Panoramic 3D Video

A. R. L. Travis
Engineering Department
University of Cambridge
Cambridge CB2 1PZ, UK

Abstract:- The problem with crude autostereoscopic displays is that the transition between adjacent views is noticeable to the eye as the observer moves around the autostereoscopic image. This problem can be eliminated if the autostereoscopic image has limited depth by making the angle (in radians) between adjacent views equal to the reciprocal of the depth of the image (in pixel widths). But what happens if the depth of the image is infinite? It is shown here that to televise seamless 3D images of a scene of infinite depth, the angle in radians between adjacent views must equal the lateral resolution of the display times the ratio between the display field of view and the camera field of view.

1. Introduction

Although a three dimensional image can be displayed by a cubic array of isotropic light sources, it is difficult to devise a camera which would capture such an image. Furthermore such a system could not televise opaque or mirror images. An autostereoscopic three dimensional image however can be formed by pointing several conventional video cameras at a solid object. The view from each camera is sent to the autostereoscopic three dimensional display which is a screen that displays each view so that it is visible to a direction parallel to the axis of the camera which formed the view.

Because it reconstitutes the three dimensional image in a similar way to a hologram, an autostereoscopic display with enough views can reproduce any real three dimensional image. But if an observer looks at an image with too few views they will notice a transition in picture content as their head moves between adjacent views. This is the autostereo equivalent of "staircasing" in two dimensional displays. How fine must be the angle between adjacent views to eliminate this effect?

The minimum perceptible angle between adjacent views has been found to be 0.0003 radians[1] which would suggest that a display with 120° field of view would need approximately 7 000 views. However the minimum perceptible pixel diameter of a two dimensional image (viewed from a distance of one metre) is a few tens of microns, whereas displays with much larger pixel diameters of several hundred microns prove acceptable. It is probable that the relaxation of resolution which is found acceptable with two dimensional images will also apply to three dimensional images.

The approach taken here is to define an autostereoscopic three dimensional display to be acceptable if it can display without transitions a cubic array of isotropic light sources whose resolution in each dimension equals that of an acceptable two dimensional image. It remains to translate between autostereoscopic pixellation and cubic pixellation.

2. Time-sequential autostereoscopic 3D display

The 3D display which was the background to this work comprises a cathode ray tube, a pair of lenses and a liquid crystal shutter[2].

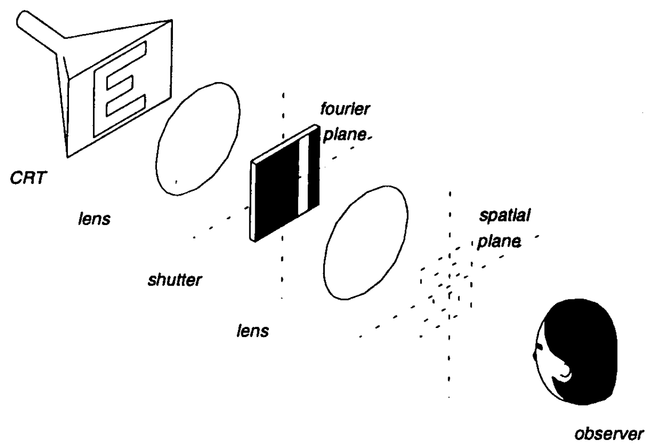

Figure 1: A CRT image is made visible to a single direction by filtering the image with a slit in the fourier plane. A three dimensional image is built up by displaying separate views of the object one by one on the CRT, and making them visible to different directions by moving the slit.

The two lenses are placed coaxially at a distance from one another such that they share a focal plane. It is well understood that if a screen is placed in the rear focal plane of the rear lens of such a pair, the optical fourier transform of the screen will be formed between the lenses, and a spatially identical image of the screen (albeit up-side-down) will be formed in the front focal plane of the front lens[3].

Suppose that an optical stop is now placed in the fourier plane. The effect of this will be to restrict the direction of rays leaving the spatial plane to a narrow angle of directions, the size of the angle being proportional to the size of the optical stop. If an observer were to look at the image of the screen, the direction from which that image was visible could be altered by moving the optical stop around in the fourier plane.

Instead of an optical stop a liquid crystal shutter is used, and a cathode ray tube is used in place of the screen. The cathode ray tube can cause a bundle of rays of a given intensity to leave any point in the spatial plane, and all rays but those travelling in a desired direction can be blocked by closing all but one of the shutters in the fourier plane. The result is a device which can emit a ray of light with a controllable intensity from any given position on the screen to any given direction. The three dimensional image is written by scanning through each coordinate of autostereoscopic space in turn.

The advantage of this system over alternatives using lenticular screens is that the pixellation is flexible. This makes it particularly adaptable for the development of new uses for three dimensional television, in particular for experimenting with different image formats. The two most formats most popular on the system are 8 views at 640x480 pixels per view, and 16 views at 320x240 pixels per view. With the former the transitions between adjacent views are noticeable; with the latter they are barely perceptible once the image has been anti-aliased. Can the transitions be eliminated altogether?

