Computer Science Department of The University of Auckland
CITR at Tamaki Campus (http://www.citr.auckland.ac.nz)

CITR-TR-19 February 1998

Color Anaglyphs for Panorama Visualizations

Shou-Kang Wei, Yu-Fei Huang, and Reinhard Klette*

Abstract

Traditional panorama visualization techniques give only the single eye's depth cues to the audience for the 3D world navigation. We introduce a new panorama visualization approach so that the depth cues for both eyes are provided in the panoramic image. The full color anaglyphs stereoscopic technique is adopted to produce and visualize the stereo panorama, which exploits human visual capability so the binocular depth and synthesized colors can be perceived and understood. Because both panoramic and anaglyphic image acquisition setups are incompatible intrinsically, we measure the errors of anaglyphic panorama image and deliver an error controlling formula that allows a better camera position to be determined at setup time. A real-time interactive player is developed for navigation in the stereo panoramic virtual world. In addition to the standard functions provided in the player, three new features are equipped to the player (i.e. a new driving method, manipulable objects embedded in the stereo panorama and a directional indicator).

* The University of Auckland, Tamaki Campus, Computing and Information Technology Research, Computer Vision Unit, Auckland, New Zealand

Color Anaglyphs for Panorama Visualizations

Shou-Kang Wei, Yu-Fei Huang, and Reinhard Klette

Computing and Information Technology Research
Computer Science Depatrment, The University of Auckland
Tamaki Campus, Auckland, New Zealand
Email: {swei002, yhua002, rklette} @cs.auckland.ac.nz

Abstract

Traditional panorama visualization techniques give only the single eye's depth cues to the audience for the 3D world navigation. We introduce a new panorama visualization approach so that the depth cues for both eyes are provided in the panoramic image. The full color anaglyphs stereoscopic technique is adopted to produce and visualize the stereo panorama, which exploits human visual capability so the binocular depth and synthesized colors can be perceived and understood. Because both panoramic and anaglyphic image acquisition setups are incompatible intrinsically, we measure the errors of anaglyphic panorama image and deliver an error controlling formula that allows a better camera position to be determined at setup time. A real-time interactive player is developed for navigation in the stereo panoramic virtual world. In addition to the standard functions provided in the player, three new features are equipped to the player (i.e. a new driving method, manipulable objects embedded in the stereo panorama and a directional indicator).

Keyword: Color anaglyphs, anaglyphic panorama, environment maps, real-time player, manipulable objects, directional indicator.

1 Introduction

All animals with two eyes have binocular vision. The animals with laterally placed eyes, such as humans, are able to integrate the information from the large area of the binocular overlap. This overlap is referred to as angle ϕ in Fig. 1.1, and humans use it to analysis depth. As in three-dimensional structures of the world, the depth cues can be realized by one or two eyes. The former can senses depth cues including perspective, image overlap, shading and others as discussed in [1], whereas the latter perceives binocular disparity exclusively. The term *stereopsis* is used for the impression of depth arising from binocular disparity.

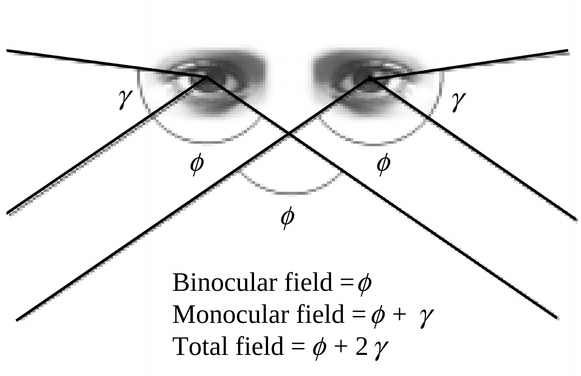


Figure 1.1: The human visual fields.

In order to convey such impression to the audience with available media, many stereoscopic vision techniques have been developed for exploring such characteristics of the human vision system. The key idea among those techniques is to simulate human binocular disparity by allowing only the left eye to see the left image, and the right eye to see the right image. Here left and right images refer to the same scene with slightly different camera positions.

Many techniques, such as the *mirror stereoscope*, *prism stereoscope* and *polaroid stereoscope* require not only extra hardware such as mirrors, prisms, partition cards, polarizing filters and projectors, but also accurate setups to accomplish the task. Details are described in [1]. Therefore they are neither applicable to visualize stereo pairs on the computer monitor nor suitable for panorama scene exploring.

There are at least two techniques applicable with modern computer devices, one is the *shutter stereoscope*, and the other is the *anaglyph technique*. In the shutter system, the left and right images are presented in rapid alternation on a computer monitor. People view the display through electro-optic shutters that alternately occlude the two eyes in phase with the alternation of the display. Flicker may appear due to the slow frame speed used on most computers, so more expensive systems may be required.

In contrast to it, the anaglyph technique suggests an affordable approach to stereo visualization. The only extra hardware needed for such technique is a pair of 3D eyeglasses. They are made of color filters, conventionally red for the left eye, and green or blue for the right eye. The left and right images are superimposed into a single image, called *anaglyphic image*, by replacing the red component of the right image with the one in the left. The trick played here is allowing only red component in the anaglyphic image through red filter to the left eye while the green/blue component through green/blue filter to the right eye. Our mind merges these two illusions into one and gives us understanding of depth relationships.

In the remainder of this report, we describe the full color anaglyphs stereoscopic technique in the Section 2. Section 3 discusses the problem occurred for combining both panoramic and anaglyphic techniques and follows our approaches to the problem. In Section 4, we explain our design and development for the stereo panorama player. The conclusions and future work are addressed in the Section 5.

2 Full Color Anaglyphs

Anaglyph technique has been commonly described as *"has the disadvantage that no color is used"* [2], and *"this method cannot be used for color images"* [3] and some similar statements. Those comments were based on the original design of anaglyphic technique in which only two primary colors were used (i.e. red-green eyeglasses and red-green superimposed images).

Here a new approach, or its variant, is proposed such that the full color anaglyphic image can be visualized. Idea is to offer more color cues to our eyes, and let our mind resolves the color information. Instead of only two primary colors, three can be configured as red-bluegreen eyeglasses with red-bluegreen superimposed image. Hence full color information is transmitted to our eyes with red to the left eye and bluegreen to the right, then the work of how to combine them and produce a full color stereo imagery is in fact mentally synthesized in our mind.

As Helmholtz wrote [4],

We therefore learn that two distinct sensation are transmitted from the eyes, and reach consciousness at the same time and without coalescing; that accordingly the combination of these two sensations into a single perceptual picture of the external world is not produced by any anatomical mechanism of sensation, but by a mental art.

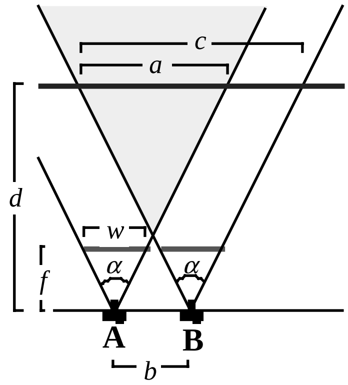


Figure 2.1: Top view of binocular camera model. The variables are explained in the text.

The result of this experiment is convinced as we easily perceive most of the original colors from this new configuration. However, there are some loss of resolution and some color rivalry in anaglyphic picture. For its applications, the color accuracy should not be important.

