Synthetic stereoscopic panoramic images

Paul Bourke
WASP, University of Western Australia, WA 6009, Australia
Email: pbourke@wasp.uwa.edu.au

Abstract. Presented here is a discussion of the techniques required to create stereoscopic panoramic images. Such images allow interactive exploration of 3D environments with stereoscopic depth cues. If projected in a surround display environment they can engage the two characteristics of the human visual system responsible for immersion, namely stereopsis and peripheral vision.

1. Introduction

Panoramic images are a natural way of capturing a wide visual field of view without the distortion introduced by wide angle or fisheye lens. In their ultimate form they capture the whole 360 degrees about one axis, normally the vertical. Depending on whether they are full spherical panoramas or cylindrical panoramas their vertical field of view will range from ±90 degrees to perhaps as low as ±10 degrees. The earliest panoramic images date back to 1850 and the early days of photography [1]. By 1880 there were a number of commercially available panoramic cameras, the most common variety consisted of a stationary cylindrical strip of film and a rotating camera head. In the early 90's panoramic and stereoscopic panoramic capture was explored in the context of capturing the whole visual field for applications in machine vision [2,3] and scene reconstruction. More recently, interactive digital panoramic images were popularised in 1995 with the advent of QuickTime VR [4] on the early Apple Macintosh computers as a way of interactively viewing a photographically captured environment [5]. These digital panoramic images could be created from dedicated cameras, by stitching together a number of images from standard cameras, or by computer graphics rendering processes. QuickTime VR and now a number of similar software playback solutions allow the user to rotate their viewpoint, zoom in/out, and jump to another panorama that is located at a different position.

In this paper a more recent development will be discussed, namely stereoscopic panoramic images. These consist of two panoramic images, one for each eye, such that when presented correctly to our visual system will give a stereoscopic 3D effect irrespective of which part of the panoramic image is being viewed. The discussion will be targeted at cylindrical panoramic images and how they can be created synthetically. The result will be imagery that, in a suitable immersive environment, can simultaneously satisfy two important characteristics of our visual system, namely stereopsis (humans have two eyes separated horizontally) and peripheral vision (up to 160 degrees horizontally).

2. Stereoscopic panoramic rendering

Monoscopic panoramic images are created by rotating a single camera about its optical axis (single viewpoint projection), shown on the left of figure 1. It is easy to show that a stereoscopic panoramic image pair cannot be constructed with simply two offset cameras rotating about their individual optical axes. For example, if the cameras start off pointing parallel to each other for a correct stereo pair then by the time they have rotated by 90 degrees there will be no parallax and effectively two different fields of view. The correct method of capturing stereoscopic panoramic images either photographically or synthetically is shown on the right in Figure 1. Two cameras are separated by the so called interocular distance, the camera view direction is tangential to the circle about which they rotate through 360 degrees. Each camera forms an image from what it “sees” along the view direction through a narrow vertical slit, these are often called circular projections and are a special type of multiple viewpoint projection. The interocular distance relates more to the scale of objects being captured rather than the actual human eye separation, for example, if the virtual model is on the scale of the solar system then the interocular separation, for a satisfying depth effect, may well be on the scale of a few Earth radii.