3. Parallel Projection of a Cubic array

If the slit in the fourier plane of the 3D display is made very small, it corresponds to an impulse in the fourier plane and hence parallel rays of light in the spatial plane. The view which should simultaneously be displayed on the CRT corresponds to a parallel projection from the solid object.

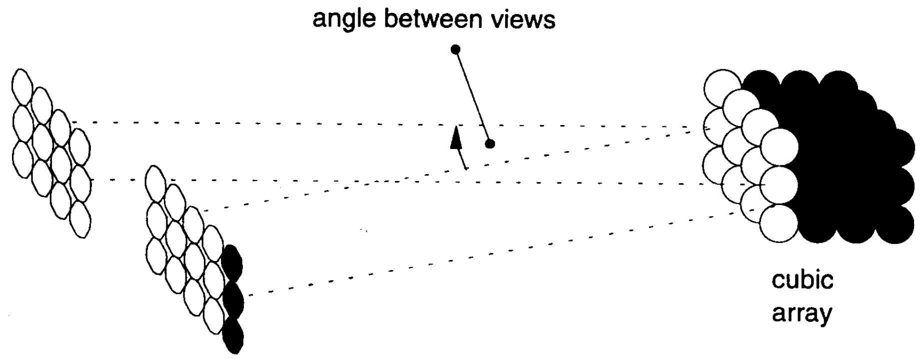

Figure 2: Two views of a cubic array are formed by parallel projection. The minimum angle between the views required for there to be a distinct difference in view content is that required to make one column of rear pixels fully visible

Suppose that the solid object is a cubic array. Through what angle can the direction of projection be rotated before the projected image changes? It has been shown[4] that if one rotates about the centre of the front of the cubic array, the limit on rotation without change is set by the rear pixels of the cubic array. There will have been an unambiguous change in the projected image once the direction of projection has been changed sufficient to translate the image of the rear pixels by one full pixel distance. The angle in radians required to do this equals the width between two pixels divided by the depth of the cubic array. Because the array is comprised of cubes this angle equals the reciprocal of the number of depth planes of the cubic array.

The problem with this result is that if the depth of the three dimensional scene tends to infinity, for example if the three dimensional image is a scene through a window, then the angle subtended by each view needs to be infinitesimal.

4. Perspective Projection

In reality the view of a three dimensional scene formed by each camera is not a parallel projection but a perspective projection, the degree of perspective depending on the distance of the camera from the front of the scene. The camera can resolve fine resolution close to, but coarse resolution far away. The smallest object which can be resolved at any distance from the camera is equal to the width of view visible at that distance divided by the number of pixels per line in the camera. In recognition of this it is suggested here that rather than a uniform cubic array, a more appropriate pixellation shceme with which to compare autostereoscopic pixellation is an array of cubic pixels in which the pixel dimension is proportional to the distance of the pixel from the camera: i.e. a distorted cubic array

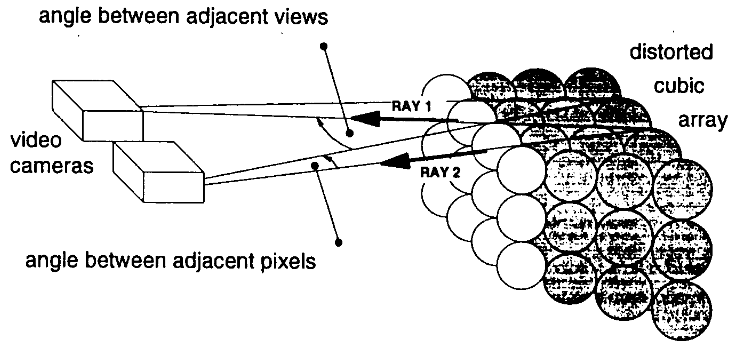

Figure 3: In the left hand video camera only the front pixels of the distorted cubic array are visible, while in the right hand camera one column of the rear pixels has become visible. As the depth of the distorted cubic array tends to infinity, RAY 1 and RAY 2 tend towards parallel, so the angle between adjacent views tends to the angle between adjacent pixels

Through what angle can the direction of projection be rotated before the projected image changes? If one rotates about the centre of the front of the cubic array, the limit on rotation without change is set by the rear pixels of the cubic array. There will have been an unambiguous change in the projected image once the direction of projection has been changed sufficient to translate the image of the rear pixels by one full pixel distance.

Figure 3 shows a pair of cameras separated by the minimum required for an unambiguous change. Simple geometry shows that as the depth of the array tends to infinity, the angle between the two cameras and the front of the array equals the angle between two adjacent pixels at the rear of the array and either camera.

