubic splines in $[S_t - 0.5\omega, S_t + 0.5\omega]$ and $[S_b - 0.5\omega, S_b + 0.5\omega]$. Each cubic spine has ends with values s and $r(x)$ and a slope of 0 at both ends.

The parameter ω controls the continuity at the source curves. For example, if the scene is discontinuous at the source curves, one can choose a very small ω without noticeable artifacts. In the special case when $s = r(x)$, $g(y)$ becomes a constant which is what we assume in deriving Equation 7.

Figure 8: Top: Viewing a panorama of a ball room with perspective view . Bottom: After applying SVU scaling.

5 Correcting for Panoramic Image Viewing

Swimming motions can be thought of as a special type of depth mis-perception. As the scene rotates, the viewer, with a smaller field of view, sees unexpected changes in the size of objects as they cross the field of view. An object looks too large as it first comes into view due to geometric distortion, then it is too small in the center of the view due to exaggerated foreshortening, and finally is too large again as it leaves the view. The result is that objects appear move away from the viewer and then move back towards the viewer (i.e., it appears to swim).

The SVU scaling function helps correct for the problems of depth mis-perception, and not surprisingly removes much of the swimming motions. We use a symmetric SVU scaling function for viewing panoramic images. The bottom source curve is the mirror of the top source curve. The source curves pass through the image corners and have a minimum (maximum) at the center vertical scanline. One parameter specifies the height of the center points and a second parameter is the α as before.

It should be noted, that a simple cylindrical projection also removes the apparent swimming motion. However, it curves straight lines more than necessary, and also removes, almost completely, the illusion of turning one's head.

The top row of Figure 8 shows five snap shots during a panorama viewing of a ball room with a perspective projection (field of view is 100 degree), while the bottom row shows the same views after SVU scaling. We can see that the images are less distorted after SVU scaling. More importantly as we rotate the scene with a perspective projection, we see serious stretching and swimming motions on the wall and the tables. After applying SVU scaling, we see very little stretching and swimming motions are basically gone. In other words, the scene looks much more stable while it is rotating.

6 Warping for Teleconferencing

Teleconferencing systems need to be able to display a wide set of views to allow remote viewers to see all the participants, typically seated around a table. Figure 2 shows a meeting room installed with a teleconferencing system. Most current systems use a tilt/pan/zoom camera to provide the possibility of viewing different portions of the room. However, these systems cannot provide a complete overview. They also suffer from latency when panning to cover a new speaker. Instead, we have constructed a system combining a novel camera with our warping function.

In an overview image such as the one shown in Figure 2, the images of the people at the far end of the table are too small due to the wide-angle projection. This can cause serious communication barriers.

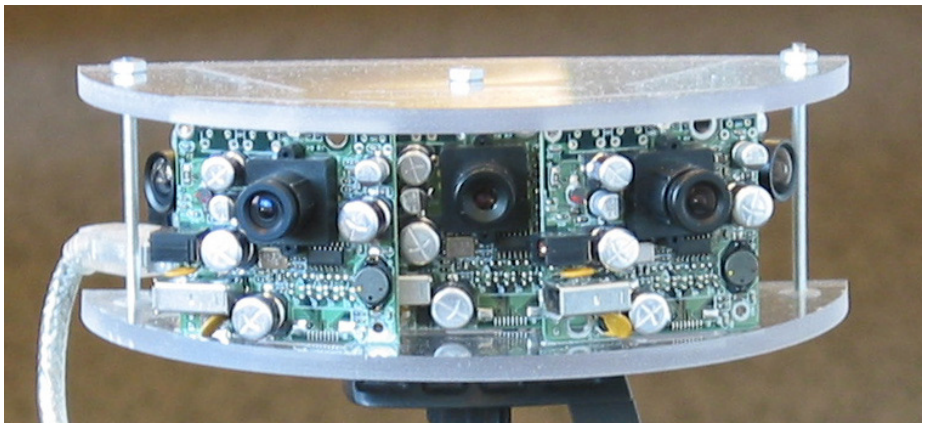

Figure 9: The half-ring camera.

Figure 10: Images from the five cameras.

