2007 International Symposium on Information Technology Convergence

A New Algorithm for Depth Perception in 3D Screen and Its Implementation

Woonchul Ham
Div. of Electronics and Information Eng
Jeonbuk National Univ, Jeollabuk-do
wcham@chonbuk.ac.kr

Seunghwan Kim
Div. of Electronics and Information Eng
Jeonbuk National Univ, Jeollabuk-do
melodyoflove@nate.com

Abstract

In this paper, we present a new smoothing algorithms for variable depth mapping for real time stereoscopic Image for 3D displays. Proposed algorithm is based on the physical concept, called Lapacian equation and we also discuss the mapping of the depth from scene to displayed image. The approach to solve the problem in stereoscopic image which we adopt in this paper is similar to multi-region algorithm which was proposed by N.Holliman. The main difference thing in our algorithm compared with the N.Holliman's multi-region algorithm is that we use the Laplacian equation by considering the distance between viewer and object. We implement the real time stereoscopic image generation method for OpenGL on the circular polarized LCD screen to demonstrate its real functioning in the visual sensory system in human brain. Even though we make and use artificial objects by using OpenGL to simulate the proposed algorithm, we assure that this technology may be applied to stereoscopic camera system not only for personal computer system but also for public broad cast system.

1 Introduction

Computer graphics algorithms and their hardware implementation have progressed spectacularly. Photon mapping, ray tracing and image based rendering techniques reuse precomputed or pre-acquired global illumination solutions to produce compelling visual experiences at interactive rates. However, when the surface appearance depends decisively on the view point as in the case of, for example, refractive or reflective surfaces, when the scene objects are dynamic, or when the lighting conditions change, interactively computed imagery falls short of realism. In such conditions rendering images that are mistaken for photographs still require more time than available in an interactive rendering system. We breifly review the recent technology in 3D stereoscopic image device. An emerging technology is based on holography. At this time, there are no commercially available devices that create moving holographic imagery. There are, however, devices for creating holographic still images. Holographic moving images demand bandwidth and computational power that appears to exceed current technology. While the LCD technology is quite mature, the technology is still in its infancy stage, equivalent to the pre-VGA stage of 2D displays. The SD line of stereo displays from Planar is the specialized stereoscopic monitors. In this implementation, they use the innovative Stereo mirror technology and implement the high quality Planar monitors to give viewers the ultimate in stereo viewing pleasure. By far the most common stereoscopic device is based on presenting two separate images to each eye that account for parallax. Sharp research team developed a new 3D viewing without the need for 3D glasses, otherwise known as autostereoscopic displays. They use the parallax barrier which is a series of vertical slots which are carefully designed to focus pixels in different directions. Sharp parallax barrier is placed between the LCD and the back light and can be switched on and off. This kind of 3D technology delivers stunning images for use in applications such as medical imaging , satellite/aerial photogrammetry, computational chemistry; complex modeling visualization and computer game. The perceived depth seen in a stereoscopic images varies not only with many geometric parameters but also with the many parameters with human eye such as separation between two eyes and focusing of each eye. It is well known fact that there is a limit to the range of perceived depth. Human can not detect the depth difference between two objects if two objects are located beyond 350m from the viewer. So it is useless to consider the scene depth of object if its distance is very far from the viewer. This is the reason why we find the relationship and optimal mapping between scene depth and perceived depth for the best implementation of stereoscopic image. N.Hollman address this problem and propose multi-region algorithm and this algorithm allows different regions of the scene to be mapped to the different ranges of perceived depth on a target 3D display[1]-[3]. The main idea of N. Holliman is to finding a smooth mappping in order to give more dynamic range of perceived depth for the region

