

STEREOSCOPIC VOLUME RENDERING OF MEDICAL IMAGES

by

MARTA KERSTEN

A thesis submitted to the
School of Computing
in conformity with the requirements for
the degree of Master of Science

Queen's University
Kingston, Ontario, Canada

August 2005

Copyright © Marta Kersten, 2005

Abstract

Recent advances in commodity graphics hardware provide new capabilities in the study of stereoscopic volume rendering. In the medical community, volume rendering is used to create 3D anatomical models for diagnostic purposes, surgical planning, and surgical guidance. A digitally reconstructed radiograph (DRR) illumination model can be used to simulate X-ray images. Unlike surface-rendered medical images, however, a volume-rendered DRR lacks depth information.

This work studies the use of stereopsis and aerial perspective as depth cues in the human perception of volume-rendered images. Two experiments and one preliminary study were conducted to evaluate the effectiveness of stereopsis and simulated aerial perspective on the depth perception of DRRs. The results of these experiments suggest that both stereopsis and simulated aerial perspective can improve relative depth perception in purely absorptive media. These results provide new ways to visualize complex volumetric data and to explore the capabilities of the human visual system.

Acknowledgments

In the process of writing this dissertation, I learned that research has its ups and downs, and that it always requires more time and effort than one plans at the outset. What I cherish most, however, is that I learned that every research project brings forth an opportunity to work with people, to help others and to receive help and support from them. There are always times when nothing works well and everything breaks down when it shouldn't and it is at these times that the presence and concern of supervisors, colleagues, friends and family becomes the most important thing in life.

For their support, understanding and guidance, I would like to take this opportunity and thank my advisors, Dr. Randy Ellis and Dr. James Stewart. I thank them for showing me that research is not a walk in the park, but that it is hard work, long discussions and coping with challenges, and that this is what gives satisfaction and brings forth results. Without their inquisitive minds and their active presence in my struggles I would neither understand nor appreciate this.

I met Dr. Randy Ellis over four years ago when he gave me the opportunity to work with him. I am greatly indebted to him for trusting and believing in me, at a time when I needed it most.

To Dr. James Stewart, I am thankful for his taking me on as a Master's student and becoming such an integral part of this work. His door was always open, he was always

there to listen, and to discuss and shed light on the road blocks we encountered along the way. Through it all, his sincerity and his sense of humour have made this a great experience.

Thank you James and Randy.

I would like to extend my thanks to the entire Medical Computing Lab at Queen's University, who have been extremely helpful, supportive and have made my time at Queen's a wonderful experience. I cannot imagine a better group of individuals; it's been a privilege to work (and play) alongside such bright, talented, and wonderful people. While finishing this thesis, I realized how amazing you all are and that I miss you already.

Our many discussions with Dr. Troje led to the design and analysis of the psychophysical experiments presented in this dissertation. For all of his help and guidance, I am very thankful to Dr. Niko Troje.

Then there are those, some of whom I have known my whole life and some who I have met along the way, my family and friends – you have listened, encouraged me, showed your support and your concern, you have offered your advice, your suggestions, and your help. You have *also* nagged me, continually asking, “Are you done yet?”. To this question, I can finally say “YES!”. And if you are reading this dissertation now, then you are certainly among those - so thank you.

Contents

Abstract	i
Acknowledgments	ii
Contents	iv
List of Tables	vii
List of Figures	viii
1 Introduction	1
1.1 Digitally Reconstructed Radiographs	2
1.2 Stereopsis	3
1.3 Aerial Perspective	4
1.4 Contributions	5
1.5 Organization of the Dissertation	5
2 Perception	7
2.1 Physiological Cues	7
2.2 Binocular Disparity	8

2.3	Psychological Cues	11
3	Volume Rendering	15
3.1	Volume Rendering: Theory	15
3.1.1	Optics of a Participating Medium	16
3.1.2	Volume Rendering Integral	19
3.2	Volume Rendering in Practice	22
3.2.1	Ray Casting	22
3.3	Texture Mapping	24
3.3.1	Alpha blending	25
3.3.2	Transfer Function and Classification	25
3.4	Limitations	26
4	Stereoscopic Volume Rendering	27
4.1	Stereoscopic Visualization Hardware	27
4.2	Generating Stereo Pairs	30
4.3	Rendering Stereo Pairs	31
4.4	Stereo Viewing in Medical Imaging	32
4.4.1	Stereo Imaging in Mammography	33
4.4.2	Laparoscopy Imaging	33
4.4.3	MIP and X-ray Rendering	34
4.4.4	Transparency in Stereo	34
4.4.5	3D Volume Rendering of CT Data	35
5	Experiments	36
5.1	Hypotheses	36

5.1.1	Stereopsis of Absorptive Media	36
5.1.2	Aerial Perspective with Absorptive Media	38
5.2	Method	40
5.2.1	Subjects	40
5.2.2	Apparatus	41
5.2.3	Design Issues: Ghosting and Flicker	41
5.2.4	Stimuli	42
5.2.5	The Illumination Model	44
5.3	Experiment 1: Stereopsis with DRRs	45
5.3.1	Results and Discussion	46
5.4	Experiment 2: Simulated Aerial Perspective in DRRs	52
5.4.1	Results and Discussion	54
5.5	Measuring Acetabular Coverage using Stereo DRRs	55
5.5.1	Stimuli and Measurements	56
5.5.2	Results and Discussion	58
5.6	Summary of Findings	59
6	Conclusion	61
6.1	Future Work	62
	Bibliography	65
	A Consent Form	72
	B Perlin Noise	74

List of Tables

5.1	Measurements as percentage covered of each quadrant region and as a total measured from pre-operative CTs for each method. The four quadrants are anterolateral (AL), posterolateral (PL), anteromedial (AM), posteromedial (PM).	58
5.2	Measurements as percentage covered of each quadrant region and as a total measured from post-operative CTs for each method.	59
6.1	Summary of Results	62

List of Figures

1.1	Digitally Reconstructed Radiograph	3
2.1	Stereopsis	10
2.2	Stereopsis	13
2.3	Motion Parallax	14
3.1	Volume Data	16
3.2	Computed Tomography	17
3.3	Participating Medium Optics	17
3.4	Volume Rendering	17
3.5	Ray Tracing	22
3.6	Trilinear Interpolation	23
3.7	Phong Illumination Components	24
4.1	Active Stereo	29
4.2	Toe-in Method	30
4.3	Off-axis Method	31
5.1	Stereopsis of Absorptive Media	37
5.2	Aerial Perspective in Absorptive Media	39

5.3	The Stimulus	42
5.4	The Cylinder Model	46
5.5	Classification Correctness as a Function of Opacity	47
5.6	Classification Correctness as a Function of Spatial Frequencies	48
5.7	Monocular Classification Correctness as a Function of Opacity	49
5.8	Stereoscopic Classification Correctness as a Function of Opacity	49
5.9	Monocular Classification Correctness as a Function of Persistence	50
5.10	Stereoscopic Classification Correctness as a Function of Persistence	51
5.11	Mean Response Time as a Function of Opacity	52
5.12	Mean Response Time as a Function of Spatial Frequency	53
5.13	Classification Correctness as a Function of Contrast	55
5.14	Outline Picking of Acetabulum and Femoral Head	56
5.15	Quadrant Divisions with Coverage Boundaries	57

Chapter 1

Introduction

Advances in commodity computer and graphics hardware, software, and display devices have made possible the rapid and effective visualization of volumetric data by means of volume rendering. Volume rendering is a direct technique of reconstructing a volumetric data set, obtained through sampling, simulation or modelling techniques, into a three-dimensional model. Since its introduction in the 1980s [26], volume rendering has gained widespread acceptance in the medical community, where large data sets are collected by means of Computed Tomography (CT), Magnetic Resonance Imaging (MRI), or Positron Emission Tomography (PET). The data collected are reconstructed as three-dimensional models that can be rendered and used for visualization of anatomical entities, for diagnostic purposes, in surgical planning, and for surgical guidance.

Perceptual cues used by humans in every day life to perceive depth are very well understood. However, what cues and the way in which these cues should be combined in order to convey depth in computer-generated images remains an ongoing topic of research [49]. With standard volume-rendering techniques, it is difficult to understand the three-dimensional structure of a volume. Adding perceptual cues to a three-dimensional

volume may enhance a user's ability to understand the structure of the volume. In this dissertation, we attempt to bring together the fields of computer graphics, medical image visualization, and psychophysics by performing a series of experiments to study the effect of two perceptual cues, stereopsis and aerial perspective, on volume-rendered data.

1.1 Digitally Reconstructed Radiographs

We focus on purely absorptive media. In a purely absorptive medium, light passes through the medium with no reflection or scattering. There are no solid surfaces. This makes it especially difficult to understand the structure of the volume [1]. Images from the X-ray domain not only have no surfaces, they also have no depth cues for observers to use, requiring them to rely on *a priori* knowledge. To study the effects of different perceptual cues in purely absorptive media we look at digitally reconstructed radiographs (DRRs).

DRRs are synthetically computed X-ray images that are calculated by integrating the attenuation of light as it passes through a volume toward the viewer (Figure 1.1). Each pixel value of a DRR image is a function of the CT values encountered along the projection rays. DRRs are of particular interest because they accurately simulate plain radiographs [37] and, therefore, are used in many medical applications. They provide information, such as internal fracture lines and intra-articular features, that is not visible via surface rendering. In intensity-based 2D-to-3D intra-operative registration, DRRs are computed from many different viewpoints to find one that most closely matches a fluoroscopic image [25]. For radiotherapy treatment, DRRs provide reference images to assist in patient positioning [47]. In planning of orthopaedic surgery, DRRs can show intra-articular features not visible in a surface-rendered CT image [32].



Figure 1.1: A digitally reconstructed radiograph (DRR) is a synthetically computed X-ray image.

1.2 Stereopsis

Perceptual aspects of stereopsis, or the idea that slightly different images are projected to each retina, have been recognized since the 4th century B.C. [40]. It was only in 1838, however, that Wheatstone showed that the horizontal disparity between left and right eye images was sufficient to produce the perception of depth [9]. Stereopsis is the impression of depth created by comparing the two different retinal images for horizontal disparity.

As early as 1898, stereoscopic X-ray imaging techniques were studied for illustrating medical and scientific work [6]. The extra time and costs of producing stereoscopic pairs, the inconsistent quality of stereo display devices, and the complex stereoscopic mounting

process all contributed to the decline of interest in stereoscopic X-ray techniques [13]. Recent advances in computer and graphics technology have made possible interactive and realistic high-quality 3D rendering of large three-dimensional data, allowing the renewed study of stereoscopic 3D rendering.

With purely absorptive media, it is not obvious that the stereopsis cue is sufficient to give three-dimensional perception. In order to use disparity information, the visual system must match points from one retinal image with similar points from the other retinal image [9]. In surface-rendered data, the left and right eye see the same feature of an image horizontally displaced on the retinae. In purely absorptive media, however, the eye can focus on a point within a volume and see a different image of that point on each retina. The difference in the images occurs because the two different light rays through same focal point towards each eye are integrated along different paths as they pass through the volume. These different integrals make it more difficult to match points of one retinal image to points of the other retinal image. This might make the stereopsis cue less effective in purely absorptive media. The primary research goal of this work was to answer the question: *Does stereopsis affect depth perception in purely absorptive media?*

1.3 Aerial Perspective

Aerial perspective is the perception of depth due to scattering of light in the atmosphere [38]. Light reflected on a near object does not scatter as much before reaching the eye as that of light reflected of a distant object. The amount of scattering of a distant object causes it to appear hazy and lighter, reducing the contrast between distant objects. In purely absorptive media, there is no scattering, so this cue is not present. We can reduce the contrast of more

distant parts of a volume to simulate aerial perspective. This contrast reduction may increase accuracy of relative depth perception. The secondary research goal of this work was to answer the question: *Does aerial perspective affect depth perception in purely absorptive media?*

1.4 Contributions

In the course of conducting the research, using specific hypotheses and experimental procedures, the following was determined:

- 1: Stereopsis improved relative depth perception in purely absorptive media, although it did not provide unambiguous information. Neither spatial frequency, nor opacity, had an effect on improving relative depth perception.
- 2: Simulated aerial perspective improved relative depth perception in purely absorptive media. The effectiveness of reducing contrast of distant features — by only a small amount — matched that of stereopsis as a depth cue.
- 3: Use of stereoscopic DRRs was effective in measuring acetabular coverage (i.e., how well the hip socket covers the femoral head). The study showed that stereoscopic viewing of DRRs to measure acetabular coverage gave similar results to a currently used method, whereas monocular viewing did not.

1.5 Organization of the Dissertation

We begin by giving a background of human perception in Chapter 2. The different depth cues used by the human visual system are presented and described.

In Chapter 3, we begin with the theoretical foundations for volume rendering. The components of the optical model used in volume rendering are discussed and the volume rendering integral is developed. In the second section of the chapter, we discuss volume rendering in practice. Ray casting, the gold standard for volume rendering, is presented, followed by an overview of volume rendering using texture-mapping hardware. In Chapter 4, we present current techniques for stereoscopic viewing on commodity hardware and discuss a number of medical studies which use stereoscopic viewing.

Chapter 5 presents the psychophysical experiments we conducted. The hypotheses are described, the methodology is presented, and the results are analyzed for statistical significance. These experiments studied the effectiveness of using stereopsis and simulated aerial perspective to perceive depth in absorptive media images.

Chapter 6 concludes the dissertation and outlines some potential avenues for future work.

Chapter 2

Perception

In this chapter, we present the cues used by the human visual system to perceive depth. When viewing a 3D scene, the scene is projected onto the retinae as 2D images. From these 2D images, we reconstruct a 3D world in which we perceive and experience depth [11]. To get a unique 3D interpretation of the visual input, a number of visual cues — relative size, linear perspective, occlusion, shading, motion parallax and others — are used.

2.1 Physiological Cues

Physiological cues are those that depend on the anatomy of the eye. Tension of the muscles around the eye and the shape of the eye lens changes with the relative distance of objects that we observe.

Accommodation

Accommodation, physiologically, is a message sent to the brain about the tension of muscle that changes the focal length of the lens of the eye [22]. This change in lens length allows

us to focus on objects at different distances. Accommodation is a weak depth cue which is only effective with other cues and at short viewing distances of zero to two meters [9].

Convergence

Convergence is the movement of the eyes to point slightly inwards (i.e., to converge) when watching a close object and to point straight ahead when looking at an object farther away [22]. The angle of convergence is used to estimate distance. Convergence is only effective for distances of less than two meters [9].

2.2 Binocular Disparity

Binocular disparity is the horizontal difference between retinal images in the left and right eye [8]. The slight offset between our left and right eyes (about 6 cm) creates two slightly different images of the world; points that lie at different depths in a 3D scene are separated by different distances in the left and right eyes' retinal images [40]. Disparity allows us to determine whether an object is located in front or behind our fixation point [38]. If an object is farther away than our fixation point we have to uncross, or diverge, our eyes to look at the farther object; hence, the retinal disparity is *uncrossed*. If an object is located closer than our fixation point, we have to cross our eyes to look at it; therefore the retinal disparity is *crossed*.

The correspondence between the images formed on the left and right retinae vary with the relative position of objects [38]. For objects on the horopter there is a non-disparate correspondence (i.e., objects fall on the same part of the retina in each eye). The *horopter* is the surface at which we have non-disparate correspondence [40]. For an object that is closer than the horopter (crossed disparity) or for an object that is further than the horopter

(uncrossed disparity), the two images of the object fall on non-corresponding parts of the retina in each eye [11].

Stereopsis

Stereopsis is the impression of depth created by comparing the two different retinal images for horizontal displacement of corresponding parts. This is the most powerful depth cue which, in the absence of all other depth cues, is sufficient to give us depth information [20] and which, when used in combination with other depth cues, can provide absolute distance information [9].

