DOCTOR OF PHILOSOPHY

Improving the impression of depth perception

Easa, Haider

Award date:
2015

Awarding institution:
Bangor University

Improving the Impression of Depth Perception

PRIFYSGOL
BANGOR
UNIVERSITY

Haider Khalil Easa

School of Computer Science

Bangor University

Thesis submitted in candidate for the degree of

Doctor of Philosophy

January 2015

Declaration and Consent

I hereby agree to deposit the following item in the digital repository maintained by Bangor University and/or in any other repository authorized for use by Bangor University.

Author Name: Haider Khalil Easa.

Title: Improving the Impression of Depth Perception

Supervisor/Department: Dr Ik Soo Lim / School of Computer Science.

Funding body (if any): Ministry of higher education and scientific research of Iraq.

Qualification/Degree obtained: PhD.

This item is a product of my own research endeavours and is covered by the agreement below in which the item is referred to as "the Work". It is identical in content to that deposited in the Library, subject to point 4 below.

Non-exclusive Rights Rights granted to the digital repository through this agreement are entirely non-exclusive. I am free to publish the Work in its present version or future versions elsewhere. I agree that Bangor University may electronically store, copy or translate the Work to any approved medium or format for the purpose of future preservation and accessibility. Bangor University is not under any obligation to reproduce or display the Work in the same formats or resolutions in which it was originally de-

posited.

Bangor University Digital Repository I understand that work deposited in the digital repository will be accessible to a wide variety of people and institutions, including automated agents and search engines via the World Wide Web.
I understand that once the Work is deposited, the item and its metadata may be incorporated into public access catalogues or services, national databases of electronic theses and dissertations such as the British Library's EThOS or any service provided by the National Library of Wales.
I understand that the Work may be made available via the National Library of Wales Online Electronic Theses Service under the declared terms and conditions of use http://www.llgc.org.uk/.php?id=4676. I agree that as part of this service the National Library of Wales may electronically store, copy or convert the Work to any approved medium or format for the purpose of future preservation and accessibility. The National Library of Wales is not under any obligation to reproduce or display the Work in the same formats or resolutions in which it was originally deposited.

Statement 1:
This work has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree unless as agreed by the University for approved dual awards.

Signed ..... (Haider Easa)

Date : January 2015 .

Statement 2:
This thesis is the result of my own investigations, except where otherwise stated. Where correction services have been used, the extent and nature of the correction is clearly marked in a footnote(s). All other sources are acknowledged by footnotes

and/or a bibliography.

Signed ..... (Haider Easa)

Date : January 2015 .

Statement 3:

I hereby give consent for my thesis, if accepted, to be available for photocopying, for inter-library loan and for electronic repositories, and for the title and summary to be made available to outside organisations.

Signed ..... (Haider Easa)

Date : January 2015 .

NB: Candidates on whose behalf a bar on access has been approved by the Academic Registry should use the following version of **Statement 3:**

Statement 3 (bar):

I hereby give consent for my thesis, if accepted, to be available for photocopying, for inter-library loans and for electronic repositories after expiry of a bar on access.

Signed ..... (Haider Easa)

Date : January 2015 .

Statement 4:

Choose **one** of the following options.

a) I agree to deposit an electronic copy of my thesis (the Work) in the Bangor University (BU) Institutional Digital Repository, the British Library ETHOS system, and/or in any other repository authorized for use by Bangor University and where necessary have gained the required permissions for the use of third party material.	
b) I agree to deposit an electronic copy of my thesis (the Work) in the Bangor University (BU) Institutional Digital Repository, the British Library ETHOS system, and/or in any other repository authorized for use by Bangor University when the approved bar on access has been lifted.	
c) I agree to submit my thesis (the Work) electronically via Bangor University's e-submission system, however I opt-out of the electronic deposit to the Bangor University (BU) Institutional Digital Repository, the British Library ETHOS system, and/or in any other repository authorized for use by Bangor University, due to lack of permissions for use of third party material.	

Options B should only be used if a bar on access has been approved by the University

In addition to the above I also agree to the following:

- That I am the author or have the authority of the author(s) to make this agreement and do hereby give Bangor University the right to make available the Work in the way described above.
- That the electronic copy of the Work deposited in the digital repository and covered by this agreement, is identical in content to the paper copy of the Work deposited in the Bangor University Library, subject to point 4 below.
- That I have exercised reasonable care to ensure that the Work is original and, to the best of my knowledge, does not breach any laws including those relating to defamation, libel and copyright.

- That I have, in Instances where the intellectual property of other authors or copyright holders is included in the Work, and where appropriate, gained explicit permission for the inclusion of that material in the Work, and in the electronic form of the Work as accessed through the open access digital repository, or that I have identified and removed that material for which adequate and appropriate permission has not been obtained and which will be inaccessible via the digital repository.

- That Bangor University does not hold any obligation to take legal action on behalf of the Depositor, or other rights holders, in the event of a breach of intellectual property rights, or any other right, in the material deposited.

- That I will indemnify and keep indemnified Bangor University and the National Library of Wales from and against any loss, liability, claim or damage, including without limitation any related legal fees and court costs (on a full indemnity bases), related to any breach by myself of any term of this agreement.

Signature: ..... Date : January 2015 .

I dedicate this thesis to

My lovely family, Asima, Manar, Zahraa and Rossul for their love and inspiration.

And also to my beloved mum, brothers and sisters in Iraq for their constant support and unconditional love.

I love you all dearly.

Acknowledgements

I wish to express my utmost thanks to many people whose input has been invaluable throughout the course of this study. First and foremost, I have to thank *Dr Rafal Mantiuk*. Without his assistance and dedicated involvement in every step throughout the process, this thesis would have never been accomplished. I would also like to show gratitude to *Prof. Nigel John* for his invaluable support through writing this thesis. A special thanks to *Dr Sa'ad Mansoor* for his valuable notes and great assistance in preparing this thesis. I also thank *Dr Ik Soo Lim* who taught me how to depend on myself.

I want to express my deepest gratitude to *Thomas Christy* who I am very honoured to be able to call my friend. He is one of the most honest people I met. I cannot find any appropriate word to express my thanks for everything he made for me to accomplish this thesis and for his invaluable support and encourage with all what he can. I could not have survived without the encouragements and supports of my best and kind friends *Robert Wanat* and *Yasser Al-Ghuraibi*. I cannot begin to express my gratitude and appreciation for their friendship. Also I would like to express my thanks to *Dr Catrin Plumpton*, *Dr Marc Edwards*, *Dr David Hughes* for their support and help. Most importantly, none of this could have happened without my mum, brothers and sisters and their unconditional love and care. They always offered their encouragement through phone calls. I would like to thank my friends in *Shaqlawa technical institute*, especially *Sabah Rasool*, for their encouragement and assistance through the years of study. A great thanks the ministry of higher education and scientific research of Iraq, Iraqi embassy and cultural attaché in London, ministry of higher education and scientific research of the Kurdistan region/Iraq to have given me this opportunity to get the PhD degree.

Abstract

Computing and visual display technologies have made great advances in the last few decades in terms of the interface between man and machine. Many computer applications today make use of 3D interactive computer graphics. Ocular depth is perceived when our visual system combines various different sources of information, which are normally called depth cues, around a scene. This information plays an important role in our life. For example, receiving accurate information about the depth of a scene will save time and effort to a surgeon when performing a surgical operation. The main aim of the dissertation is to improve the impression of depth perception when visualise a multiple-layered computer-generated imagery (CGI).

Three experiments were achieved: (1) to investigate which monocular cue provides a better impression of depth when other cues are not available, (2) to perform an evaluation to find which type of translucency among the two common types has a stronger effect on the depth perception, and (3) to study whether the shape of an incision made on the outer layer of a multiple-layered image would have an influence on depth perception.

The outcomes of the study were promising. The results of the first experiment demonstrated that brightness, contrast and relative size cues having better impression of depth than other monocular cues. According to the results of the second experiments, we were unable to decide which type of translucency had a stronger effect on depth impression. It showed that there is no statistically significant difference between the two types. In conclusion, the consequences of the third experiment confirmed that the shape of an incision, made on the outer layer, influences the depth perception.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Generating images that simulate realism is still one of the goals in computer graphics, since its founding in the 1960s. Portraying realism in computer graphics has made great advances in the last decade. Images made by computer are normally far more accurate than those produced by hand and are produced faster. These features (accuracy and speed) have inspired designers of computer graphics to find and develop techniques, and algorithms to produce images that look as real as possible. Other than utilising both eyes, the simplicity of perceiving the world in three dimensions is assisted by our visual system, which uses all available sources of information about the objects around us, known as *depth cues*. Depth cues describe the relationship between objects spatially.

Artists understood that depth cues were important long before the appearance of computer graphics. They realised that the impression of depth improved when adding more depth cues, making their paintings appear more lifelike. Despite the fact that paintings were created on a two dimensional canvas, many portrayed a strong impression of being created in three dimensions. In many cases, it is difficult to discern the

difference between handmade paintings and photographs taken by camera. This form of paintings are called *photorealistic*. Conversely, an insufficient amount of depth cues causes the wrong impression and an inaccurate perception of depth, particularly so with abstract or unfamiliar paintings.

It is difficult to study the visual effect found in nature on a computer screen. This is because these natural phenomena are modelled in geometry (triangles and pixels) before they are rendered. The process of rendering also makes assumptions over how the user will perceive the scene, and on most display technologies, shadows, lighting and even depth are created using local calculations, rather than global style calculations that would model every light particle in that scene. In essence the computer uses "models" and heuristics to display the data. Natural phenomena such as clouds and fire (Kajiya & Von Herzen, 1984) further add complications, because they are transparent and translucent effects, with shadows and internal reflections that makeup by internal appearance. Volume rendering is therefore essential to create more realistic renderings of such three-dimensional datasets. VR forms the basic modality in most scientific and engineering applications that deal with three dimensional (3D) visualization.

In the last decades, to render a 3D model of a real scene, a huge amount of memory and fast computation were needed (Strothotte & Schlechtweg, 2002). However, medium sized volumes can be rendered in real time on today's hardware. With new trends towards Big Data, the computational demands on real time VR continue to grow. The aim of these volume renderings is to visualise the 'hidden'. Surface rendering techniques merely display the surround of the object, while VR techniques allow the user to see inside. There are several techniques used to visualise a dataset. For instance, removing the less important layers, usually external layers. Transparency is another way to reveal the hidden parts in a multiple-layered images. However, the

main point in most of the studies that deals with visualisation is to reveal the hidden layers.

One big challenge with such volume renderings is to perceive depth. The resultant image, after visualisation, gives an impression as if it latched on the outer surface and it is difficult to determine the depth order of the resultant parts (Viola et al., 2005; Coffin & Hollerer, 2006; Liang et al., 2005; Correa et al., 2006; Krüger et al., 2006). The main reason for this misperception is the lack of depth cues. Therefore, by adding depth cues to these images, it could result in enhanced and more informative renderings. For example, in medical imaging, information inside the body is very important to the surgeon before entering theatre. Providing depth information gives the surgeon a better overview about the internal anatomy, which reduces both time and effort during an operation.

1.1 Problem Description

There are two broad categories describing depth cues, monocular and binocular. This study focuses on monocular cues only. One of the advantages of monocular cues is that these cues provide extra information about the relationships of the objects in a scene, in terms of location and size. Each individual cue provides different information, and they contribute by providing depth information to the user. Therefore they are known as *depth cues*.

Usually, the importance of depth cues is affected by ones understanding of the objects in a scene. Thus, cues are less important when the objects are known to the viewer. Consequently, the importance of using cues increases when the objects in the scene are unknown or abstract. For instance, information about location, dimension,

or the distance between the objects is needed.

Normally, this information is not added to a computer generated image automatically. Extra computation is needed, according to the cue that is being applied to the image. For example, the computational cost required to apply ambient occlusion on an image, producing a shadow effect, differs from the cost required to apply blurring. In fact, the time required to produce a blur cue is less than the time required to generate a shadow cue using ambient occlusion. Consequently, most researches have focused on how to reduce the computational time needed to apply a cue. As a result, it would be advantageous for a user to appropriately select the most visually effective cue, before doing any computer processing.

Browsing the internal structures of a volume is one of the main objectives of three dimensional (3D) volume rendering. Several studies have been conducted to produce and design techniques for volume rendering, for use in volumetric illustration systems (Pflesser et al., 2002; McGuffin et al., 2003; Diepstraten et al., 2003; Owada et al., 2004; Burns & Finkelstein, 2008). Making the outer surface of a multi-layered three dimensional model transparent, is one of the methods used in visualisation to show the internal layers of the model (Krüger et al., 2006). Direct volume rendering (DVR) techniques are based on this method (Levoy, 1988). Such techniques are useful when visualising known datasets, such as human anatomy. The reason is that most displayed internal structures may be known to the viewer. However, it would be difficult to know or understand unknown or abstract models using this technique. One reason for this misconception is the lack of depth information about the inner contents, especially when they have multiple layers. In addition, the order of these layers is usually unknown or difficult to assess. Another reason, is the ambiguity that goes with transparency. Two out of three cases (Adelson & Anandan, 1990), which are usually

accompanied by transparency, produce ambiguous ordering. Blurring is one of the techniques used to make a distinction on transparent surfaces to give the perception of two distinct surfaces (Singh & Anderson, 2002a).

Another technique to browse the inner parts of a volumetric model is to cut a part of its outer surface. By cutting part of the surface, it is therefore possible for a user to see internal structures of the model. Most of the models that were used in the previous studies were based on *non-photorealistic rendering* (NPR) model, which usually lack detail. Moreover, the models used were known to the user, such as a part of the human body, cars, or machines, etc. Therefore, the results from the cut in their study revealed expected data. Thus, the shape of the cut in the studies was less important. Conversely, there are few studies that take into account the shape of the cut (Liang et al., 2005; Coffin & Hollerer, 2006). We were not able to find a study that considered the effect of the shape of the cut on depth perception, in 3D models. According to psychologists (Bertamini & Lawson, 2008), a shape that consists of a closed line in two dimensions (2D) impacts ones impression, on perceiving it as a cutaway or an overlaid figure.