0-7695-3045-1/07 $25.00 © 2007 IEEE
DOI 10.1109/ISITC.2007.73

of interest in the scene dept. In his algorithm, he considers the scene outside of the region of interest without neglecting it at all because he thinks that it is important to see the context around the region of interest in some applications. He also consider the visible discontinuity at the boundary of the multi-region algorithm and stress on the continuity of the perceived depth for more comfortable depth feeling. In this paper, we specifically suggest a new mapping between scene depth and perceive depth based on Lapacian equation. We also comments on the resolution of human retina and derive the relationship between perceived depth and real scene depth. We propose compensation algorithm which post process the right and left image for handling focusing problem. In here, focusing problem means that the focusing object is felt to be located at the same depth level of real 2D LCD screen and this make human viewer feel more comfortable when he watch the 2D LCD screen because the focusing objects and real 2D LCD screen has the same distance of focus in front of human viewer.

2 Perception of Human Eye

In this section, we discuss the limitation of human eye for the detection of depth when the object is located far from the viewer. The following figure shows the relationship between the distance d of object from the viewer and the corresponding mapping angle θ on the retina. Let the variable d_1 and d_lim be the distance between two eyes and the distance limit that the human can perceive the depth difference, that is, human can't tell the depth difference if two objects are located over the distance limit from the viewer. From the recent researches, the average value of d_1 is 0.03m and the d_lim is about 350m. from this values, we can guess that the angular resolution of human retina is about $(\pi/2 - tan^{-1}(350/0.03)$ radian or 0.0049 degree. Let us assume that the distance range of human's focusing capability is from 0.1m to infinite. We can guess that the magnitude of the dynamic angular range of retina is $(\pi/2 - tan^{-1}(0.1/0.03))$= 20 degree. We can guess from this value that the retina has two dimensional visual capability about 3500x3500 pixel resolution.

The surface of space on which human perceive same depth can be described by the following equation.

$$x^2 + y^2 + (z - M)^2 = d_1^2 + M^2$$

where M is any positive number. From the above equation and angular resolution of human eye, we obtain Figure 2 which indicates equi-depth surfaces on the three dimensional space.

From figure 2, we can see that there exists a nonlinear mapping between scene depth and perceived depth in human eye and we must consider this fact when we realize

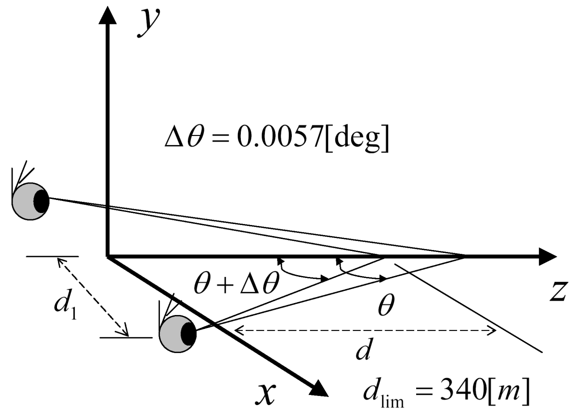

Figure 1. Geometric View and Perceived Depth of Human Eye.

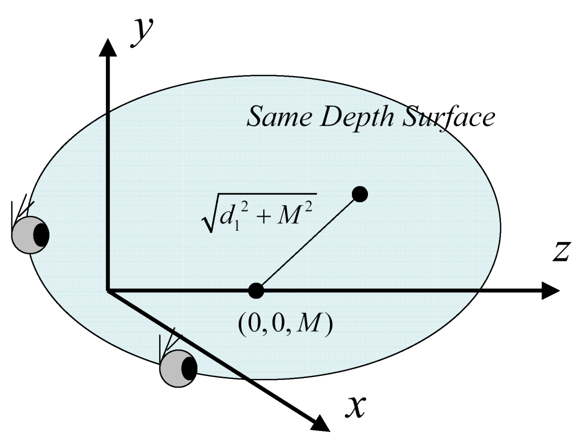

Figure 2. Perceived Equi-depth Surface.

3D image in a 2D plane screen such as polarized 3D LCD monitor.