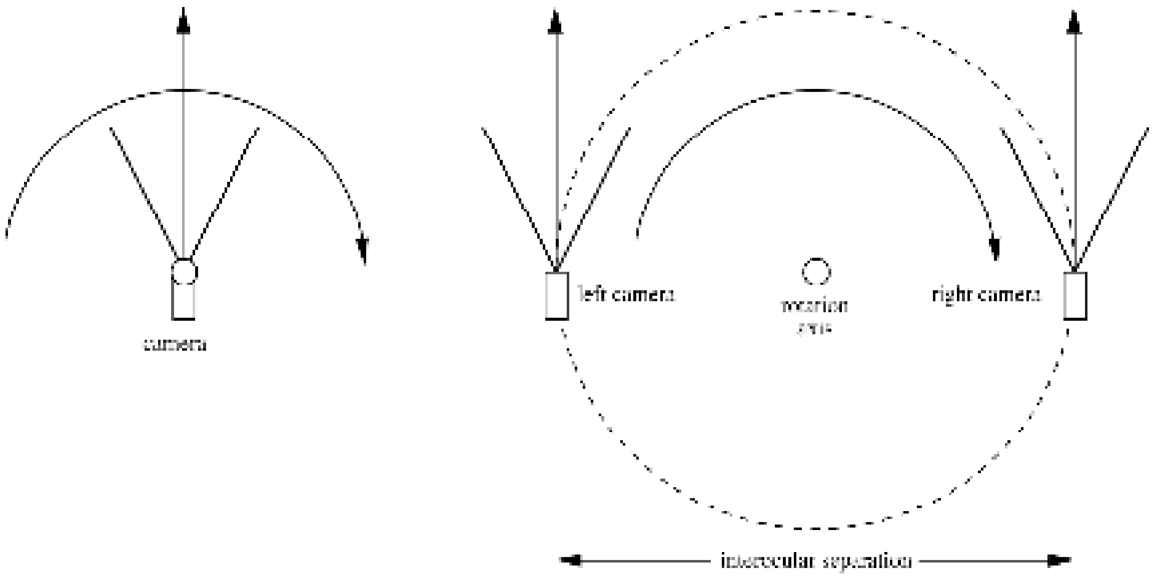

Figure 1. Idealised camera rig for monoscopic panoramic capture (left) and stereoscopic panoramic capture (right).

Ideally the two cameras would rotate continuously, imaging as they do so. There are some physical cameras [6] designed to capture stereoscopic image pairs that function exactly this way, an example of a stereoscopic pair captured with such a camera is shown in figure 2. Alternatively there are other approaches that form each of the panoramic image pairs by stitching together slices from a large number of images from a camera that rotates in discrete steps.

Correct stereoscopic creation in commercial rendering packages is unusual in itself, and while monoscopic panoramic rendering is fairly common, there are no packages that the author is aware of that create stereoscopic panoramic images in a continuous fashion. While some rendering packages have a sufficiently powerful virtual camera model and scripting language to support continuous capture, that is also relatively unusual. The approach that can be implemented in most packages is a finite slit approximation, the resulting panoramic images get progressively better as the slit width becomes narrower. Typically, while 5 degree slit widths may be suitable for distance imagery, it is usually necessary to use at most a 1 degree slit width and often even narrower angular slit widths. For example, if ½ degree slit widths are chosen, the final panoramic images for each eye will be built up from 720 separate images that are very narrow and as tall as the final panoramic image height. With this approach the cameras are also rotated in discrete steps, the stepping angle is identical to the horizontal aperture of the camera ensuring the image strips forms a continuous composite image, this is illustrated in figure 3 for the left eye camera.

Figure 2. Stereoscopic panoramic pair from Hampi in Inidia captured continuously on a roll of film using the RoundShot camera. Left eye (upper) and right eye (lower). Vertical field of view is 60 degrees. Credits: PLACE-Hampi: Sarah Kenderdine, Jeffrey Shaw & John Gollings.

An interesting limitation to the forming of stereoscopic panoramic images compared to traditional monoscopic panoramic images is that in the later one can form the panoramic approximately and even with definite errors without the artefacts necessarily being obvious to the viewer. This is particularly so for photographic panoramic pairs [7,8] that can often include arbitrary stitching and blending of a relatively small number of separate photographs. Because of the additional depth perception in stereoscopic panoramic pairs, approximations and imperfections are usually very noticeable. Additionally, while most panoramic playback applications do not allow camera roll, it was at least possible with monoscopic panoramic images and has been useful for panoramic images where there is no preferred up direction. It is of course not possible to roll (rotation about the view direction vector) inside a stereoscopic panoramic viewer because there is an implied up direction.