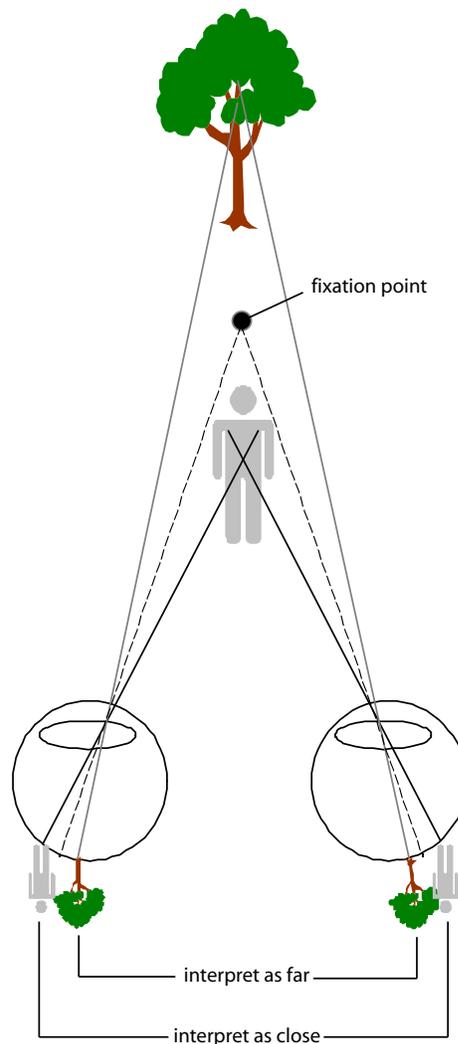


Figure 2.1: Stereopsis is the comparison of the two different retinal images for horizontal displacement of corresponding parts to judge the relative distances of objects. When the eye is fixed on the point, there is an uncrossed disparity for the distant object and a crossed disparity for the close object. The disparity between the image elements on each retina allows us to perceive the depth of different image elements.

The Correspondence Problem

In order to use disparity information, the visual system must match points from one image with similar points from the other image [9]. The problem of comparing two images for

matching points is termed the correspondence problem. To date, there is no one accepted theory as to how the visual system solves the correspondence problem.

2.3 Psychological Cues

Both physiological and psychological cues are used by the human visual system to interpret depth in a 3D scene. All psychological cues are monocular, so they are available when looking at images with just one eye.

Occlusion

Occlusion occurs when one object partially obscures another in view. The object that is obscuring the other is perceived to be closer [38]. This depth cue provides no information about the absolute distance of an object but, rather, information about whether one object is closer than another.

Shading

Shading is a dark region in the view where light cannot fall because light rays are blocked by an obscuring object. If the location of a light source is known, an object that casts a shadow on another object is perceived to be closer to the light source [38]. When illumination is coming from above, shading information is used to resolve ambiguities in the relative distances of objects [8].

Retinal Image Size

Retinal image size is used by the brain to gather information about the distance of an object by comparing the object's known real size to the sensed size of the object [38]. If an object

has a larger retinal image (i.e., if it takes up more of the visual field) it is perceived to be closer. For this cue to be reliable, a viewer must be familiar with the actual size of the objects involved. Otherwise, an object with a larger retinal image may be perceived as closer when it is actually farther away but is physically very large [8].

Linear Perspective

Linear perspective is the convergence of parallel lines as they recede into the distance. Parallel lines appear, with distance, to merge together to a single vanishing point on the horizon [8].

Texture Gradient

Texture gradient is a surface pattern which provides information about the distance, depth and shape of an object. As a surface gets further away, the fineness of detail of the surface pattern decreases and the surface appears smoother [8]. Thus, smoother objects tend to be perceived as being farther away. Conversely, objects with more detailed textures, are perceived as being closer.

Aerial Perspective

Aerial perspective is the cue caused by scattering of light in the atmosphere, which makes distant objects, such as mountains, look bluish and hazy [38], as in Figure 2.2. Scattering is the reflection of light in all directions by small water and dust particles in the atmosphere. The farther away an object is the more light is scattered between the object and the observer, making it appear hazy. A distant object will also appear bluish because blue wavelengths are scattered more easily by molecules in the atmosphere.



Figure 2.2: Scattering of light in the atmosphere makes distant objects, such as the mountains in this picture, appear hazy in comparison to near objects, such as the skier.

Motion Parallax

There is retinal disparity between the 3D scene projected on the right eye retina and that on the left eye retina. The same is true for an image on either retina when the head is moved several centimeters to the left or right [8]. Motion parallax is the apparent relative motion of objects as a viewer moves his head [8]. It provides a viewer with the ability to distinguish the depth of objects due to movement.

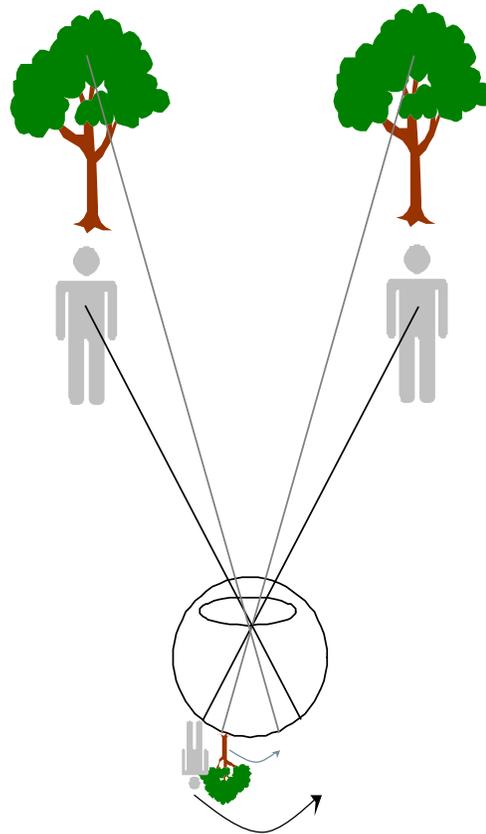


Figure 2.3: The two objects, person and tree, lie at different distances. If the viewer moves his head or, conversely, the objects move at equal speed, the image of the close object moves further across the retina than does the image of the distant object allowing a viewer to perceive the relative depths of objects. If a single eye moves laterally by about 6cm (the inter-ocular distance) the changes in retinal image between the start and finish of the movement will be the same as those in each eye when viewed simultaneously.

Chapter 3

Volume Rendering

3.1 Volume Rendering: Theory

A volume consists of a three-dimensional array of *voxels* (Figure 3.1). Analogous to pixels, which are picture elements that surround a point in a two-dimensional space, voxels are volumetric units of graphic information that surround a point in three-dimensional space. Voxels are typically small rectilinear volumes.

A spatial scalar field is represented as $f(x, y, z)$ and defined over a subset of \mathbb{R}^3

$$f : \mathbb{R}^3 \rightarrow \mathbb{R} \tag{3.1}$$

Data from this field may be obtained through discrete sampling, simulation, or modelling techniques. The most popular application of volume rendering has been medical imaging [43] in which large, discretely sampled, data sets are collected by means of Computed Tomography (CT), Magnetic Resonance Imaging (MRI) or Positron Emission Tomography (PET). In CT, for example, $f(x, y, z)$ is the X-ray absorption coefficient at position (x, y, z) , measured in Hounsfield units [10]. CT acquisition is depicted in Figure 3.2.

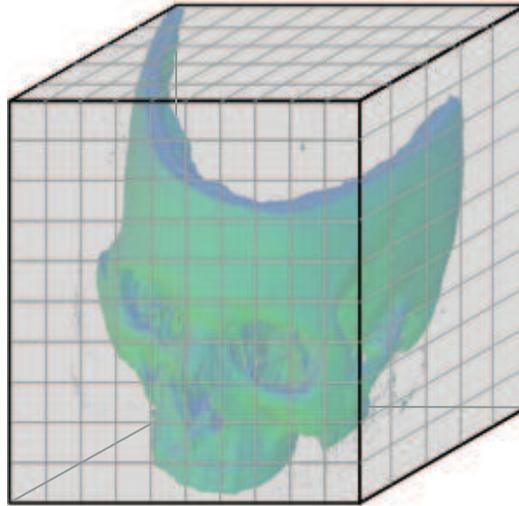


Figure 3.1: A human skull: volumetric data consists of a three–dimensional array of voxels.

There are two methods for visualizing volumetric datasets: indirect volume rendering (IVR), where the data is first converted into a set of polygonal isosurfaces [27], and direct volume rendering (DVR) where the data is directly rendered without any intermediate steps [31]. We discuss the latter because this is the method we used to render DRRs.

3.1.1 Optics of a Participating Medium

The *participating medium* is the material that affects the transport of light through its volume. Direct volume rendering algorithms use an optical model that take into account the physical way in which light rays interact with the medium. The light is absorbed, emitted and scattered [29] (Figure 3.3).

A volume is modelled as a particle-filled slab of width L , through which a ray, r , is cast in a direction of the viewer. Figure 3.4 depicts this process.

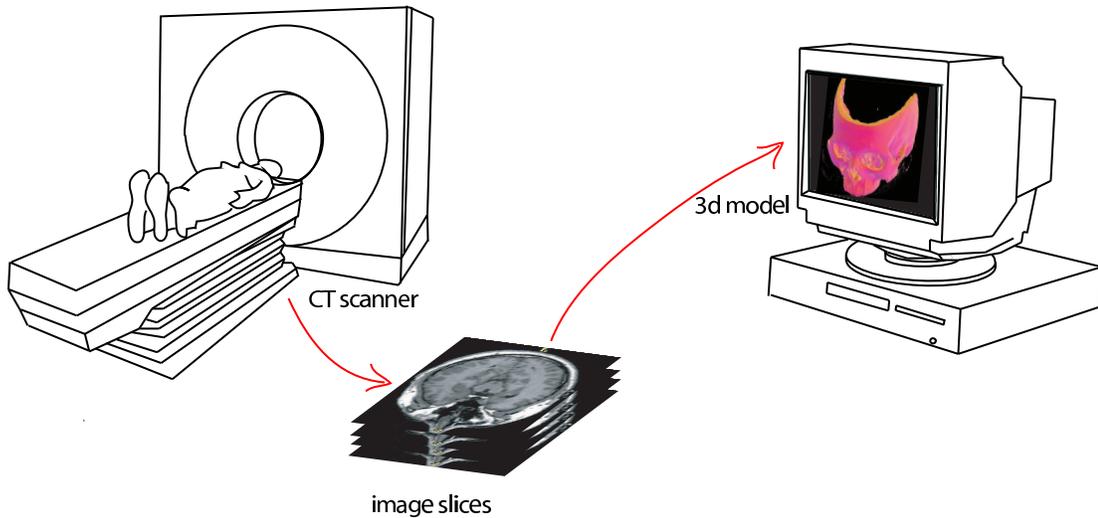


Figure 3.2: Computed Tomography (CT) scanners are a typical and widely known source of voxel data.

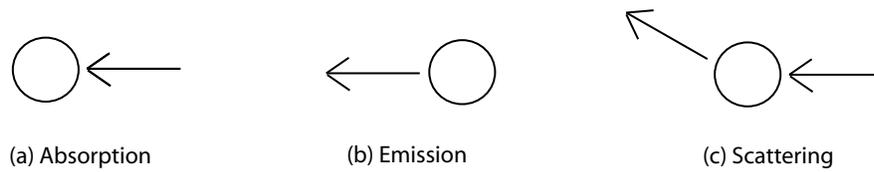


Figure 3.3: Direct volume rendering takes into account the physical way in which light is (a) absorbed, (b) emitted and (c) scattered by a participating medium.

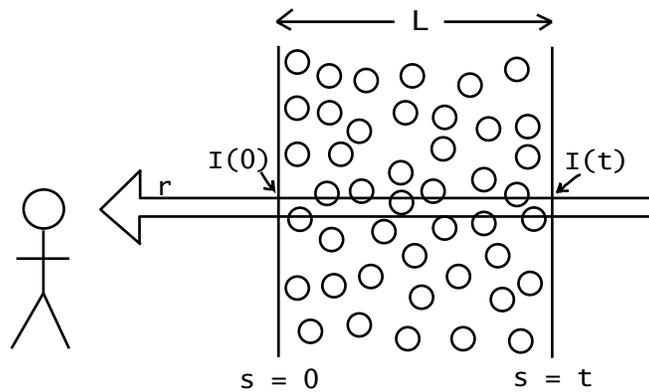


Figure 3.4: A slab of width L of a participating medium. A ray r is cast through the medium and $r(s)$ is a point along the ray.

Absorption

Light travelling along ray \mathbf{r} will be *absorbed* by particles that it hits. Let $\tau(s)$ be the probability density (i.e., the probability per unit distance) of light being absorbed at location s . $\tau(s)$ is proportional to the particle density at $\mathbf{r}(s)$. A fraction $\tau(s)\Delta s$ of $I(s)$ is absorbed over a distance Δs at location s :

$$\Delta I = -\tau(s) \Delta s I(s) \quad (3.2)$$

where $-\Delta I$ is the amount of light absorbed.

An expression for the light intensity leaving the volume at position t is readily derived considering the corresponding expression using infinitesimal quantities dI and ds in place of ΔI and Δs .¹

$$\begin{aligned} -d I &= \tau(s) ds I(s) \\ \frac{d I}{I(s)} &= -\tau(s) ds \\ \int_0^t \frac{d I}{I(s)} &= \int_0^t -\tau(s) ds \\ \ln I(t) - \ln I(0) &= -\int_0^t \tau(s) ds \\ \ln \frac{I(t)}{I(0)} &= -\int_0^t \tau(s) ds \\ \frac{I(t)}{I(0)} &= e^{-\int_0^t \tau(s) ds} \\ I(t) &= I(0) e^{-\int_0^t \tau(s) ds} \end{aligned} \quad (3.3)$$

¹Derivations are adapted from James Stewart's CISC 454 Computer Graphics class notes, Queen's University, 2004

Emission

Particles can *emit* light. We model the emission of light at position $\mathbf{r}(s)$ as

$$C(s) \tau(s) ds \quad (3.4)$$

where $C(s)$ is the intensity of light per particle. Because $\tau(s)$ is proportional to the particle density per unit length, $C(s) \tau(s) ds$ is the [infinitesimal] intensity emitted at $r(s)$.

Scattering

Particles can *scatter* light by deflecting incident light. Scattering is typically not taken into account in volume rendering because it has the effect of blurring the rendered image, despite the fact that plain X-ray radiographs can exhibit substantial scattering artifacts.

3.1.2 Volume Rendering Integral

In volume rendering, the scalar field of the volume is mapped to the optical properties of colour and opacity. This is done using a transfer function that maps $f(x)$ onto colour, $C(x)$, and opacity, $\alpha(x)$. Colour and opacity are integrated along viewing rays, which are cast through the volume from the viewpoint.

Only the emission and absorption of light are taken into account when we determine the the final intensity $I(0)$. We sample along the ray at a fixed interval step size, $\Delta(s)$. At each point, $\mathbf{r}(s)$, along the ray the contribution to $I(0)$ is

$$C(s) \tau(s) ds e^{-\int_0^s \tau(u) du} . \quad (3.5)$$

The total contribution of all points is then

$$I(0) = \int_0^t C(s) \tau(s) ds e^{-\int_0^s \tau(u) du} . \quad (3.6)$$

We approximate $I(0)$ by discretizing the ray into n pieces, each of length $\Delta(s)$:

$$I(0) \approx \sum_{i=0}^{n-1} C(s_i) \tau(s_i) \Delta s \prod_{j=0}^{i-1} e^{-\int_{s_j}^{s_{j+1}} \tau(u) d(u)}$$

where $s_i = i \Delta s$.