2.1 Anaglyphic Image Acquisition

Instead of using a more expensive special-purpose device for acquiring stereo image pairs, here we discuss the way using single normal camera with a 35 mm film.

For an anaglyphic image pair, two camera positions should be setup obeying the following rules:

- Their optical axes are parallel.
- Their image planes are coplanar.
- They are apart from each other in a proper distance.
- For each conjunction pair of epipolar lines it holds that both lines have identical row coordinates.

The idea of giving a proper displacement of the cameras is to obtain the desired depth cues from the target object(s), such as a tree, in the scene. If it is too small, then there may be no depth information given in the resulting image; or if it is too large, it produces a smaller resulting image due to the less overlapping area. Let us look at a simple geometric relation as shown in Fig. 2.1.

The distance between the focal points of two cameras, say **A** and **B**, is called the base distance denoted by *b*. The line joining **A** and **B** is a base line. Suppose there is a plane perpendicular to both camera optical axes, which is *d* away from the base line, see Fig. 2.1. We can then describe the relation between *b* and *d* in terms of the percentage of overlapping area, denoted by *p*, in both images. The value *p* is defined as follows:

$$p = \frac{a}{c}. \quad (2.1)$$

Then, we have

$$p = \frac{2d\tan(\frac{\alpha}{2}) - b}{2d\tan(\frac{\alpha}{2})}, \quad (2.2)$$

where α is the camera's horizontal field of view.
So the base distance *b* is equal to

$$b = 2d(1-p)\tan(\frac{\alpha}{2}). \quad (2.3)$$

The formula derived is an approximation to the base distance because normally *d* is not precisely determined, for instance, the scene objects are not coplanar. In general, a base distance of 6.5 cm is suggested because human beings perceive binocular vision in such base distance, which is the average distance of the human eyes. Hence it should produce a fairly good depth approximation to what we see. Let us answer the question, what actually *d* values result for the base distance *b* = 6.5 cm if *p* is between 70% to 99% for a given focal length.

We first convert the field of view α to the effective camera focal length *f* based on *w* = 35 mm film size and then compute *d* in terms of other parameters,

$$d = \frac{bf}{w \times (1-p)}. \quad (2.4)$$

The computed data are listed in Tab. 2.1.

Usually we use *p* in the range between 75% to 98% depending on the preferred proportion of the overlap in the image. The 70% and 99% of *p* in Tab. 2.1 give the lower and upper bounds of range *d* required.

	f = 20 mm	f = 50 mm	f = 80 mm	f = 120 mm
p = 0.70	12.0	30.1	48.1	72.2
p = 0.75	14.4	36.1	57.8	86.7
p = 0.80	18.1	45.1	72.2	108.3
p = 0.85	24.1	60.2	96.3	144.4
p = 0.90	36.1	90.3	144.4	216.7
p = 0.95	72.2	180.6	288.9	433.3
p = 0.96	90.3	225.7	361.1	541.7
p = 0.97	120.4	300.9	481.5	722.2
p = 0.98	180.6	451.4	722.2	1083.3
p = 0.99	361.1	902.8	1444.4	2166.7

Table 2.1: The distances *d* (in cm) between the camera and the object(s) are calculated concerning the image's overlapping percentages between 70% to 99% for four different focal lengths, 20,50,80 and 120 mm.

For example, if we use 20 mm lens and the target object is definitely more than 360 cm away from the camera, then there is no point to take 2 pictures that almost identical (i.e. > 99% overlap). In this case, a wider lens should be used. Similarly if 120 mm lens is used and the proportion of subject to the whole image is expected at least greater than 70%, then the distance between the object and the camera should be at least equal to or greater than 72 cm.

Although the *d* values change subject to the different camera's parameters, the data of *d* in the Tab. 2.1 show a reasonable range of distances. We can conclude that the base distance 6.5 cm is acceptable for a wide range of applications.

2.2 Anaglyphic Image Processing

After the acquisition and digitalization of a photo pair, the next is to superimpose them onto a single anaglyph image. One can use image processing software to do the following steps:

1. Create a new image of the size equal to the left and right images.
2. Copy red channel from the left image to the new image.
3. Copy green and blue channels from the right image to the new image.
4. Select the red channel of the new image and then offset it horizontally to a proper distance.
5. Crop the non-overlapping area.

The reason for the horizontal offset of the red component is to adjust the disparity in the anaglyphic image not too big for our eyes to fuse. If the disparity displayed on the anaglyph image is very large, i.e. over the limitation of eyes, then we see two images - referred to as *diplopia* [1]. Since each individual has its own fusing capability, there is no fixed disparity that works for all cases. Instead, we adjust it accordingly so our eyes feel comfortable with it.

In fact, it is possible that the 3D effect can be realized immediately while just put on the 3D eyeglasses. It is because the disparity just matches up eyes' fusing capability. However sometimes our eyes need more time to adapt to the anaglyphic image. We call it fusing delay. It usually takes longer when the disparity is relatively large with respect to fusing capability of our eyes. So we can adjust the disparity slightly that there is no fusing delay. But the quantity of the adjustment may vary for each individual. The farthest object in the scene, where the greatest disparity occurs, is the major dependence to make the adjustment. If one has no fusing delay on the farthest sign, then everything nearer should have no problem.

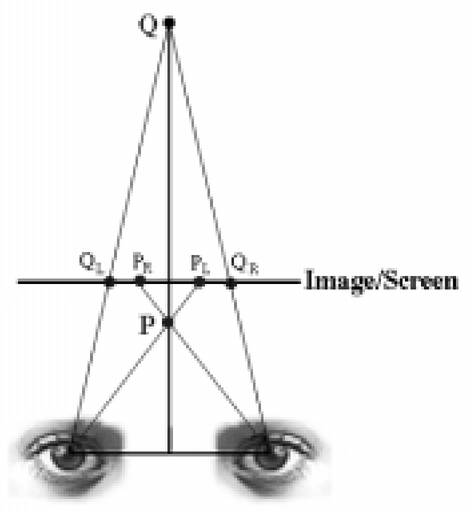


Figure 2.2: The geometry of a parallax stereogram.

Many anaglyphic images have made the left image on the right and the right image on the left, i.e. crossed disparities. For example, see Fig. 2.2, the point **P** is the sight intersection point corresponding to the point $\mathbf{P}_L$ of the left image on the right and the point $\mathbf{P}_R$ of the right image on the left. So we should see the object of a scene standing out of the image/screen. It works nicely if the main feature(s) of scene present completely or consistently in the image area, such as an animated object active constantly within the image or just a still image. In our case, a panoramic anaglyph image, see the details in the Section 3, will be very poor to look in this way. Because as looking around, panning to the right says, the new scene from the right will stretch out toward the viewer while the scene on the left shrinks back from the viewer and disappears from the view port then. So we suggest placing the left image on the left and right image on the right, which makes sight intersection behind the image/screen, e.g. the point **Q** converged from the image points $\mathbf{Q}_R$ and $\mathbf{Q}_L$ in the Fig. 2.2. Therefore we have more consistent movement in the scene just as we see in the real world.

2.3 Color Filters for 3D Eyeglasses

A color filter is used for filtering out a range of color components from the emitting or reflected light input. For example, red filter only allows red component of input passing through the filter. There are many different filters made for serving various applications.

In our case, we use it to extract the desired color component from the anaglyphic image for each eye. Basically, to be able to separate the left/right image from the anaglyphic image for the left/right eye, one needs to take a good care of choosing proper color filters for the 3D eyeglasses. Following we explain what is meant by "good" filter and how to choose it.