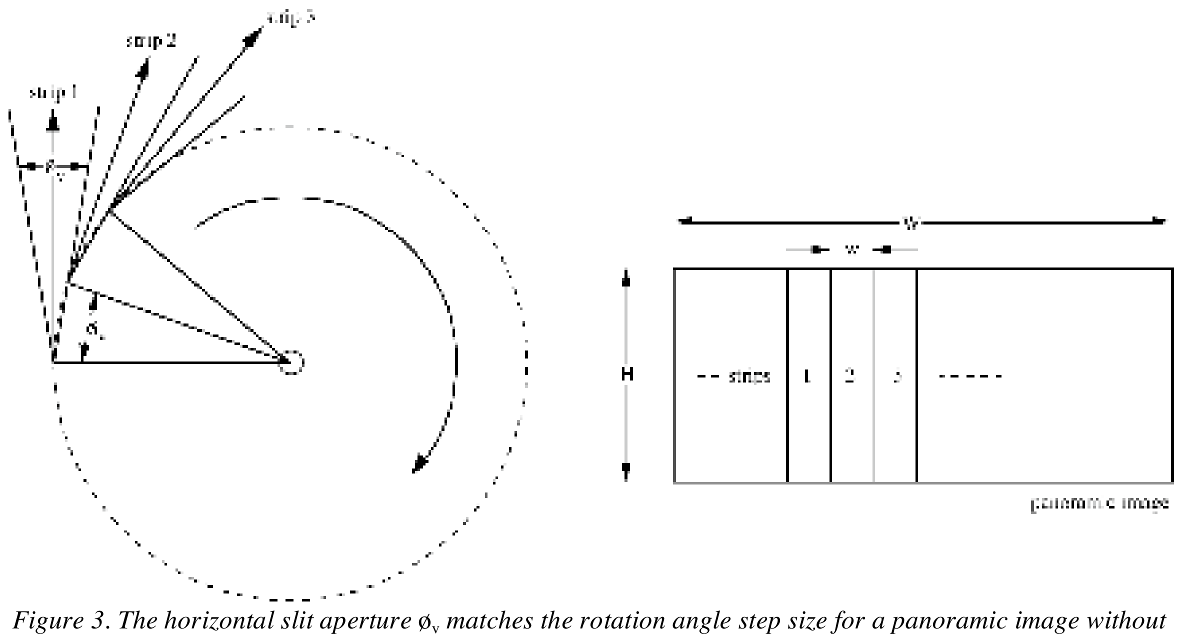

Figure 3. The horizontal slit aperture ϕ_v matches the rotation angle step size for a panoramic image without creating gaps in the imagery.

Common to all cylindrical panoramic images is the relationship between the dimensions in pixel units of the panoramic and the vertical angular field of view ϕ_v. This is given by

$$\phi_v = 2 \operatorname{atan}(H \pi / W)$$

where H and W are the height and width respectively of the panoramic image measured in pixels. When rendering the panoramic in strips the horizontal field of view of the strip ϕ_h in degrees needs to be an integer divisor of 360 degrees in order to achieve a perfect wrap around at the 0 and 360 degree boundary (left and right edge) of the panoramic. The width (w) of the strip in pixels is given by

$$w = W \phi_h / 360$$

where w is also needs to be an integer. This equation can only be solved for quantised values of w, ϕ_h, and W. The relationship between the width of each strip, the height of the panoramic image, and the horizontal and vertical field of view of each strip (assuming square pixels) is given by

$$H \tan(\phi_h / 2) = w \tan(\phi_v / 2)$$

The above relationships define the horizontal and vertical field of view and the aspect ratio, totally defining the frustum of the virtual camera required to render each slit image.