1.2 Research Hypotheses

There are three principal hypotheses that will be investigated throughout this dissertation:

(H1) Each monocular depth cue has a different impact on depth perception. Finding a cue that provides an intuitive perception of depth, in the absence of other cues, will be useful in enhancing the realism of a rendered image. Then we can use the found cue automatically at the time that the image is being rendered. Or we can

add the cue as an enhancement function to the system, that is used to generate images. Furthermore, a high dynamic range (HDR) display device can show hidden information, which exists in HDR images. This data cannot be seen with typical low dynamic range (LDR) displays. Thus, this study hypothesizes that HDR devices will be beneficial in visualization, since they will be able to show hidden information, which might be important.

(H2) Previous research revealed that blurring is useful for layer discrimination in multi-layered images, as it can be considered as a type of translucency. There are two common types of translucency, uniform and non-uniform. This study hypothesizes that the latter will provide a better result in depth-ordering tasks within multi-layered images. This will enhance the understanding of the scene, with multi-layered images.

(H3) This study hypothesises that depth perception is influenced by the shape of an incision or cut on the outer layer or surface, on multi-layered 3D models. Altering the shape of a cutaway can alter the perceived depth ordering of the layers within an image. Therefore, finding the best shape for the cut will enhance the understanding of the rendered image, more accurately.

1.3 Outline of Tasks

The following steps will be implemented to verify these hypotheses:

- A comparison will be made among the commonly used cues in the area of depth perception. This will be done to find the cue that provides the most intuitive depth impression, when it is being used individually. The comparison will be

conducted using a HDR display device, rather than a typical LDR. In addition to finding the most effective cue on depth perception, the results will show how important the use of a HDR display device is on depth perception tasks used in visualisation.

- Another comparison will be made between the two common types of translucency, uniform and non-uniform. This will attempt to find which has the strongest effect on depth ordering in multi-layered images. Of the multiple layers, only the foremost layer will be translucent.

- Three geometrically shaped cuts are compared, to determine which shape will give the strongest impression on depth perception. This study will focus on the outer surface of a three dimensional multi-layered volumetric model. It considers two previously used shapes and introduces a third.

1.4 Papers

The work in this dissertation has been published, as follows:

1. Haider K. Easa, Rafał K. Mantiuk,Ik Soo Lim. " Evaluation of monocular depth cues on a High-Dynamic-Range Display for Visualisation". ACM Transactions on Applied Perception (TAP), 15(6), 2013, pp 1555–1562.

2. H Eesa, I S Lim, D Hughes, M Jones, B Spencer. "Contour shape and perception of holes on 3D surfaces"; Abstract. Presented at European Conference on Visual Perception — 2-6 September 2012, Alghero, Italy.

3. H Eesa, I S Lim, D Hughes, M Jones, B Spencer. "Non-uniform image blur

and perceptual transparency"; Abstract. Presented at European Conference on Visual Perception — 2-6 September 2012, Alghero, Italy.

1.5 Thesis Outline

The following chapter provides an overview of the main computer graphics techniques, which have been utilized and investigated during this study. These techniques are used to enhance depth perception. The chapter highlights the depth cues and how important they are at increasing the realism of an image or scene. The development of HDR imaging and devices that are used to display such images are also described. It also discusses volume visualisation advantages, types, and areas of use. Finally, the chapter considers the most commonly used techniques in volume rendering.

Chapter three provides a literature review about previous related research. It begins by highlighting research that focus on the importance of depth cues on depth perception. It also considers research that make cue comparisons. In addition, it looks at visualisation research that has focused on cutaway techniques, relating to internal components of 3D models.

In chapter four, we consider individual monocular depth cues, looking to determine the most effective cue for depth impression. The cues considered are applied to three dimensional images. The techniques for how these images are prepared are also explained. The images are converted to HDR, and presented on a HDR display device. The cues compared are blur, brightness, contrast, overlapping, shadow, relative size and transparency. The chapter also investigates the importance of HDR images, in improving realism in computer graphics.

Chapter five looks to find whether the common types of the translucency techniques

have the same effect on depth impression. Uniform and non-uniform types are tested. A Gaussian blur function is used to simulate both types of translucency.

In chapter six, we investigate how the shape of the cut, on the outer layer of a 3D multi-layered image, effects depth perception. Three primitive shapes are tested. These shapes are circular, rectangular and an Astroid curve. We use the term "convex" to refer to the Astroid curve shape throughout this study. This is because when considered from the inside viewpoint, each side of the Astroid curve appears convex.

Finally, chapter seven summarises the study and presents the results, conclusions and future works.

Chapter 2

Background

This chapter provides an overview of the main computer graphic techniques that have been utilized and investigated during this research. We begin by reviewing how depth cues can be used in computer graphics rendering. The development of high dynamic range imaging is set to improve the realism of computer graphics and we will investigate this technology in the context of our research. Volume visualization and volume rendering techniques are then discussed. Finally we provide a synopsis of non-photorealistic rendering (NPR).

2.1 Depth cues

The human visual system perceives the world in three-dimensions (3D). It is dependant on a pair of 2D projections, one for each eye. The brain adjusts the difference between the two images and produces new information to better understand the resulted image. This information is commonly known as a *depth cue*. In other words, the brain integrates depth cues to reconstruct 3D information from two 2D images. Many depth

cues have been discovered and they differ from each other in their nature, and their impact when they are present in a scene.

Typically, the traditional depth cues can be classified into the following categories (Boynton, 2008; Harris & Jenkin, 2011; Okoshi, 2012; Sima, 2014):

1. **Oculomotor cues**: concern the cues that are based on sensing the position of the eye and muscle tension, which includes:

   (a) *Convergence*: the angle between the two pupil centres and a fixated point is changed according to the location of that point, whether it is close or far from the eyes (Richards & Miller, 1969), as shown in Figure 2.1. This angle is controlled by a set of respective muscles. The effect of this cue can be seen clearly in short distances.

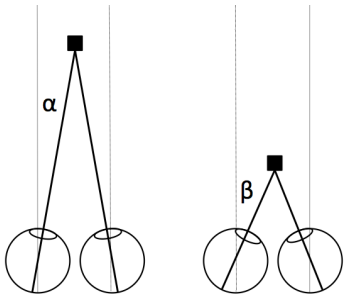


Figure 2.1: Convergence cue illustration. α and β represents the angle between a fixated point, the black square, and the direction of the eyes. The size of β is greater than the size of α since the object is closer.

   (b) *Accommodation*: when focusing on a fixed point, the eye's focal lens changes in according with the distance between the eye and the point. This cues also does not work well in relation to far objects.

2. **Binocular cue**: includes only one cue:

(a) *Binocular Disparity*: also known as *Binocular Parallax*, which means the difference between the views of both eyes. The location of our eyes plays an important role in seeing the world in 3D. Since, each eye takes an individual view, the location of each eye allows us to generate two similar, but distinct views with a slightly differing viewing angle. The brain analyses these disparate images to determine true stereoscopic depth.

3. **Monocular cues**: deal with cues that are derived from one eye, but also available using both eyes (Mather, 2006). These cues are divided into two main categories:

   (a) **Static cues**: also known as *pictorial*, because these cues have been used by artists to draw realistic 3D scenes (Solso, 1996). The source of depth information comes from 2D images, and it includes the following cues:

      i. *Overlap*, also known as *occlusion* and *interposition*: it appears when a nearest object partially occludes the furthest object.

      ii. *Relative height*: objects that appear in higher regions of a scene are further, in distance, than objects that appear in the lower regions of the scene.

      iii. *Relative size*: with equal size objects, the smaller one in an image is the further in distance.

      iv. *Perspective convergenc*e: parallel vertical lines converge into a single point after a certain distance.

      v. *Familiar size*: the size of objects in a scene reflects their distance information. This cue is mainly based on knowledge,

      vi. *Atmospheric perspective*, also known as *Aerial perspective*: further objects look less saturated and also less sharp. In other words, the

distant object looks hazier and has a blue tint.

vii. *Texture gradient*: closer objects look sharper, i.e. the details of the objects surface texture is seen clearer. Contrarily, a smoothed texture indicates an object is further in the distance.

viii. *Brightness*: closer objects normally appear brighter than the furthest objects.

ix. *Shadow*: when the source of light is known, objects that cast shadows on other objects are closer in distance to the light source.

(b) **Motion-based cues**: these cue can be seen clearly through movement, and include the following:

i. *Motion parallax*: in the direction of movement, near objects move rapidly, whilst far objects appear to move slowly.

ii. *Deletion and accretion*: objects are covered or uncovered according to our movement in relation to them.

More recently, psychological studies have introduced *blurring* as a pictorial cue Mather (1996). It is produced when using modern images that are generated by photographic equipment. In addition, O'Shea et al. (1994) claimed that the *contrast* cue, which may be thought of as a consequence of atmospheric perspective, be useful to estimate distance, and the furthest objects display lower contrast values.

All these cues help us to estimate size and distance of objects portrayed in a 3D world, by giving a more accurate interpretation, regarding the perceived 2D images (Held & Cooper, 2010). Depth cues have also been classified as physiological and psychological cues (Cutting & VIshoton, 1995). For example, when judging distance using monocular cues, we depend more on cognitive psychological responses. This

is why we see different reactions to a group of people observing the same image. Therefore, all monocular cues are considered as psychological cues. However, with oculomotor cues, we use physiological mechanisms, such as pupil dilation and eye convergence to judge distance.

In this study, we focused on the monocular cues, especially the static cues. The reason was because they are generally utilised in computer graphic imagery, and it is difficult to test the oculomotor cues without using special devices. As previously mentioned, the depth cues that are investigated in this study are: blur, contrast, shadow, relative size, brightness, overlap and transparency. Although transparency is not one of the depth cues, it is commonly used in visualisation of multiple-layered images and in many computer graphic applications. We chose these cues because they can be used for depth ordering in a multi-layered images. In the following sections, a brief introduction is given for each type of these cues.

2.1.1 Blur cue

Two main types of blurring are common (Davis, 2008): defocus, also called *out-of-focus* (Liu et al., 2008), and motion blur. Defocus is usually caused by the incorrect focus of an object, while the motion blur type is the result of taking a picture of moving objects, as shown in Figure 2.2. Blurring is used to reduce sharpness and contrast of on object. In motion blur, we can infer the direction of movement from the blurred portion. The defocus blur types includes uniform and non-uniform. In the uniform type, every part of the image has the same amount of blurring. Conversely, each part of the image has a different amount of blurring using the non-uniform type. Frosted glass is an example of the latter.

Figure 2.2: Motion blur example. Image re-produced from Flicker as London bus.

2.1.2 Contrast cue

As a simple definition, contrast is a ratio of the difference between patch luminance and its background over average luminance. This means that if the average is high then a small difference will be ineffective, whilst the same amount of difference will be effective when the average is low (Legge, 1981). The following equation, which is known as Weber contrast, is used to find the contrast of the images that include a large uniform background:

$$C = \frac{I - I_b}{I_b}, \tag{2.1}$$

where I represents the intensity of a part in the image and, I_b is the intensity of the background of the image. Michelson (1995) produced another way to calculate the contrast of an image, known as Michelson contrast, which is defined as follows:

$$C = \frac{I_{max} - I_{min}}{I_{max} + I_{min}}, \tag{2.2}$$

where I_{max} and I_{min} represent the maximum and the minimum intensity values of the image respectively. The equation 2.2 is used to find the contrast of images that contain periodic patterns and also in the models that describe transparency (Singh &

Anderson, 2002b; Kasrai, 2004). Ichihara et al. (2007) classified contrast into two main types, *area contrast* and *texture contrast*. Area contrast can be defined as the difference between the average luminance of the surface area of a part of an image and the average luminance of the background. The left column of the Figure 2.3 shows a high area contrast while the right column presents a low area contrast. The texture contrast is the difference between the luminosity of the characteristics of an object. The first row in the Figure 2.3, has the low texture contrast compared with the second row, which has high texture contrast.

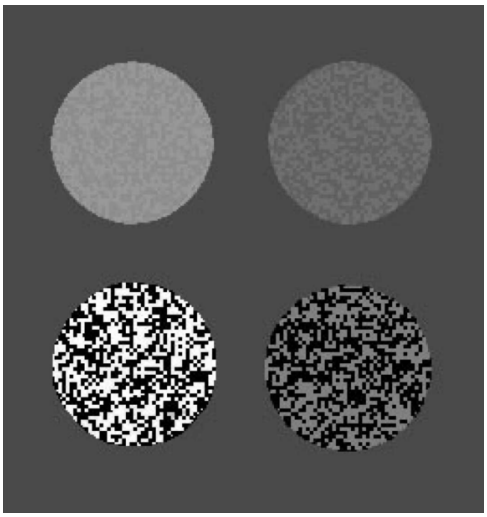


Figure 2.3: Contrast types. High area contrast can be seen in the left column and high texture contrast can be seen in the second row. Image re-produced from (Ichihara et al., 2007).

2.1.3 Shadow cue

Shadows can be defined as the consequence of light occlusion triggered by an object while shading is the results from a change in the angle of a 3D surface normal, with respect to the light direction (Kingdom, 2008).

Consequently, shading models are used extensively in computer graphics applica-

tions to provide the interaction of light with surfaces. They can be used to express 3D shape on flat 2D display screens (O'Shea et al., 2008). Shading models consist of four basic components (Ware, 2013):

- **Lambertian shading**: the light is reflected from a surface equally in all directions.
- **Specular shading**: the light is reflected from a glossy surface.
- **Ambient shading**: the light comes from the surrounding environment.
- **Cast shadows**: the shadows cast by an object, either on itself or on other objects.

Lately, ambient occlusion (AO) has been used in many computer graphics applications to simulate shadows and shading. It increases the realism in computer based imagery, see Figures 2.4. AO is a visual effect that was originally developed by Industrial Light and Magic ILM (GEFORCE, 2014). AO does not depend on the direction of light, therefore, it can be pre-computed for static objects.

(a) (b)

Figure 2.4: A pair of screenshots, from the HALF Life 2 game, to illustrate the effect of applying ambient occlusion AO. The Ao was not enabled in the image (a), therefore, there is no sense of depth between the phone and the wall against. The same in the image (b), a realistic shadow effect can be seen clearly. Image re-produced from (GEFORCE, 2014).

2.1.4 Relative Size cue

The size of an object serves as a cue to determine its location in a scene (Ittelson, 1951). Hochberg & McAlister (1955) distinguished between *familiar size* and *relative size*. They argued that familiar size requires only one image of stimulus to determine the distance, whilst in the relative size, two or more similar or identical shapes of different size are required. This is because this cue depends mainly on previous experience. Thus, larger shapes tend to be nearer to the viewer. However, the exact opposite result is seen in the "corridor illusion" phenomenon (Fineman, 1981), as shown in the Figure 2.5. This phenomenon appears when two or more objects of the same size are located in the near and far end of a corridor; without accounting for the perspective projection. Aks (1996) claimed that the linear perspective of the corridor makes the object in the near end of the corridor appear larger, even though the geometric size of all objects is exactly the same.