If we register the pixellation in the image plane of the cameras so that each pixel in the image plane of a camera maps a pixel at the rear of the distorted cubic array, then the angle between the two cameras and the array equals the angle between two adjacent pixels in any camera image plane and the camera aperture.

$$\text{angle between views} = \text{angle between camera pixels} \tag{1}$$

So how many views are needed for an autostereoscopic 3D image to be free of transitions? The number required is equal to the field of view of the display divided by the angle between views.

$$\text{number of views} = \frac{\text{display field of view}}{\text{angle between views}} \qquad (2)$$

It has been shown that the angle between views must equal the angle between camera pixels. This in turn must equal the field of view of the camera divided by the number of pixels per line in the camera image plane.

$$\text{angle between views} = \text{angle between camera pixels} \qquad (3)$$

$$= \frac{\text{camera field of view}}{\text{number of pixels per line}} \qquad (4)$$

The number of pixels per line is the same for the camera and the display, so the number of views required is the number of pixels per line of the display times the ratio of display field of view to camera field of view.

$$\text{number of views} = \frac{\text{display field of view}}{\text{camera field of view}} \times \text{number of pixels per line} \qquad (5)$$

5. Modifying the Display

In its original form the three dimensional display described in section 2 projects each view so that rays from all pixels on the screen are emitted in parallel. If the display is to produce a three dimensional image of an object without distortion then views of the object must be captured by parallel projection, i.e. by telescopic cameras far from the object.

However in the pixellation scheme defined in section 4 the cameras form perspective projections of the object. If the display is to produce a three dimensional image of the object without distortion from views formed in this way, then for each view rays from all pixels on the screen must be emitted so as to converge. Expressed specifically, the angle at which rays from the left and right hand edges of the display screen converge must equal the camera field of view.

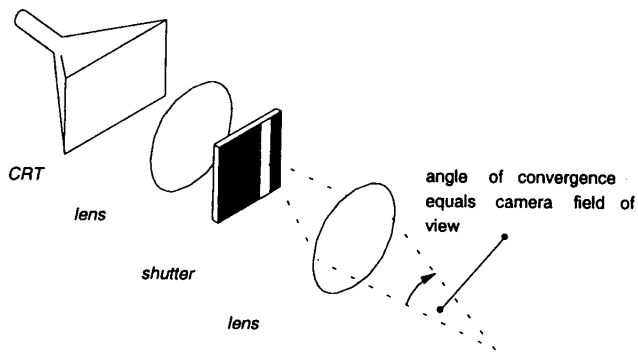

Figure 4: The three dimensional display can be reconfigured for convergent projection by moving the shutter to the rear. Rays will converge to the image of the slit formed by the second lens

So for the display of a single view on the CRT, rays should be selected by the liquid crystal shutter so that they approximately converge to a single point. Arranging this selection of convergent rays is simple. One merely moves the liquid crystal shutter back from the fourier plane, to a position where an image of the shutter is formed in the plane of points at which rays of each view are to converge.

6. Resolution

The arguments of section 4 can be used to calculate the number of views required to display a seamless autostereoscopic image with views of 160x120 pixels and a field of view of 16^{o}. This would be the three dimensional equivalent of a two dimensional display with quarter VGA resolution. If a camera with a field of view of 48^{o} is used then the number of views required follows as:

$$\text{number of views} = \frac{\text{display field of view}}{\text{camera field of view}} \times \text{number of pixels per line} \quad (6)$$

$$= \frac{16}{48} \times 160 \quad (7)$$

$$= 53.3 \quad (8)$$

It has been found that what limits the resolution of the autostereoscopic display is the line rate. Assuming an interlaced field repetition rate of 50 Hz, the line rate required to display such an image equals the product of the number of views, the number of lines per view and 25 Hz:

$$\text{line rate} = \text{number of views} \times \text{number of lines} \times 25 \quad (9)$$

$$= 53 \text{ x } 120 \text{ x } 25 \quad (10)$$

$$= 159 \text{ kHz} \quad (11)$$

Some allowance must be made for the time required for vertical flyback, so a line rate of 170 kHz might be necessary. This is greater than the 150 kHz line rate for which the CRT in the autostereoscopic display is specified, but it is hoped that the CRT line rate can be increased sufficiently without too much difficulty.

7. Conclusions

It has been proposed that an autostereoscopic 3D display will have satisfactory resolution if it can display a 3D array of cubic pixels without any projected edge moving by more than one pixel dimension between adjacent views. If the display effects parallel projection, then the angle between adjacent views should equal the reciprocal of the number of depth planes of the cubic array. If an infinitely deep 3D image is to be displayed then the display should effect convergent projection and the angle of convergence should equal the field of view of the cameras. For an image free of transitions the angle between adjacent views should equal the angle between any pair of adjacent camera pixels and the camera aperture. Looking into such a display with sufficient resolution will be virtually the same as looking through a window. The Cambridge display should be able to produce such a three dimensional image with resolution equivalent to a two dimensional display with quarter VGA resolution.

---

8. References

1 Pastoor, S., and Schenke, K., "Subjective assessments of the resolution of viewing directions in a multi-viewpoint 3DTV system", Proc SID, Vol 30/3 pp217 to 223, 1989

2 "A 16-view time-division-multiplexed autostereoscopic display", Moore, J. R., Travis, A. R. L., Lang, S. R., and Castle, O. M., SID '93 Digest pp700-703

3 "Optics", Hecht, E., Addison Wesley, Second Edition p566 (1987)

4 "3D Autostereo Display - are 16 views enough?", Travis, A. R. L., 4th European Workshop on Three-Dimensional Television, Rome, pp137-140, 1993