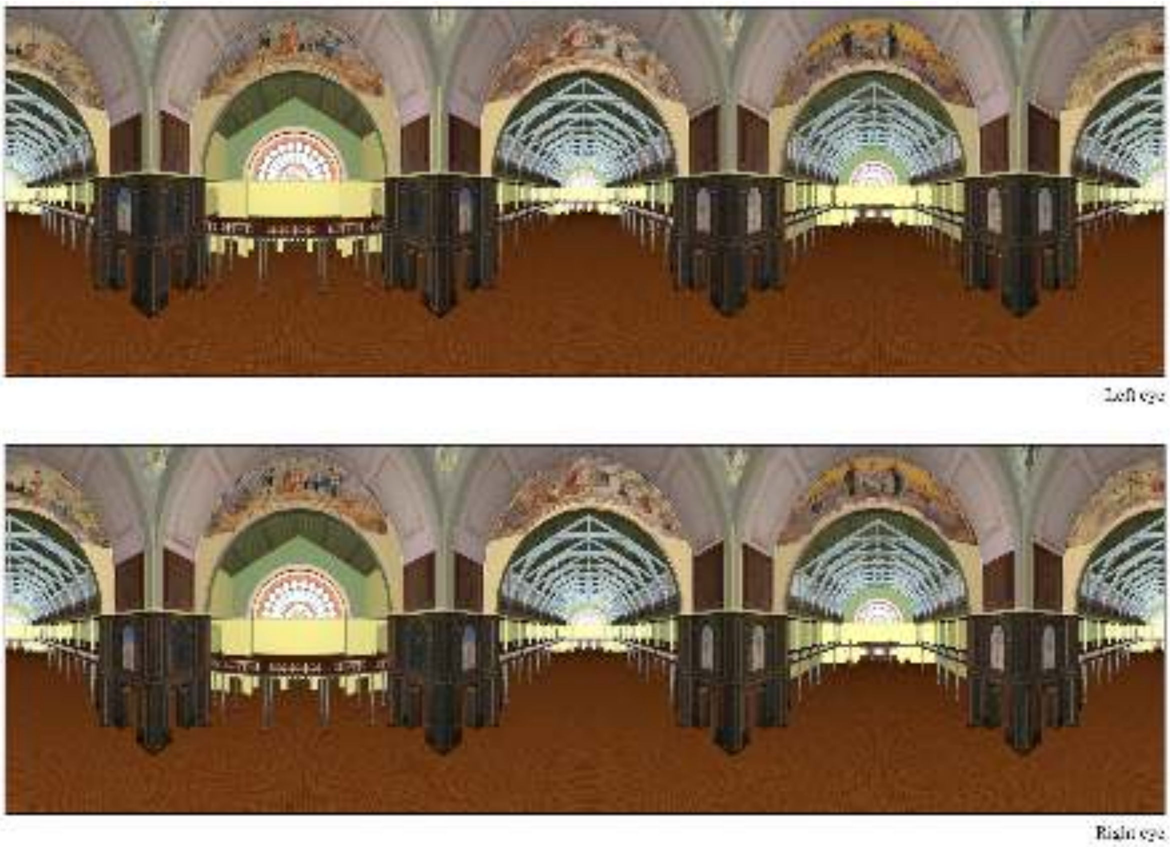

Figure 4. Example of a synthetic stereoscopic panoramic image pair [11] of the Royal Exhibition Building in Melbourne, Australia. Generated using 3DStudioMax and 1 degree slit width. Vertical field of view is 90 degrees.

3. Control of zero parallax

Of central importance to any stereoscopic image generation and viewing is how to control the distance to zero parallax f_o, that is, at what distance in the rendered scene will objects appear to be at the projection screen depth. Objects closer than f_o will appear to be in front of the screen (negative parallax), objects more distant than f_o will appear behind the screen (positive parallax). Unlike perspective stereo pairs where horizontally shifting the images with respect to each other to control zero parallax adds occlusion errors, stereoscopic panoramic pairs can be shifted without adding any additional error. For the camera arrangement proposed here the view direction rays are parallel so the final stereoscopic pairs would have zero parallax distance at infinity. The geometry for calculating the relative shift between the image pairs in order to achieve a particular f_o is shown in figure 5. The angle ø between the left and the right camera such that an object is the desired zero parallax distance will be coincident on the left and right eye panoramic image is given by

$$ø = 2 \operatorname{asin}(r / f_o)$$

where r is half the camera separation (the radius of the rotation circle). The number of pixels (D) to shift the final panoramic images by horizontally is then

$$D = W ø / (2\pi) = W \operatorname{asin}(r / f_o) / \pi$$

where W is the width of the whole panoramic image measured in pixels. Obviously, when shifting the panoramic image it is performed in a circular fashion, those parts of the image that are shifted past the right hand edge of the panoramic image reappear on the left hand edge and visa-versa.