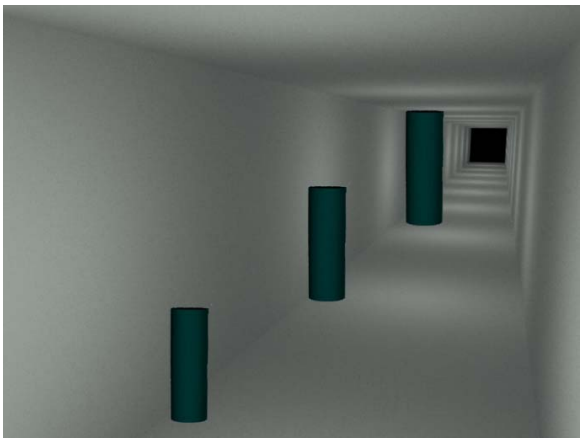
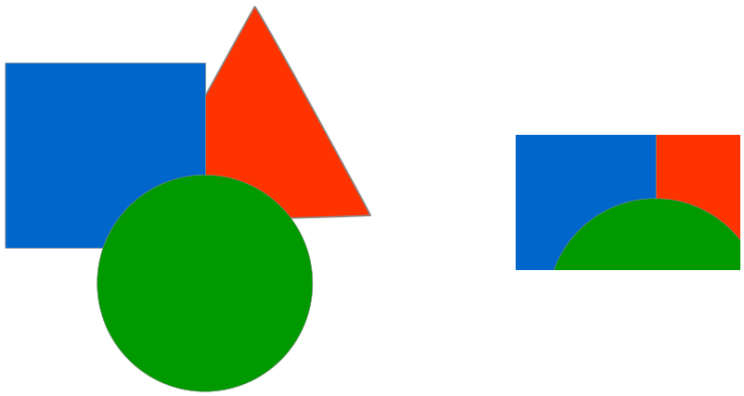


Figure 2.5: The Corridor illusion phenomenon, all bars are the same size, but the furthest bar seems bigger because of the linear perspective of the corridor.

2.1.5 Overlapping cue

Overlapping, also known as *interposition* or *occlusion*, is one of the monocular cues that occurs in an image when it contains many objects, which overlap fully or partially. The occluded object appears to be further away from the observer. Therefore, one of our visual systems task is the completion of clutter or occluded object boundaries and depth assignment of overlapped boundaries (Hund & Mertsching, 2009). Figure 2.6(a) illustrates the effect of the overlapping cue on depth ordering tasks. Since all the shapes in the image are well known, it is easy to order these shapes using the overlapping cue. The circle shape is fully displayed, so it should be the closest one to the observer. The rectangle looks to be the second closest shape to the observer, because a part of it is occluded by the first (the circle shape) and it, in turn, occludes the triangle shape, which seems to be the furthest shape from the observer. This cue works perfectly with known objects or shapes, but with random, abstract or unknown shapes the ordering task will be difficult, as shown in Figure 2.6(b).

Psychological studies have tried to determine a distinction between specific shapes and their relation with depth ordering tasks. It is claimed that closed-line shapes can be perceived as an object (also called a figure) or as a hole (referred to as ground). This process is called *figure-ground segmentation* which, in turn, can determine whether a shape's order is in front or at the back (Kanizsa & Gerbino, 1976; Kanizsa & Kanizsa, 1979; Koenderink et al., 1984). They argue that the convexity is the main factor of figure-ground segmentation and convex regions are perceived as a *figure*.



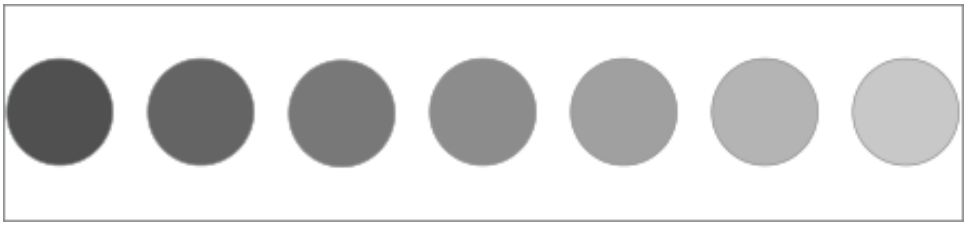
(a) (b)

Figure 2.6: An illustration of the overlapping cue. In (a), it is easy to order these shapes using this cue. The green circle is the nearest shape to the viewer since it is shown completely, and because a part of the rectangle shape occludes the triangle shape, the rectangle will be the second closest shape, and the triangle is the furthest shape. In (b), it is difficult to tell the order of these *unknown* shapes.

2.1.6 Brightness cue

An object can be described according to its light reflection as: bright, dim and dark. Increasing or decreasing the amount of the brightness of a colour can be achieved by adding black or white to that colour (Cognates, 2013). Figure 2.7(a) shows that the brightness of an object also can be altered by changing its background intensity (Vladusich et al., 2007). The furthest right disk looks brighter than the others, whereas, in fact, all disks in the image have the same luminance value. This occurs because of the effect of the background luminance, which is varied along a horizontal gradient. The background colour ranges from the high brightness, on the left, to the low brightness. The term “high” brightness refers to a colour that is close to white on a brightness scale. In Figure 2.7(b), the disks have different luminance values while the background remains unchanged.

(a)



(b)

Figure 2.7: Brightness can be acheived either by changing the background luminance in a gradient and keeping the objects' luminance unchanged (a) or, by changing the luminance value of the objects (b).

2.1.7 Transparency and Translucency

A property of a surface that allows a certain amount of light to pass through is called transparency (Kasrai, 2004). It gives an impression of multiple layers or multiple surfaces when it occur in an image. The first quantitative model to assess visual perception of transparency was Mettlli's episcotister model (Metelli et al., 1985). Based on this model, and in addition to the modification made on this model by Singh & Anderson (2002b),Chan et al. (2009), they developed a system to optimize transparency in volume rendering automatically, to enhance the impression of a transparent structure. The transparency phenomenon may be divided into the following main classes: *media*, *substances* and *layers*. The media class can be seen in fog and water. Substances can be seen in coloured liquid and glasses. And finally, layers, sometimes called *filter*, have been studied widely because of the simplicity in generating stimuli for psychophysical research (Gilchrist, 2013). A transparent surface differs from a translucent surface,

in that the translucent surface diffuses the light that is passed through it. The ground layer underneath the translucent surface, in a multi-layered images, looks distorted or blurred (Singh & Anderson, 2002a).

Transparency is one of the techniques that is used in visualization of 3D volumes (Bruckner & Groller, 2005). It is also known as the *Ghost* technique. In this technique, rather than removing the outer surfaces of a multi-layered image by cutting, its opacity is decreased to become semi-transparent which in turn reveals the hidden parts. Such a technique is useful in visualising known objects, such as the human body. The outer surface (the skin) is well-known and it is considered as less important when comparing with the internal parts of interest. Moreover, we might use our knowledge about the anatomy to guess the location of the internal parts without focusing on its exact depth locations. Transparency and colouring are the basic concepts of volume rendering which are used consistently in visualization of 3D medical images.

Reflection of light on object surfaces differ depending on the material that the object is made of. On an ideal reflective surface, the surface reflects the light ray without allowing it to pass through. However, when a light ray bleeds through a translucent material, it gives the surface a special visual characteristic such as a visual softness and glow (Thompson et al., 2011), such as wax and frosted glass. Normally, the amount of lights to pass through a translucent object depends on how thick the object is. Thinner translucent objects allow more light to pass through and it look more transparent. Dealing with translucent surfaces, there is a phenomenon called *subsurface scattering* (Fleming & Ulthoff, 2005). When light strikes a real translucent surface, it passes through, scattering the light beam many times with the internal material, and then emerging to another location on that surface or passes through it, as shown in Figure 2.8.

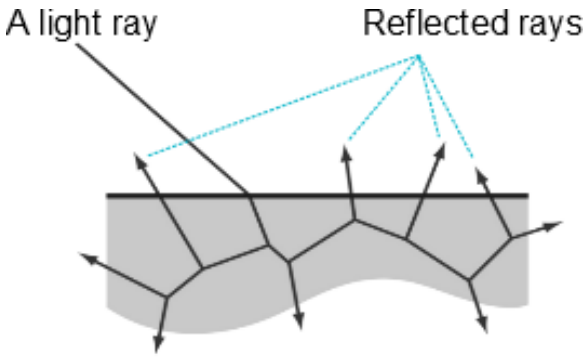


Figure 2.8: A translucent surface, the light passes through the first layer of the surface and scattes many times with internal material which produces new rays. Some of these rays fully pass through the surface and others reflect in different locations from the surface. Image reproduced from (Thompson et al., 2011).

2.1.7.1 X-Junctions

According to psychological studies, transparency and luminance play an important role in perceiving the depth ordering of multiple semi-transparent layers. When two or more layers are partially overlapped and the front layer is semi-transparent they give rise to X-junctions, whereas T-junctions appear as if the front layer is opaque (Watanabe & Cavanagh, 1993) as shown in Figure 2.6. X-junctions are seen where the contours of a transparent material cross contours of the surfaces behind (Sayim & Cavanagh, 2011). Based on a model in Adelson & Anandan (1990), the type of transparency depends on the type of X-junction. Figure 2.9(a), shows a typical X-junction model. Let p, q, r and s represent four regions surrounding the X-junction which is indicated as a red point in the image. Let l_p, l_q, l_r and l_s be the luminance values for the four regions respectively.

In Figure 2.9 (b), $l_p > l_q$ and $l_r > l_s$ which means that the vertical edge retains the same sign in both halves of the X-junction. Likewise, $l_p < l_r$ and $l_q < l_s$, the horizontal edge also retains the same sign in both halves. This type of X-junction called "Non-

reversing" because both edges retain their sign. Transparency can be seen in this type of X-junction but it is difficult to tell the depth order.

In Figure 2.9 (c), the horizontal edge retains the same sign whilst the vertical edge sign is changed. This type is called "single-reversing", since the sign is changed only on one edge. In this type of X-junction, the transparency can be seen and the depth order is also clear. Finally, in Figure 2.9 (d), the X-junction is called "double-reversing", since the sign of both edges are changed. In the double-reversing type, the transparency cannot be seen and it is difficult to see the depth order. The correct depth order can be determined only for the single-reversing type (Anderson, 1997; Singh & Anderson, 2002b; Delogu et al., 2010).

2.2 High Dynamic Range (HDR)

Nowadays, there are many applications that are focused on digital images. These images are normally created using either digital devices, such as modern cameras and scanners, or using drawing programs. A standard imaging tools cannot record the whole range of luminosity that is found in everyday scenes and because of the restricted capability of standard image capture apparatus, lots of information is lost (Mann & Picard, 1994). Different lighting environments have different ambient luminance levels. For example, starlight has an illumination of 10^{-3} cd/m^2, moonlight of 10^{-1} cd/m^2 and sunlight has 10^5 cd/m^2 (Wandell, 1995; Ferwerda, 2001). Figure 2.10 shows an image shot with only daylight luminance using both a standard and HDR image capture device. The HDR image reveals information that was invisible in the standard shot. This loss of information may effect the full understanding of the shaded part of the image and using HDR imaging would be a positive solution. HDR images contain

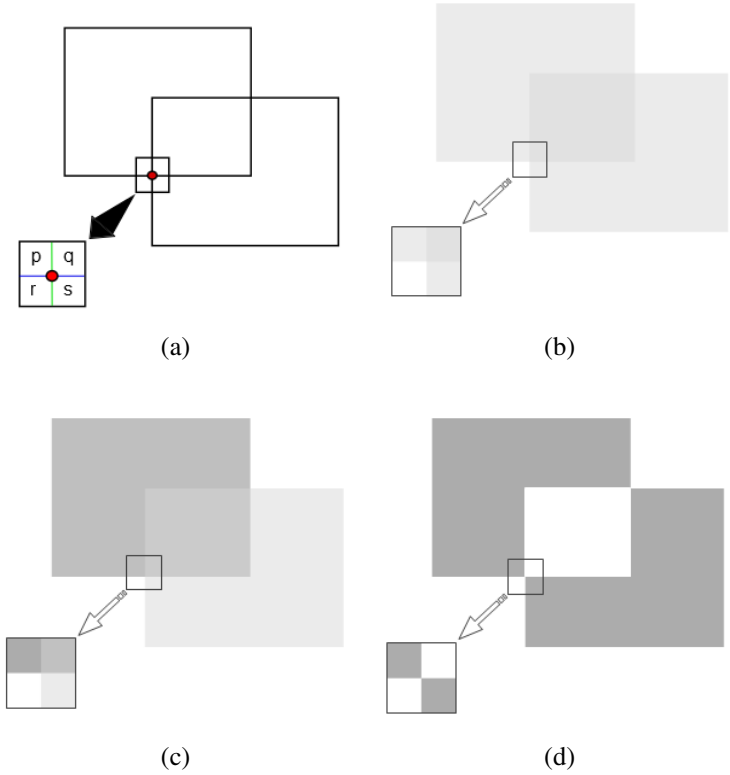


Figure 2.9: X-junctions occur when two semi-transparent layers are partially overlapped. (a) An illustration of X-junction marked by a red point. (b) *non-reversing* junction produces an ambiguous depth order but the transparency can be seen. (c) *single-reversing* junction provides the correct depth order and a clear transparency. (d) *double-reversing* junction does not evoke the impression of transparency and an ambiguous depth order.

an immense range of intensity level compared to standard images. In this study, an experiment is conducted using a HDR display, as seen in Figure 2.11, to determine whether HDR can provide a better impression of depth order when visualizing abstract data on 2D displays.

In order to approximate the appearance of a HDR image in a standard range display, a technique called *tone mapping* was used to map one set of colours to another (Ledda et al., 2005; Mantiuk et al., 2006b). While conventional displays are usually restricted by the peak luminance of 200-500 cd/m^2 and the contrast 1 000:1,

(a) (b)

Figure 2.10: A lot of detail is revealed when using an HDR image (b) compared to a standard image. Image re-produced from (Reinhard et al., 2010).