6.1 Half-ring Camera Array

The SVU-scaling function addresses this problem. However, if we directly apply our warping function to this wide-angle image, the extreme enlargement of the far people will be very blurry due to the limited resolution of the image in this area. To solve this problem, we have built a special "half-ring" video camera consisting of five inexpensive ($< \$50 each$) fire-wire video cameras daisy-chained together (See Figure 9). A single IEEE 1394 fire-wire delivers five video streams to the computer. The resolution of each camera is 640 by 480. Each camera has a different lens. Figure 10 shows the five images directly from the five video cameras. The center camera has the smallest field of view (about 25 degrees) to provide enough resolution for the distance. The field of view of the two cameras next to the center are 45 degrees, with the outer having the largest field of view (60 degrees). Together, they cover 180 degrees with enough overlap between neighboring cameras for calibration and image stitching.

We use well-known techniques to calibrate these cameras and compute the homography between the cameras [14, 7, 9, 13]. We then stitch the individual images together to generate a 180 degree cylindrical image (see Figure 2). Computation overhead is reduced at run time by pre-computing a stitch table that specifies the mapping from each pixel in the cylindrical image to pixels in the five cameras. For each pixel in the cylindrical image, the stitch table stores how many cameras cover this pixel, and the blending weight for each camera.

Blending weights are set to one in most of the interior of each image with a rapid fall off to zero near the edges. Weights are composed with an *over* operator where the higher resolution pixel is composed over a lower resolution one. At run time, we use a look up the table to perform color blending for each pixel.

6.2 SVU Scaling the Stitching Table

Applying the SVU scaling function to the stitched image would result in a loss of resolution. Instead, we warp the stitch table itself, and generate a new table. During this offline warping, we use bilinear interpolation to fill in zoomed-up regions to avoid losing resolution. At run time, we generate the warped images by a simple look-up in the pre-warped stitch table. The complete stitching, blending, and warping are computed in a single frame time.

6.3 Teleconferencing Results

Figure 11 shows a 110 degree perspective projection. The person at the far end of the table is extremely small. Figure 2 is the cylindrical projection. While the geometric distortions around the edges are corrected by cylindrical projection (consistent with the work by Zorin and Barr [15]), there is little improvement in terms of depth misperception. People at the far end of the table still looks very small compared to the people near the camera. Figure 12 shows both the source and target curves with $\alpha = 0.3$. Figures 13 through 16 on the final page show the results of using the SVU scaling function. Figure 13 shows the result of applying the SVU scaling function without correcting horizontal distortion. Figure 14 shows the result after correcting for horizontal distortion. Finally we show some results with different α. Figure 15 shows the result with $\alpha = 0.2$, and Figure 16 shows the result with $\alpha = 0.4$.

During live meetings, we store multiple tables corresponding to different α's so that one can change levels in real time. The size of the stitched image is approximately 300 by 1200 pixels. During warping, we keep the image width the same, and as a result, the image height decreases as we zoom up. The frame rate is about 10 frames per second on a CPU with a single 1.7GHZ processor.

Finally, Figure 17 shows a snapshot of a remote person conferencing with people in the meeting room where the meeting room view is generated by using our half-ring camera and SVU scaling functions.

7 Conclusion

We have presented a class of warping functions to address perceptual problems when viewing wide-angle images. One important application of these SVU scaling functions is teleconferencing. We have shown their effectiveness at reducing the exaggerated depths between near and distant participants while limiting distortion of each individual. We have described a special five-lens video camera to capture and warp the images in real time. In addition, we have shown that

Figure 11: 110 degree perspective view.