We define opacity as $\alpha_i = \tau(s_i) \Delta(s)$. If we assume that $\tau(s)$ is constant over interval $\Delta(s)$, and simplify the exponential by approximating it with the first two terms of its Taylor expansion ($e^x = 1 + x + \mathcal{O}(x^2)$), we get

$$\begin{aligned} I(0) &\approx \sum_{i=0}^{n-1} C(s_i) \alpha_i \prod_{j=0}^{i-1} e^{-\alpha_j} \\ &= C(s_i) \alpha_i \prod_{j=0}^{i-1} (1 - \alpha_j) \end{aligned} \tag{3.7}$$

Equation 3.7 is known as the discrete volume rendering integral.

$$\begin{aligned}
 I(0) &\approx C_0 \alpha_0 \\
 &+ C_1 \alpha_1 (1 - \alpha_0) \\
 &+ C_2 \alpha_2 (1 - \alpha_0)(1 - \alpha_1) \\
 &+ C_3 \alpha_3 (1 - \alpha_0)(1 - \alpha_1)(1 - \alpha_2) \\
 &+ \dots \\
 &= C_0 \alpha_0 \\
 &\quad + (1 - \alpha_0) (C_1 \alpha_1 \\
 &\quad \quad \quad + (1 - \alpha_1) (C_2 \alpha_2 \\
 &\quad \quad \quad \quad \quad + (1 - \alpha_2) C_3 \alpha_3 \\
 &\quad \quad \quad \quad \quad \quad \quad \quad + \dots) \\
 &= C_0 \alpha_0 \\
 &\quad + (1 - \alpha_0) (C_1 \alpha_1 \\
 &\quad \quad \quad + (1 - \alpha_1) (C_2 \alpha_2 \\
 &\quad \quad \quad \quad \quad + (1 - \alpha_2) \\
 &\quad \quad \quad \quad \quad \quad \quad \quad + \dots)
 \end{aligned}$$

Given I_n , the intensity of light arriving from behind the volume position s_i , I_{i-1} is evaluated by computing the inside term first:

$$I_{i-1} = C_i \alpha_i + (1 - \alpha_i) I_i \quad (3.8)$$

This equation suggests a method of evaluating the discrete VRI: Set I_n , evaluate I_{n-1} , evaluate I_{n-2} , and so on to I_0 . This is known as “back-to-front” evaluation. I_n is the intensity of light arriving from behind the volume. With a pure DRR, C_i is zero.

3.2 Volume Rendering in Practice

We discuss ray casting and texture mapping, the two principal methods of volume rendering.

3.2.1 Ray Casting

Ray casting [26] is a DVR method in which a straightforward numerical calculation of the VRI (Equation 3.6) is done. A ray, r , is cast from the eye, through each pixel, $x_{i,j}$, on the screen, and through the volume, f . The volume is represented as a matrix of scalar values, $f_{i,j,k}$. Samples are taken along the rays at intervals of size Δs and composited, as shown in Figure 3.5. For each of the samples, $r(s)$, along the ray, we find interpolated values for $f(s)$ and $\nabla f(s)$, the local gradient. The local gradient is used as the “surface normal” when we simulate the reflection of light at $r(s)$.

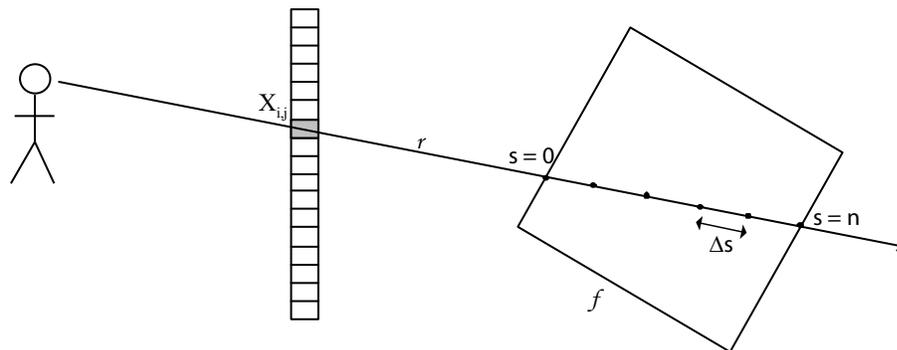


Figure 3.5: In ray tracing, a ray is cast from the view point through each pixel $x_{i,j}$ and the VRI is evaluated using evenly spaced samples, Δs , within the volume, f .

For each sampling location, trilinear interpolation of the discrete $f_{i,j,k}$ values is done to obtain a smooth field. Trilinear interpolation is the process of linearly interpolating a point within a cube given the values for each vertex of the cube. For a sample point, $r(s)$, the

nearest eight discrete neighbours, which form a unit cube in the volume data set (as in Figure 3.6), are considered. If (x, y, z) is the position and \mathbf{p}_x is the fraction (in x) of the way across the cube, the weighted sum of the densities of the eight samples is calculated as follows:

$$\begin{aligned}
 f(x, y, z) = & f_{0,0,0}(1 - \mathbf{p}_x)(1 - \mathbf{p}_y)(1 - \mathbf{p}_z) \\
 & + f_{1,0,0} \mathbf{p}_x(1 - \mathbf{p}_y)(1 - \mathbf{p}_z) \\
 & + f_{0,1,0}(1 - \mathbf{p}_x)\mathbf{p}_y(1 - \mathbf{p}_z) \\
 & + f_{0,0,1}(1 - \mathbf{p}_x)(1 - \mathbf{p}_y)\mathbf{p}_z \\
 & + f_{1,0,1} \mathbf{p}_x(1 - \mathbf{p}_y)\mathbf{p}_z \\
 & + f_{0,1,1}(1 - \mathbf{p}_x)\mathbf{p}_y\mathbf{p}_z \\
 & + f_{1,1,0} \mathbf{p}_x\mathbf{p}_y(1 - \mathbf{p}_z) \\
 & + f_{1,1,1} \mathbf{p}_x\mathbf{p}_y\mathbf{p}_z
 \end{aligned}
 \tag{3.9}$$

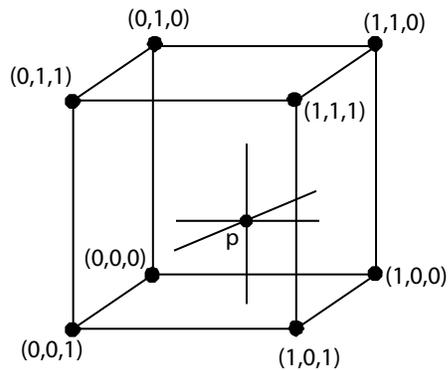


Figure 3.6: A point, \mathbf{p} , within a unit cube.

To calculate $C(x, y, z)$, the Phong illumination model [29] is used. Given L , the incoming light direction, N , the surface normal, R , the ideal reflection direction and V the viewer direction (Figure 3.7), C is calculated as

$$C = k_d N \cdot L + k_s (R \cdot V)^n \quad (3.10)$$

where k_d is the diffuse reflection coefficient (i.e., the RGB surface colour) and k_s is the specular reflection coefficient (i.e., the RGB light colour).

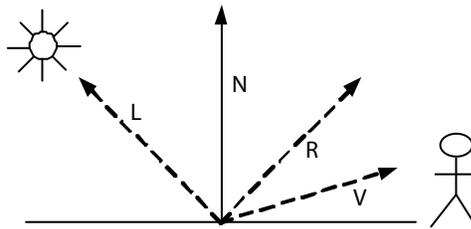


Figure 3.7: The Phong illumination model calculates the colour, $C(x, y, z)$, given L , the incoming light direction, N , the surface normal, R , the ideal reflection direction and V , the viewer direction.

3.3 Texture Mapping

With the new programmability offered by graphics hardware, a number of optimized volume graphics algorithms have been developed in the last few years [7]. By allowing the graphics hardware to do the work of blending, texture mapping, and trilinear interpolation, volume rendering speed is greatly accelerated.

The volume data is stored on the graphics card. A volume is stored in a single 3D texture, where a single *texel*, or texture element, corresponds to a single voxel. Given a 3D texture representing a 3D data set, graphics hardware allows slices through the dataset to be drawn onto polygons spaced Δs apart. These polygons are composited to form a

final image. The texture-mapping hardware supports trilinear interpolation to map the 3D texture to each of the sliced polygons. Blending hardware composites the slices back-to-front, as required by Equation 3.7.

3.3.1 Alpha blending

Catmull and Smith [39] developed the notion of *integral alpha*, which suggests that transparency, or conversely, opacity, are as fundamental as colour and should be made part of an image. In order to do this, they used a fourth channel for opacity (called “alpha”, or A) that is added to the three colour channels to form the RGBA standard. This was also the origin of the term “alpha blending”. Alpha blending allows us to create the effect of transparency by combining a translucent foreground colour, C_F , with a background colour, C_B as:

$$C = \alpha C_F + (1 - \alpha) C_B \quad (3.11)$$

As seen in Equation 3.8, alpha blending is used in ray casting. When rays are cast through pixels into the volume, the voxel colours are composited from back to front using alpha blending. In texture mapping, alpha blending is used to composite the 2D polygon slices that are taken through the volume.

3.3.2 Transfer Function and Classification

Mapping the scalar data of a volume, $f(x, y, z)$, to optical properties is achieved via a transfer function. The application of a transfer function is also known as classification because it classifies voxels according to the transfer map. A transfer function is simply a mapping from scalar volume values to an optical model consisting of surface colour (RGB) and opacity (α).

A transfer function can be used to isolate or enhance parts of a volume. For example, for CT data we can isolate bone by mapping low density material (such as skin and soft tissue) to be transparent and conversely high density material (such as bone and vascular calcifications) to be opaque.

3.4 Limitations

A main drawback of volume rendering techniques is that they do not allow a viewer to easily perceive the three-dimensional structure of the volume. This is especially true when the volume does not have solid surfaces [1]. One solution is to use motion parallax to relay the depth information; showing the volume in motion (perhaps by rotating) provides a motion parallax cue, which allows the viewer to perceive depth. Another solution is to generate a stereoscopic pair to exploit the binocular disparity cue [1]. In the next chapter, we discuss stereoscopic volume rendering on commodity graphics hardware, which can help to enhance perception in volumes.

Chapter 4

Stereoscopic Volume Rendering

In Chapter 2, stereopsis was presented as one of the strongest cues for depth perception. This cue is exploited in computer graphics by using hardware (i.e. graphics card and stereo glasses) to show each eye a slightly different image of the scene. A viewer with functioning stereoscopic sight can merge these two images to reconstruct a 3D scene.

4.1 Stereoscopic Visualization Hardware

There are two main types of stereo hardware: passive and active. Passive stereo hardware includes anaglyphic stereo, horizontal or vertical split-screen stereo, and polarized-light stereo [48]. For anaglyphic stereo, the red and blue channels of an image are split. The left eye image is typically projected as red and the right eye image as blue. The viewer wears 3D glasses in which the left lens is red (blocking the right-eye data) and the right lens is blue (blocking the left-eye data); the channels are reassembled by the brain so that the image appears to be a 3D image. In the polarized-light method, the left and right eye images are projected onto a single screen through filters that orthogonally polarize the two

views. The viewer wears glasses with polarized lenses that match the polarization of the projected images. In split-screen stereo, two images are projected side-by-side and the viewer must cross or defocus the eyes in order to get a perception of depth.

Special stereo graphics cards support *active stereo*. Active stereo hardware uses LCD shutter glasses that are connected to the graphics card. The glasses are synchronized with the graphics card so that, when rendering the left image, the right lens of the glasses is made opaque, as shown in Figure 4.1. Similarly, when the right eye image is rendered, the left lens of the glass is made opaque. By allowing each eye to see only the image intended for it, each eye receives the correct perspective and the scene appears to have true depth [48]. Each rendering is updated at half the monitor refresh rate, so for optimum interactivity a high screen refresh rate should be used.

Typical rendering uses double-buffering, where the contents of the front buffer are visible to the viewer, whereas those of the back buffer are not. The successive frames are drawn into the back buffer and to show the next frame, the front and back buffer are swapped. To support stereo display, stereo enabled graphics cards allow drawing into left and right drawing buffers, each of which has a front and back buffer. Because of the four buffers, the cards are often described as quad-buffered stereo cards [48]. Quad-buffered stereo, as opposed to “consumer stereo”, allows us to develop stereo applications using four buffers with double-buffered left and right views that avoid flickering and other artifacts associated with single-buffered rendering techniques.

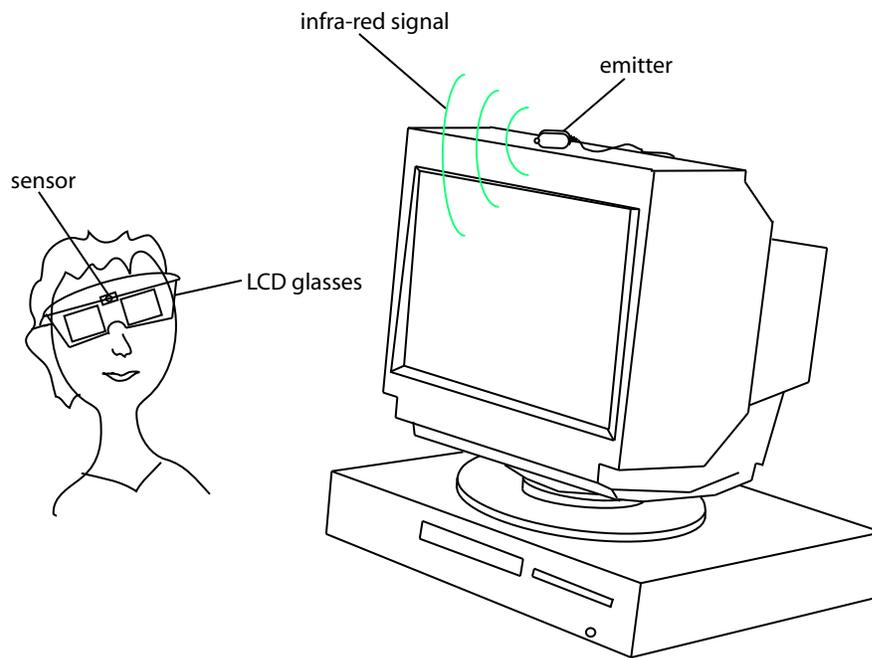


Figure 4.1: Active stereo uses LCD shutter glasses which are connected to the graphics card. The glasses are synchronized with the graphics card so that when rendering the left image the graphics card sends a signal through an infrared connector to the glasses to make the right lens of the glasses opaque.

4.2 Generating Stereo Pairs

When the left and right eye images are created, they must be projected onto an image plane. There are two ways of doing this: toe-in and off-axis. With the *toe-in* method, each camera is pointed at a common focal point as shown in Figure 4.2. The right and left viewpoints are separated but are aimed inward toward a single focal point. Conversely, the *off-axis* method shown in Figure 4.3 has left and right cameras with different projection parameters. In the off-axis method, parallel fields of vision are used to define a vertical projection plane. Each of the views is aimed directly forward and the intersection of the viewpoints forms a virtual projection screen.

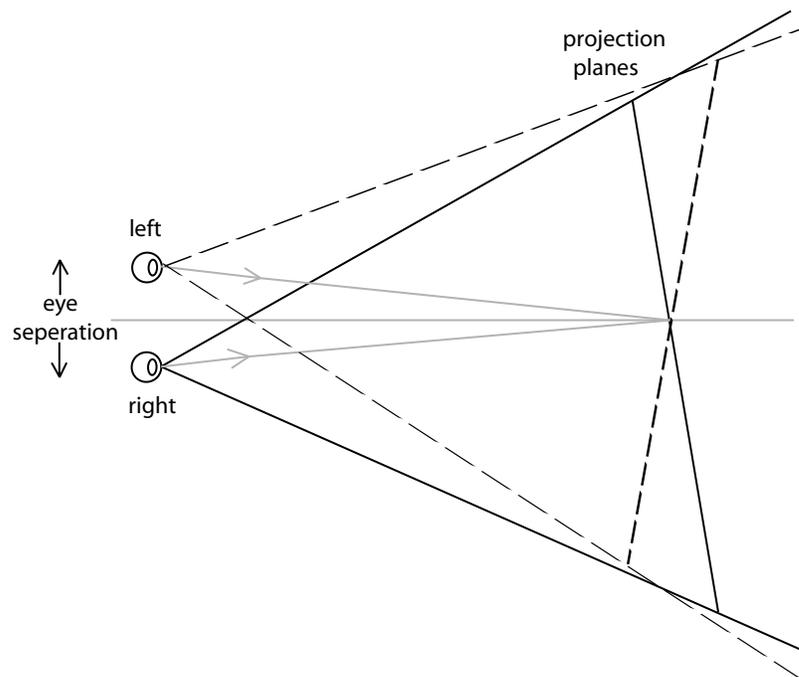


Figure 4.2: With the toe-in method, each camera is pointed at a common focal point.