2.3.1 Good Filters for Anaglyphic Image

The key to a quality stereo visualization for anaglyphic image, firstly one needs to have a pair of good filters. The good filters should totally cover up the whole range of colors presented in the given anaglyphic image yet mutually exclusive one and the other. For instance, if you have an anaglyphic image that composite only two primary colors, say red and green for the left and right image, then the pair of filters are evaluated as good if all red components can be seen only through the left filter and all the green components can be seen only through the right filter.

What happens if the color range of filter is smaller than the one distributed in the image? Furthermore, what happens if the color ranges of filters are overlapped each other? In the first situation, the covered color range will work fine as expected, but the uncovered part of colors can not get through the filter, so viewer simply sees black color in the corresponding place of the image. For the second situation, the color within the overlapping range will pass through both filters and we see two images, the phenomenon is referred to as *double-image*.

So far, we have shown the meaning and importance of good filters for the anaglyphs technique, now let us discuss how to choose them.

2.3.2 How to Choose the Filters

Usually the match either red-green or red-blue is chosen for the left and right eyes. Other matches are also possible but less common. Green on the right is more preferable than blue, due to our eyes have higher response to the green components [5].

One can start searching from three primary colors first and then their variants. The typical practice is bring along target anaglyphic picture and testing it whether only desired image can be seen (i.e. no double image) through selected filters. However, if the target medium is not portable, such as anaglyphic image on the computer monitor, then other approaches are suggested.

(A) With Technical Information Description Provided

Many color filter vendors provide a technical information for each filter. The specific information may include a describable color name, a wavelength-transmission distribution graph, effect description, transmission, absorption, chromatics and so forth.

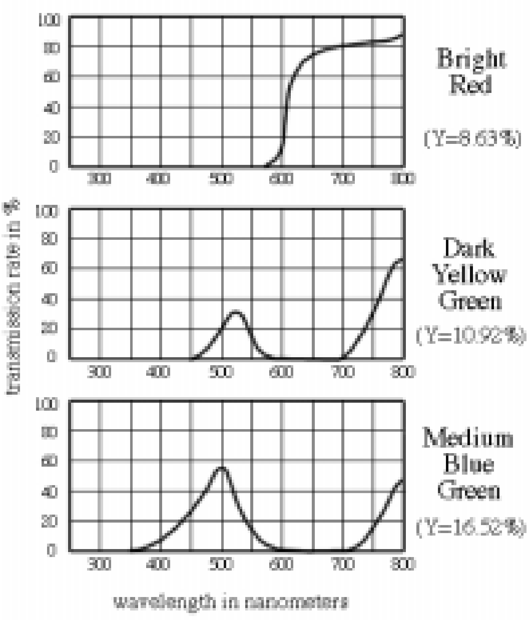


Figure 2.3: Wavelength-transmission distribution graph of lighting filters.

See three swatches of lighting filters in Fig. 2.3, also [6] for more details.

In wavelength-transmission distribution graph, the areas under the curve(s) indicate in what range of wavelength the percentage of transmission is allowed to pass through the filter. So if the approximation range of wavelength for the desired colors is known, e.g. 550-700 nm for the red, then the curve of filters should approximately cover such range. For a set of possible pairs, overlap two graphs, such as red and green. The pair with the least overlapping area, or even better totally separated, is the one we are looking for. The Fig. 2.3 is the result of such selection for red, green and bluegreen.

Note, the wavelength exceeds 700 nm is infrared that is not usually visible by human eyes, so one should not take it into account as overlapped area.

(B) Without Technical Information Provided

In case the technical information is not available, we also have an approximation to deal with it. After picking out the desired color filters according to our own color sense, a simple verification can be done by overlapping two filters, such as red and green, and check if we can see through it or not. If we can see through the overlapped filters, it means one or both of the filters are not appropriate (i.e. their color ranges are overlapped). Otherwise, we should see the world in black or the reflection of our eyes from the filters.

This just follows a simple fact that the red components passing through the red filter will not be possible to pass through green or blue filters and vice verse.

Although two approaches should be sufficient for most of the cases, it does not guarantee that the selected pair will work fine with all anaglyphic image. Because different materials, media or devices interpret the colors differently and our eyes are good at detecting such tiny difference.

2.4 Reducing the Double Image

Here, to simplify the problem, let us assume the anaglyphic image only displayed on the fixed medium which is a computer monitor. In the previous section, we have mentioned that the double-image is caused by the overlapped color ranges of filter pair. We have also suggested and shown the wavelength-transmission distribution graph of good filters in Fig. 2.3. However, one may still notice that even we have optimal eyeglasses, our fuzzy eyes can still tell slightly double-image appearing. What are those colors that pass through both filters? Let us find out those problematic colors.

Assume we have filters with no overlapping area under curves of the graph, then the color with full saturation is not possible to pass through both filters, regardless of the factor of mirror production errors of filters. So now the only possible colors causing the problem should be non-saturated colors, such as pink color. It can be realized that less saturated color is mixed its dominant color with other colors' component and those components are very likely passing through the filter that it should not be.

Increasing the saturation of problematic colors in certain degree may solve the problem. To simplify the task, one can adjust the saturation of the image as a whole and that gives the similar effect. The reason for that is as those colors become more distinguishable than they were after the adjustment, the higher probability they can only pass through one filter but the other. There is no exact degree of the adjustment should be due to the colors of filters and anaglyphic image are relative dependent. But make them all full-saturated, of course, is over-done! The aim of this operation is to make colors more distinguishable without changing its original contents too much. So normally, in our experience, increasing about 15-20% saturation will be adequate.

In general, one should do it with the 3D eyeglasses, use image processing software with saturation adjustable function to tune-up the saturation of anaglyphic image until the double-image disappears or is less visible.

3 Panoramic Anaglyphs

Next we consider combining the full panoramic scene with anaglyph technique to make a stereo visualization of the walk-through [7]. To visualize the result, all it needs is just one pair of low cost and conveniently uses red-bluegreen spectacles, and a software to render it on normal personal computer. Everything we see in the panoramic environment is no longer a flat picture but in a stereo form. All the depth impressions preserve almost the correct proportion to the real 3D synthesis. Since anaglyph method is based on the simulation of human vision, we feel like actually walking into a 3D world rather than just looking at 2D images.

We show up the basic problem occurred in acquiring anaglyph panoramic image and measuring the errors for such problem. A formula is also derived for errors control so a better camera position can be determined at setup time. Two possible camera setups and two post-processing of panoramic anaglyph images are also discussed.

3.1 The Problem of Acquiring an Anaglyph Panoramic Image

A very important rule for making a full view panoramic image is that the nodal point (or optical center) of the camera must pass through the rotating center while taking series of photos in 360 degree. It is to prevent the error in images stitching. To produce stereo effect we need at least two photos taking from slightly different positions but the same viewing direction for each picture scene. Hence it is necessary to produce a pair of photos for every rotation angle. However, according to the rule we have mentioned, this task sounds impossible.

Only if the camera's optical center (nodal point) is on the rotating axis, every 3D point **P** which lies on both successive images may share the same y value in image's coordinate system. Otherwise the point **P** might project onto different row of the images due to the slightly difference of photo captures distance, shown in Fig. 3.1 where $\mathbf{P}_1 > \mathbf{P}_2$. The problem that causes error matching in stitching process is also illustrated in Fig. 3.1 where $|\mathbf{Q}_1 - \mathbf{P}_1| > |\mathbf{Q}_2 - \mathbf{P}_2|$.