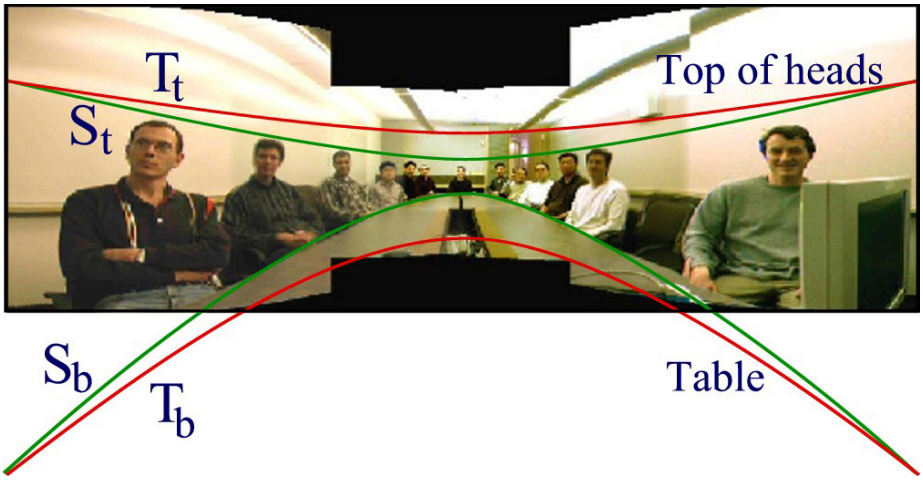

Figure 12: The source curves and the target curves with $\alpha = 0.3$.

the SVU scaling functions can also be used to remove swimming motions when viewing panorama images.

There are a number of enhancements we plan for the camera and image stitching. Our next version of the camera will pack the lens array tighter to minimize parallax (ghosting) and help the calibration process. We will use better cameras and lens as they become available. We are also working on automated color balancing to get the images to match better. Finally we are integrating this work with current efforts on head-tracking and speaker detection to provide close-ups of the speaker (without the need to pan the camera). We expect to deliver both the overview and the close-up to remote participants.

Figure 13: SVU scaling without horizontal distortion correction.

Figure 14: SVU scaling with horizontal distortion correction.

Appendix

In this appendix, we show that a projection onto a generalized cylindrical surface does not produce locally uniform scaling. Let us consider the example of projecting a regular cylindrical view to an elliptic cylinder (the cross section is an ellipse). Figure 18 shows a 2D diagram of the cross sections of the regular cylinder and the elliptical cylinder. Denote a and b to be the short and long axes of the ellipse. Then the equation of the ellipse is $x = a\,cos(\theta)$ and $z = b\,sin(\theta)$. The equation of the circle is $x = a\,cos(\theta)$ and $z = a\,sin(\theta)$. Denote s' and s to be the arc lengths of the ellipse and the circle, respectively. Then

$$\frac{ds'}{d\theta} = \sqrt{(\frac{dx}{d\theta})^2 + (\frac{dz}{d\theta})^2} \quad (8)$$

$$= \sqrt{a^2 sin^2\theta + b^2 cos^2\theta} \quad (9)$$

$$\frac{ds}{d\theta} = a. \quad (10)$$

Given an infinitely small $d\theta$, the arc lengths on the circle and the ellipse are ds and ds', respectively. Therefore the local horizontal scaling is $ds'/ds = \sqrt{a^2 sin^2\theta + b^2 cos^2\theta}/a$. When $\theta = \pi/2$, the local horizontal scaling $ds'/ds = 1$ while the vertical scaling (y-axis) is b/a. Therefore, the vertical scaling is larger than the horizontal scaling and objects will appear stretched vertically. The exact opposite is true when $\theta = 0$. Its vertical scaling is 1 while its horizontal scaling is b/a. Therefore objects close to $\theta = 0$ will appear stretched horizontally.

References

[1] M. Aggarwal and N. Ahuja. High dynamic range panoramic imaging. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2001.

Figure 15: SVU scaling with $\alpha = 0.2$.

Figure 16: SVU scaling with $\alpha = 0.4$.

[2] S. Coorg, N. Master, and S. Teller. Acquisition of a large pose-mosaic dataset. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 872–878, 1998.

[3] R. Cutler, Y. Rui, A. Gupta, J. Cadiz, I. Tashev, L.-W. He, A. Colburn, Z. Zhang, Z. Liu, and S. Silverberg. Distributed meetings: A meeting capture and broadcasting system. In *ACM Multimedia*, 2002.

[4] G. Glaeser. *Fast algorithms for 3-D Graphics*. Springer-Verlag, New York, 1994.