The disadvantage of the toe-in method is that it may distort images around the edge

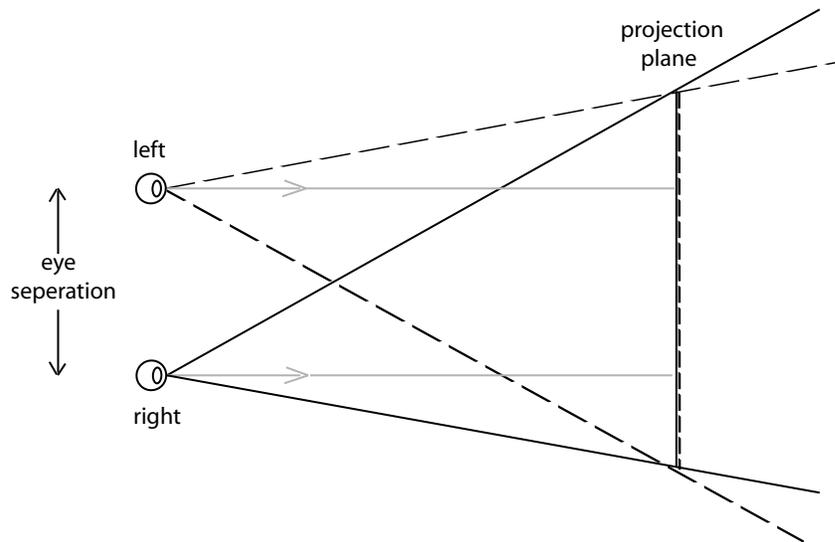


Figure 4.3: With the off-axis method, the left and right cameras have different projection parameters.

of the field of vision because it introduces vertical parallax [3, 48]. The off-axis method however, does not work well with small objects because they become too separated for the brain to merge, resulting in a distorted image.

4.3 Rendering Stereo Pairs

Bethune [2] developed a volume rendering method which increased interactive rendering rates. The speed-up was obtained by separating the volumetric data set into regions of empty and non-empty voxels. Those voxels which were non-empty were contained within Axis Aligned Bounding Boxes (AABBs). During rendering, view-aligned slices were drawn back-to-front and only those parts of the slices which were within an AABB were rendered.

The slices were rendered back-to-front, computing the intersection of each slice with

the individual boxes from the set of AABBs. The Adaptive Slice volume rendering algorithm not only yielded considerable overall performance increase, but also was nearly optimal in eliminating the rendering of empty regions in the volume data. The performance gains were due to factors such as the percentage of empty voxels within a data set, the distribution of the data, and the maximum number of AABBs. Testing these different factors showed an increase in performance. Frame rates of up to a factor of ten were achieved with no degradation of image quality.

We extended the adaptive slice rendering algorithm, which takes advantage of texture-mapping hardware, to allow stereoscopic rendering. This algorithm was used in our experiments.

4.4 Stereo Viewing in Medical Imaging

Numerous studies [15, 16, 17, 28, 42, 49] have examined the use of psychological depth cues in computer-generated images to enhance depth perception. Of these, some [15, 16, 17, 49] have shown that stereoscopic rendering can aid in perceiving depth, grasping, recognizing, or understanding the shape of computer-generated objects. Other studies, however, have shown that the benefits of stereoscopic viewing are task dependent [45] and that, for certain tasks there is no benefit to using stereopsis [41, 46].

In the medical community, the use of stereoscopic viewing seems promising [5, 35], particularly for volume rendering. In the next section, we review some of the results of studies of stereoscopic volume rendering for medical visualization.

4.4.1 Stereo Imaging in Mammography

Radiologists have few depth cues when working with X-ray images. One example is in the detection of abnormalities in mammograms. Mammograms, the standard imaging modality for screening and diagnosis of breast cancer, have no depth cues so it can be difficult to detect abnormalities in soft tissue. A study by Hsu *et al.* [13, 14] looked at the effectiveness of stereoscopic imaging in the detection of abnormalities, such as tumors, in computer-generated mammograms. Results showed that stereoscopic mammography aided in the detection of certain tissue abnormalities. This can lead to early detection and better diagnosis of breast cancer.

4.4.2 Laparoscopy Imaging

Laparoscopy is a minimally invasive surgery in which surgeons work with 2D video pictures provided on a monitor. In traditional laparoscopic systems, one of the major limitations is the loss of depth perception as the surgeons work from 2D images.

Studies have provided a mixed picture of the usefulness of stereoscopic imaging for laparoscopy, a surgical technique used to diagnose and treat a range of abdominal or pelvic problems. It is known that, because of the need for video imaging, there are performance limits to traditional laparoscopic systems. Thus, there is a need to develop some mechanism to improve depth perception [41]. It is debatable whether a stereoscopic laparoscopic system can provide a significant advantage over monocular systems. Some studies [12, 41] have demonstrated no statistical advantage, whereas others [4, 19] showed a significant improvement in speed and outcome of laparoscopic tasks. In monocular systems, other available cues (such as motion parallax, relative position, occlusion, perspective, and lighting) may be used to compensate for the loss of depth perception produced by the use of 2D

images [41].

4.4.3 MIP and X-ray Rendering

Mora and Ebert [33] consider several “order independent volume rendering” methods, including maximum intensity projection (MIP) and X-ray projection (DRR). MIP is a method where the maximum or brightest value seen along a viewing ray is used for a given pixel. Mora and Ebert consider whether stereoscopic rendering with these methods can provide enough information to understand the volumetric data. While no formal user studies were done, the authors experience was that stereoscopic rendering with these methods allows the viewer to better understand the volumetric data.

4.4.4 Transparency in Stereo

Kasrai *et al.* [21] propose an optic array model for the perception of multiple transparent surfaces. They perform psychophysical experiments to study transparency in stereo, transparency and spatial frequency, and multiple surface transparency. Transparency is often used in medical imaging to integrate images from different modalities. For example, 2D anatomical CT data can be overlaid with functional fMRI data in such a way that one can see through the functional image to the structural image. With volume rendering, opacity values can be assigned so that surfaces closer to the viewer are rendered as more transparently.

The Kasrai *et al.* experiments of transparency in stereo and multiple surface transparency show good agreement with their optical array model [21]. The results of their spatial frequency and transparency experiment show that under the viewing distances and display resolution tested there is no effect of spatial frequency on transparency perception.

However, more tests to explore higher frequencies are underway.

4.4.5 3D Volume Rendering of CT Data

A number of important rendering parameters (such as transfer function, opacity, brightness and percentage classification) have an effect on the appearance of a rendered image [5]. Preliminary results [5] show that both radiologists and non-radiologists prefer stereoscopic viewing of volume-rendered medical data sets to 2D displays.

Stereoscopic viewing provides a realistic representation of 3D relationships and better visualization of complex anatomy [5]. Further research, however, must be done to evaluate the efficacy of stereoscopic viewing of computer-generated images, particularly in novel medical applications. In the next chapter, we present a number of experiments performed to study whether or not stereoscopic viewing of X-ray type images aids in depth perception.

Chapter 5

Experiments

The lack of depth cues in 2D X-ray images makes it difficult for orthopaedic surgeons and radiologists to interpret depth in such images. In this chapter, we describe two experiments that studied the effects of using stereopsis and simulated aerial perspective on depth perception in DRR images. We also present a study that examined the usefulness of stereoscopic viewing in estimating acetabular coverage for patients with hip dysplasia.

5.1 Hypotheses

We discuss stereopsis and aerial perspective in absorptive media, and present our hypotheses.

5.1.1 Stereopsis of Absorptive Media

It is not obvious that the stereopsis cue is effective in purely absorptive media. With surface-rendered objects, the left and right eye see the same surface features slightly displaced. The same feature is present on both retinae and the disparity of the feature on the retinae allows

the viewer to determine the feature's depth.

In a purely absorptive medium, where there is no reflection or scattering, the eye can focus on a point within the volume and see two different images of the point. These different images occur because the retinal images are produced by the integral along two different rays that pass through the focal point toward each of the eyes, as shown in Figure 5.1.

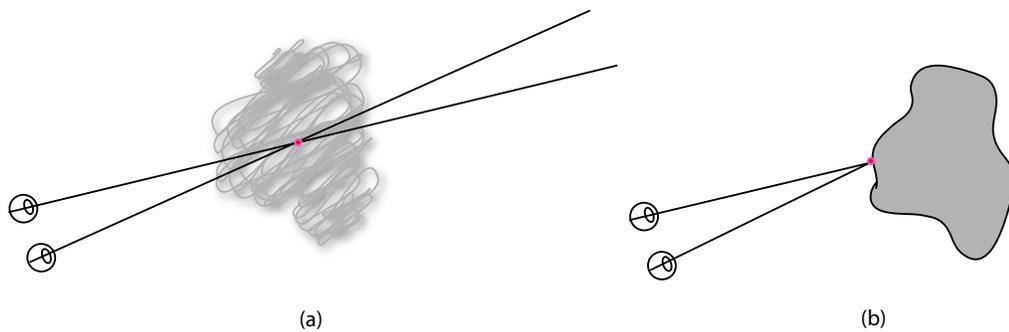


Figure 5.1: With absorptive media (left) the eyes see different images of a particular point because light from behind is attenuated differently along each path. With surface rendered objects (right) the left and right eye see the same feature.

Opacity and spatial frequency are two factors that may affect stereopsis in absorptive media. As overall opacity increases, distant features will become more obscured by close features; this may mimic the perceptual depth cue of occlusion. Low spatial frequency may make it easy to detect and track large clusters; conversely, high spatial frequency will introduce edges that may aid the brain in finding correspondences between the retinal images.

The first experiment described in this chapter explored how stereoscopic rendering helps in relative depth perception (i.e. the ability to perceive *relative* distances), and how opacity and spatial frequency affect the accuracy and speed of stereoscopic depth perception.

We consider the accuracy of depth perception to be the fraction of times a subject correctly classifies the relative depth of two surfaces. The following hypotheses were considered:

Hypothesis 1: Accuracy of depth perception in purely absorptive material is better with stereoscopic viewing than with monocular viewing.

Hypothesis 2: Accuracy of depth perception in purely absorptive material with stereoscopic viewing is not perfect.

Hypothesis 3: Accuracy of depth perception varies with opacity in purely absorptive material.

Hypothesis 4: Accuracy of depth perception varies with spatial frequency in purely absorptive material.

Hypothesis 5: The time taken to decide relative depth is less with stereoscopic than with monocular viewing in purely absorptive media.

We use the term “monocular” to refer to viewing without 3D glasses. Subjects actually used both eyes, but the terms “non–stereoscopic” and “2D” would not adequately convey the differences in viewing modes.

5.1.2 Aerial Perspective with Absorptive Media

Aerial perspective is the perception of depth due to the scattering of light in the atmosphere. Light that reflects off of a close object does not scatter much before reaching the eye, but the same light that reflects off of a distant object undergoes much scattering, causing the darker portions of the distant object to appear lighter, which reduces the contrast between

distant objects [8]. Objects that exhibit more contrast are perceived to be closer than objects that do not.

In a purely absorptive illumination model, however, there is no reflectance and no incident illumination. Rather, light is simply attenuated as it passes through the medium, as shown in Figure 5.2. In purely absorptive media, a distant difference in translucency is indistinguishable from an equal, but physically closer, difference in translucency.

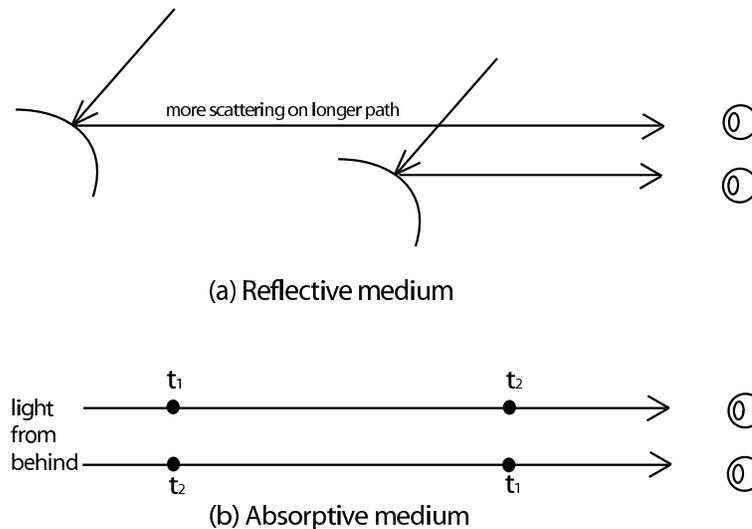


Figure 5.2: (a) In a reflective medium, incident light reflected off a close object or feature does not scatter as much as that off a more distant object. Scattering causing dark portions of the distant object to appear lighter. (b) In a purely absorptive medium, there is no incident light and the difference in adjacent translucencies (e.g. $|t_1 - t_2|$) provides no depth cue.

We simulated the aerial perspective effect in absorptive media by reducing contrast in more distant parts of the volume. Our experiment considered how contrast reduction affected relative depth perception accuracy:

Hypothesis 6: Contrast reduction with distance increases accuracy of depth perception in purely absorptive material.

5.2 Method

In this section, we detail the design of the experiment, specifically the apparatus used, stimuli tested, and the subject population.

5.2.1 Subjects

Stereoscopic acuity varies among subjects because it is dependent in part on the sensitivity of retinal receptors and the sharpness of the focus of the image on the retina [14]. Stereoscopic acuity can be improved with practice, so subjects who have experience with stereoscopic viewing tend to have better depth perception [14]. It should also be noted that about eight percent of the population cannot fuse stereoscopic pairs. There is, therefore, a wide range of stereoscopic perception ability among the population. In our experiments, all subjects had normal vision or vision corrected to nearly normal. Pre-testing showed that all of our subjects could fuse stereoscopic pairs. Different visual abilities and, specifically, stereoscopic acuity among our subjects, should be representative of depth perception in the general population.

Ethics approval for the experiments was obtained from the Queen's University Graduate Student Ethics Board and all subjects signed an informational consent waiver (see Appendix A) before participating in the study.

In all, there were twenty-one subjects. Fifteen subjects performed Experiment 1. Thirteen of the twenty-three subjects performed Experiment 2, so eight subjects performed both Experiment 1 and Experiment 2. Experiments 1 and 2 were each done in a single session of 30 to 40 minutes, and of 15 minutes, respectively. The task was explained to the subject and they were trained to perform the task on a preliminary trial.

5.2.2 Apparatus

Images were generated at greater than 15 frames per second on a 2.8GHz Pentium 4 processor. They were displayed on a ViewSonic Professional Series P95f+ CRT display and viewed by the subjects through e-Dimensional Wireless 3D LCD shutter glasses (Florida, USA). The window size was 300x410 pixels, with a pixel pitch of 0.255 mm and a refresh rate of 100Hz (50Hz per eye). The stimulus occupied 260 by 285 pixels in the center of the window. Subjects sat approximately 50 cm away from the monitor.

5.2.3 Design Issues: Ghosting and Flicker

The rapid alternation between the left and right eye images caused by the LCD shutter glasses, can lead to ghosting or flicker [14]. Flicker, which can cause uncomfortable viewing, may occur when the switching rate of the lens is too slow. In order to present flicker-free images to subjects who participated in our experiments, two things were done. First, the maximum monitor refresh frequency (100 Hz, or 50 Hz per eye) was used. Second, each experiment was conducted in a dark room with no fluorescent lights, which could have interfered with the infrared receptor of the glasses to cause flickering.

Ghosting is the perception of an object which is not present [14]. For example, because of imperfect shuttering by the glasses or excessively slow screen phosphor decay, a ghost of the right-eye image may appear when viewing with the left eye, or vice versa. The phosphor type of the monitor, the ViewSonic Professional Series P95f+, was B22 and the monitor used an enhanced phosphor treatment to induce fast decay.

5.2.4 Stimuli

In our experiments, a cylinder model was rotated about its vertical axis and was rendered on the screen. Test subjects were asked to determine the direction of rotation, that is, whether the front surface was rotating right-to-left or left-to-right. Subjects pressed one of two buttons (the left or right arrow keys on the computer keyboard) to indicate the classification. Pressing the key recorded the response and, after a time interval of 8 seconds (at which time subjects were shown a blank screen), the next stimulus was presented. Subjects were timed and asked to respond as accurately, but also as quickly, as possible.