HDR displays can produce as high luminance as 4 000-8 000 cd/m^2 and contrast exceeding 10 000:1 (Seetzen et al., 2004). Such a luminance range is much closer to that found in natural physical scene. The images shown on such displays are strongly preferred (Daly et al., 2013) and there are also indications that the expanded contrast range can enhance contrast perception (Rempel et al., 2011).

2.3 Volumetric Visualization

The main goal of a volume visualization is to see the interior parts of a volumetric dataset. One of its benefits is to give a quick look at the contents of the internal structure of a 3D model, which reduces thousands of words that might be used to describe the internal parts. The field of visualization started in the nineteen eighties, after a report was published in the U.S. National Science Foundation (McCormick, 1988; Viola,

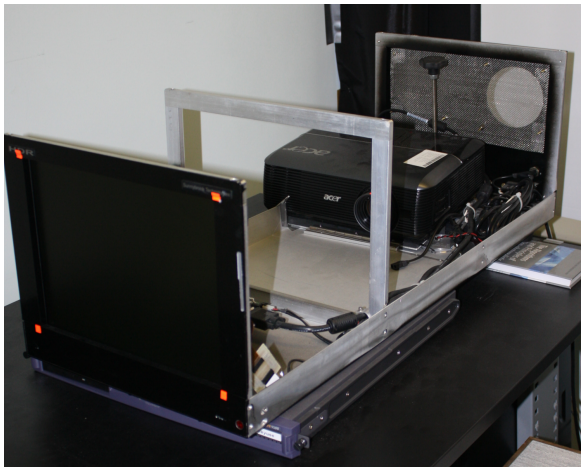


Figure 2.11: HDR display device at Bangor university which consisted of a 4 000 lumen projector and a 15" LCD panel with a resolution of 1024×768 pixels. The device's peak luminance was 2 446 cd/m^2 and the black level was 0.01 cd/m^2.

2008; Moreland, 2013), which stated the new challenges in using scientific visualization. It was used in many different areas, for example, education, medicine, industry and applied sciences. As a consequence, it has been developed quickly to keep pace with the urgent need to make use of it.

Volumetric data are sets of 3D entities, called *voxels* that represent 3D scalar field. These data sets are usually comprised from a group of 2D slices that are obtained using different scanning techniques such as, computed tomography (CT), magnetic resonance Imaging (MRI), or positron emission tomography (PET) (Meißner et al., 2000). These techniques are used at most to scan the internals of the human body. The main objective of volume visualisation is to represent the full data set at the same time (Deserno, 2011) and then, extract the significant information from the volumetric data Hansen & Johnson (2005). There are two basic specialities in volume visualization; realistic rendering and scientific visualization (or non-photorealistic rendering NPR) (Hughes & Lim, 2009). Realistic rendering attempts to visually simulate how

an optical phenomenon translates form real life scenarios, such as how light interacts with materials and how fluid particles move in fluid dynamics, simulating pressure and viscosity (Yagel et al., 1991; Deussen et al., 1998; Shirley & Morley, 2008), while scientific visualization simply tries to give a visual view of data with more focus on portraying the information of interest. Volume visualization plays an important role in many modern sciences, medicine, education, and industry. For instance, getting high-resolution optical slices of a microscopic object, oil exploration and analysing geo seismic data. Volumetric medical data has seen much research over many years. Last but not least, a photorealism product can be viewed, tested and improved before it is actually manufactured (Yagel et al., 1991; Meißner et al., 2000; Hadwiger et al., 2008). In general, there are two major approaches, which typically utilised in volume visualisation; indirect and direct volume rendering (Deserno, 2011; Bruckner et al., 2005; Hansen & Johnson, 2005; Engel et al., 2001).

2.3.1 Indirect and Direct Volume Rendering

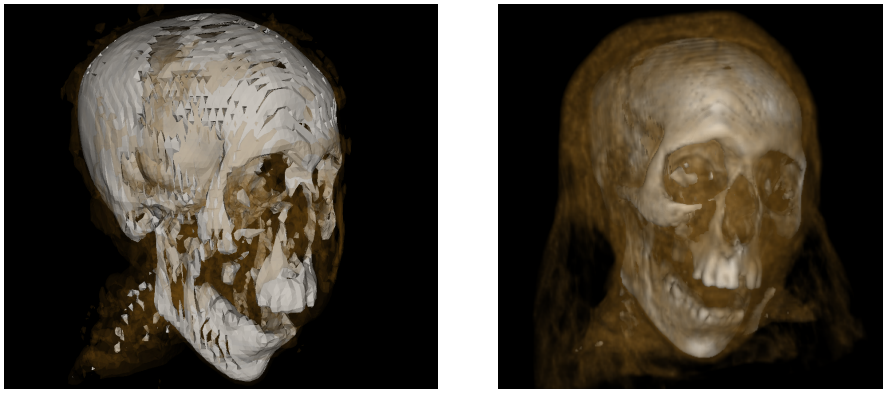
One of the first studies in the topic of volume rendering was in 1988, when Marc Levoy (Levoy, 1988) claimed that volume rendering is an effective method for the display of surfaces from sampled scalar functions of three spatial dimensions. His technique developed form his earlier work on the use of points as a rendering primitive (Levoy & Whitted, 1985). Indirect volume rendering (IVR), also called *surface* - or *isosurface* - rendering, extract an intermediate polygonal representation from the volume dataset and then render this representation. "marching cubes" and *seed-sets* are the widely used algorithms for the extraction (Lorensen & Cline, 1987; Hansen & Johnson, 2005; Newman & Yi, 2006; Yang et al., 2008; Wenger, 2013). The main limitations of the isosurface methods are that they can only be used with specific blend-

ings, they are view dependent and only boundaries can be represented (Boucheny et al., 2009). It is suited to a particular type of data sets (CT scans), but it is not always appropriate to use, for example, when simulating fire.

Direct volume rendering (DVR) generates the visual representation of the entire structure without an intermediate process, by projecting the voxels directly onto the image plane. It can be defined as a process of creating a 2D projection directly from 3D volumetric data set, this is why it called *direct volume rendering*. The DVR's main idea is to represent data in a semi-transparent manner, so the volume is used as a whole, which means the structure inside of the volume can be seen. Thus, the gaseous phenomena can be simulated using this type of rendering. DVR provides greater flexibility than the isosurface method. Every voxel in the volume dataset should contribute to the image and its colour and opacity can be controlled. Therefore, a specific region can be shown or focused on using DVR. Volume ray casting, splatting, shear-warp, and texture mapping are the main algorithms that are being used in DVR. Figure 2.12 shows skin and bone of a mummy dataset rendered using isosurface and DVR.

2.4 Non-Photorealistic Rendering (NPR)

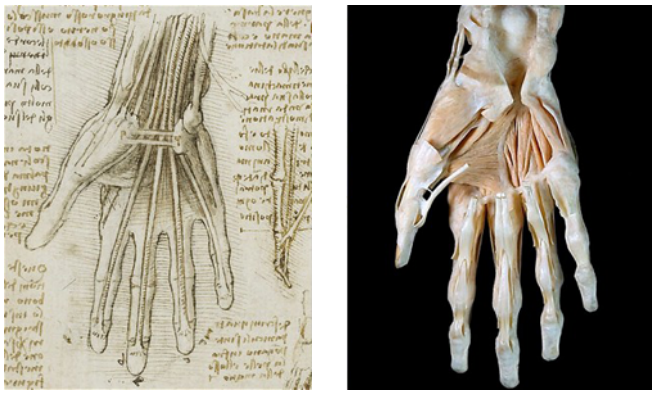
Illustrative visualization is one of the most important topics in computer graphics, which provides an effective way to display 3D models with the ability to focus only on the important features, or parts of interest (Gooch & Gooch, 2001; Rheingans & Ebert, 2001; Strothotte & Schlechtweg, 2002; Chen et al., 2011). Artists were the first who tried visualising the internal contents of volumetric datasets. For example, Leonardo Da Vinci [1452-1519] might be considered as one of the first artists who



(a) (b)

Figure 2.12: Mummy dataset rendered using (a) isosurfave, and direct volume rendering DVR (b). Image re-produced from (Scheidegger, 2007).

tried to reveal the interior parts of the human body and also the internal structure of his inventions, when he started to draw anatomical sketches of the human body centuries ago, as shown in Figure 2.13(a).



(a) (b)

Figure 2.13: Examples of a hand visualisation. (a) A hand drawn by Da Vinci, Image re-produced from ©Royal Collection Trust HM Queen Elizabeth II 2013. (b) A hand scanned and drawn by a computer. Image reproduced from ©Mark Mobley, West Midlands Surgical Training Centre.

Da Vinci's anatomical drawings have shown amazing accuracy when compared with new medical scans, as shown in Figure 2.13(b). Painters have been relying on

the illustrations in various fields of science until the advent of computer graphics in the nineteen sixties. Some of the paintings have extremely high precision, which are difficult to distinguish from photographs that are captured by camera. This type of painting is called *photorealism* (Gooch & Gooch, 2001).

Rendering 3D landscape scenes in real-time is considered to be one of the most challenging task in computer graphics (Coconu et al., 2006). Deussen et al. (1998) rendered a realistic outdoor plant model, but to achieve real-time rendering they used a non-photorealism models (Deussen & Strothotte, 2000; Coconu et al., 2006) rather than photorealism, as shown in Figure 2.14. Non-photorealistic is simple and it can be produced quickly compared to photorealistic. It is suitable for the topics that deal with explanations and illustration, whilst photorealistic is appropriate to the topic that deal with documentation and simulation (Adam, 2005). Bruckner & Gröller (2007) reproduced the anatomical illustration style found in medical books using non-photorealistic rendering.

2.5 Summary

In this chapter, the foundations of all of the topics that will be covered in this research are highlighted. Every topic is defined briefly, focusing on its main uses and types, if there are more than one type. The chapter begins with an explanation of depth cues and their main types. Followed by a brief description about the development of HDR imaging and devices, with an explanation of how these can increase the enhancement of computerised imaging. The chapter gives a synopsis about volumetric visualisation, as well as the main volume rendering types used. Finally, an outline definition about NPR is also provided.

(a)

(b)

Figure 2.14: Landscape rendering uisng,(a) photorealistic, and (b) non-photorealistic. Images re-produced from (Coconu et al., 2006).

Chapter 3

Related Work

This chapter discusses research that relates to this dissertation. The chapter begins highlighting research that focused on the importance of using depth cues to improve depth perception. In addition, the studies that clarify the benefit of these cues to increase the understanding of a scene, are reviewed. It also highlights the studies that encourage the use of HDR technologies. This chapter also discusses studies that use the cutaway techniques, which are used to reveal the internal parts of 3D volumes; emphasising only studies that used geometrical shapes for the cutaway process.

3.1 Depth cues

As previously mentioned, the depth cues that are investigated in this dissertation are blur, contrast, shadow, relative size, brightness, overlap, and transparency. Although the transparency cue is not classically known as one of the depth cues, it is commonly used in visualisation of multiple layer modelling. These cues are usually utilised in computer graphic imagery. They have been chosen based on the results of previous

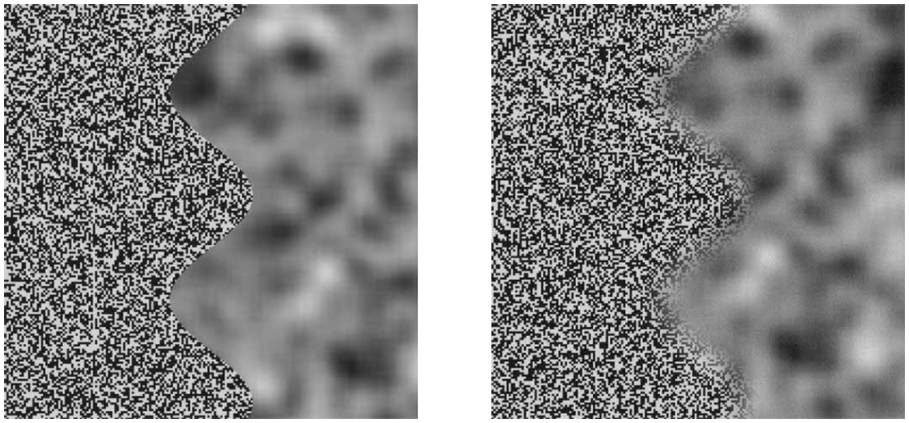
research, which present their effect on depth ordering. The following sections highlight the studies that focus on the depth cues that will be used in this study, which in turn clarifies the reason of selecting these cues.

3.1.1 Blur Cue

Mather (1996) claimed that the blurring feature can be used as a depth cue, and it can be useful in depth ordering tasks (Mather & Smith, 2000). Mather contended that the converging contour, when blurred and focused parts converge, will play an important role in determining which part is closer to the viewer. If the converging contour is sharp, the focused part will be seen as closer, and if it is blurred, then the blurred part, in this situation, will be seen as the closer, as shown in the Figures 3.1(a) and (b) respectively. To support his idea, Mather conducted an experiment (Mather & Smith, 2002) and asked observers to choose the closest region to them between two regions; focused and blurred. The results confirmed the studies expectations, where most of the participants chose the blurred part as being the closer part, when the merging contour line was blurred.

Marshall et al. (1996) discussed how the blur cue can be utilized to establish depth-order in a multi-layered image. They hold the same point of view as Mather, in which the edge that separates the blurred and focused parts (of two adjacent parts) plays an important role in depth order judgement. O'Shea & Govan (1997) claimed that the blur cue is more effective on depth perception, as a pictorial cue comparing with the contrast cue because of its usual association with other depth cues, particularly in pictures and photographs. Held & Cooper (2010) claimed that the blur cue can be used in conjunction with other pictorial cues to estimate the absolute distances between the

objects in a scene.



(a) Sharp edge (b) Blurred edge

Figure 3.1: Mather & Smith (2002) argued that the edge between two regions; blurred and sharp, plays an important role in depth judgement. In (a) most of the observers selected the sharp region as being closer, whilst in (b), most of the observers selected the blurred part as being the closest. Images re-produced from (Mather & Smith, 2002).

3.1.2 Contrast Cue

O'Shea et al. (1994) claimed that the contrast cue is an effective depth cue, in the absence of other depth information, and it has a significant effect on depth perception. They argued that when two areas that have a different contrast value, the area that has a lower contrast looks further. O'Shea & Govan (1997) tried to ascertain whether blur and contrast cues have a separate effect on depth perception when they used in an image. Both cues (blur and contrast) influence depth perception. However, the contrast of an image is affected when the image is blurred. They found that both cues influence depth perception separately. However, the greatest effect of the blur cue was observed when using a moderate contrast.