As we already mentioned, the binocular image pairs require certain camera displacement (base distance) to have depth information. It is then impossible to have camera's nodal point passing through the rotating axis for all acquisitions. However, the error caused by this problem seems harmless to the resulting panoramic image if the distance between the closest object and the camera is far enough.

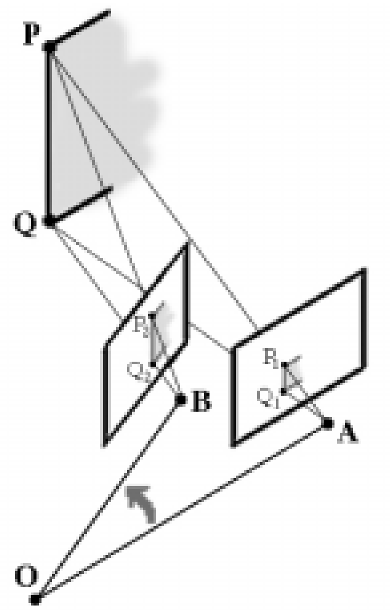


Figure 3.1: The corresponding points, **P** and **Q**, lie on the different rows of two images.

In practice, the small amounts of unmatched pixels can be blended up during the stitching process so the result may be acceptable. For avoiding an unacceptable big error occurred we measure the maximum possible error and come up with a formula allowing error controlling which may be required in different applications.

3.2 The Error Measurement and Control

What is the minimum distance required between the camera and the closest object in the scene for acquiring panoramic images to produce a reasonably good result? In other words, what the maximum error tolerance can be accepted? In this section, we measure these mismatch errors in two successive rotating images and followed by the result of experiment for the digital camera.

Let us define the error e to be the vertical pixel difference of a corresponding point at two images. To be more precise in the calculation we use cm as our measurement unit. The conversion between unit cm and pixel is trivial. Moreover, the assumption of undistorted image for lens distortion is made through the error calculation.

Now two successive photographs are taken without fixing camera's nodal point at the rotation axis, but equal distance away from the rotation center. More importantly, camera's up-vector must be parallel to the rotation axis and it's optical axis must be perpendicular to the line joining camera and the rotating center.

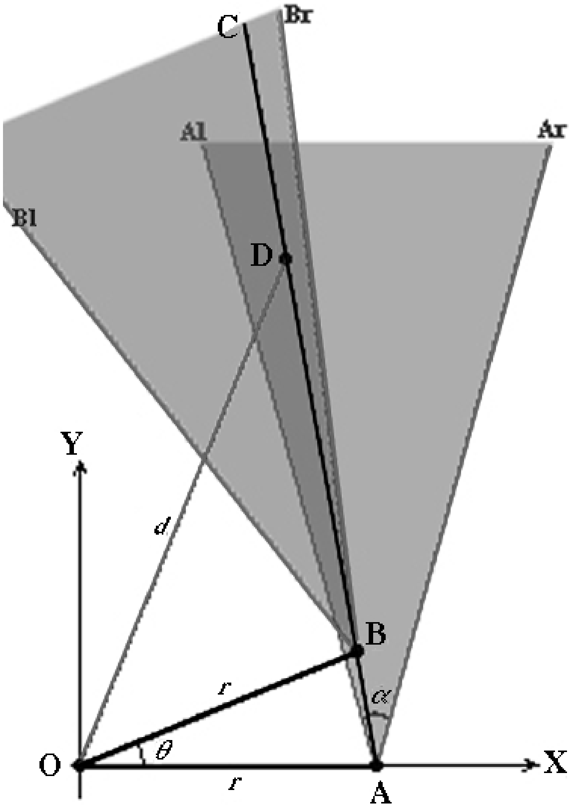


Figure 3.2: Two successive camera positions in XY-plane. The variables are explained in the text.

The settings are described in 2D XY-plane shown in Fig. 3.2 Let rotation center **O** be the origin of our coordinate system and two successive camera's optical center (nodal point) positions are denoted by **A** and **B**, both *r* cm away from **O**. It is clear that point **A** = (*r* , 0), and point **B** = (*r* cosθ , *r* sinθ) with respect to **O**, where *q* is the rotation angle. Symbol α there denotes the camera's horizontal or vertical field of view depending on how it is placed. Since $\alpha \geq \theta$ is required to ensure some overlapping between images, point **B** must always lie inside the area bounded by line **AAl** and line **AAr**.

The darker gray area shown in Fig. 3.2 is the overlapping section of two central projections from camera **A** and **B** respectively. Within the section, the greatest error must occur on the plane that vertically cut through XY-plane along line **AC** passing through point **B**. Let **D** be an arbitrary point on line **AC**. Point **D** stands for a point in 3D environment which can be seen by both camera **A** and **B** and is *d* cm away from the rotation center. We are interested in the error that happen along the vertical-up line **DF** -- a line perpendicular to XY-plane, passing through point **D**.

Due to the distances of **DA** and **DB** are different, a point on line **DF** might project to different row of image A and B. Therefore there will be a problem when matching up two overlapping section in the image A and image B. The maximum error *e* should occurs on the column corresponding to line **DF** in both images. Let us first discover the largest possible error *e* in terms of the known parameter *d*.

First of all, we need to find the coordinate of point **D**. It is the intersection point of the straight line **AC** and the circle centered at **O** with radius *d*. Solve the following two equations simultaneously to get **D** = (x_D , y_D).

$$x_D{}^2 + y_D{}^2 = d^2, \text{ and} \tag{3.1}$$

$$y_D = (x_D - r)\tan(\frac{\theta}{2} + 90). \tag{3.2}$$

To reduce the messy symbol involvement in calculation process, the equation (3.2) is represented as $y_D = m\, x_D + c$, where $m = \tan(\theta/2 + 90)$ and $c = -r\tan(\theta/2 + 90)$.

First square both sides of equation (3.2), then substrate it from (3.1) we get

$$(1+m^2)x_D{}^2 + 2mcx_D + (c^2 - d^2) = 0. \tag{3.3}$$

Calculate x_D from (3.3). It follows that

$$x_D = \frac{-mc \pm \sqrt{(1+m^2)d^2 - c^2}}{1+m^2}.$$

Substitute original *m* and *c* back in the solution, the coordinate of **D** becomes

$$x_D = \frac{rt^2 \pm \sqrt{(d^2 - r^2)t^2 + d^2}}{1+t^2}, \text{ and}$$

$$y_D = \frac{-rt \pm t\sqrt{(d^2 - r^2)t^2 + d^2}}{1+t^2},$$

$$\text{where } t = \tan(\frac{\theta}{2} + 90).$$

Now, we plot the setting in YZ-plane, i.e. a side view, as shown in Fig. 3.3. Let *a* and *b* be the distance of **AD** and **BD** respectively. Calculating values a and b is straight forward since all coordinates of **A**, **B** and **D** are known. Suppose **G** is an arbitrary point along vertical line **DF** with h cm apart from **D**. Moreover, point **I** and **J** are the projections of point **G** on image B and A respectively, where **I** is vertically i cm away from the center of image B and similarly **J** is j cm away. The value e then is defined as (i - j) cm. In order to find the maximum e, value i must also be maximum. So point **I** happens to lie on top/bottom boundary of image B.