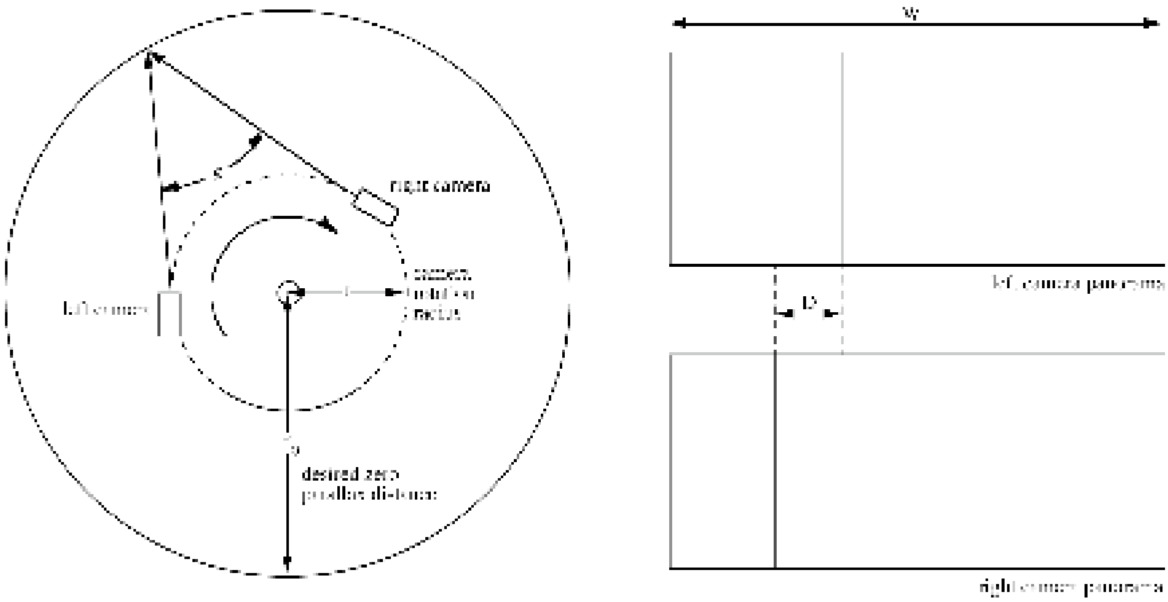

Figure 5. Geometry for determining the distance to shift the panoramic pairs with respect to each other in order to achieve a particular zero parallax distance.

An alternative to parallel cameras is to use what is known in stereoscopic rendering as as "toe-in" cameras. The view direction vector of the two cameras are no longer parallel to each other but rotated inwards and the intersection of their view direction rays determines the distance to zero parallax, see figure 6. While toe-in cameras do not yield correct stereo pairs in traditional perspective stereoscopic image generation (it introduces vertical parallax towards the corners of the image plane), it does not have the same problem when rendering narrow strips for stereoscopic panoramic formation and is a perfectly acceptable solution. As with perspective stereoscopic pairs this approach results in panoramic images with a desired zero parallax distance with no extra post processing required.