Ichihara et al. (2007) claimed that the contrast cue can be helpful for identifying the depth of an object. For example, in Figure 3.2, and because the *texture contrast* in the lower row is higher than those in the upper row, the lower row seems closer. Likewise, because of the *area contrast* in the left column is higher than those in the right column, the left column looks closer. As a conclusion in both cases, if an object has a high contrast level, it will seem closer in distance to the observer. Rempel et al. (2011) tried to examine the effect of contrast on depth perception by conducting a series of experiments. They found that the observers perceived increases in contrast to correspond with increases in perceived depth. They argued that according to their results, they can simulate sensations of depth by manipulating contrast.

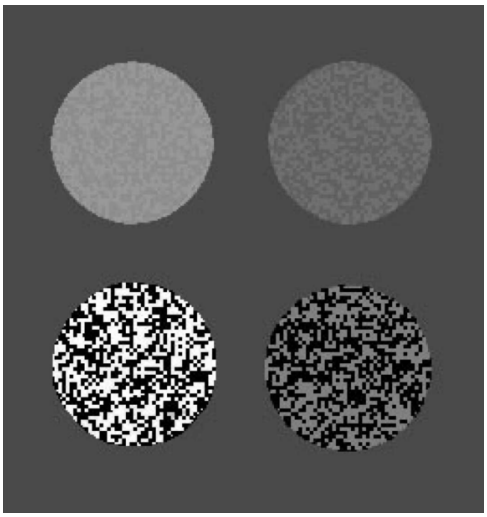


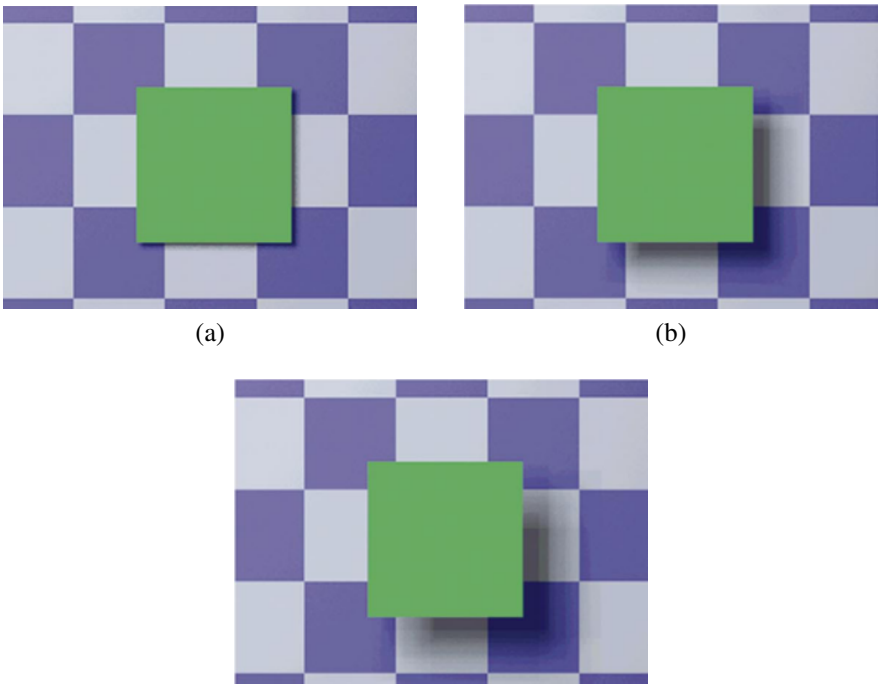
Figure 3.2: Illustration of the influence of *texture contrast* and *area contrast* on depth perception. The left column looks closer because of its *area contrast* is higher than the second column, and the lower row seems closer because its *texture contrast* is higher than the upper row. Image re-produced from (Ichihara et al., 2007).

3.1.3 Shadow Cue

The depth of objects can be estimated according to its cast shadow Mamassian et al. (1998). Sugano et al. (2003) argued the effect of the shadow cue of virtual objects plays an important role in Augmented Reality (AR). Elder et al. (2004) tried to examine the role of shadows in the rapid discrimination of scene properties. They found that depending on the type of shadow (cast or attached), the direction of the shadow, and the displacement of the shadow, the visual system is capable of rapid discrimination of depth. Rensink & Cavanagh (2004) also claimed that cast shadows can have a significant influence on the speed of visual search. Figure 3.3 illustrates the effect of casting shadow on the impression of depth. The green rectangle is in the same position in all images. Because each image has a different amount of shadow spread, the green rectangle looks to be at different distances from the background layer.

3.1.4 Overlapping Cue

Convexity can be considered as a factor that affects figure-ground segmentation (Kanizsa & Gerbino, 1976). Figure-ground refers to the visual perception that a contour separating two regions belonging to one of the regions (Fowlkes et al., 2007). According to a prior expectation that favours convexities as a figure, Bertamini & Lawson (2008) claimed that they have an objective evidence that supports the idea that convexity affects figure-ground assignment. They argued that the response in the figure-ground segmentation task is faster when the surface in front is bounded by a convex contour, rather than a straight contour. Zheng et al. (2013) thought that occlusion is an intuitive monocular depth cue, but it is inappropriate in volume rendering because the structures are rendered semi-transparently. We named this particular cue a *convex cue*, which is



(a) (b)

(c)

Figure 3.3: The effect of a cast shadow on depth perception. The green rectangle is in the same location in the all images but it looks to be at a different distances from the background, because of the different amount of shadow spread. Images re-produced from (Mamassian et al., 1998).

featured in the experiment in chapter Four of this thesis.

3.1.5 Brightness Cue

In general, and with multiple objects, the nearer object to the light source looks brighter than the further object. The brightness cue could play an important role in depth ordering tasks. Swain (2000) claimed that increasing the luminance of an object of interest will make that object look closer. Farnè (1977) had a different viewpoint about the brightness cue. He argued that increasing the brightness of an object is not enough to make it looks closer. He claimed that the high contrast between an object and its

background influences more in distance perception than increasing its brightness. This will makes the object seems closer, even if it is dark. In Figure 3.4, the black (dark) disk looks closer to the viewer over the brighter disk.

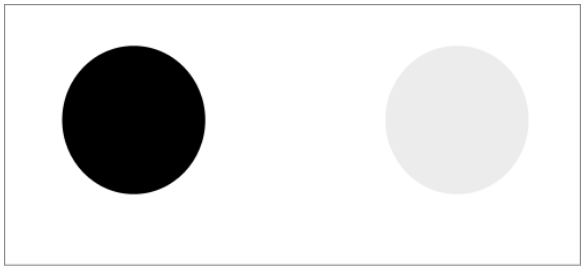


Figure 3.4: According to Farnè (1977) the high contrast between an object and the background makes the objects looks closer, the black disk appears closer while the brighter disk looks farther from the observer. Image re-produced from (Farnè, 1977).

3.1.6 Transparency Cue

According to previous psychological studies, transparency and luminance can be considered as important cues for depth ordering (Adelson & Anandan, 1990). The ambiguity that accompanies transparency in some of the X-junction cases is a problem that can be overcome. Everitt (2001) presented a technique called *depth peeling* to sort the depth of multiple transparent surfaces or objects. Zheng et al. (2013) produced an approach to enhance the perception of depth order in volume rendered images without using additional cues.

3.1.7 Relative Size

It is obvious that for familiar sized objects or objects that have an equal size, the closer object appears larger than the furthest one. Ittelson (1951) argued that the size of an object serves as a cue to its localisation in visual space. Epstein (1961) claimed that

the *known size* cue can be utilised to determine the distance. He argued that discrete changes in the size of an object, whose known size remains constant, are perceived as corresponding changes in distance. We utilised this cue during this study as a control condition.

3.2 Depth Perception

Depth perception has been studied extensively. Many of these studies are mainly focused on how to improve the depth effect, using depth cues. However, some studies used a combination of depth cues to enhance the impression of depth. The number and the type of depth cues are not equally considered in these studies. However, there has been no evaluation of depth cues. The following section illustrates some samples of such studies.

Landy et al. (1995) argued that several depth cues available in a scene encourage the visual system to combine them. They tried to provide a way to analyse depth cue combinations. Swain (1997) argued that depth perception in a 2D image can be improved by adding monocular depth cues. He used four monocular depth cues in his study. They were blur, shadow, brightness, and overlap (occlusion). He claimed that a *pseudo* 3D image can be generated from a single 2D image using these cues. Thus, only a single 2D image was required in his study. However, to apply his algorithm, the original image should be first segmented into objects. Depth cues are applied to segmented images then the result is combined with the original image.

Mather & Smith (2004) claimed that the number of depth cues available in a multi-layered image will increase both speed and accuracy in depth-ordering tasks. They used three depth cues in their study; blur, contrast, and overlap. They found that the

observers responded faster when all cues were presented, over a single cue at a time. They proved that more depth cues in an image will give a stronger impression on depth perception than only one.

Saxena et al. (2005) argued that depth estimation is a challenging problem. They tried to solve the problem by making a combination between binocular and monocular cues to enhance depth estimation (Saxena et al., 2007). They claimed that by adding monocular depth to stereo vision (stereopsis) significantly improves depth estimation, over using each type of cues alone. They used texture, blur, and haze monocular depth cues in their study. Based on their results, they tried to reconstruct 3D depth from a single still image (Saxena et al., 2008).

Luft et al. (2006) presented a simple method to enhance the depth perception quality of images that contain depth information. Their idea was motivated by artwork. They used a low-pass filter to produce a shadow-like effect, which then was utilised to determine information about spatially important areas of the scene.

There are many studies that consider combining depth cues, detailing them all exceeds the scope of this thesis. However, there are a few studies that try to evaluate depth cues. Surdick (1994) conducted an experiment to test several depth cues to find which one provides effective depth information. They tested seven depth cues; brightness, relative size, relative height, linear perspective, foreshortening, texture gradient, and stereopsis. They also tried to find whether the effectiveness is altered when the viewing distance is being changed. Two viewing distances were used; one and two meters. The stimuli were viewed dichoptically through a modified WeatStone stereoscopic device, constructed with first-surface mirrors (Arditi, 1986). They found that the effectiveness of the perspective cues (linear perspective, foreshortening, and texture gradient) were superior to the other cues. Relative brightness was the inferior

cue in their experiment. The stereopsis was the only cue effected by distance. It was among the more effective cues at a one meter distance, while its effectiveness was less at two meters. Their study was restricted to standard LDR displays using a Mac IIfx computer with a high resolution grey scale monitor.

3.3 Visualisation of HDR Data

The high range of colour and intensities that HDR provides motivated researchers to use these types of data in their studies. Yuan et al. (2006) claimed that high resolution volumes, comprised of huge datasets generated by a supercomputer systems for large fluid simulation, require high precision composition to preserve detail. Yuan used HDR visualization to produce a technique for volume rendering in large volume data, such as in large fluid dynamic simulations, that are performed on supercomputer systems. Yuan et al. (2007) argued that using HDR datasets is vital for understanding complex geophysical phenomena, such as earthquakes, mantle temperature fluctuations, etc. Dinesha et al. (2012) claimed that fine details in an uncertainty distribution can be revealed when HDR mapping is used.

3.4 Illustrative Visualisation

According to psychologists, our brain can perceive a fully enclosed line in a 2D image either as an object or as a hole (Arnheim, 1954; Elder & Zucker, 1993; Nelson et al., 2001; Bertamini, 2006), as shown in Figure 3.5. It is hard to determine whether the white part is a hole or just a white disc on a black surface. They tried to realise what makes some shapes appear as holes or cuts and others as objects or figures, which is

known as *figure-ground* phenomena.

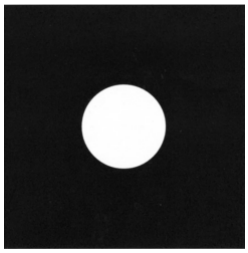
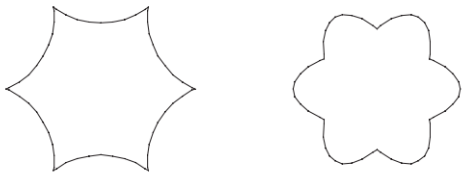


Figure 3.5: The white circle provides an ambiguous impression, since it might be perceived as a hole or as a white disc on the top of a black background. Image re-produced from (Nelson et al., 2001).

Arnheim (1954) claimed that the convexity and the concavity of the shape's contour encourage the shape to be perceived as a hole or as an object. For example, he claimed that the shape in 3.6(a) tends to be seen as a hole because it is concave, and the shape in 3.6(b) as a figure because it is convex. Bertamini (2006) wanted to verify this conclusion by applying an empirical experiment, based on the images in the Figure 3.6(a) and (b). He asked twenty-two naive observers to select the shape that looked more like a hole, when comparing the two shapes. More than 85% of the observers chose shape (a).

In fact, he made several studies to strengthen the idea that convexity and concavity plays an important role in perceiving a shape as hole or as a figure (Bertamini, 2000; Bertamini & Croucher, 2003; Bertamini & Mosca, 2004; Bertamini, 2006; Bertamini & Lawson, 2008). One of his results was that the shape can determine the objects ownership of the contour (Bertamini, 2006). This result can be utilised to enhance the perception of a cut made on the outer surface of 3D multi-layered images.

The cutaway technique, when used to remove a part of the outer surface of a 3D volume to visualise its inner contents, is one of the methods used in technical illustration and visualisation (Liang et al., 2005). It is also known as volume clipping (Weiskopf



(a) (b)

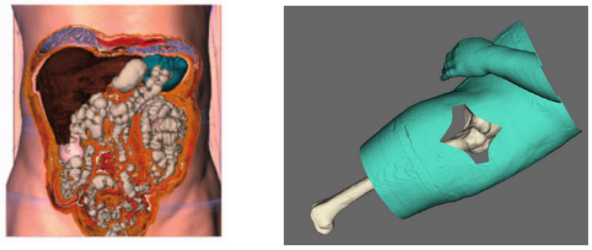
Figure 3.6: Arnheim (1976) claimed that the shape in(a) tends to be seen as a hole because it is concave, and the shape in (b) as a figure because it is convex. Image re-produced from (Arnheim, 1954; Bertamini, 2006).

et al., 2002, 2003; Bruckner et al., 2005). The first mention of the cutaway illustration, in the field of computer graphics, was in the SIGGRAPH 99 advanced OpenGL rendering course (McReynolds et al., 1998). However, the effect of the shape in a cutaway on depth perception in 3D is rarely studied. Most research that deals with the cutaway technique focus mainly on presenting internal volume contents (Diepstraten et al., 2003; Viola & Gröller, 2005; Li et al., 2007). The shape of the cutaway is less important in such studies. The resulting images can appear as if they are pasted on the outer surface Viola & Gröller (2005).