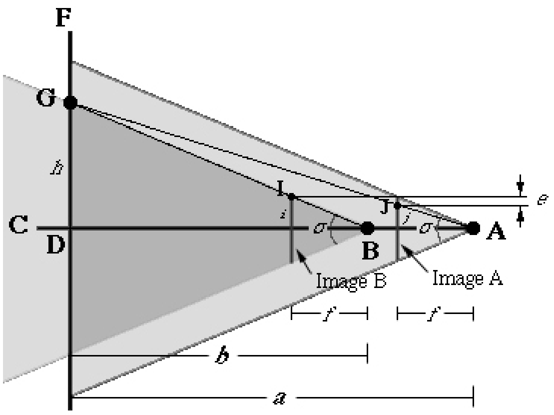


Figure. 3.3: Two successive camera positions in the YZ-plane. The variables are explained in the text.

From the triangle **BGD**,

$$h = b\tan(\frac{\sigma}{2}), \text{ and}$$

$$i = \frac{hf}{b} = f\tan(\frac{\sigma}{2}).$$

Then from the triangle **AGD**, we get

$$j = \frac{hf}{a} = \frac{bf\tan(\frac{\sigma}{2})}{a}.$$

Hence, the maximum error equals

$$e = (i - j) = (1 - \frac{b}{a}) f\tan(\frac{\sigma}{2}). \qquad (3.4)$$

Normally, the film size is known in advance, so the relation between effective focal length and field of view can also be calculated. For example, the most common film is 35 mm, so the relation between f and fov becomes,

$H.fov = 2\tan^{-1}(18/f)$ - horizontal field of view, and

$V.fov = 2\tan^{-1}(12/f)$ - vertical field of view.

Then formula (3.4) can be written as,

$$e = 18 \times (1 - \frac{b}{a}).$$

Let us illustrate this by an example with following setting parameters:

- Length of base line r = 6.5 cm, based on human eyes' distance.
- The rotation angle θ = 22.5 degrees, so 16 iterations are required in total.
- Focal length f = 20 mm.
- Vertical field of view σ = 39.6 degrees.

The possible maximum errors with respect to different distance between the closest object and camera are list in Tab. 3.1. Where, d stands for the distance between the closest object and camera in cm; e stands for the maximum possible error that can happen in every two successive acquiring images, unit in pixels. The conversion from cm to pixel for the particular digital camera is 1 cm = 1150 pixels. In order to see the change of error e in precise way, we use real number even though they will be round up for the discrete image. The relation between the error and distance is plotted in Fig. 3.4.

d (cm)	e (pixel)
50	21.12
100	10.62
150	7.10
200	5.33
250	4.27
300	3.56
350	3.05
400	2.67
450	2.38
500	2.14
550	1.95
600	1.78
650	1.65
700	1.53
750	1.43
800	1.34
850	1.26
900	1.19
950	1.13
1000	1.07
1050	1.02
1100	0.97

Table 3.1: The maximum possible errors for different distances between the closest object and the camera.

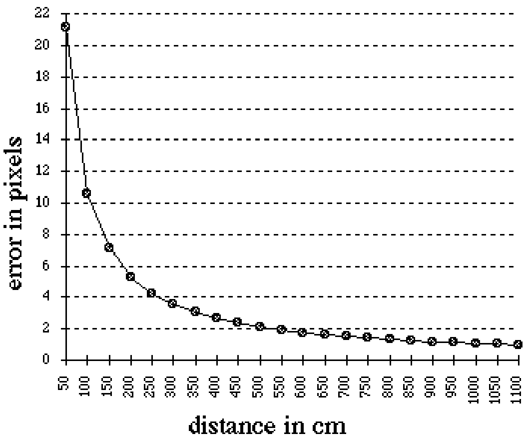


Figure. 3.4: The plot of error vs. distance.

Figure 3.4 shows the relationship that the errors decrease dramatically as distance increasing. As well as from the table it shows that if the distance between the closest object and camera is large enough such as 1100 cm, i.e. 11 m, then the maximum error can happen is less than 1 pixel. Of course, if the application does not require such high accuracy, 1.5 and 2 m may be acceptable for just 7 and 5 pixels error in maximum.

By given *d*, *a* and *b* can be found, then *e* is found. Conversely, if we specify the maximum error *e* allowed in the stitching process for certain quality, the shortest distance *d* can also be calculated. Since *e* can be determined in advance and camera can also be placed at *d* cm away from the closest object in the scene, we refer this scheme as errors controlling.

As the position of **A** and **B** are both known, the distance between them becomes trivial, so we have the following relation between values *a* and *b*.

$$a = b + r\sqrt{2 \times (1 - \cos\theta)}. \qquad (3.5)$$

Combine equation (3.4) and (3.5) we get,

$$a = \frac{rf}{e}\tan(\frac{\sigma}{2})\sqrt{2 \times (1 - \cos\theta)}, \qquad (3.6)$$

where $e > 0$ as long as $\mathbf{A} \neq \mathbf{B}$.

Again, *fov* can be expressed in terms of *f*. The equation becomes,

$$a = \frac{18 \times r}{e}\sqrt{2 \times (1 - \cos\theta)}.$$

Since we have the equation of line **AC** and the value of *a* which is the distance between **A** and **D**, therefore we are able to find the exact position of point **D**:

$$x_D = \frac{r(1+t^2) \pm a\sqrt{1+t^2}}{1+t^2}, \text{ and}$$

$$y_D = \frac{\pm at\sqrt{1+t^2}}{1+t^2},$$

where $t = \tan(\frac{\theta}{2} + 90)$ and refer to a in (3.6).

Finally the distance between rotating center **O** and point **D** should be at least equals to

$$d = \sqrt{r^2 + a^2 \pm \frac{2ra\sqrt{1+t^2}}{1+t^2}},$$

where $t = \tan(\frac{\theta}{2} + 90)$ and refer to a in (3.6).

The value *d* is driven in terms of parameters: *e* - the error, *r* - the distance between camera and the rotation center, *f* - the camera focal length, θ - the rotation degree, and σ - the camera vertical field of view. Since they are known in advance and keep constantly for making one complete panoramic images, by using the above formula for *d* we know the closest object in 360 degrees panorama should be at least *d* cm away from the camera rotation center.

3.3 Panoramic Anaglyph Processing and Camera Models

In stitching stage, different stitchers may make different decision to deal with those errors, the result of decisions may leave partial unmatched or make objects in the scene fatter or thinner than they should be. In the first situation, for example, it can match $\mathbf{P}_1$ and $\mathbf{P}_2$ but $\mathbf{Q}_1$ and $\mathbf{Q}_2$ as in Fig. 3.1; and for the second, it may push two images toward or pull away from each other for the best match. Here we are more concerned with the second approach due to the belief in that most stitchers try their best matching. Now we have two camera models and two options for processing the panoramic anaglyph images and we want to make choice among them to have resulting image least influenced from those errors, i.e. less noticeable by human eyes.

The two options of processing the panoramic anaglyph images are shown in Fig. 3.5. Option 1 is to make anaglyph image first for each stereo pair then stitch them all to form a complete panorama image. Option 2 is to stitch left images set and right images set separately and combine the resulting two to form the final anaglyphic image. In Option 2, two separate panoramas have at least one having the errors mentioned before, so after combining two panoramas it might give incompatible disparities in the resulting anaglyphic image, and therefore an incorrect depth cues to our eyes. For the option 1, every single anaglyph image, A_1 ... A_n, has the desired disparities as it is made of. After stitching them all, the designed disparities are still preserved as the original while for Option 2 the incompatible disparities might influence to the others. Therefore Option 1 should be chosen.