3 A New Mapping from Scene Depth to Perceived Image Depth

There has been many researches for the mapping from the scene depth to image depth. One of the famous researchers in this research area is N. Holliman. he suggested many algorithms for this research topic and the famous one is called multi-region algorithm. We also adopt his idea to propose a new mapping algorithm but in here, we utilize the famous physical law called Laplacian equation expressed by

$$\nabla^2 V(x, y, z) = 0 \tag{1}$$

The physical meaning of above Laplacian equation is that the brightness function $V(x, y, z)$ at a point P in a space must satisfy the above equation under the assumption that the light source does not exists at a point P. Usually the light sources are located far from the object on which our eyes are focusing. From the Laplacian equation, we can obtain averaging method expressed as follow for lightness function $V(x, y, z)$.

$$V(x, y, z) = \frac{1}{8} \sum_{i=-1,1} \sum_{j=-1,1} \sum_{k=-1,1} V(x - i, y - j, z - k) \tag{2}$$

We can obtain following fact from the physical meaning of Laplacian equation, when the light sources are located near the viewer and far from the objects we are focusing on.

Fact 1 *Light intensity is inversely proportional to R^2, where R denotes distance of object from the viewer.*

Fact 2 *When we process the 3D vision image, we must consider a nonlinear mapping between scene depth and perceived depth about which we will discuss bellow.*

Turning our attention to this nonlinear mapping between scene depth and perceive depth, it is interesting to note that the same perceived depth is proportional to the inverse square of the distance approximately. Figure 3 shows the relation between perceived depth for the same resolution of retina of human eye and distance under the assumption that the unit of x-axis denotes angle of retina on which an object is mapped and we obtain the following equations between them.

$$\begin{aligned} d &= d_{1/tan}(\theta) \qquad (3) \\ d_1 &= 0.03[m] \\ \theta_0 &= tan^{-1}(0.200/0.03) \\ \triangle\theta &= 0.0049[deg] \end{aligned}$$

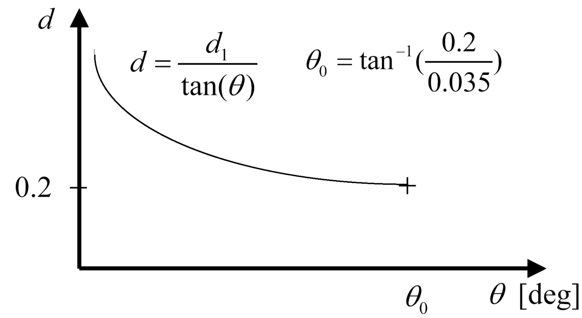

Figure 3. Perceived Equi-depth Surface.

From the above fact, it may be better to preprocess nonlinear mapping between scene depth and perceive depth and then we process all the regular imaging processing introduced by N. Holliman, such as camera frustrum at al. We recommend a nonlinear scaling of the $z-$component of the object as a nonlinear mapping by considering Laplacian equation expressed as follow.

$$\begin{aligned} x_v &= x_c/R^2 \\ y_v &= y_c/R^2 \\ z_v &= z_c/R^2 \end{aligned} \tag{4}$$

where x_c, y_c, z_c denotes the coordinates of the objects in camera space, R denote the distance of the object from the camera system and x_v, y_v, z_v is the virtual coordinate system on which all the regular image processing is done for the corresponding 2D image. Let us remain as a next research topic how to utilize this kind of nonlinear mapping for better perceived depth

4 Perceived Depth and Its 2D Implementation

Now we discuss 3D image implementation on the 2D LCD panel of the captured 3D image by stereo vision, that is by using two CCD cameras. From the figure 4, we should consider the nonlinear mapping between 3D real stereo image and its 2D implementation. In most case, 3-axis scaling calibration is need for the comfortable feeling of perceived depth in 2D LCD panel. This nonlinear relation between 3D image and its 2D implementation is shown in figure 4. The main key point of 3-axis scaling calibration is companding or compression each axis to main the same relation between human eyes and 2D LCD screen, such as ratio of d_1 and f_1.