Moreover, only a few studies tried to use a geometric shape for the cut. Using geometrical shapes will assist automation of cutting a volume (using a computer), over using arbitrary shapes. For instance, Weiskopf et al. (2003) used the cutaway technique in their study. They used cuboid and spherical shapes to make the cut. Coffin & Hollerer (2006) used an interactive method to apply a cut. They used circular and rectangular shapes, as well as user defined (arbitrary) shapes in their study. Knödel et al. (2009) used four 3D geometrical shapes to be used as a cut. They were cubic, spherical, wedge, and tube. Sigg et al. (2012) used three geometrical shapes in their study. Theses shapes were cuboid, sphere and cylinder. Their main objective was to

develop a method which places parametrized cutaway objects automatically.

Other studies tried to use a specific shape to be the shape of cutaway, as in Liang et al. (2005) and Correa et al. (2006). The shape they used was similar to an incision made by a surgeon. They used this shape without considering its effect on depth perception.



(a) (b)

Figure 3.7: Cutaway Samples. (a) The cutway seems more as if it is being pasted on the outer surface of the skin. Image re-produced from (Viola & Gröller, 2005). (b) A specific shape for a cutaway that looks as though it is an incision made by a surgeon, because it is convex. Image re-produced from (Liang et al., 2005).

3.5 Summary

This chapter illustrates the reasons for using monocular depth cues in this study. It begins with highlighting the studies that showed the importance and influences of these cues on depth perception. In addition, it discusses the studies that used a combination of these cues in order to improve depth perception. Also highlighted in this chapter are studies that try to evaluate depth cues. The benefit of using HDR techniques is also discussed. The cutaway technique that is usually used in technical illustration and visualisation is discussed.

The following is a summary of the main conclusions of this chapter:

1. The work done by Surdick (1994) was the only study in the field of depth cue evaluation. In addition, this study was based on classic LDR images and display devices.

2. HDR techniques are being used increasingly in different fields of science. Visualising volume data using HDR is one of these fields.

3. A few studies tried to utilise specific shapes to be the shape of cutaway in multiple-lyered images (Coffin & Hollerer, 2006; Correa et al., 2006; Liang et al., 2005).

Chapter 4

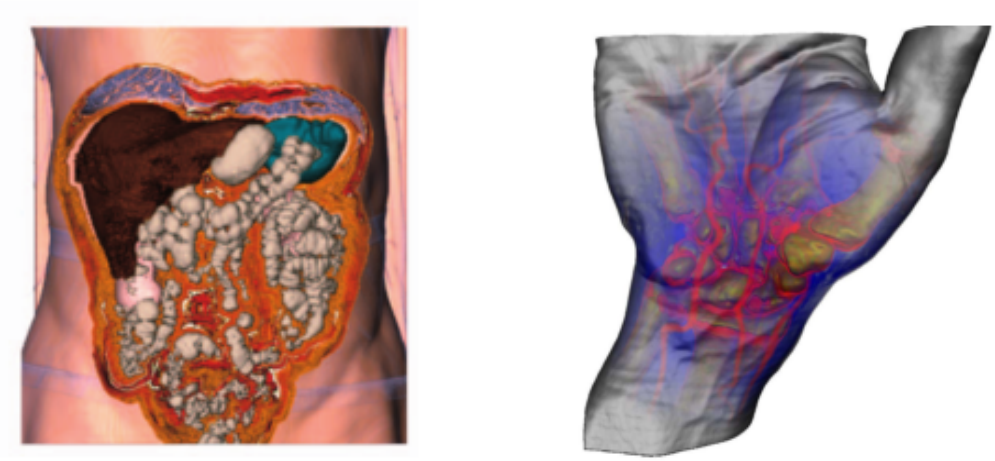
Evaluation of Monocular Depth Cues on a High-dynamic-range Display for Visualisation

The main objective of this chapter is to find an intuitive depth cue, among several monocular depth cues, which provides a better impression of depth ordering when no other cues are available. To do so, an experiment was conducted based on showing images to observers, asking them to select the closest part of the image. Four computerized tomography (CT) volumes were used in the experiment. The experiment was conducted using a HDR display, for the reason of the high contrast and brightness.

4.1 Introduction

Multiple layered abstract data is difficult to visualize on two dimensional displays, especially when the spatial arrangement of depth needs to be shown. One of the meth-

ods used is to remove less important layers to reveal underlying layers (Viola et al., 2005). Although the layers are revealed, they often appear as if they are painted on the skin's outer surface (Figure 4.1(a)). Transparency is another method which is used for visualising volumes (Krüger et al., 2006). In this case, the front layer is partially transparent, which provides the user with an intuitive depth cue. However, as shown in Figure 4.1(b), the depth order is not easy to see and it often needs to be deduced using knowledge of the visualized phenomena, which is the human anatomy in this case.



(a) (b)

Figure 4.1: Examples of visualizing depth layers. (a) The revealed layers appear as if they were painted on the skins outer surface. Image re-produced from (Viola et al., 2005). (b) It is difficult to determine the depth order. Image re-produced from (Krüger et al., 2006).

4.2 The Procedure

We conducted an experiment to determine which monocular cues provide intuitive depth ordering when all other cues are reduced or eliminated. Four CT and MRI volumes[1] were used in this experiment. Each volume consists of four slices of 512×512

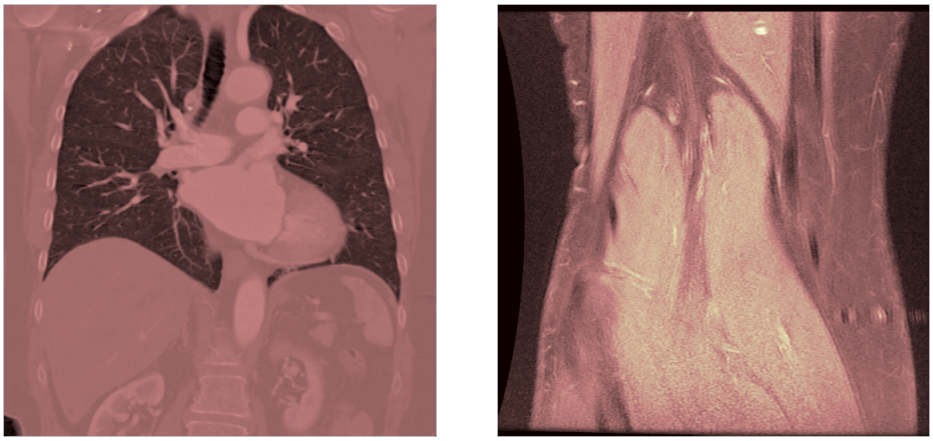
[1]The volumes were downloaded from http://www.osirix-viewer.com/datasets/

pixels (see Figure 4.3).

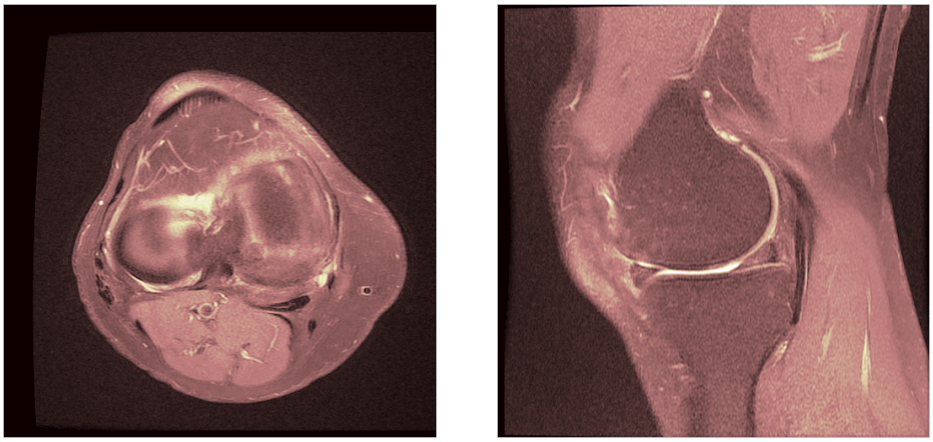
To reveal more detail and also to reduce the need for large luminance contrast, we used a false-colour map instead of gray-scale images. The fifth volume was generated from the superposition of 3D Gaussian functions with randomized parameters. Then, we took four slices of the resulting 3D function. Gaussian noise was added to the volume in order to better see the effect of blur and transparency cues. Such an abstract volume was introduced to see the effect of each cue without any contextual information (see Figure 4.2).

We visualized a volume by slicing it diagonally, starting either from the left or the right side (Figures 4.4(a) and 4.4(b)). This simulates a practical scenario in which it is necessary to reveal the internal layers of a volume. Also, since the visualized volumetric data is mostly abstract, it is almost impossible to guess the correct depth ordering from the content of the volumetric data and without any depth cues.

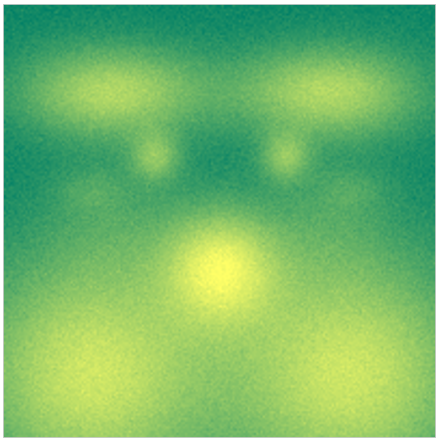
To prevent any possibility of guessing the depth order from the volume content, a set of "flipped" volumes was created, in which each slice was flipped horizontally and in depth (along z-axis). Figure 4.7 shows two pairs or original and flipped volumes. Note that the choice of the stimuli was dictated by the necessity to eliminate bias in the experiment rather than a common practice in volume visualization.



(a) (b)



(c) (d)



(e)

Figure 4.2: Selected slices from the four volumes containing CT and MRI scans (a-d) and from the fifth "abstract" volume (e).

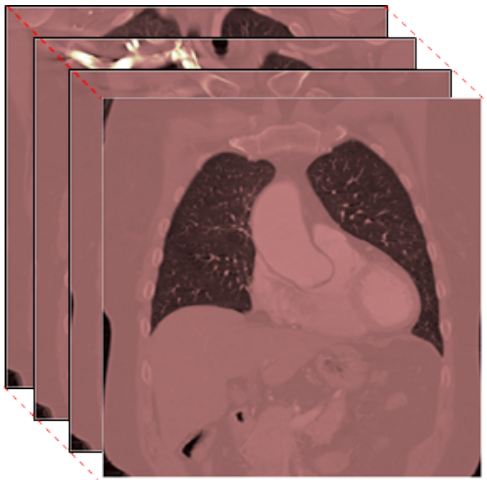
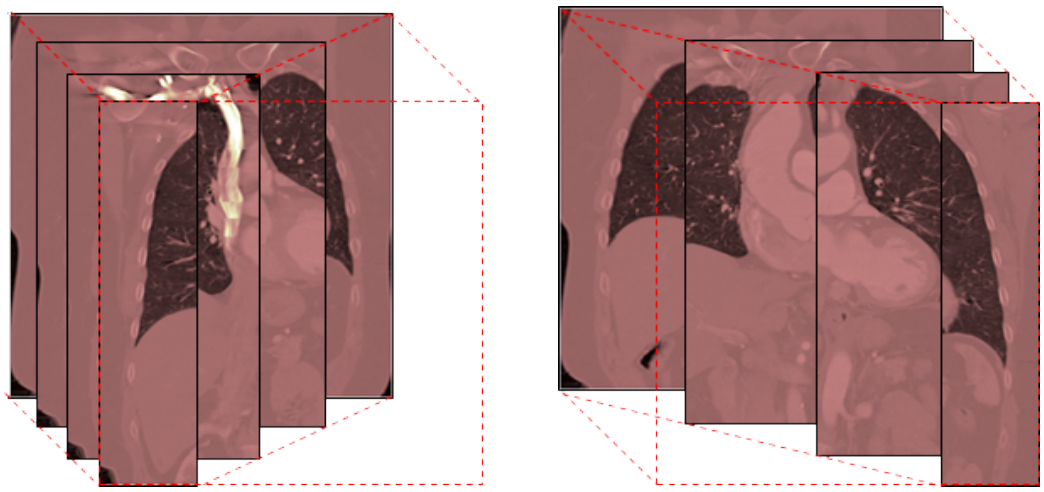


Figure 4.3: Four slices of the CT volume of the lung.



(a) (b)

Figure 4.4: Examples of slicing a volume. A cut has been made either (a) from the left, or (b) from the right side of the volume.

The stimuli used in the experiment were samples of medical images, which were taken using CT and MRI scanners. These scanners are used to scan a specific part of a body. They generate images that are more detailed than standard X-rays scans. The reason for choosing these types of images was because first, the images are usually unknown to most people, making it hard to understand what they represent. So, the observer cannot use previous knowledge to make a decision throughout the experiment. The observer will focus more on the depth cues available, included in the image. The second reason, there is a slight difference between each sequential pair of such images, which is normally caused by the internal organs inside the body moving during a scan. This causes a difficulty in deciding the order of these images based upon these differences. Which in turn, makes the observer depend more on the depth cues, that are located in the image. The resultant images are normally of type DICOM (*Digital Imaging and Communications in Medicine standard format*). Usually, such a format needs a special system to read the file. Therefore, to simplify the work, we transformed these images into another format using following steps:

- Read the source image using a Matlab function called *dicomread*, which is a part of *Image Processing* Toolbox.

- Normalise the image, since the range of intensity of the source image values exceeds the allowed range, as seen in Figure 4.5(a). To do so, the following equation was used:

$$I' = \frac{I - min}{max - min}, \tag{4.1}$$

  where, I' represents the new intensity value, I is the old intensity, min and max are the minimum and the maximum intensity values of the image respectively. The equation was used to make the intensity values range from 0 and 1, so the

image detail can be seen more clearly, as shown in Figure 4.5(b).