Two camera models are shown in Fig. 3.6. Model I has right images taken with camera whose nodal point correctly passing through the rotating axis while left images taken certain distance away from the rotation center. Model II places the rotation center at the middle of the camera displacement so that both left and right images are taken from incorrect setup, but about half errors in each side.

Recall from the discussion in Section 2, we have red filter for the left eye and the bluegreen filter for the right eye in order to perceive the full color panoramic anaglyph image. For the model I green

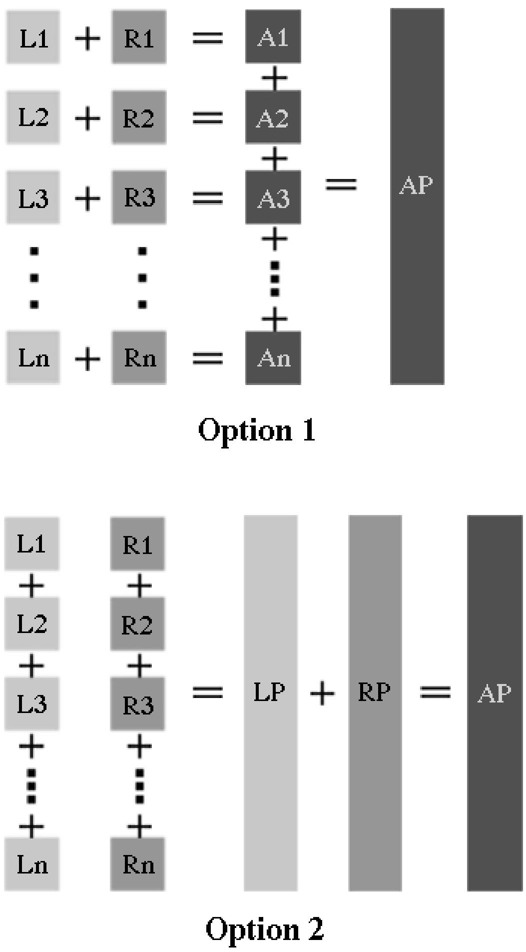


Figure 3.5: Two options of processing the panoramic anaglyph images, where L stands for left; R for right; A for anaglyph image; P for panorama image.

and blue components are perfectly matched in stitching processing, which contributes two third of primary colors to final panoramic image. In other words, only red component image might have the errors. In model II, red, green and blue components all have the errors even though each one of them has only about half size of errors compared with the red component in Model I. Combining each model with Option 1 processing method chosen above, the model I only has one third chance that error may be noticed by our eyes while Model II has 3 times more chance than Model I does. In addition, our eyes have less sensibility with red component than bluegreen components. Therefore the model I should be adopted.

4 Stereo Panorama Player

A stereo panorama player is a navigation tool for exploring the photo-realistic 3D world in 360 degree panoramic view with a pair of 3D eyeglasses. It allows the user to interact with the stereo scene dynamically according to viewer's own preference in

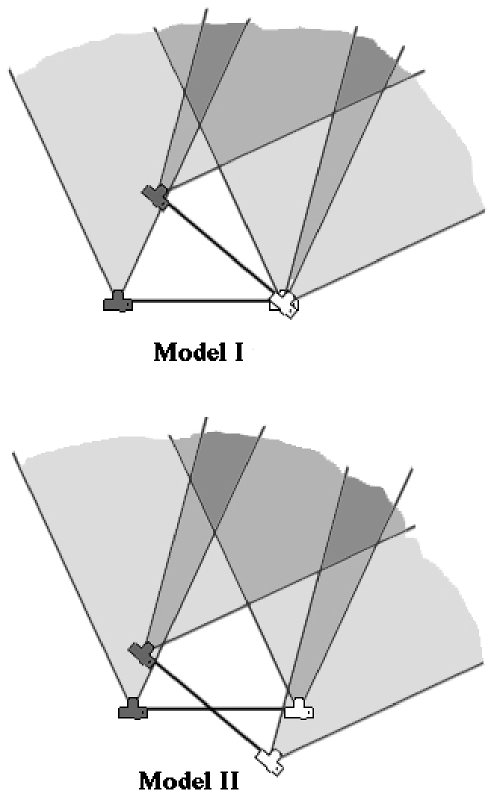


Figure. 3.6: Two camera setup models.

the play-time, just like play station. The basic functions currently provided in the player include 360 degree horizontal and less than 180 degree vertical panning, zooming in/out and hopping between multi-nodes.

In fact, this player is similar to most panoramic players except of the different considerations involved for its design. Besides, there are three new proposals for the player enhancements. First, we like to have the player spin around the scene automatically with specified speed without toggling the mouse. So we can carry the 3D eyeglass close to the eyes with one hand and the other free for other operations such as zoom in/out and so forth. Second, we like to manipulate the object which is part of scene directly within the scene so user plays with object more intuitively. Third, we want to have a graphical *directional indicator* which tells the user about her/his current position and viewing direction as well as where they can go next for multiple panoramic nodes environment.

4.1 Reprojecting an Environment Map

The real-time reprojection of an environment map is used to visualize surrounding scene and to create interactive walk-through. An environment map is a projection of a scene onto a simple shape, where the

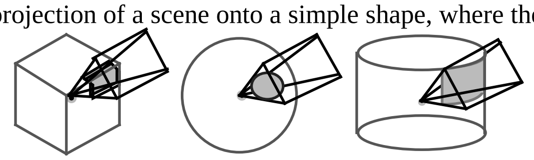


Figure 4.1: Reprojecting a cubic, a spherical and a cylindrical environment map.

scene is of panoramic anaglyph image. Typically the shape can be a cylinder, a sphere or a cube. The decision of choosing a shape is done in the design time and to implement the projection of the scene onto the chosen shape is held in the stitching processing [7]. The resulting environment map then can be reprojected onto the view port of the player according to the type of the environment map. The reprojection of environment map for a cylindrical, spherical and cubic shape are shown in Fig. 4.1.

4.1.1 Choose Environment Map Type

Since we are dealing with the anaglyphic panoramas, it is important to choose a proper type of the environment map so the correctness and quality of reprojection can keep above the certain standard. We choose a cylindrical environment map for the following reasons:

- If the reprojection of an environment map, such as the spherical shape, warps the vertical line/border of object in the scene, then it distorts the disparities where the line is bent. The stereo form therefore may no longer keep in correct proportion relatively. Thus it causes the closer part of scene further away from the viewer, and vice verse. However, in the reprojection of a cylindrical environment map the vertical line keep straight invariantly.

- Normally an ordinary camera has less vertical field of view than the special-purpose camera such as a fisheye camera. So we do not have much top and bottom views with the ordinary camera in 360 degree panorama acquisition. Therefore it is pointless to use the cubic or spherical type which is intended for wider vertical view.