Then we adopt projection matrix in terms of new coordinate system which we obtain by using above 3-axis scaling calibration. In here, we use the same project matrix derived

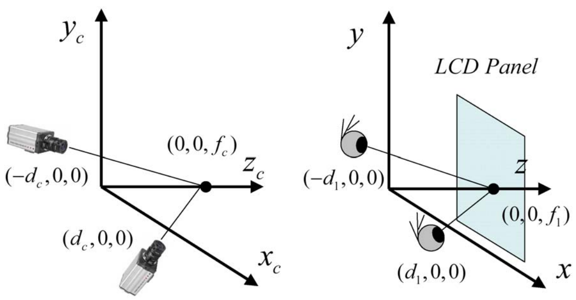

Figure 4. 3D Stereo Image and Its 2D Implementation.

by N. Holliman for handling the region of interest mapping and we briefly review his research. He derive the camera parameters to generate the left and right partial image for the region of interest using asymmetric camera frustrum[1]. He derive the following equations for the camera parameters of the right camera. We can use the parameters when we simulate the synthesized 3D image by using OpenGL. We can design the perspective matrix for OpenGL by using these parameters. Geometrical view of these parameters are shown in figure 5.

$$Left = -\frac{n}{d}(w - d_1) \quad Right = \tfrac{n}{d}(w + d_1) \tag{5}$$

$$Top = -Bottom = (h)\frac{n}{d} \tag{6}$$

$$Near = n \qquad Far = f \tag{7}$$

Even though his idea is basically correct, there is one thing he didn't consider and this is the focusing problem in his algorithm. When we simulate his algorithm, there exists separating problem of the focusing point in left and right image and this makes uncomfortable feeling of depth. So we propose the following compensation that is to move the projection images to the right for right image and to the left for left image so that focus point is mapped on the center point of the 2D LCD screen.

$$LFB(i,j) = LFB(i + \triangle l, j)$$
$$RFB(i,j) = RFB(i - \triangle l, j)$$

$$\triangle l = (d_1/2w) * N_h \tag{8}$$

where LBF and RBF denote the frame buffers for left and right image respectively and N_h is the number of pixels for the horizontal line of 2D LCD screen.

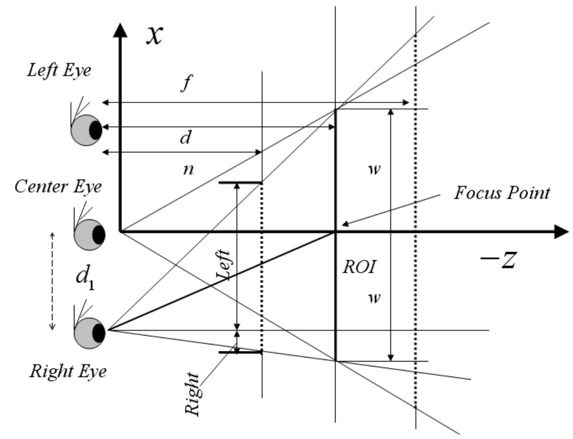

Figure 5. The Scene Geometry for Right Eye.

5 Simulation

We use the special circular polarized 3D monitor and special circular polarized eye glasses developed by Pavonine company for implementing the 3D depth stereo image. We use the OpenGL real time graphics as a software tool. The test scene consists of same size of 9 cubes and 3 axes. The width, length and height of cube are 4cm each and surrounding 8 cubes are located 6cm far from the centered cube on each axis. We assume that distance between 3D LCD screen and human viewer is 50cm and the distance between the centered cube and human viewer is about 50cm. So we set the near and far clipping planes into the z-axis as follows.

$$z_n = -50cm \quad z_f = -200cm$$

Next we move the cubes 100cm in the direction of negative z-axis by using glTtranslatef() API function of OpenGL.