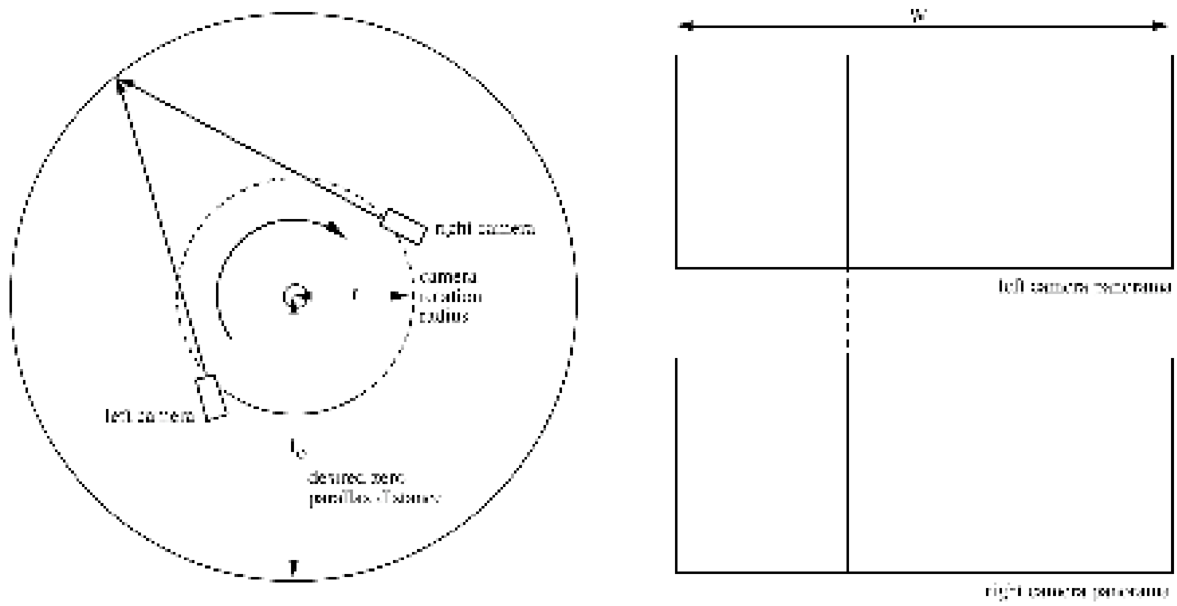

Figure 6. Toe in camera model for direct zero parallax generation. The panoramic images so created have a preset zero parallax distance.