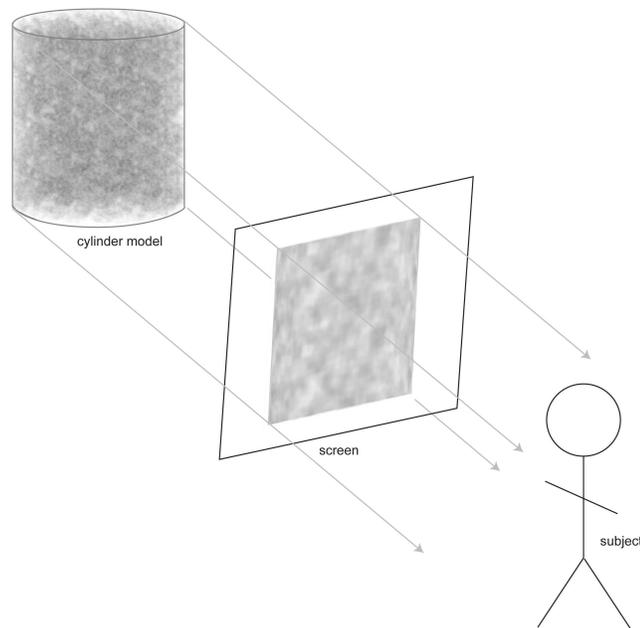


Figure 5.3: The cylinder model was orthographically projected onto the screen and subjects were asked to determine the direction of rotation.

Orthographic projection using the “toe-in” method was used to create stereoscopic pairs. In a monocular orthographic view without depth cues, the rotation direction should

be ambiguous because there is no ordering of attenuation coefficients, $\tau(s)$, in the expression for $I(0)$ of Equation 3.6. Subjects would therefore be expected to have 50% accuracy in determining the direction of cylinder rotation.

The attenuation coefficient, τ , at point x inside the volume, was defined by Perlin noise [36], which is a sum of scaled harmonics of a predefined random noise function, $N(x)$:

$$\text{Perlin}(x) = \sum_{i=0}^{n-1} \frac{N(b^i x)}{a^i} \quad (5.1)$$

where $1/a$ is the persistence (the relative amplitude between adjacent harmonics) and b is the relative frequency between adjacent harmonics. $N(x)$ is created with seeded random number generator. Implementation details are provided in Appendix B. For the purpose of our experiments we set $n = 2$ and $b = 2$. To study the effect of spatial frequency, the persistence value a was varied from one to five. With higher persistence, more weight is given to higher frequencies.

The cylinder surface was modelled separately from the interior so that the interior opacity could be varied without changing the surface opacity. The opacity, α , was varied using a parameterized transfer function, $f_v(\tau)$:

$$\alpha = f_v(\tau) = \begin{cases} \tau + 2 & \tau(v - \frac{1}{2}) & \text{for } v \leq \frac{1}{2} \\ \tau + 2 & (1 - \tau)(v - \frac{1}{2}) & \text{for } v > \frac{1}{2} \end{cases} \quad (5.2)$$

$$(5.3)$$

To study the effect of opacity, we used values of v from 0.0 (completely transparent) to 0.9 (almost opaque). The same Perlin noise was used for the surface and for the interior

volume. Perlin noise and a cylindrical object were used to avoid bias from any domain-specific knowledge that the test subjects might have. Subjects with a medical background, for example, might perform better with anatomical shapes than would other subjects.

For the experiments, we measured each subject’s classification correctness and decision response time, then analyzed the data using an analysis of variance (ANOVA). ANOVA tests hypotheses about significance of differences between two or more class means by analyzing the variances with respect to the sample sizes. The comparison between the actual variation of the class averages and that which is expected is expressed in terms of the F-ratio. If the F-ratio is greater than 1 then it is likely that differences between class means exist. The results are tested for statistical significance, measured with the p value, which is the probability that a variate would assume a value greater than or equal to the value observed due to chance alone. If p is less than 0.05, there is less than a 5% probability that the means differ due to chance alone, meaning there is a highly significant difference between classes. In our experiment, we summarized the means and standard errors (SE, the standard deviation divided by the square root of the sample size), and used the F-ratio and p value to determine significance.

5.2.5 The Illumination Model

For DRR rendering, recall that the radiance visible to the eye coming along ray r is defined as

$$I(s_0) = I(s_1) e^{-\int_{s_0}^{s_1} \tau(s) ds} \quad (5.4)$$

where the ray through the volume is parameterized between s_0 and s_1 (the closest and farthest points, respectively), the attenuation coefficient is $\tau(s)$, and $I(s_1)$ is the radiance of the backlight. To implement contrast reduction, we used a variant of the general volume

rendering integral (Equation 3.6) in which the light was set to white where it entered the volume from the back:

$$C(s_1) \tau(s_1) ds = 1$$

$C(s)$ was zero (or black) everywhere inside the volume.

Contrast reduction was achieved by shifting radiance densities, $C(s)$, toward white from their usual solid black. At distance d from the viewer, normalized so that $d \in [0...1]$ for all points in the volume, we set

$$C(s) = (1 - k) d \tag{5.5}$$

where k was a “contrast factor”.

For $k = 0$, points near the back of the volume emitted the same radiance as the back-light, resulting in complete loss of contrast at the back. For $k = 1$, points near the back emitted no radiance, resulting in no change in contrast. For any value of k , points at the front did not change in contrast. To study contrast reduction, k , was varied from 0.75 to 1.0 in increments of 0.05.

5.3 Experiment 1: Stereopsis with DRRs

Experiment 1 tested Hypotheses 1 through 5. Fifteen subjects were each presented with 100 instances of rotating cylinders. Five opacity factors, $v = 0.0, 0.1, 0.3, 0.6$ and 0.9 (from Equation 5.2), and five persistence values, $a = 1, 2, 3, 4$ and 5 (from Equation 5.1), were considered. Opacity was varied by altering the density of the Perlin noise inside the surface cylinder. Each combination of opacity and spatial frequency was shown to the subjects, twice in mono and twice in stereo. The order in which instances were presented to subjects was randomized.

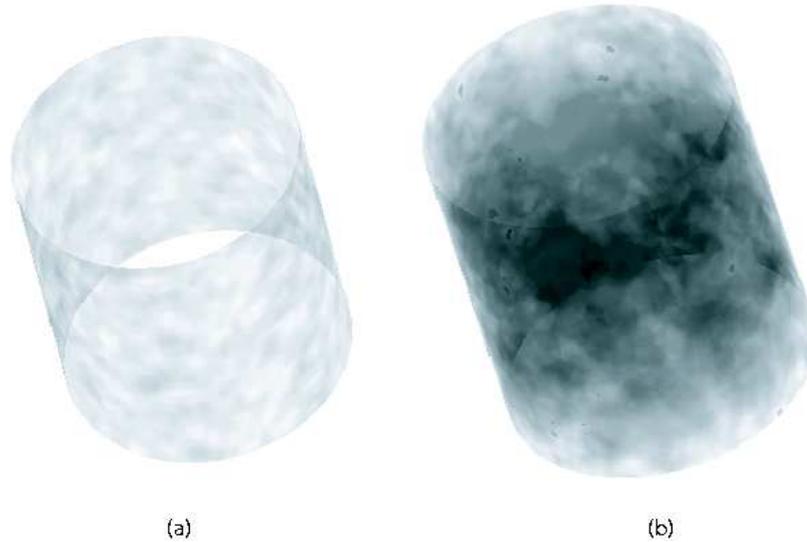


Figure 5.4: The cylinder surface (a) was modelled separately so that the interior opacity could be varied without changing the surface opacity. The cylinder surface was texture mapped with Perlin noise and the attenuation coefficients of the volume inside the cylinder (b) were generated with a Perlin noise function.

5.3.1 Results and Discussion

Hypothesis 1: Accuracy of depth perception in purely absorptive material is better with stereoscopic viewing than with monocular viewing.

A one-way ANOVA test showed that viewing mode (monoscopic or stereoscopic) significantly affected classification correctness (F -ratio = 28.526, $p < 0.001$). Mean classification correctness overall persistences and opacities for monocular viewing was 51.7% (SE 1.5%), which as expected was near the chance value of 50%; stereopsis improved overall classification correctness to 80.1% (SE 1.6%). The results for mean classification correctness are plotted in Figure 5.5, where correctness is a function of opacity. The results for mean classifications where correctness is a function of persistence is plotted in Figure 5.6.

These two results suggest that stereoscopic viewing of purely absorptive media is associated with better depth perception.

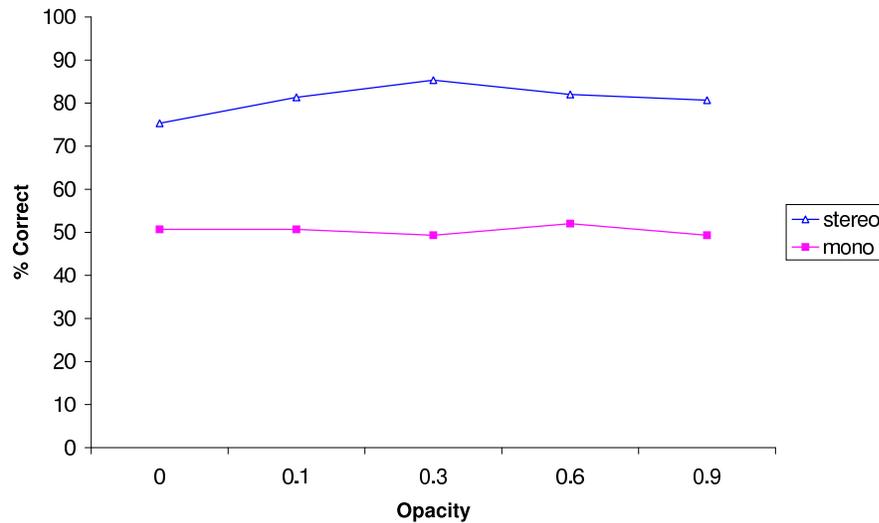


Figure 5.5: Mean classification correctness as a function of opacity. As expected, mean classification correctness for monocular viewing was near 50%. Mean classification correctness increased to 80.1% with stereoscopic viewing. The standard errors, 1.5% and 1.6% respectively, are too small to appear clearly in the plots.

Hypothesis 2: Accuracy of depth perception in purely absorptive material with stereoscopic viewing is not perfect.

The results from Hypothesis 1 showed that stereopsis provided a mean value of 80.1% in classification correctness. This imperfect classification is possibly because, when the eyes focus on a point within the volume in purely absorptive media, it sees two slightly different images of the point due to the different corresponding integrals.

Hypothesis 3: Accuracy of depth perception varies with opacity in purely absorptive material.

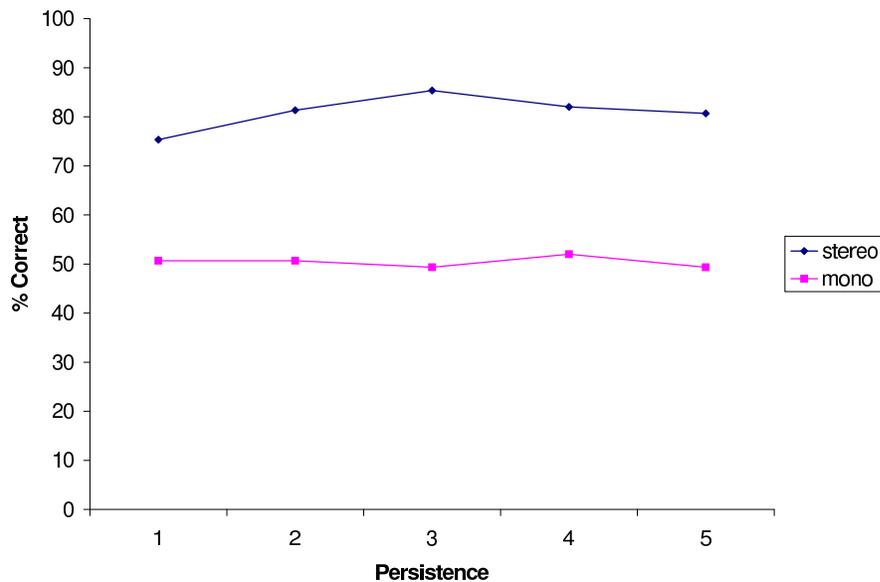


Figure 5.6: Mean classification correctness as a function of the persistence of spatial frequencies. Mean classification correctness increased from 51.7% (SE 1.5%) for monocular viewing to 80.1% (SE 1.6%) for stereoscopic viewing.

A one-way ANOVA test showed that opacity has no significant effect on classification correctness within absorptive materials (F -ratio = 0.411, $p = 0.8$). This is apparent in Figure 5.5, where the mean classification correctness is essentially unchanged for various opacity values. We analyzed monocular viewing and stereoscopic viewing separately, to determine if there are differences for the two presentation methods.

For monocular viewing, a one-way ANOVA test showed that opacity has no significant effect) on classification correctness, regardless of the persistence values used (F -ratio = 0.092, $p = 0.985$). These results are plotted in Figure 5.7.

For stereoscopic viewing, a one-way ANOVA test showed that opacity also has no significant effect on classification correctness, regardless of the persistence values used (F -ratio=1.092, $p = 0.369$). These results are plotted in Figure 5.8.

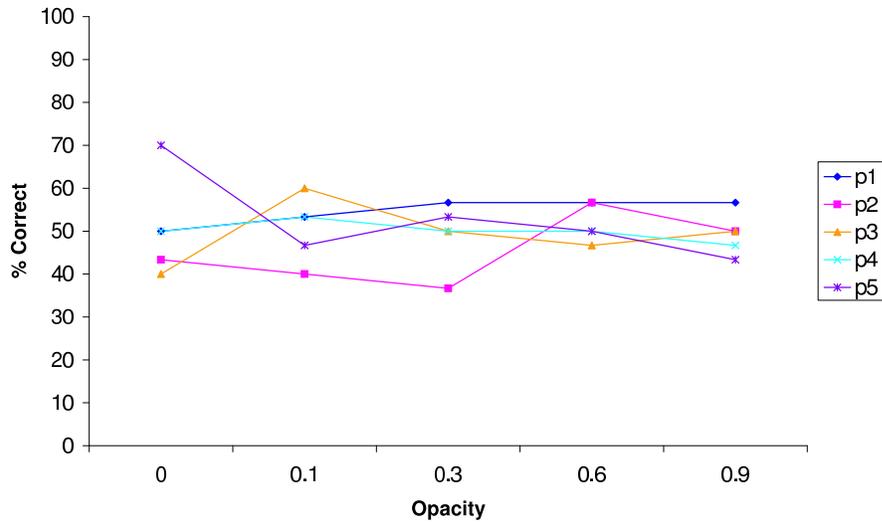


Figure 5.7: Mean classification correctness as a function of opacity, for various persistence values, in monocular viewing. Opacity has no significant effect on classification correctness, and persistence is also not well correlated with correctness.

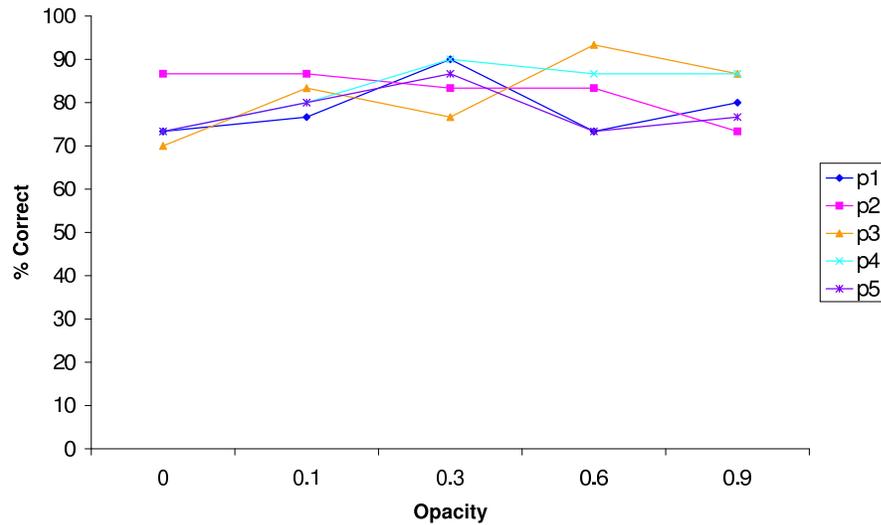


Figure 5.8: Mean classification correctness as a function of opacity, for various persistence values, in stereoscopic viewing. As for monocular viewing, neither opacity nor persistence have significant effects on classification correctness.

Hypothesis 4: Accuracy of depth perception varies with spatial frequency in purely absorptive material.