4.1.2 Reprojection Formula

Figure 4.2 shows the reprojection of a cylindrical environment map onto a view port. The reprojection formula can be described as a relation between two coordinate systems. One is the panoramic image system, another is the view port coordinate system. Figure 4.3 shows the mapping

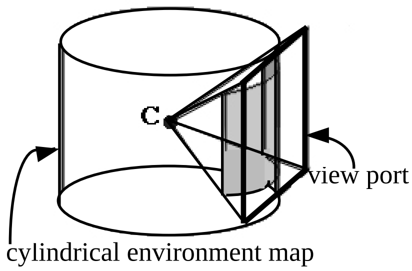


Figure 4.2: Reprojecting a cylindrical environment map onto the view port.

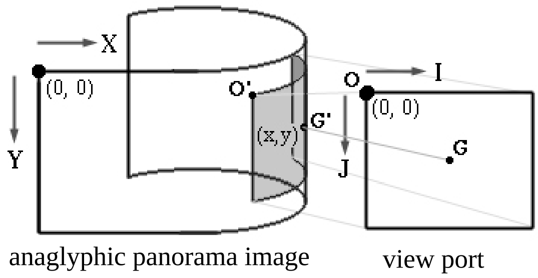


Figure 4.3: The mapping from panoramic image coordinate system to the view port coordinate system.

from panoramic image coordinate system to the view port coordinate system. Their origins are both at the top left corner.

We use XY for the panoramic image coordinate system and IJ for the view port, both in pixel unit. In the following contents, $(x, y)_P$ denotes the XY coordinate system, and similarly $(i, j)_v$ for the IJ's.

Assuming the view port is the tangent plane to the side face of cylinder and its vertical axis is parallel to the vertical center line of the cylinder. For the panoramic image with the height h and width l, view port v pixels high and w wide, and a point **G**, the center of IJ, maps to point **G'** = $(x, y)_P$, an arbitrary point **P** = $(i, j)_V$ then maps to $(x+a, y+b)_P$, where a and b are :

$$a = \frac{l}{360}\tan^{-1}\left(\frac{i-\frac{w}{2}}{r}\right), \text{ and}$$

$$b = \frac{r\left(j-\frac{v}{2}\right)}{\sqrt{r^2+\left(i-\frac{w}{2}\right)^2}}.$$

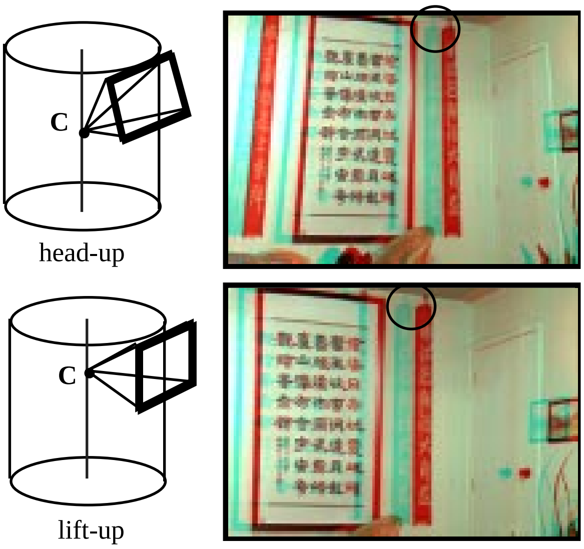


Figure 4.4: Two reprojection approaches: lift-up and head-up.

4.1.3 A Lift-up Reprojection Approach

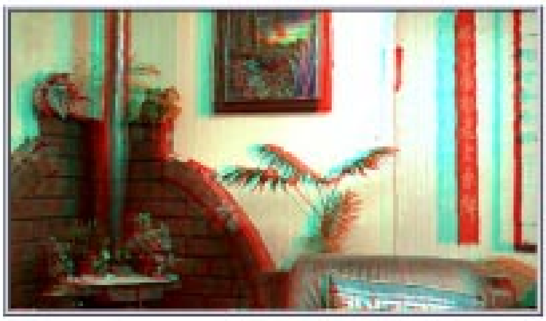
Instead of normal cylinder reprojection which fix the projection center at the center of cylinder and heading up/down for the top/bottom view, our reprojection allows the projection center lift up and down. The reason for doing that is to preserve the relatively correct disparity proportion in the reprojection area.

Fig 4.4 shows where the problem occurs for the head-up approach, and the improvement from the lift-up approach. In the head-up reprojection, the top of bamboo artwork circled has smaller disparity than the one of its bottom. Therefore, the top of bamboo artwork looks closer to the viewer while the bottom is relatively further away form the viewer. That is wrong stereo impression to the actual scene. To avoid that, we suggest the lift-up approach, which gives the correct visualization in the stereo form with less distortion.

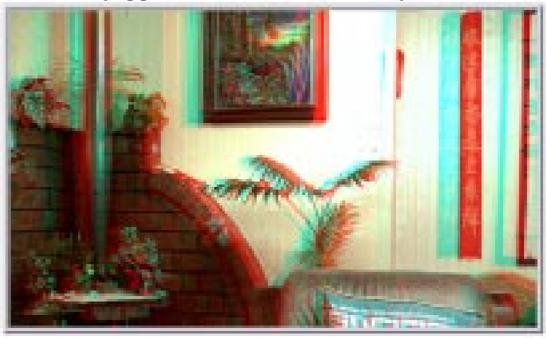
4.1.4 The Result

The final result of our player for map reprojection is shown in Fig. 4.5, and the comparison is made to the commercial panoramic player.

The main difference between the two, beside of the disparity proportion correctness problem mentioned in the previous section, is the vertical component of object border. The commercial player has jagged border lines for objects in the scene while we have crisp vertical borders, which should be attributed to the choice of environment map and reprojection method.



jagged vertical border of objects



crisp vertical border of objects

Figure 4.5: Two snapshots of rendered anaglyphic panorama in two players; the top one is from the commercial product, the bottom comes from our player.

4.2 Auto Spin

To design a panoramic player, one very interesting issue, beside of reprojection, is how to define the interface and interactions between viewer and the player. Conventionally, a modest size window represents as view port for rendering the panoramic scene. Viewer controls the viewing direction via toggling the mouse over the view port to simulate the camera panning and types the specified keys in the keyboard for other actions, such as zoom in, zoom out, hopping ...etc..

Here, we introduce a slightly different way to drive around in the virtual world -- Auto Spin. The interaction rules are defined as follows:

- The mouse toggling direction indicates the panning direction.
- The toggling distance (in pixel) determines the panning speed where short distance for slow and long for fast. Once the speed is calculated, it keeps constantly for that panning.
- Player continues the panning if the toggling distance is greater than 1 pixel, otherwise the scene stays static.

It preserves the conventional interaction rules used in other players with one more option allowing camera panning in the desired direction continuously without further mouse toggling. So viewer can look around the scene relaxedly with one more free hand to invoke other action. It is especially useful for the anaglyphic technique since usually one hand is required for carrying the 3D eyeglasses.

4.3 Manipulable Objects

When we walkthrough the virtual world, there should be many objects around us, e.g. a book on the table, a plant beside the door or a globe on the cabinet. So far with a panoramic player, we can view those objects in the different direction (for multi-nodes only), and probably having a close-up via zoom-in function. But we cannot manipulate them, for instance, open the book, exam the plants or lookup New Zealand from the globe on the cabinet. In this project, we present one of possible object manipulations, which allow viewer directly examining the object with predefined behaviors.