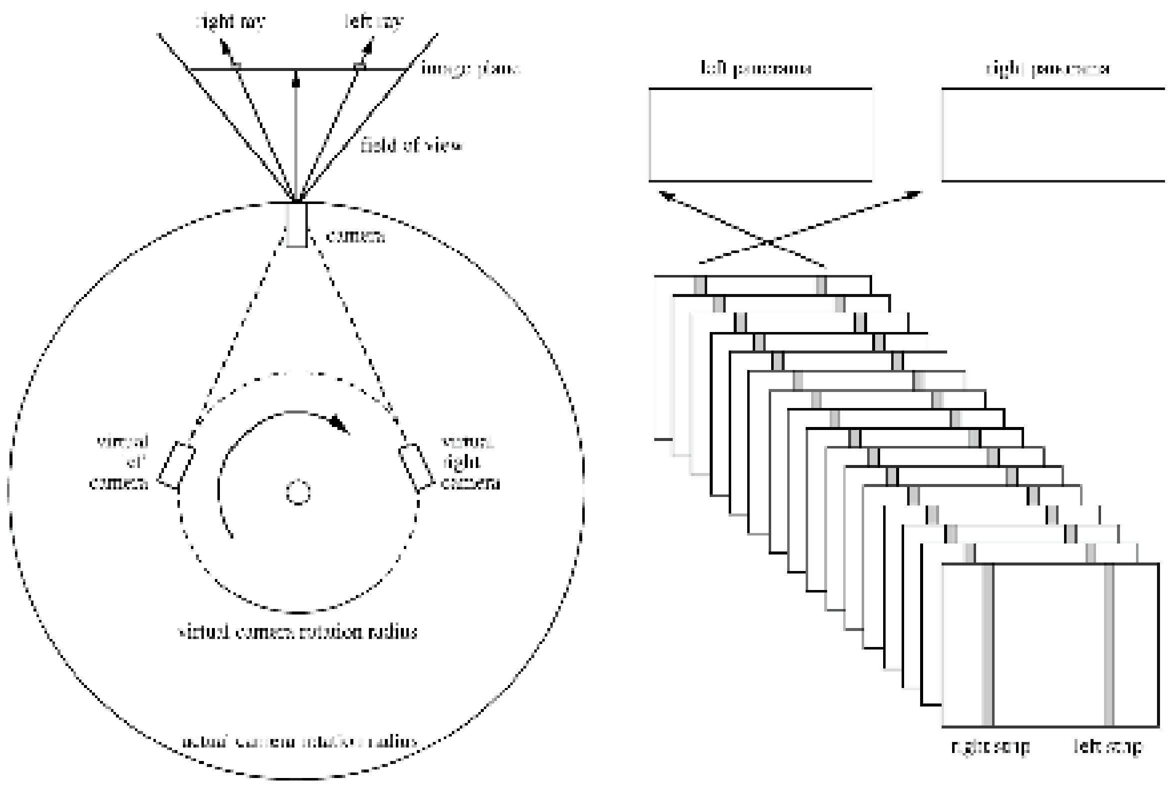

Figure 7. Single camera capture of stereoscopic panoramic images, more common for real camera capture than computer generated images for efficiency reasons.

4. Alternative approach

There is a technique that is sometimes used for capturing stereoscopic panoramic image pairs with a single real camera [10] and forming the pairs using vertical strips from each side of the resulting images, see figure 7. While there is little reason to employ this method for synthetic panoramic capture, it is mentioned here for completeness. The obvious disadvantage of this approach for synthetic stereoscopic panoramic production is inefficiency, that is, only a very small area of the rendered images is actually employed in the final panoramic images.

5. Viewing

There are a number of options for viewing stereoscopic panoramic images. On a flat screen such as a monitor or a projected image onto a flat screen, the equivalent to a QuickTime VR player has been developed [12]. The standard algorithm uses hardware acceleration such as OpenGL to map the two panoramic images as textures onto two cylinders. Each cylinder is centered on a coincident left and right eye virtual camera position, each camera renders a perspective projection, see figure 8. Each image is presented to the appropriate eye, the details are dependent on the particular stereoscopic projection environment.

More immersive environments exist where the entire 360 degrees is presented simultaneously surrounding the viewer or audience [13] as a large cylindrical display. Unlike most stereoscopic surround environments these cylindrical displays do not necessarily need head tracking and are even suited to viewing by multiple simultaneous participants. The image generation and projection details are outside the scope of this discussion but projection into these environments usually involves capturing multiple perspective projections as described above and then applying geometry correction such that the projection of those modified images appear correct on a section of the cylindrical display surface. Each of these sections may additionally need to be edge blended with the neighbouring sections. In such cylindrical environments it is interesting to note that the stereopsis errors that occur towards the periphery of ones vision need not be problematic because they are hidden by the limited field of view of the glasses being worn, but the peripheral vision of the observer is still engaged.