Spatial frequency had no significant effect on classification correctness (F-ratio = 0.366, $p = 0.832$). These results were presented in Figure 5.6 as part of the test of Hypothesis 1, showing the mean classification correctness as a function of persistence. We analyzed the presentation modes separately and determined that there were no differences for either monocular viewing (F-ratio = 1.012, $p = 0.409$) or for stereoscopic viewing (F-ratio = 0.920, $p = 0.459$).

Plots of classification correctness as a function of persistence, for all values of opacity, are shown in Figure 5.9 for monocular viewing and in Figure 5.10 for stereoscopic viewing.

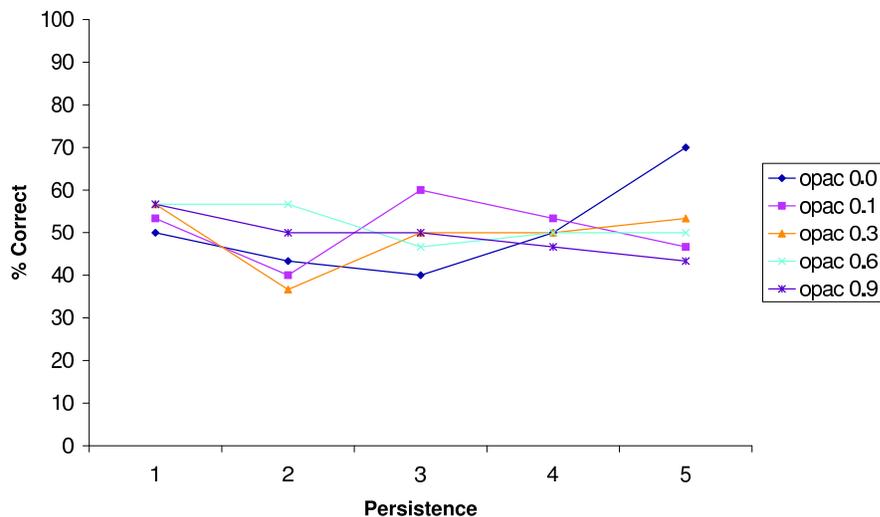


Figure 5.9: Mean classification correctness as a function of persistence, for various opacity values, in monocular viewing. Opacity has no significant effect on correctness.

Hypothesis 5: The time taken to decide relative depth is less with stereoscopic than with monoscopic viewing in purely absorptive media.

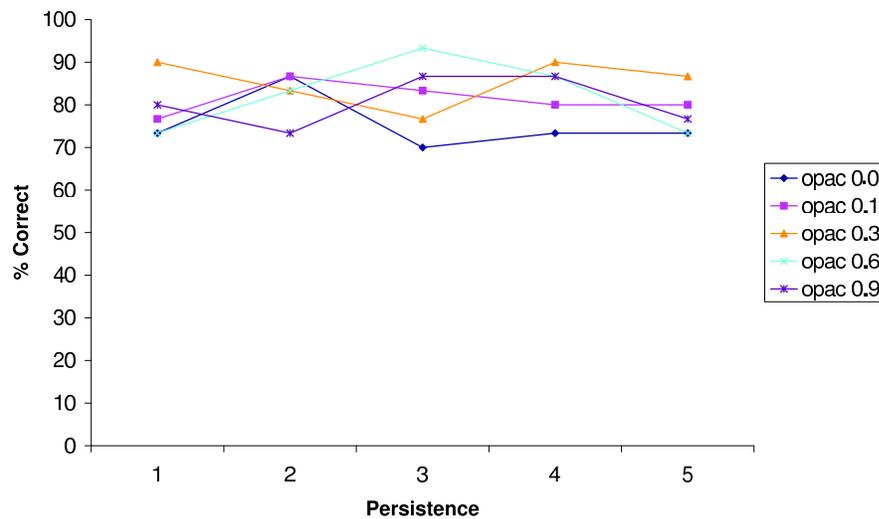


Figure 5.10: Mean classification correctness as a function of persistence, for various opacity values, in stereoscopic viewing. Persistence has no significant effect on correctness.

Interestingly, stereoscopic response times (mean = 6.23 secs, SE = 0.24) were similar to those of monocular response times (mean = 5.93 secs, SE = 0.26) although a reasonable assumption would be that stereo viewing would reduce response times, as it should make the decision task easier. ANOVA test results showed that there was no significant difference between response times for the two viewing groups (F -ratio = 0.255, $p = 0.642$). We posit two possible explanations for similar response times. One possibility is that there was a latency in fusing the stereo pairs when they first appeared. Another possibility is that the monocular cases were ambiguous and thus led subjects, upon recognizing the presentation mode as being monocular, to quickly choose a direction based on personal bias.

The ANOVA test showed that both opacity and spatial frequency had an effect on response time. Opacity had a significant effect on response time (F -ratio = 3.582, $p = 0.011$). Increasing opacity was associated with longer response times, as plotted in Figure 5.11. We speculate that this was because, with increased opacity, it was more difficult for the subject

to find features to follow, either in the volume or on the surface, in order to determine the direction of rotation.

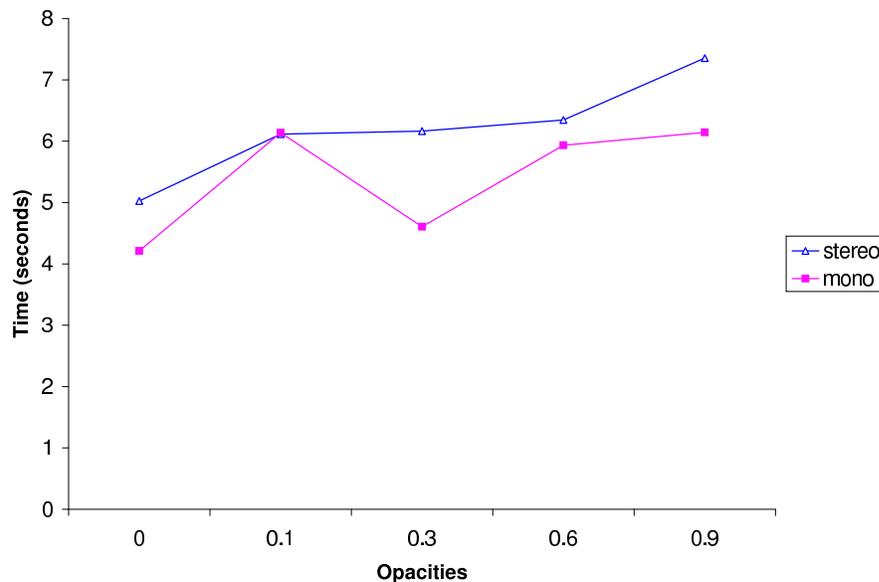


Figure 5.11: Mean response time as a function of opacity. Opacity significantly affected response times of the two viewing modes.

There was a highly significant effect of spatial frequency on response time (F -ratio = 4.639, $p = 0.003$). Higher persistence (i.e., more high frequency noise) was associated with longer response times, as shown in Figure 5.12. This may occur because low spatial frequency allowed the observer to more readily detect and track large clusters of texture.

5.4 Experiment 2: Simulated Aerial Perspective in DRRs

Preliminary testing without simulated aerial perspective showed that viewers were able to distinguish the direction of rotation with only monocular viewing. It was found that this occurred because of the unexpected presence of a second cue. Consider two points on a

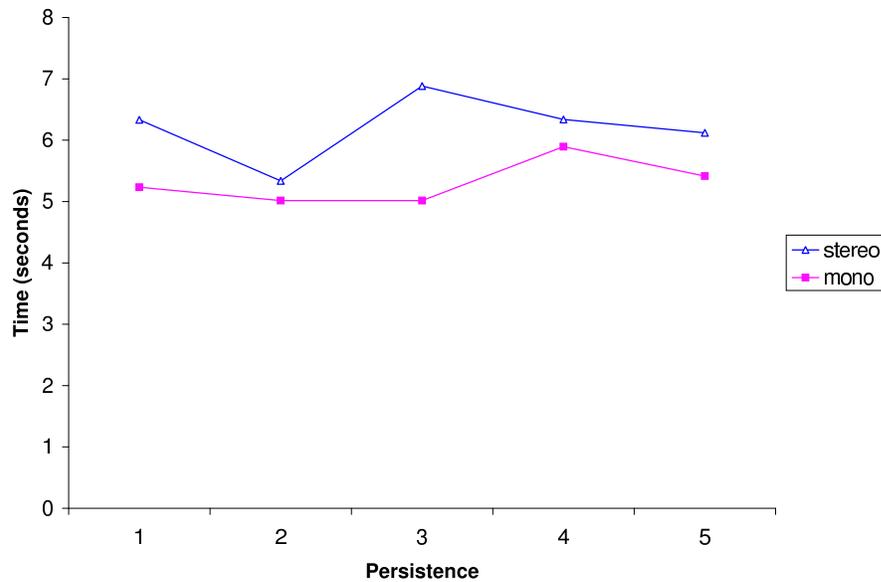


Figure 5.12: Mean response times as a function of spatial frequency. Higher persistence was significantly associated with longer response times.

volume cylinder that project to the same pixel. The opacity of the back point is α_1 and of the front point is α_2 . Recall that, for light L , the absorptive light using hardware blending in the graphics card is

$$(1 - \alpha_1) (1 - \alpha_2) e^{-\int_0^s \tau(u) du} L \quad (5.6)$$

where $e^{-\int_0^s \tau(u) du}$ is the attenuation due to the cylinder interior.

Equation 5.6 is the same when the back and front α values are switched. In practice, this means that an image of a volume cylinder, and an image of the same cylinder rotated 180 degrees and flipped left to right, should be identical. This, however, was not the case: *points at the back of the cylinder had a reduced contrast*. We believe that this was due to round-off error. As light is attenuated through the volume (back-to-front) the accumulated opacity was rounded off with each slice through the volume. After 200 slices, it is possible that the accumulated round-off errors caused contrast reduction. It should be noted that because

hardware–assisted volume rendering was used, contrast reduction was present (especially at the higher opacities) during Experiment 1.

To study this cue of simulated aerial perspective, thirteen subjects were each presented with 32 cases of rotating cylinders where contrast at the back of the cylinder was reduced, as described in Equation 5.5. The cylinder was empty (opacity=0.0), so we rendered the surface as a set of polygons. There was thus no accumulated round–off error because no volume rendering was done. Six contrast values (k in Equation 5.5), ranging from 0.75 to 1.0 in 0.05 increments, were evaluated. All tests used the same, low frequency Perlin noise ($a = 1, b = 2, n = 2$), on the cylinder surface.

5.4.1 Results and Discussion

Hypothesis 6: Contrast reduction with distance increases accuracy of depth perception in purely absorptive material.

At contrast factors between 0.75 and 0.85, the effectiveness of contrast reduction matched that of stereopsis as a depth cue. Mean classification correctness with contrast reduction between factors of 0.75 and 0.85 was 78.8% (SE=5.7%), and with stereoscopic viewing was 80.1% (SE=1.6%). In this range of contrast reduction, all parts of the volume seem to be visible. At even lower contrast factors (i.e., with more contrast reduction), the effectiveness of this depth cue exceeded that of stereopsis. Seven subjects performed the test for $k = 0.65$ and $k = 0.70$, resulting in mean percentages of correctness of 89% and 86% respectively.

A one–way ANOVA test showed a significant effect of contrast reduction on classification correctness (F–ratio = 7.987, $p < 0.001$). Classification correctness was about 50% with no contrast reduction ($k = 1.0$) but quickly improved with even slightly reduced contrast. Figure 5.13 shows a clearly increasing trend in classification accuracy as the contrast

of the back surface was decreased. Lower contrast factors, however, resulted in a perceptible loss of information in the distant parts of the volume.

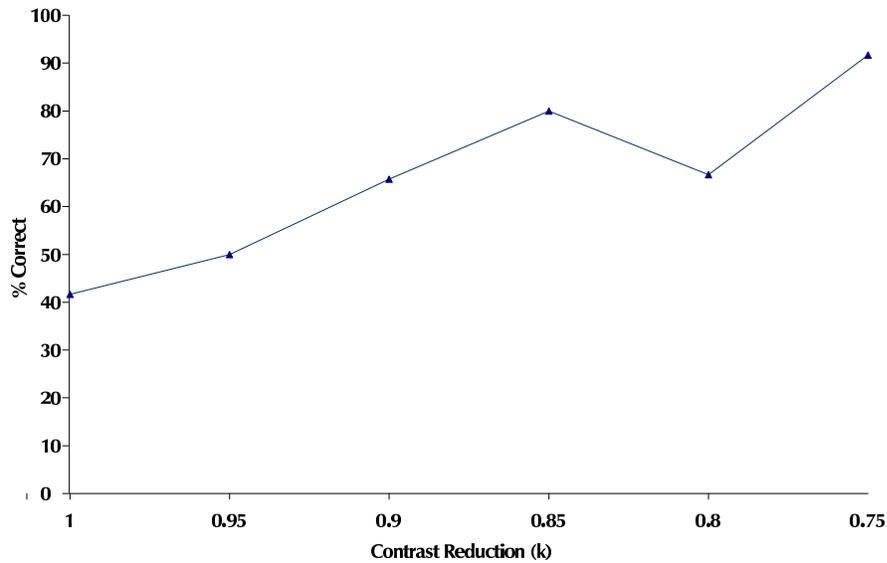


Figure 5.13: Classification correctness as a function of contrast. Even slightly reduced contrast improved classification correctness.

5.5 Measuring Acetabular Coverage using Stereo DRRs

DRRs can provide projection images not obtainable by conventional X-ray imaging, one example of which is a top-to-bottom projection of the hip socket (the acetabulum). Acetabular dysplasia is a condition in which the acetabulum is too shallow for the femoral head. Traditionally, assessment of acetabular dysplasia has been done by measuring distances or angles from plain radiographs. A number of different metrics and methods [18, 23, 24, 30, 44] have been explored for quantifying the amount of acetabular coverage of the femoral head. All suffer from the same problem: 2D images lack the 3D information needed to get a true picture of the shape of the femoral head, of the acetabulum, and of the degree and

direction of the acetabular coverage deficiency. Attempts at using 3D information from CT data [23, 34] have either been labour-intensive or required extensive pre-processing of the data prior to visualization. We propose that a fast and easy method for measuring acetabular coverage is to use stereoscopic DRRs; we compared stereoscopic DRRs, monoscopic DRRs, and an accepted alternative method [30] to investigate this proposition.

5.5.1 Stimuli and Measurements

One non-expert observer viewed 20 pelvic CT scans (10 pre-operative and 10 post-operative) of patients treated using computer-assisted periacetabular osteotomy. No further ethics approval was needed as the CTs were taken for the purposes of research and treatment with image guided surgery, prior to the conception of our research. The pelvic CTs were displayed as DRRs in a craniocaudal view (looking down from the patient's head to the toes).

The observer picked points on the DRRs, separately outlining the femoral head and the acetabulum as shown in Figure 5.14. The outlining was performed separately in monocular and stereoscopic rendering modes for each of the 20 CT data sets.

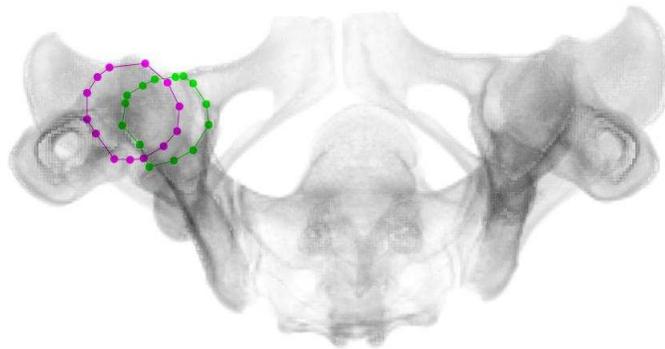


Figure 5.14: Craniocaudal view DRR of the pelvis. The medial set of picked points outlines the acetabulum and the lateral set outlines the femoral head.

The outlines were projected onto the horizontal plane, with the center of the femoral

head defined as the geometric center of the bounding box of the femoral outline. This center was used to divide the outlined area of the femoral head into four quadrants: anterolateral (AL), posterolateral (PL), anteromedial (AM), posteromedial (PM), as done by Mechlenberg et al. [30]. Each quadrant was further subdivided into covered and uncovered portions as in Figure 5.15. The ratio of covered area to total area was calculated for each quadrant, and for the entire femoral outline.