- Colouring the resultant image, Figure 4.5(c). This allows small changes of the intensity values to be seen more easily, than seen in grey-scale levels. The image was then saved in a 16-bit portable network graphics (png) format.

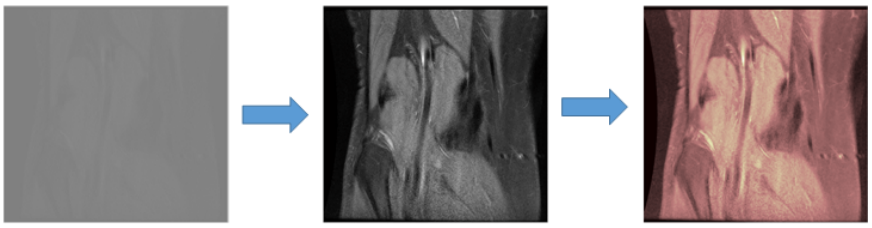


Figure 4.5: Image transformation operation used to convert images from *dicom* format to *png* format. (a) Shows the source image (*dicom* format), and (b) through normalising the source image, and finally (c), colouring the image to reveal details that cannot be seen or are difficult to see through grey-scales.

Among several available colour maps in Matlab, the *Pink* colour map was used. This was because the Pink colour map provides just more noticeable difference levels, which meant the details could be seen more clearly, as shown in Figure 4.6. However, for the last group of stimuli, which is the abstract group, the *summer* colour map was used. The reason for this was because the majority of pixel intensity values were low. Therefore, using any other colour map would make these pixels invisible or difficult to see.

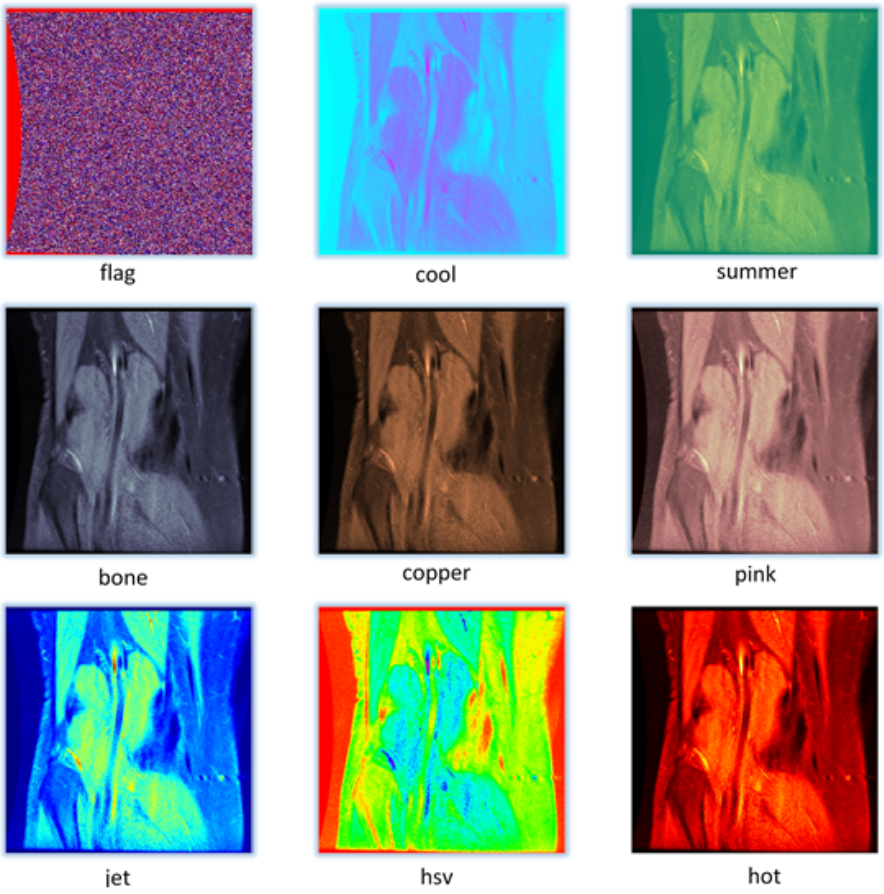


Figure 4.6: Different colour maps that available in Matlab

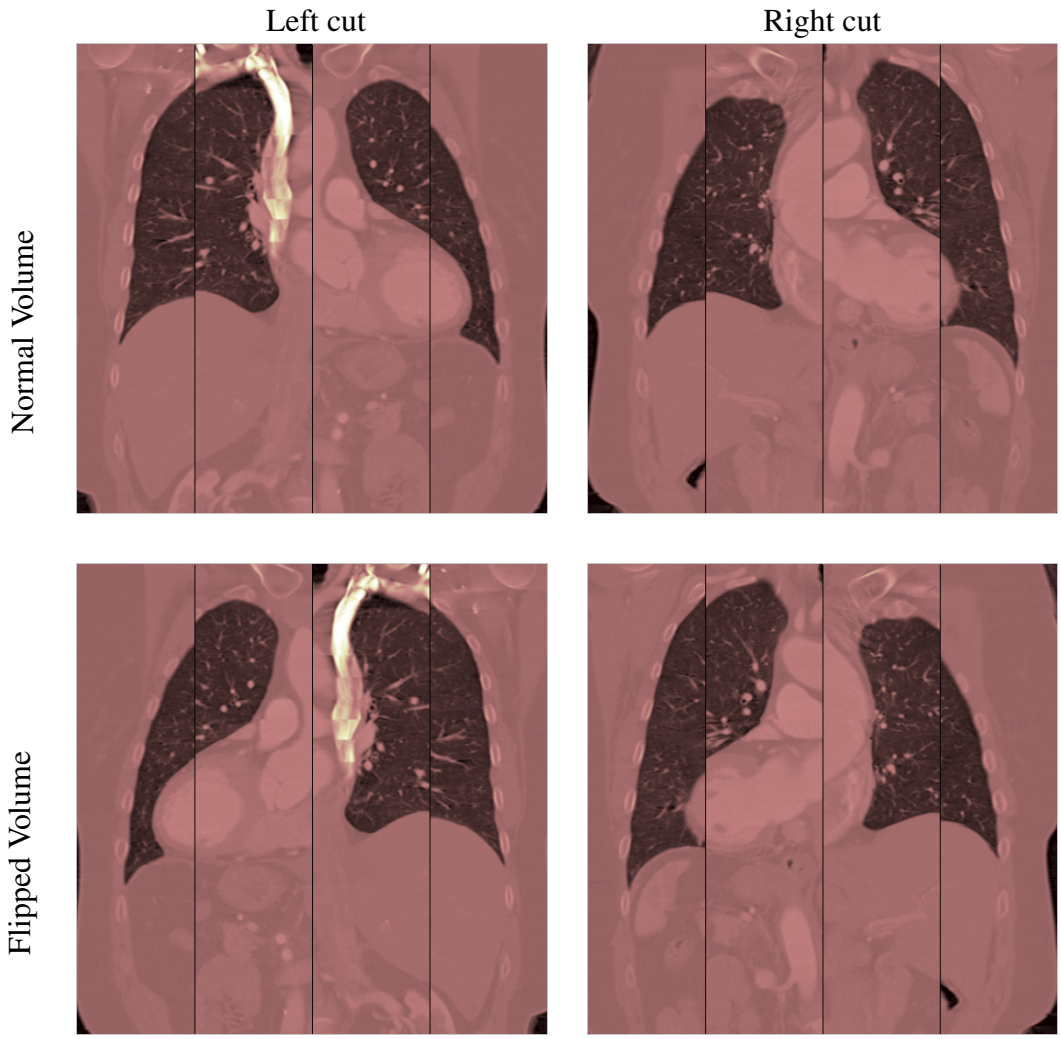


Figure 4.7: An orthogonal projection of the volume after cutting. These images contain no cues therefore it is impossible to tell whether the cut was made from the left or the right.

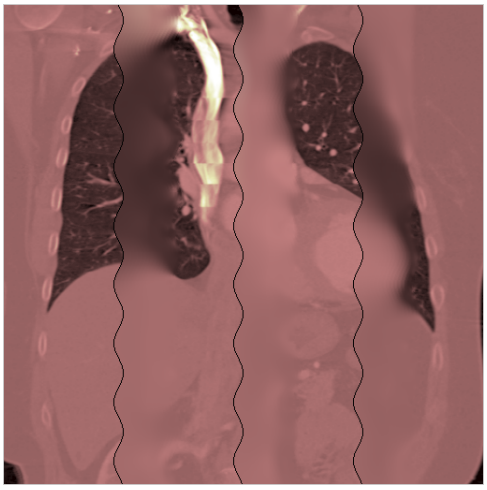
4.2.1 Stimuli

Seven cues were used in the experiment: blur, Overlap, contrast, shadow, relative size, brightness and transparency. In the following paragraphs we explain how these cues were generated.

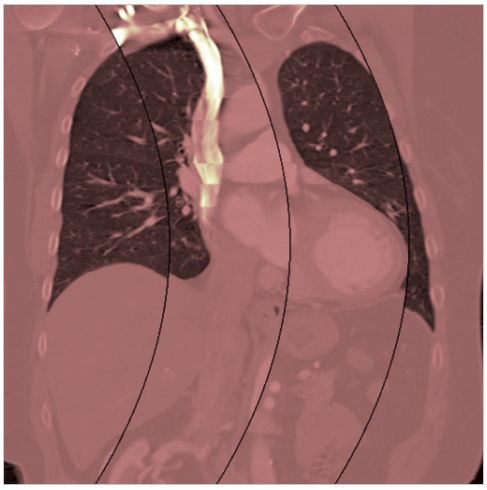
Blur cue Marshall et al. (1996) argued that the appearance of an edge, that separates blurred and sharp region, is important for determining the depth order. They demonstrate that if the separating edge is sharp, the sharp region (in focus) is perceived as closer. Figure 4.8 (a) shows an example of the blur cue stimulus.

To generate an image with the blur cue, the left-most column was kept sharp while the other three columns were partially blurred, starting from their left side towards the middle of the column. Sinusoidal lines were used to separate the columns and provide a stronger cue. This procedure was performed to produce a blur cue similar to that employed by Marshall et al. (1996). The blur was using a Gaussian kernel with the standard deviation ranging from $\sigma = 20$ on the side corresponding to the depth discontinuity, to $\sigma = 0.1$ in the middle of the column. According to Marshall et al. (1996), we expected the observers to choose the leftmost side as the closest.

Overlap cue (also known as interposition cue) Bertamini & Lawson (2008) showed that because of convexity, observers responded faster when asked to report on depth order, when two surfaces were overlapped by a curved line. Therefore, to generate an overlapped cue stimulus that corresponds to a left cut, Figure 4.8(b), we used curved lines to separate the columns. This way the curve direction opposed the cut direction (for example; for a left cut, the direction of the curves will be towards the right). According to Bertamini & Lawson (2008) we expected the observers to choose the leftmost side as the closest. Throughout this study we called this cue "*convex* ".



(a) Blur cue stimulus



(b) Overlap (Convex) cue stimulus

Figure 4.8: Blur and overlap cues stimuli for the left cut of the volume. (a) The slices are blurred at the edge to visualise the discontinuity in depth. (b) Curved lines suggested the depth ordering for overlapping slices.

Contrast cue To generate a contrast cue stimulus, each of the slices had its contrast modified so that the closest slice had the highest contrast. Before modifying slices, we converted them from the gamma corrected sRGB into a linear RGB colour space and computed the relative luminance value for each pixel. Then, we adjusted the luminance contrast using the equation:

$$L_{out} = \left(\frac{L_{in}}{P}\right)^{c} \cdot P, \tag{4.2}$$

where L_{in} and L_{out} are input/output luminance, P is a luminance that should remain unchanged (usually background luminance) and c is the contrast modification factor. Since contrast change affects the perceived saturation of colours, it is necessary to correct for that. This can be achieved using the colour transfer equation Mantiuk et al. (2009):

$$C_{out} = \left(\frac{C_{in}}{L_{in}}\right)^{s} \cdot L_{out}, \tag{4.3}$$

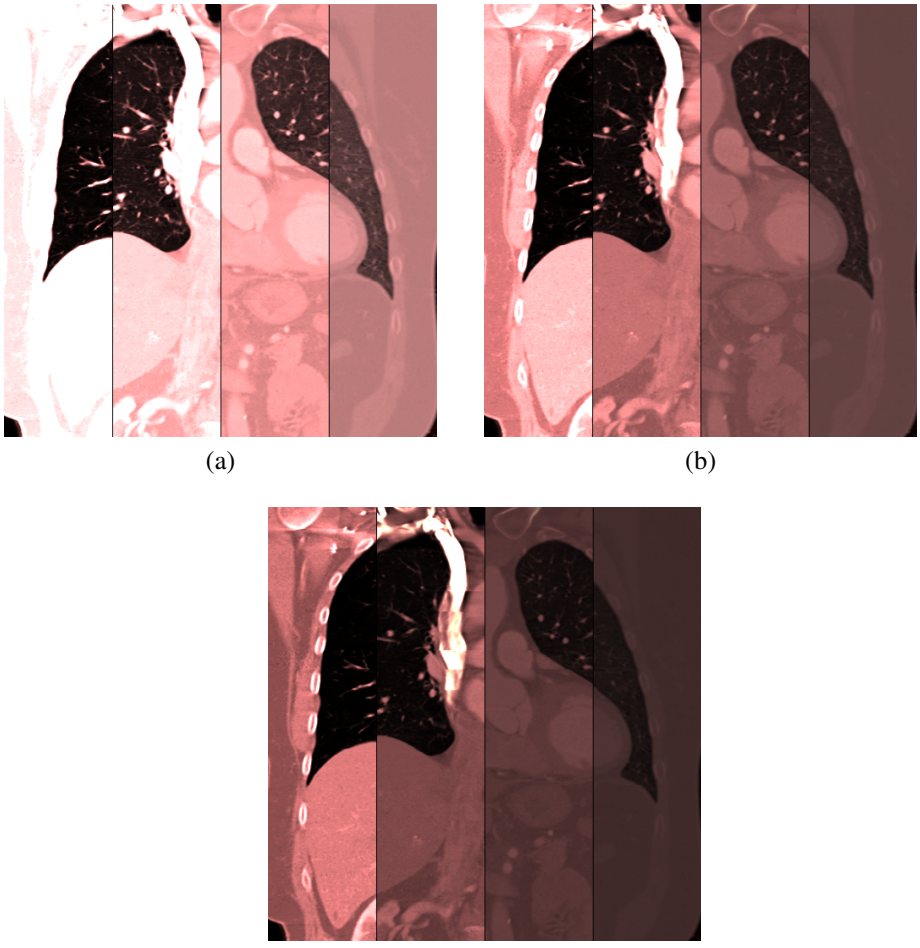
where C denotes one of the colour channels (red, green, or blue), and s controls colour saturation. Mantiuk et al. (2009) provide an empirical equation for finding the proper saturation correction factor:

$$s(c) = \frac{(1+k_1)c^{k_2}}{1+k_1 c^{k_2}}, \tag{4.4}$$

where k_1 and k_2 are constants with the values 1.6674 and 0.9925 respectively.