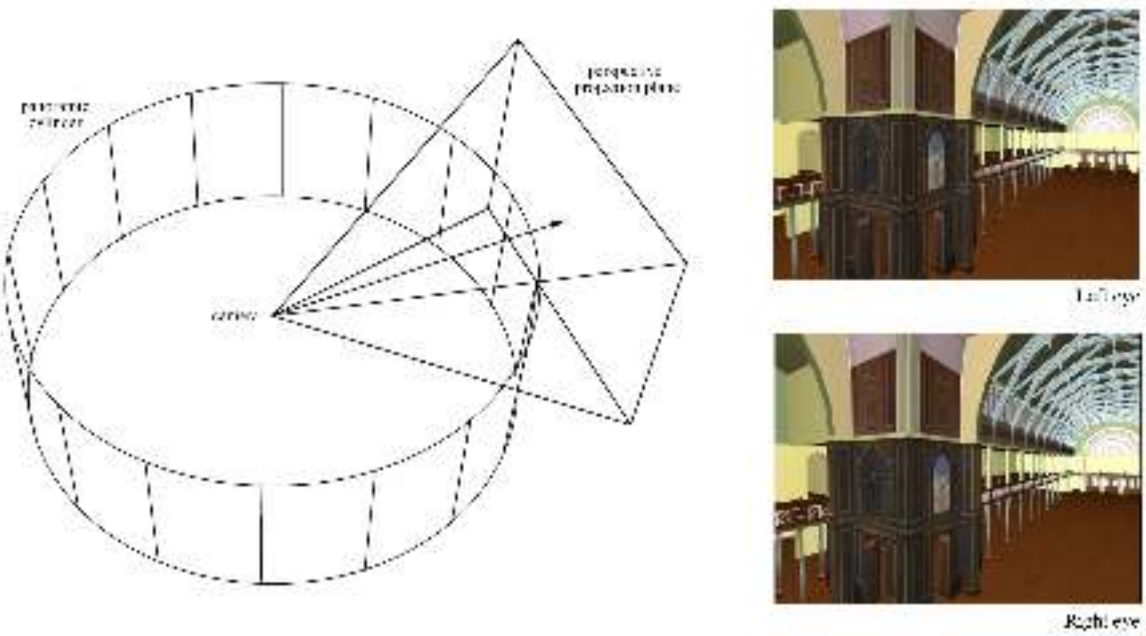

Figure 8. Perspective projection plane and textured cylinder for flat screen display (upper), examples of resulting stereo pairs (lower).

6. Summary

A method of creating synthetic stereoscopic panoramic images that can be implemented in most rendering packages has been presented. If single panoramic pairs can be created then stereoscopic panoramic movies are equally possible giving rise to the prospect of movies where the viewer can interact with, at least with regard to what they choose to look at. These images can be projected so as to engage the two features of the human visual system that assist is giving us a sense of immersion, the feeling of "being there". That is, imagery that contains parallax information as captured from two horizontally separated eye positions (stereopsis) and imagery that fills our peripheral vision. The details that define how the two panoramic images should be created in rendering packages are provided, in particular, how to precisely configure the virtual cameras and control the distance to zero parallax.

References

1. C. Johnson. *Panoramas of Duluth, Minnesota*. History of Photography, Vol. 16 (Summer 1992) pp141-146.
2. H. Ishiguro, M. Yamamoto, S. Tsuji. *OmniDirectional Stereo*. IEEE Transactions on Pattern Analysis Machine Intelligence 14(2), pp 257-262, 1992.
3. J. Y. Zheng, S. Tsuji. *Panoramic Representation for route recognition by a mobile robot*. International Journal Computer Vision, Vol. 9, no. 1, pp. 55-76, 1992.
4. S. Chen. *Quicktime VR – an image based approach to virtual environment navigation*. In *SIGGRAPH '95*, Pages 29–38, Los Angeles, California, August 1995. ACM.
5. H. H. Sham, K. T. Ng, S. C. Chan. *Virtual Reality Using the Concentric Mosaic: Construction, Rendering and Compression*. Microsoft Research, China, 2000.
6. Roundshot camera, Super 200 VR. Seitz Phototechnik AG, Hauptstr 14, 8512 Lustdorf Switzerland. [http://www.roundshot.ch/]
7. H. C. Huang and Y.P. Hung. *Panoramic stereo imaging system with automatic disparity warping and seaming*. Graphical Models and Image Processing, 60 (3) pp 196–208, May 1998.
8. S. Peleg, Y. Pritch, and M. Ben-Ezra. *Cameras for stereo panoramic imaging*. IEEE Conference on Computer Vision and Pattern Recognition, pp 208--214, 2000.
9. S. Peleg and M. Ben-Ezra. *Stereo panorama with a single camera*. In IEEE Conference on Computer Vision and Pattern Recognition, pp 395 Ð401, Ft. Collins, Colorado, June 1999.
10. S. Peleg and M. Ben-Ezra. *Stereo panorama with a single camera*. In IEEE Conference on Computer Vision and Pattern Recognition, pp 395–401, Ft. Collins, Colorado, June 1999.
11. Digital Model Of The Royal Exhibition Building (Melbourne, Australia), S. Bekhit, A. Kuek, H. Lie. [http://local.wasp.uwa.edu.au/~pbourke/modelling/reb/]
12. Stereo-capable panoramic viewer. P. D. Bourke. [http://local.wasp.uwa.edu.au/~pbourke/stereographics/stereopanoramic/]
13. AVIE stereoscopic 360 degree panoramic environment, ICinema, UNSW. [http://www.icinema.unsw.edu.au/projects/infra_avie.html]