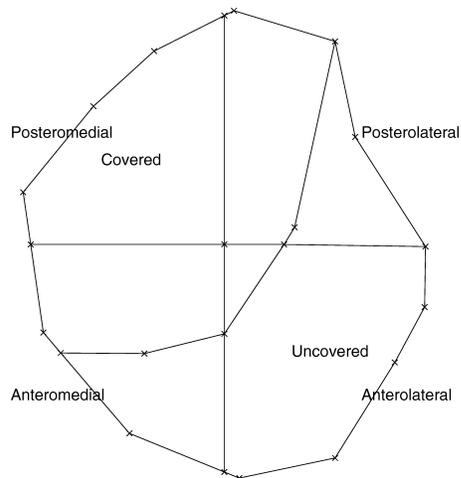


Figure 5.15: Quadrant divisions with coverage boundaries.

The same CT data were also used to assess coverage using an alternative method [30]. A second non-expert observer viewed the same 20 pelvic CT scans, off the original scans rather than on-screen. To assess coverage in this method [30], the hip area is resliced sagittally and the coverage in the sagittal plane is calculated for every second slice. The covered area is assumed to be a rectangle of the length measured and the width of two slices. The total area is the sum of the rectangular areas. For this study, the femoral head was assumed to be spherical, and the ratio of acetabular coverage was calculated from the raw measurements.

5.5.2 Results and Discussion

For the purpose of this study, we present only our analysis of monocular viewing mode versus stereoscopic viewing mode with respect the accepted sagittal slice method. The results of the coverage measurements are shown in Table 5.1 and Table 5.2.

Table 5.1: Measurements as percentage covered of each quadrant region and as a total measured from pre-operative CTs for each method. The four quadrants are antero-lateral (AL), posterolateral (PL), anteromedial (AM), posteromedial (PM).

Patient	Mono					Stereo					Sag. Total
	AL	AM	PL	PM	Total	AL	AM	PL	PM	Total	
1	19.2	94.9	1.1	71.8	48.2	28.1	93.4	40.4	100	63.5	37.2
2	1.0	71.6	17.6	94.6	48.0	8.1	78.9	66.1	100	59.7	35.7
3	6.6	50.2	60.5	100	56.1	7.2	59.9	63.2	100	55.3	22.9
4	34.7	100	13.2	98.7	59.0	29.1	99.3	21.6	97.9	60.6	42.3
5	22.5	100	5.9	100	55.4	39.0	100	7.7	93.3	59.2	28.7
6	31.2	90.5	71.3	100	73.5	11.9	99.5	37.7	100	67.7	33.2
7	0.0	20.8	0.8	81.4	29.6	13.9	82.1	10.4	92.6	49.9	26.8
8	10.9	96.4	8.9	99.4	51.6	3.3	96.5	3.6	98.0	50.2	28.4
9	32.4	99.6	57.2	100	69.6	17.4	92.8	16.6	92.6	55.0	27.9
10	6.3	86.2	14.1	95.4	49.4	14.2	95.1	28.4	100	57.4	33.4

A two-way independent-sample t-test showed that measuring acetabular coverage using monocular viewing was significantly different from measuring using the sagittal slice method ($t = -2.148$, $p = 0.038$). Conversely, a two-way independent sample t-test showed that stereoscopic measuring and sagittal measuring methods did not differ significantly from one another ($t = -0.825$, $p = 0.415$). These results suggest that stereoscopic viewing of DRRs is a viable technique for measuring acetabular coverage. Furthermore, there was a benefit to using stereoscopic viewing over monocular viewing.

Table 5.2: Measurements as percentage covered of each quadrant region and as a total measured from post-operative CTs for each method.

Patient	Mono					Stereo					Sag. Total
	AL	AM	PL	PM	Total	AL	AM	PL	PM	Total	
1	37.0	94.1	26.5	100.0	64.0	84.8	100.0	67.6	98.5	88.0	100.0
2	10.5	95.0	23.6	98.0	56.9	45.9	100.0	1.8	86.9	59.7	76.0
3	14.4	58.6	51.1	99.3	52.1	24.0	54.4	83.0	100.0	65.3	58.0
4	30.7	98.2	2.2	82.7	55.2	51.7	99.8	62.8	93.9	78.5	97.7
5	76.0	100.0	37.9	96.8	77.5	71.0	100.0	70.9	100.0	84.8	74.6
6	47.7	100.0	38.3	100.0	72.8	58.1	95.5	45.8	100.0	76.2	79.3
7	71.9	83.0	100.0	100.0	87.3	27.3	77.3	48.4	100.0	63.9	63.3
8	63.8	100.0	51.5	91.0	76.5	25.9	98.4	23.7	100.0	60.5	81.6
9	31.6	100	37.6	100.0	66.7	73.8	99.6	73.2	100.0	85.6	60.5
10	13.1	86.6	56.4	100.0	64.8	76.2	100	77.9	100.0	88.2	78.4

Our stereoscopic DRR method also allows fast, easy measurement of acetabular coverage; in all cases but one, the outlining of both the femoral head and acetabulum was completed in under a minute.

5.6 Summary of Findings

In this chapter, we have presented two experiments one which suggested that simulated aerial perspective is a useful cue for judging relative depth perception in purely absorptive media, and the other that there is a benefit of stereoscopic viewing of purely absorptive media. Based on these results, we conducted a study of using stereoscopic viewing in a particular medical application.

The results of this study showed that measurements using stereoscopic viewing of DRRs was not significantly different from an accepted method for measuring acetabular

coverage, whereas measurements using monoscopic viewing of the same DRRs was significantly different. This suggests the possible benefit of using stereoscopic viewing for medical applications. However, further testing with expert subjects (e.g., radiologists or orthopaedic surgeons) is yet to be done. Initial meetings with experts suggested an eagerness to use stereopsis for computer-assisted planning and surgical guidance. One subsequent phase of this work could be to examine how stereo DRRs can be applied to other orthopaedic interventional and assessment applications.

Chapter 6

Conclusion

This dissertation presents a discussion of stereoscopic volume rendering of purely absorptive media, an examination of the perception of digitally reconstructed radiographs (DRRs) using stereopsis and simulated aerial perspective, and psychophysical experimental studies to investigate the effectiveness of these cues.

Our results showed that both stereopsis and simulated aerial perspective are useful depth cues for rendering with a purely absorptive local illumination model. Stereoscopic viewing provided an advantage over monocular viewing, although it did not provide 100% accuracy, as might be expected for opaque surface rendering. Neither opacity nor spatial frequency had an effect on classification correctness. Interestingly, simulated aerial perspective provided a cue that nearly equalled that of stereopsis, at levels of contrast reduction that seemed not to obscure parts of the data. Our results are summarized in Table 6.1.

Results from the stereoscopic DRR measuring of acetabular coverage study showed that this was a fast and easy method for measuring acetabular coverage. It was comparable to an accepted traditional method ($t = -0.825$, $p = 0.415$), whereas measuring with monocular viewing of the same DRRs was not ($t = -2.148$, $p = 0.038$). This suggests the possible

Table 6.1: Summary of Results

Hypothesis	Result	Significance
1. Stereoscopic viewing of absorptive media provides better depth perception then monocular viewing	true	0.016
2. Stereoscopic viewing does not provide 100% accurate depth perception	true	0.000
3. Opacity affects depth perception in purely absorptive media	false	0.800
4. Spatial frequency affects depth perception in purely absorptive media	false	0.487
5. Stereoscopic viewing reduces response time	false	0.642
6. Simulated aerial perspective helps depth perception in purely absorptive media	true	0.000

benefit of using stereoscopic viewing for medical applications.

6.1 Future Work

One of the best indicators of usefulness of new methodologies is often the enthusiasm from the potential user group. Continued work with orthopaedic surgeons will examine how stereoscopic DRRs can be applied to orthopaedic interventional and assessment applications. One specific application area may include measuring carpal bone indices for distal radius osteotomies.

For computer-guided distal radial osteotomies (a form of wrist realignment) a surgical plan is made on surface-rendered volumetric bone models. However, the conventional indices that are used to quantify abnormalities of the wrist are measured on conventional X-rays. There is no corresponding measure of these indices for surface models. It may be interesting to examine measurement of these indices using DRR images, and specifically to study whether stereoscopic measurement of the indices is comparable to the traditional

X-ray measurement and whether it is more beneficial than monocular measuring.

Further studies on the usefulness of the contrast reduction cue could also be done. An especially useful experiment would be to determine whether contrast reduction loses, obscures, or otherwise affects important information at levels that provide the same depth perception as stereopsis. If there is no loss of information, contrast reduction could be a useful cue without the need for additional hardware, such as the stereo glasses. One such experiment could involve threshold discrimination where, for example, symbolic information (such as alphabetic letters) would be present on distant features of a volume. Contrast would be reduced until the threshold at which the subject could no longer perceive the symbols is found.

Some of our preliminary studies, not presented here because of their lack of completeness, suggested that presentation mode and training may affect classification correctness in the cylinder task. Further exploration on the effects of the ways in which stimuli are presented, e.g., randomly interleaved stereoscopic and monoscopic images versus serial presentation in each mode, could be done. We speculate that, in interleaved random viewing, subjects may immediately determine whether they were seeing a monocular or stereoscopic view – and perhaps perform differently than for serial presentation. We also did not study how training may condition subjects to attend more carefully to the stereoscopic views in the mixed presentations and less carefully to the monocular views. We leave these studies as topic for future work.

New 3D technologies available at the time of writing, such as 3D laptop and desktop spatial displays, show that there is a continued interest in 3D technology. The ways in which these technologies can be effectively incorporated to support various activities in both medicine and other fields requires more study. The results of the studies presented in

this dissertation are encouraging, as they show that 3D technologies can be used to enhance the interpretation and perception of medical images.

Bibliography

- [1] S. J. Adelson and C. D. Hansen. Fast stereoscopic images with ray-traced volume rendering. In *Proceedings of the 1994 symposium on Volume visualization*, pages 3–9. ACM Press, 1994.
- [2] C. Bethune. Adaptive slice geometry for hardware assisted volume rendering. Master's thesis, School of Computing, Queen's University, Kingston, ON, 2003.
- [3] P. Bourke. 3D stereo rendering using OpenGL (and GLUT). <http://astronomy.swin.edu.au/pbourke/opengl/stereogl/>, 2002.
- [4] G.F. Buess, P. van Bergen, W. Kunert, and M.O. Schurr. Comparative study of various 2-D and 3-D vision systems in minimally invasive surgery. *Chirurg.*, 67(10):1041–6, Oct 1996.
- [5] P. S. Calhoun, B. S. Kuszyk, D. G. Heath, J. C. Carley, and E. K. Fishman. Three-dimensional volume rendering of spiral CT data: Theory and method. *RadioGraphics*, 19:745–764, 1999.
- [6] J. Mackenzie Davidson. Remarks on the value of stereoscopic photography and skiagraphy: records of clinical and pathological appearances. *The British Medical Journal*, pages 1669–1671, December 1898.

- [7] K. Engel and T. Ertl. Interactive high-quality volume rendering with flexible consumer graphics hardware. State of the art report, Eurographics 2002, 2002.
- [8] J. J. Gibson. *The Perception of the Visual World*. The Riverside Press, 1950.
- [9] B. Goldstein. *Sensation and Perception*. Brooks/Cole Publishing Company, 1999.
- [10] R. C. Gonzalez and R. E. Woods. *Digital Image Processing*. Prentice Hall, 2002.
- [11] W. E. L. Grimson. *From Images to Surfaces*. The MIT Press, 1981.
- [12] G. B. Hanna, S. M. Shimi, and A. Cuschieri. Randomized study of influence of two-dimensional versus three-dimensional imaging on performance of laparoscopic cholecystectomy. *Lancet.*, 351:248–251, 1998.
- [13] J. Hsu, D. M. Chelberg, C. F. Babbs, Z. Pizlo, and E. J. Delp. Preclinical ROC studies of digital stereomammography. *IEEE Transactions on Medical Imaging*, 14(2):318–327, 1995.
- [14] J Hsu, Z. Pizlo, C. F. Babbs, D. M. Chelberg, and E. J. Delp. Design of studies to test the effectiveness of stereo imaging truth or dare: is stereo viewing really better? *SPIE*, 2177:211–222, 1994.
- [15] H. H. Hu, A. A. Gooch, W. B. Thompson, and B. E. Smits. Visual cues for imminent object contact in realistic virtual environments. In *11th IEEE Visualization 2000 Conference (VIS 2000)*. ACM Press, 2000.
- [16] G. S. Hubona, G. W. Shirah, and D. G. Fout. The effects of motion and stereopsis on three-dimensional visualization. *International Journal of Human-Computer Studies*, 47:609–627, 1997.

- [17] G. S. Hubona, P. N. Wheeler, G. W. Shirah, and M. Brandt. The relative contributions of stereo, lighting and background scenes in promoting 3D depth visualization. *ACM Transactions on Computer-Human Interaction*, 6(3):214–242, 1999.
- [18] D.L. Janzen, S.E. Aippersbach, P.L. Munk, D.F. Sallomi, D. Garbuz, J. Werier, and C.P. Duncan. Three-dimensional CT measurement of adult acetabular dysplasia: technique, preliminary results in normal subjects, and potential applications. *Skeletal Radiology*, 27:352–358, 1998.
- [19] I.C. Jourdan, E. Dutson, A. Garcia, T. Vleugels, J. Leroy, D. Mutter, and J. Marescaux. Stereoscopic vision provides a significant advantage for precision robotic laparoscopy. *British Journal of Surgery*, 91:879–885, 2004.
- [20] B. Julesz. *Foundations of Cyclopean Perception*. Bell Telephone Laboratories, Incorporated, 1971.
- [21] R. Kasrai, F. A. A. Kingdom, and T. Peters. The psychophysics of transparency in medical images. In *Medical Image Computing and Computer-Assisted Intervention - MICCAI'99, Second International Conference, Cambridge, UK Proceedings*, volume 1679, pages 726–733. Springer, 1999.
- [22] L. Kaufman. *Sight and Mind: An Introduction to Visual Perception*. Oxford University Press, 1974.
- [23] K. Klaue, A. Wallin, and R. Ganz. CT evaluation of coverage and congruency of the hip prior to osteotomy. *Clinical Orthopaedics and Related Research*, 232:15–25, 1988.

- [24] A. Kojima, T. Nakagawa, and A. Tohkura. Simulation of acetabular coverage of femoral head using aneteroposterior pelvic radiographs. *Archives of Orthopaedic and Trauma Surgery*, 117:330–336, 1998.
- [25] D. LaRose. *Iterative X-ray/CT Registration Using Accelerated Volume Rendering*. PhD thesis, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, May 2001.
- [26] M. Levoy. Display of surfaces from volume data. *IEEE Computer Graphics and Applications*, 8(3):29–37, 1988.
- [27] W. E. Lorensen and H. E. Cline. Marching cubes: A high resolution 3D surface construction algorithm. *Computer Graphics*, 21(4):163–169, 1987.
- [28] J. A. Marshall, C. A. Burbeck, D. Ariely, J. P. Rolland, and K. E. Martin. Occlusion edge blur: a cue to relative visual depth. *Journal of Optical Society of America*, 13(4):681–688, April 1996.
- [29] N. Max. Optical models for direct volume rendering. *IEEE Transactions on Visualization and Computer Graphics*, 1(2):99–108, 1995.
- [30] I. Mechlenburg, J.R. Nyengaard, L. Rømer, and K. Søballe. Changes in load-bearing area after ganz periacetabular osteotomy evaluated by multislice CT scanning and stereology. *Acta Orthopaedica Scandinavica*, 75(2):147–153, 2004.
- [31] M. Meissner, J. Huang, D. Bartz, K. Mueller, and R. Crawfis. A practical evaluation of popular volume rendering algorithms. In *Proceedings of the 2000 IEEE Symposium on Volume visualization*, pages 81–90. ACM Press, 2000.