Let us start from image acquisition. Object photographics should be taken within the same scene for the panorama and the place it shows up, so all the complexities such as illumination, shading, reflection, refraction and geometric constraints are then possibly consistent one and the other. It is very important to have a set of consistent object images otherwise the object may appear odd in the scene, i.e. reducing the quality, and lost the point to have this enhancement. Each object rotation needs two camera shootings for stereo pair. After digitalization of acquired photos, process and place them into single array if only single rotation is required; or into 2D array if more than one rotation are enabled. Finally calculate the corresponding position in the panorama image, and place the initial frame at computed position. The frame size should correspond to the object size. The basic model is shown in Fig. 4.6.

Since the manipulable objects are embedded in the panorama, the user interaction area for objects and panorama are overlapped. We define a rule such that they can behave independently. The rule is if the starting position of the mouse toggling is within the object area, the mouse event is passed to the object handler, else to the panorama handler. Since the object position is varied in the view port of the player, the player needs to dispatch the object position dynamically. So even the panorama is in auto-spinning we should still be able to manipulate

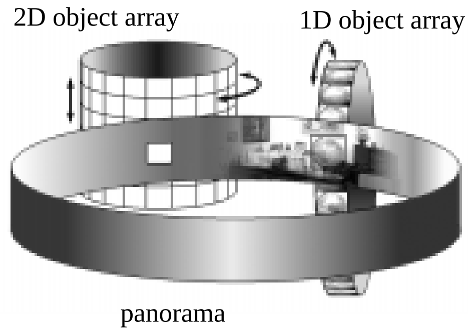


Figure 4.6: The basic model for manipulable objects embedded in the panorama.

the object desirably. The state of object is kept separately from the panorama so the object appears consistently. Furthermore, since in our player we also have the object auto-spinning (i.e. a globe on the cabinet), both object and panorama are turning around simultaneously and consistently. This animation vivifies the virtual world and increases the entertainment of navigation.

To acquire the object images in different rotation is known very difficult, especially for the big object. To acquire the stereo object images for all the rotation is even harder. Beside, the larger amount of images need to pre-load onto physical memory for real-time speed, the system resource are consumed heavily. The trade-off for that is its photographic realism attraction and rich contents.

4.4 Directional Indicator

In the real world, if we go to a new place we can go lost very easily. The same situation for the virtual world, if we have never been there before, we are also easily to get lost. A helper is required for guiding the user exploring the new world more effectively.

The helper should contain competent information to the new environment and easy to understand. The role it plays in the player should be as a reference and appears as needed. Although it is passive, it should reflect the current exploring status to the viewer dynamically. It should not contain any controls used for the navigation flow, or even take over the role of the view port where the world is explored.

We present one of possible representations for the indicator, the main goal is to show this idea. We use the top view of virtual space as background of

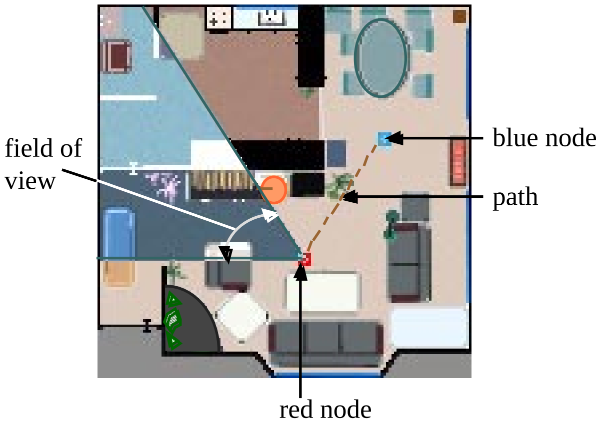


Figure 4.7: Demo directional indicator for the multi-nodes panorama player.

the indicator. Red node represents the position of the current active panorama, i.e. the user current position. The projection emitted from the red node indicates the current viewing direction and field of view, also it is rotating synchronously as the horizontal panning in the view port of the player. To indicate the connectivity between nodes, the path is plotted in desh-line. User can only hop to the connected node if the path is covered by the emitting projection from the current node, i.e. user should see where she/he is hopping to in the view port. All inactive nodes is represented by blue node, when user jump to other node, the color of nodes are updated as specified. The demo indicator for our player is show in 4.7.

Like manipulable objects embedded feature, it requires more labor work from the panorama maker. For example, the maker needs to prepare a map (i.e. top view of panoramas), finds out the reference node positions in the panorama image, calculates the registration points between connected nodes, specify the starting node and its initial viewing direction, and so on. So if the number of nodes is few or the topology of nodes is simple, such as the linear, then the indicator may not be necessaryly needed.

5. Conclusions and Future Works

In this report, we introduce a new panorama visualization which utilizes the human binocular visual characteristic and the capabilities of colors comprehension through the full color anaglyphs stereoscopic technique. The panoramic and anaglyphic stereo images acquisitions are known incompatible in their setups. We measure the errors of the problem and show the errors can be controlled at the setup time.

To improve the quality of stereo perception, we show how to choose the filters for the 3D eyeglasses and propose a method to reduce the double image. The camera model and panoramic anaglyph images processing are selected purposely in order to have less influence from the errors introduced by combining both panoramic and anaglyphic images acquisitions. The result of our approach is convinced to the audience.

We have also developed a real-time navigation tool allowing exploring the stereo panorama in arbitrary directions. Because it is written as a Java Applet, the potential of the internet accessibility and multimedia capability can be exploited. Regarding the properties of anaglyphic image, the cylindrical environment map and lift-up reprojection method are chosen so the better visualization is performed. Three additional features are added to the player, auto-spin, manipulable objects and directional indicator. To make more enjoyable and richer for the user, the multimedia potential should be exposed in the future work.

To walk closer to the object in the scene, normally we simulate this motion via zooming in the present image frame. One problem encountered for the anaglyphic panorama visualization is that the disparity becomes larger as we walk closer to the object, which is not subject to the fact that the closer to the object the smaller disparity it should be. The suggestion for the future work is that the disparity should be made adjustable dynamically according to the zoom-in factor so the proportion of disparities are kept correct relatively .

The analyphic panorama approach produces a manifest stereo impression to our vision with no extra overhead and degradation in the performance compared with the traditional panorama visualization method. The applications of this approach also inherits the benefits of normal panorama visualization, which is affordable and applicable for wide available workstations and PCs.

6. Acknowledgments

We are very grateful to the anonymous reviewers for their invaluable suggestions which made this a significantly better report. Many thanks go to Georgy Gimel'farb and Karsten Schlüns for various helpful discussions at the early stages of this work and their valuable remarks.

References

[1] Ian P. Howard and Brain J. Rogers. *Binocular Vision and stereopsis.* Oxford Psychology Series No. 29. Oxford University Press, New York, 1995.

[2] Bernd Jähne. *Digital Image Processing: Concepts, Algorithms, and Scientific Applications.* Second edition. Springer-Verlag, Berlin, 1993.

[3] John C. Russ. *The Image Processing Handbook.* Second edition. CRC Press, Boca Raton, 1995.

[4] H. Helmholtz. (Trans. by E. Atkinson). *Popular Lectures on Scientific Subjects.* Longmans Green, London, 1893.

[5] Gerald M. Murch. *Visual and Auditory Perception.* The Bobbs-Merrill Company, New York, 1973.

[6] LEE Filters.
URL: http://www.leefilters.com/

[7] Shenchang Eric Chen. QuickTime® VR - An Image-Based Approach to Virtual Environment Navigation. In *Proc. SIGGRAPH' 95*, pp. 29-38.