5.1 Rotating Transformation

The central cube acts as the region of interest and we try to test the degree of comfortable feeling of the depth of surrounding cubes compared with the central cube. The left and right image for stereo vision are shown in figure 6 and they are obtained by rotating the whole geometric space with respect to the y-axis of central cube coordinate system after transferring it 100cm on the -z-axis. The rotating angle θ can be expressed as follows.

$$\theta = tan^{-1}(3.0/100) = 1.7[deg] \tag{9}$$

As anticipated depth in the region of interest is reproduced well enough to clearly distinguish the depth difference in central cube whereas these difference are significantly less clear in front cubes our algorithm. We also feel

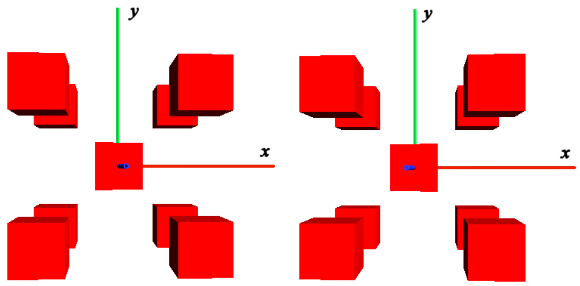

Figure 6. Left and Right Image using Proposed Algorithm.

uncomfortable depth when we are watching front 4 cubes. In our algorithm, all the objects are rendered by the same transform such as rotating appropriate angle with respect to the y-axis of central cube coordinate system. From the computer simulation, we found that this rotating transformation concept is not suitable for handling large size scene compared with focusing object. And more the object located far from the origin to the x-axis with same z-depth is felt that it seems to be located more to the -z axis.

5.2 Relative Tranlational Transformation

In this subsection, we adopt the relative translational concept for implementing left and right images. So after transferring all the cubes 100cm on the -z-axis, we move all the cubes -3.0cm and 3cm in the direction of x-axis for the right eye image and left eye image by considering the average distance of human eyes. We also consider the focusing problem and we select central cube as focusing object and we handle the camera frustum parameter as follow under the assumption we are using 17 inch LCD screen so that the central cube be located on the center point in LCD screen and with the zero depth feeling compared with frame of LCD screen. This kind of focusing handling gives human viewer very comfortable depth feeling naturally without any temporary eye focusing drill for it.

$$Left = -16 + \delta \quad Right = 16 + \delta \text{ for right image}$$
$$Left = -16 - \delta \quad Right = 16 - \delta \text{ for left image}$$

where we select δ as 1.5 by considering the fact that the the average distance of human eyes is 6cm.

We can apply the compensation method which was introduced in the previous instead of finding above frustum parameters. Anyway, we obtain more comfortable depth feeling and there exists no inconsistency in the depth feeling of the front cubes.

6 Conclusion

In this paper, we analyze and comments on the resolution of retina of human eye and derived perceive equi-depth surface function in the real scene space. We also comment on the nonlinear mapping based on the Lapalcian equation. The main proposal of this paper is implementation of 2D LCD image from the 3D real stereo image based on the new algorithm by considering not only the position of human eye but also the orientation of human eye and its implementation in OpenGL. From the computer simulation we prove the correctness of our suggested algorithm in the virtual synthesized 3D scene. We hope that the technology proposed in this paper will be utilized to the real stereo camera system such as broadcast system in near future.

Acknowledgments

This work was supported by Post BK21 Project of Chonbuk National University

References

[1] N. Holliman,"Smoothing Region Boundaries in Variable Depth Mapping for Real Time Stereoscopic Images," *Stereoscopic Displays and Virtual reality System XII*, *Proceedings of SPIE-IST Symposium on Electronic Imaging.*, (2005), vol.5664A of

[2] N. Holliman,"3D Display Systems", Handbook of Opto-electronics, IOP Press, Spring, 2005.

[3] N. Holliman,"Mapping Perceived Depth to Regions of Interest in Stereoscopic Images, " *Stereoscopic Displays and Virtual reality System XI*, *Proceedings of SPIE* 5291, 2004.