- [32] P. Messmer, G. Long, N. Suhm, M. Hehli, J. Wirth, P. Regazzoni, and A. L. Jacob. Three-dimensional fracture simulation for preoperative planning and education. *European Journal of Trauma*, 27(4):171–177, August 2001.
- [33] B. Mora and D. S. Evert. Instant volumetric understanding with order-independent volume rendering. *Computer Graphics Forum*, 23(3), 2004.
- [34] S. Nakamura, J. Yorikawa, K. Otsuka, K. Takeshita, A. Harasawa, and T. Matsushita. Evaluation of acetabular dysplasia using a top view of the hip on three-dimensional CT. *Journal of Orthopaedic Science*, 5:533–539, 2000.
- [35] J. Owczarczyk and B. Owczarczyk. Evaluation of true 3D display systems for visualizing medical volume data. *The Visual Computer*, 6:219–226, 1990.
- [36] K. Perlin. An image synthesizer. *Computer Graphics (SIGGRAPH Proceedings)*, 19(3):287–296, July 1995.
- [37] C. Robertson, R. E. Ellis, T. Goetz, W. Gofton, P. V. Fenton, C. F. Small, and D. R. Pichora. The sensitivity of carpal bone indices to rotational malpositioning. *J Hand Surg*, 27A(3):435–442, 2002.
- [38] R. Sekuler and R. Blake. *Perception*. McGraw-Hill Publishing Company, 1990.
- [39] A. R. Smith. Alpha and the history of compositing. Technical Memo 7, Microsoft Corporation, August 1995.
- [40] L. Spillman and J. Werner, editors. *Visual Perception: The Neurophysiological Foundations*. Academic Press, Inc., Harcourt Brace Jovanovich, 1990.

- [41] F. Tendick, S. Bhojrul, and L. Way. Comparison of laparoscopic imaging systems and conditions using a knot-tying task. *Computer Aided Surgery*, 2:24–33, 1997.
- [42] L. R. Wanger, J. A. Ferwerda, and D. P. Greenberg. Perceiving spatial relationships in computer-generated images. *IEEE Computer Graphics and Applications*, 12(3):44–58, 1992.
- [43] A. Watt. *3D Computer Graphics*. Addison-Wesley Publishing Company, 1993.
- [44] G. Wiberg. Studies on dysplastic acetabula and congenital subluxation of the hip joint with special reference to the complication of osteoarthritis. *Acta Chirurgica Scandinavica Supplementum*, 83(Suppl. 58), 1939.
- [45] C. Wickens and Y. Liu. Use of computer graphics and cluster analysis in aiding relational judgement. *Human Factors*, 34(2):165–178, 1992.
- [46] C. Wickens and L. Thomas. Effects of CDTI display dimensionality and conflict geometry on conflict resolution performance. In *Proceedings of the 13th International Symposium on Aviation Psychology*, 2005.
- [47] C. Yang, M. Guiney, P. Hughes, S. Leung, K.H. Liew, J. Matar, and G. Quong. Use of digitally reconstructed radiographs in radiotherapy treatment and verification. *Australasian Radiology*, 44(4):439–443, November 2000.
- [48] J. M. Zelle and C. Figura. Simple, low-cost stereographics: VR for everyone. In *SIGCSE '04: Proceedings of the 35th SIGCSE Technical Symposium on Computer Science Education*, pages 348–352, New York, NY, USA, 2004. ACM Press.

- [49] S. Zhai, W. Buxton, and P. Milgram. The “silk cursor”: Investigating transparency for 3D target acquisition. In *Proceedings of ACM CHI'94 Conference on Human Factors in Computing Systems*, volume 1, pages 459–464, 1994.

Appendix A

Consent Form

Consent: Experiment on stereo-viewing

I am volunteering to participate in a study of stereoscopic viewing that is conducted by Marta Kersten under the supervision of Dr. James Stewart, Department of Computer and Information Science, Queen's University.

I understand the following concerning my participation:

1. The experimental task involves looking at stimuli on the monitor while both while wearing and not wearing stereo glasses. The experiment will involve looking at the stimuli on the screen and deciding which way it is moving.
2. The experimental session will last about 30 minutes.
3. I am aware that under normal conditions, stereoscopic viewing is safe for any duration that one would normally view a monitor for and that the experiment has been set up in order to minimize the potential for problems or adverse symptoms when using the stereo glasses. This has been done by placing me in a room without flickering lights and by using the highest monitor refresh rate.

4. I am aware that some people may experience discomfort using stereo-glasses. As well, I am aware that people who have a history of epilepsy or those who suffer from vertigo could possibly experience epileptic seizures or discomfort. If at any time during the test I experience any discomfort, I will stop the experiment.

5. All records of my individual participation will be treated as confidential. They will only be accessed by the experimenter, Marta Kersten.

6. I may terminate my participation at any time and I am not obliged to answer any questions that I find objectionable or which make me feel uncomfortable.

The faculty member conducting this research is Prof. James Stewart. Any complaints or additional questions that I have regarding the study may be expressed to Prof. Stewart at 533-5354, or anonymously to the representative of the Ethics Committee: Joan Stevenson at 533-6288.

I have read the above and have had my questions if any answered to my satisfaction. I understand the requirements of the experiment and voluntary nature of participation.

I will sign 2 copies, one for myself and one for Marta Kersten.

SIGN:

DATE:

Appendix B

Perlin Noise

```
// Make a volume of Perlin noise
//
// Good ref:
// http://freespace.virgin.net/hugo.elias/models/m\_perlin.htm
// http://astronomy.swin.edu.au/~pbourke/texture/perlin
//
// Original code provided by James Stewart, with modifications
// by Marta Kersten

#include <cstdlib>
#include <iostream>
#include <fstream>
#include <cmath>
using namespace std;

#ifndef M_PI
    #define M_PI 3.14159265358979323846
#endif

#ifndef MAXFLOAT
    #ifdef WIN32
        #include <float.h>
        #define MAXFLOAT FLT_MAX
```

```
#else
    #define MAXFLOAT 9e99
#endif
#endif

#define SQRT2 1.414213

using namespace std;

int noiseXdim = 16; // noise volume dimensions
int noiseYdim = 16;
int noiseZdim = 16;

int xdim = 128;           // turbulence volume dimensions
int ydim = 128;
int zdim = 128;

int blurRadius = 0;

float maxRadius      = 0.45; // radius beyond which volume is empty
float surfaceRadius = 0.3;  // radius to surface
float stdDev         = 0.05; // standard deviation around surface
int   numHarmonics  = 2;    // number of Perlin noise harmonics
float persistence    = 0.5; // amplitude reduction with each
// higher-freq harmonic

char *filename = "perlin.img";

float *** data;
float *** data2;
float *** noise;
float *** smoothedNoise;
float *** blurWeights;

// Get command-line options

void getOptions( int argc, char **argv )

{
    while (argc > 1) {
        argv++; argc--;
    }
}
```

```

if ((*argv)[0] != '-')
  cerr << "Unrecognized flag " << *argv << endl;
else {
  char c = (*argv)[1];
  argv++; argc--;
  switch (c) {

    case 's':                // s: standard deviation
      stdDev = atof( *argv );
      break;

    case 'n':                // n: number of noise harmonics
      numHarmonics = atoi( *argv );
      break;

    case 'r':                // r: surface radius
      surfaceRadius = atof( *argv );
      break;

    case 'p':
      persistence = atof( *argv ); // p: persistence
      break;

    case 'd':                // d: dimensions
      xdim = atoi( *argv );
      argv++; argc--; ydim = atoi( *argv );
      argv++; argc--; zdim = atoi( *argv );
      break;

    default:
      cerr << "Unrecognized option " << *argv << endl;
  }
}
}

// Smoothed noise at grid points. Use a discrete Gaussian filter.

float smoothNoise( int x, int y, int z )
{
  int xm = (x - 1 + noiseXdim) % noiseXdim;

```

```

int ym = (y - 1 + noiseYdim) % noiseYdim;
int zm = (z - 1 + noiseZdim) % noiseZdim;

int xp = (x + 1) % noiseXdim;
int yp = (y + 1) % noiseYdim;
int zp = (z + 1) % noiseZdim;

float corners = noise[xm][ym][zm] +
                noise[xm][ym][zp] +
                noise[xm][yp][zm] +
                noise[xm][yp][zp] +
                noise[xp][ym][zm] +
                noise[xp][ym][zp] +
                noise[xp][yp][zm] +
                noise[xp][yp][zp];

float sides    = noise[x ][ym][zm] +
                noise[x ][ym][zp] +
                noise[x ][yp][zm] +
                noise[x ][yp][zp] +
                noise[xm][y ][zm] +
                noise[xm][y ][zp] +
                noise[xp][y ][zm] +
                noise[xp][y ][zp] +
                noise[xm][ym][z ] +
                noise[xm][yp][z ] +
                noise[xp][ym][z ] +
                noise[xp][yp][z ];

float faces    = noise[x ][y ][zm] +
                noise[x ][y ][zp] +
                noise[x ][ym][z ] +
                noise[x ][yp][z ] +
                noise[xm][y ][z ] +
                noise[xp][y ][z ];

float centre   = noise[x ][y ][z ];

return corners/64.0 + sides/32.0 + faces/16.0 + centre/8.0;
}

// Cosine interpolation (1st derivative continuous)

```

```

float interp( float a, float b, float t )
{
    float f = 0.5 * (1 - cos(t*M_PI));

    return a*(1 - f) + b*f;
}

// Trilinear interpolation of noise at discrete grid points

float interpolatedNoise( float x, float y, float z )
{
    int x0    = ((int) floor(x*noiseXdim)) % noiseXdim;
    int y0    = ((int) floor(y*noiseYdim)) % noiseYdim;
    int z0    = ((int) floor(z*noiseZdim)) % noiseZdim;

    float fx = x*noiseXdim - x0;
    float fy = y*noiseYdim - y0;
    float fz = z*noiseZdim - z0;

    int x1    = (x0+1) % noiseXdim;
    int y1    = (y0+1) % noiseYdim;
    int z1    = (z0+1) % noiseZdim;

    float n000 = smoothedNoise[x0][y0][z0];
    float n001 = smoothedNoise[x0][y0][z1];
    float n010 = smoothedNoise[x0][y1][z0];
    float n011 = smoothedNoise[x0][y1][z1];
    float n100 = smoothedNoise[x1][y0][z0];
    float n101 = smoothedNoise[x1][y0][z1];
    float n110 = smoothedNoise[x1][y1][z0];
    float n111 = smoothedNoise[x1][y1][z1];

    float x00 = interp( n000, n100, fx );
    float x01 = interp( n001, n101, fx );
    float x10 = interp( n010, n110, fx );
    float x11 = interp( n011, n111, fx );

    float xy0 = interp( x00, x10, fy );
    float xy1 = interp( x01, x11, fy );
}

```

```
float xyz = interp( xy0, xy1, fz );

return xyz;
}

// 3D Perlin noise

float turbulence( float x, float y, float z )
{
    float total = 0;
    float freq = 1;
    float ampl = 1;

    for (int i=0; i<numHarmonics; i++) {

        total += interpolatedNoise(x*freq, y*freq, z*freq) * ampl;
        freq *= 2;
        ampl *= persistence;
    }
    return total;
}

int main( int argc, char **argv )
{
    getOptions( argc, argv );

    // Find maximum noise value

    float maxValue = 0;
    float ampl = 1;

    for (int i = 0; i < numHarmonics; i++)
    {
        maxValue += ampl;
        ampl *= persistence;
    }

    cout << "MAX NOISE VALUE = " <<maxValue <<endl;
```

```

// Create arrays

int i, x;

data = (float***) malloc(xdim * sizeof(float**));
for (x=0; x<xdim; x++) {
    data[x] = (float **) malloc(ydim * sizeof(float*));
    for (int y=0; y<ydim; y++)
        data[x][y] = (float *) malloc(zdim * sizeof(float));
}

noise = (float ***) malloc(noiseXdim * sizeof(float**));
for (x=0; x<noiseXdim; x++) {
    noise[x] = (float **) malloc(noiseYdim * sizeof(float*));
    for (int y=0; y<noiseYdim; y++)
        noise[x][y] = (float *) malloc(noiseZdim * sizeof(float));
}

smoothedNoise = (float ***) malloc(noiseXdim * sizeof(float**));
for (x=0; x<noiseXdim; x++) {
    smoothedNoise[x] = (float **) malloc(noiseYdim *
sizeof(float*));
    for (int y=0; y<noiseYdim; y++)
        smoothedNoise[x][y] = (float *) malloc(noiseZdim *
sizeof(float));
}

data2 = (float***) malloc(xdim * sizeof(float**));
for (x=0; x<xdim; x++) {
    data2[x] = (float **) malloc(ydim * sizeof(float*));
    for (int y=0; y<ydim; y++)
        data2[x][y] = (float *) malloc(zdim * sizeof(float));
}
}

// Set up the histogram

int histo[255];

for (i=0; i<256; i++)
    histo[i] = 0;

```

```

// Create the noise

cout << "\rnoise          "; cout.flush();

for (x = 0; x < noiseXdim; x++)
  for (int y = 0; y < noiseYdim; y++)
    for (int z = 0; z < noiseZdim; z++)
      noise[x][y][z] = (drand48() * 2 - 1) * 0.05;

// Create the smoothed noise

cout << "\rsmoothing    "; cout.flush();

float minNoise = MAXFLOAT;
float maxNoise = -MAXFLOAT;

for (x = 0; x < noiseXdim; x++)
  for (int y = 0; y < noiseYdim; y++)
    for (int z = 0; z < noiseZdim; z++) {
      smoothedNoise[x][y][z] = smoothNoise(x,y,z);
      if (smoothedNoise[x][y][z] < minNoise)
        minNoise = smoothedNoise[x][y][z];
      if (smoothedNoise[x][y][z] > maxNoise)
        maxNoise = smoothedNoise[x][y][z];
    }

for (x = 0; x < noiseXdim; x++)
  for (int y = 0; y < noiseYdim; y++)
    for (int z = 0; z < noiseZdim; z++)
      smoothedNoise[x][y][z] = (smoothedNoise[x][y][z] - minNoise)
/ (maxNoise - minNoise) * 2 - 1;

// Create the turbulence

for (x = 0; x < xdim; x++) {
  cout << "\rperlin " << xdim-x << "          "; cout.flush();
  for (int y = 0; y < ydim; y++)
    for (int z = 0; z < zdim; z++)
      data[x][y][z] = turbulence( (float)x/xdim,
(float)y/ydim,
(float)z/zdim );
}

```

```

// Cut out the shape

float halfDiagonalLength = 1 / SQRT2 / ((xdim+ydim)/2);

for (x = 0; x < xdim; x++)
  for (int y = 0; y < ydim; y++)
    for (int z = 0; z < zdim; z++) {
float xx = (float)x/xdim-0.5;
float yy = (float)y/ydim-0.5;
float zz = (float)z/zdim-0.5;
float dist = sqrt( (float) (xx*xx + yy*yy) );

if (dist > maxRadius + halfDiagonalLength)
  data[x][y][z] = -maxValue;
else
  if (dist > maxRadius - halfDiagonalLength) {
// right on the surface: antialias it
float frac = 1 - (dist - (maxRadius - halfDiagonalLength))
/ (2 * halfDiagonalLength);
  data[x][y][z] = frac * (data[x][y][z] + 1) - 1;
  }

// Apply Gaussian

float diff = fabs( dist - surfaceRadius );
float gaussian = exp( -diff*diff / (2*stdDev) );
data[x][y][z] = gaussian * (data[x][y][z] + 1) - 1;
  }

#if 0
  for (int x = 0; x < xdim; x++)
    for (int y = 0; y < ydim; y++) {
      data[x][y][0] = 0;
      data[x][y][zdim-1] = 0;
    }
#endif

// Output the turbulence

cout << "\routput      "; cout.flush();

ofstream output( filename, ios::binary );

```

```
if (!output) {
    cerr << "Couldn't open " << filename << endl;
    exit(1);
}

for (x = 0; x < xdim; x++)
    for (int y = 0; y < ydim; y++)
        for (int z = 0; z < zdim; z++) {
            float val = data[x][y][z];
            if (val < -1) val = -maxValue/fabs(maxValue);
            if (val > +1) val = val/fabs(maxValue);
            unsigned char byte = (unsigned char) (val * 127 + 127);
            output.put( byte );
            histo[byte]++;
        }

    cout << "\r" << " " << "\r"; cout.flush();

// Output the histogram

ofstream histOut( "histo.dat" );

if (!histOut) {
    cerr << "Couldn't open histo.dat" << endl;
    exit(1);
}

for (i=0; i<255; i++)
    histOut << i << " " << histo[i] << endl;

return 0;
}
```