[5] M. A. Hagene. *The perception of pictures. Academic Press series on cognition and perception*. Academic Press, New York, 1980.

[6] R. A. Hicks and R. Bajcsy. Catadioptric sensors that approximate wide-angle perspective projections. In *Workshop on Omnidirectional Vision*, pages 97–103, 2000.

[7] M. Irani, P. Anandan, and S. Hsu. Mosaic based representations of video sequence and their applications. In *International Conference on Computer Vision (ICCV'95)*, pages 605–611, 1995.

[8] M. Kubovy. *The psychology of perspective and Renaissance art*. Cambridge University Press, New York, 1996.

[9] S. Mann and R. W. Picard. Virtual bellows: Constructing high quality images from video. In *First IEEE International Conference on Image Processing (ICIP'94)*, pages I:363–367, 1994.

[10] S. Nayar. Catadioptric omnidirectional camera. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 1997.

[11] S. Nayar. Omnidirectional video camera. In *DARPA Image Understanding Workshop*, 1997.

[12] S. Nayar and A. Karmarkar. 360x360 mosaics. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages II388–392, 2000.

[13] R. Szeliski and H.-Y. Shum. Creating full view panoramic image mosaics and environment maps. In *Computer Graphics, Annual Conference Series*, pages 251–258. Siggraph, 1997.

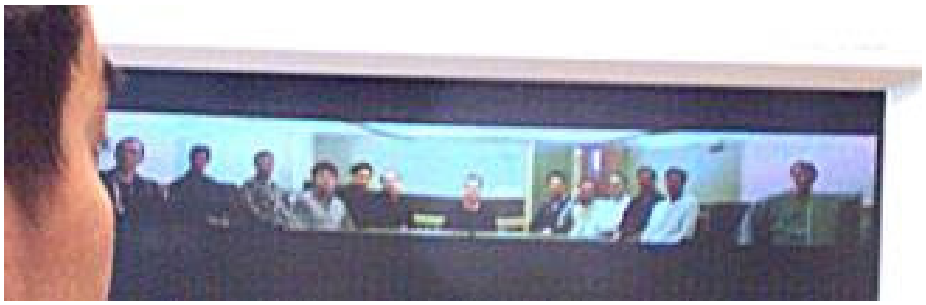

Figure 17: A remote person conferencing with the people in the meeting room.

[14] Z. Zhang. Flexible camera calibration by viewing a plane from unknown orientations. In *International Conference on Computer Vision (ICCV'99)*, pages 666–673, 1999.

[15] D. Zorin and A. Barr. Correction of geometric perceptual distortions in pictures. In *Computer Graphics, Annual Conference Series*, pages 257–264. Siggraph, August 1995.

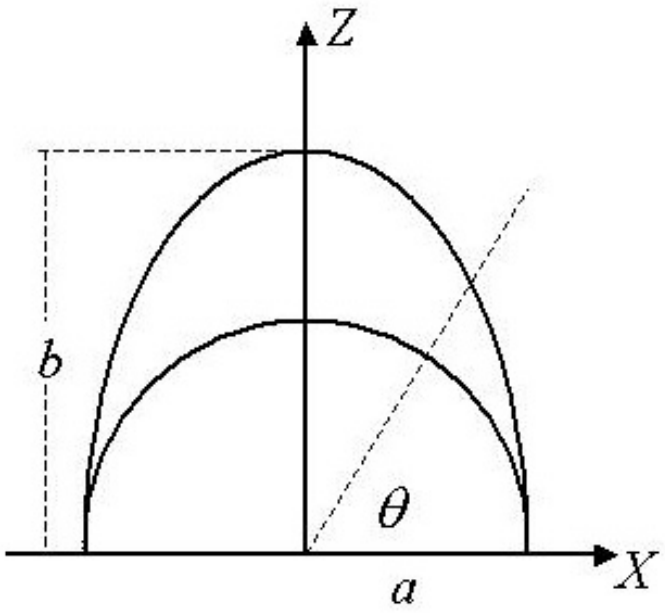

Figure 18: Projection from a regular cylinder to an elliptic cylinder.