Each slice in a volume was assigned a different c value depending on the direction of the cut. The c values (1, 1.7, 2.93 and 4.98) were selected to differ by a constant ratio, which resulted in approximately equal increase in perceived contrast. Figure 4.9 shows an example of the contrast stimulus at three virtual exposures.

Shadow cue Shadows help to determine a distance between objects. Shadows and in particular self-shadows can be efficiently approximated using the ambient occlusion method Landis (2002). Figure 4.10(a) shows a stimulus for a left-cut volume with a



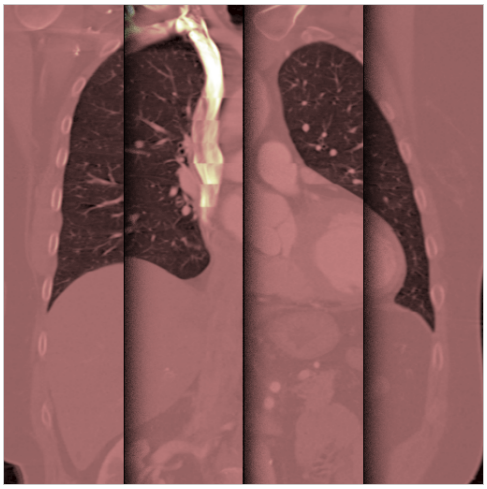
(a) (b)

(c)

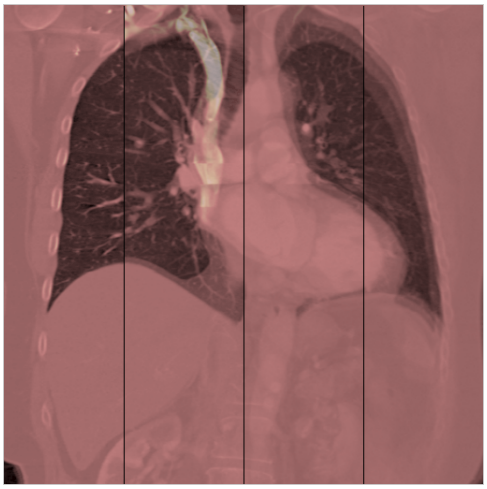
Figure 4.9: Contrast stimulus. Because the stimulus is a high dynamic range image, it is shown at three different exposure values.

shadow cue. We generated a contact shadow for the right side of the first column of the first slice and merged it with the second column of the second slice. A sample space ambient occlusion (SSAO) algorithm Mendez (2010) was used to generate shadows. We speculate that the observer will select the leftmost column as the closest side. To reproduce an equal distance between each slice, the size of the shadow spread was the same.

Transparency cue Talbot's law states that the reflectance of a fusion colour is



(a) Shadow cue stimulus



(b) Transparency cue stimulus

Figure 4.10: Shadow and transparency cues for the left-cut of the volume. (a) An ambient occlusion was used to generate a shadow-like between columns. (b) The removed parts of the slices were made transparent.

the weighted average of the reflectance of its component colour proportions mixed together (Beck et al., 1984). Talbot's law can be expressed by the following equation:

$$c = \alpha a + (1 - \alpha)b, \tag{4.5}$$

where a and b are reflectance of the component colours which are going to be mixed together, c is the reflectance of the fusion colour and α is the transparency factor (between 0 and 1). Equation 4.5 was the origin of *alpha blending* that is commonly known in computer graphics today (Kasrai, 2004). Alpha blending is a process used to produce transparency by combining two or more transparent foreground layers with a background layer using the equation (Bavoil & Myers, 2008):

$$\begin{aligned} C_n = \quad & \alpha_n C_n + (1-\alpha_n)\alpha_{n-1}C_{n-1} + (1-\alpha_n)(1-\alpha_{n-1})\alpha_{n-2}C_{n-2} + \ldots \\ & (1-\alpha_n)(1-\alpha_{n-1})..(1-\alpha_1)C_0, \end{aligned} \tag{4.6}$$

where C_n is the output colour, α_x is the alpha value of layer x, and C_x is the colour of layer x and C_0 represents the background layer.

The stimuli were generated by making the parts of the slices transparent instead of removing them (as shown in Figure 4.4). The alpha values were: $\alpha_4 = 1$, $\alpha_3 = 0.5$, $\alpha_2 = 0.333$ and $\alpha_1 = 0.25$, where α_4 corresponds to the nearest slice. The background colour C_0 was black (red, green and blue components equal to 0). Alpha values were chosen to result in identical contribution of each layer. An example of the trasparency stimulus is shown in Figure 4.10(b).

Relative size cue Given a perspective projection and equal sized objects, the closest object will appear larger than the farthest object. The relative size cues were generated by reducing the size of each consecutive slice to 75% of the previous slice. An example of the relative size cue is shown in Figure 4.11. This cue could be considered as one of the control conditions as we expect very few wrong answers for this very suggestive cue.

Brightness cue Given two similar objects, the brighter one will appear to be closer

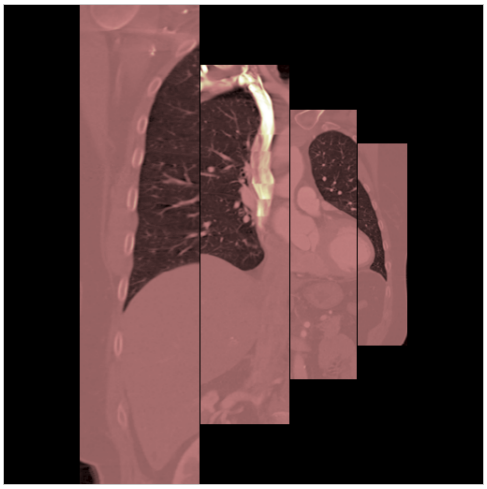
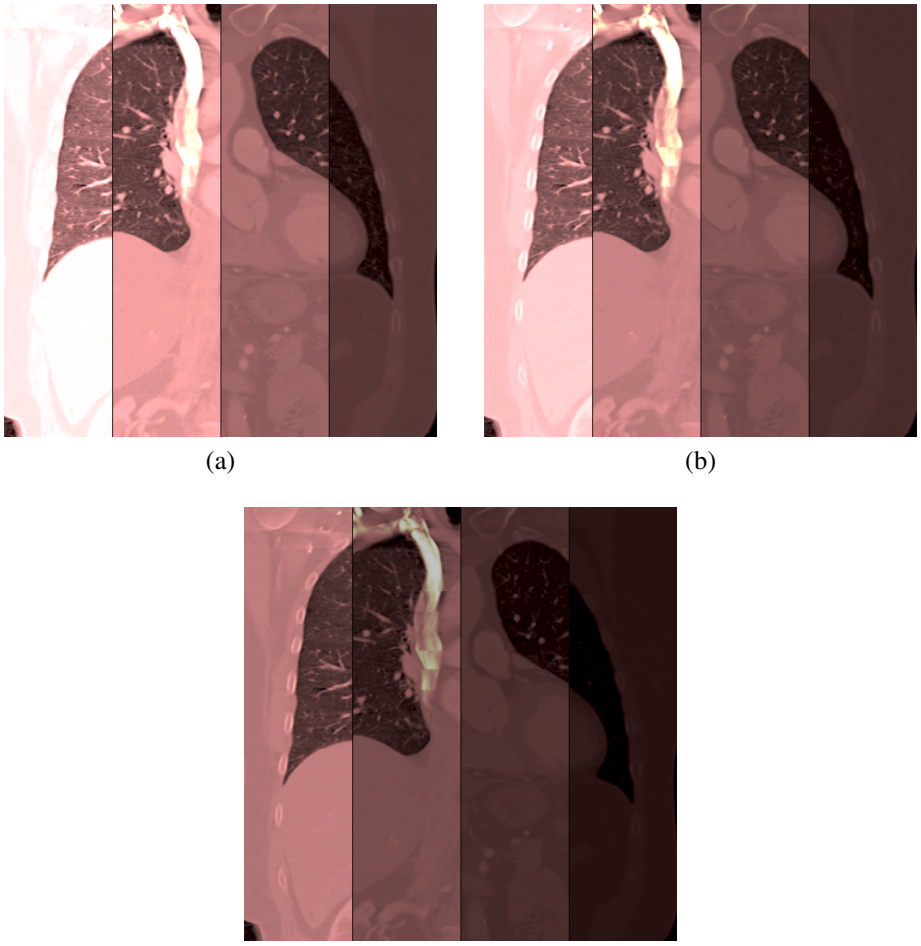


Figure 4.11: A relative size cue for the left cut of the volume. The size of the further located slices is reduced.

to the observer (Gibson, 1950). The stimulus was created by increasing the luminance of each slice in a volume with higher luminance assigned to closer slices. We set the peak luminance of each slice to 2 000 cd/m^2, 737 cd/m^2, 271 cd/m^2 and 100 cd/m^2, from the closest to the farthest. The values correspond to equidistant points in the logarithmic space, which result in approximately the same steps in perceived brightness Mantiuk et al. (2006a). Figure 4.12 shows an example of the brightness cue image as three different exposures of the corresponding HDR image.

Control condition "no cue" We also introduced images with no cues to the experiment, such as the one in Figure 4.7. They were used as a control condition for which the observers were expected to provide random answers. We use the label "No cue" for this condition.



(a) (b)

(c)

Figure 4.12: Brightness stimulus. As the stimulus is a high dynamic range image, it is shown at three different exposure values.

4.3 HDR image Preparation

After all stimuli had been prepared, the next step was converting these images so they could be displayed on an HDR display. This included the following:

1. Since these stimuli are going to be displayed on an HDR device, the colour space of the image should be changed from sRGB to linear RGB using the following equation: (Anderson et al., 1996):

$$C_{linear} = \begin{cases} \frac{C_{srgb}}{12.92}, C_{srgb} \leq 0.04045 \\ \\ \left(\frac{C_{srgb}+a}{1+a}\right)^{2.4}, C_{srgb} > 0.04045 \end{cases} \tag{4.7}$$

where C_{liner} is the resulted linear colour channel, and a=0.055. The equation was applied on all colour channels of the image, red, green and blue.

2. The resulted image was then multiplied by a number to increase its luminance. In this case 250. The reason was because the peak luminance of a typical LDR display is 250 cd/m^2.

4.4 The Experiment

Each observer was asked to read an instruction, which asked them to select the slice (the leftmost or the rightmost) of the displayed image that appeared closer. The observer had to make the best guess if he or she was unsure which depth ordering was correct (two alternative forced choice). An observer could cancel and repeat the last measurement in a rare case of pressing a wrong key. The instruction explained how the images were created by slicing volumes and projecting them orthogonally. Each experimental session was preceded by a short training session to ensure good understanding of the task. In each session, the observer judged all 160 images (5 volumes $\times$ 2 cuts $\times$ 2 flips $\times$ 8 cues).

4.4.1 Apparatus

The images were generated in the linearized RGB colour space (ITU-R BT.709 colour primaries) and displayed using an experimental HDR display. The device was a modified version of the SBT1.3 model (Seetzen et al. (2004)), which consisted of a 4 000 lumen projector and a 15" LCD panel with a resolution of 1024×768 pixels. The device's peak luminance was 2 446 cd/m^2 and the black level was 0.01 cd/m^2. The details on the display can be found in Wanat et al. (2012). The viewing distance between each observer and the display screen was approximately 80 cm. The room lights were switched off to avoid screen reflections and maximize display contrast.

4.4.2 Observers

Twenty one volunteers participated in this experiment, seven female and fourteen male, between 23 and 40 years old. All observers had normal or corrected to normal vision. All observer were naive to the objective of the experiment.

4.5 Results

To test for the statistical significance, the data were analyzed using the Binomial test assuming the null hypothesis H_0 that the cue has no effect on the perceived depth ordering and the observers make random choices. The binomial probability distribution is given by Cunningham & Wallraven (2011):

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x} \qquad x = 0, 1, \ldots n, \tag{4.8}$$

where $P(X)$ is the probability of X successes, p is the probability of success on any one trial (0.5 in our case) and n is the number of trials. We tested the hypothesis at the $\alpha = 0.05$ significance level.

For each cue, we tested for H_0 by computing the probability of observing a given number of correct answers assuming that the observers were guessing. Then, H_0 was rejected when the probability was below the critical value. Figure 4.13 shows the result for all observers. The red-dashed lines mark the critical region for the Binomial test: all results within the range limited by the dashed lines are not statistically significant (shown in red). All data in the Figure 4.13 is summarized in the Table 4.1.

The control condition cue ("No cue") has no effect on depth ordering, which means that it was impossible to tell the depth order given no cue. This confirms that the stimuli were well balanced and the depth ordering could not be deducted from the content of volumes alone.

We found no statistically significant effect for the transparency cue in the case of two volumes: *Abstract* (50%) and *Knee-2* (57.1%). This could be explained by the lack of single-reversing X-junctions in the depicted volumes (refer to Figure 2.9). Such ambiguity of depth ordering can also be seen for *Knee-1* (60.7%) and *Knee-3* (60.7%) volumes since H_0 was barely rejected. The highest correct rate for the transparency cue was achieved for the *Lung* volume (64%), which contains strong edges and thus makes the slices easier to recognize, as shown in Figure 4.10(b). Our findings showed that the transparency cue was strongly affected by the nature of the images. Moreover, the transparency cue has the lowest percentage of correct judgement among all other cues. This result is of particular interest, given that transparency is a common method of presenting multi-layered phenomena in visualization.

Another cue that is also affected by the content of the volume is blur. The H_0

could not be rejected for the abstract volume (58%). The most likely reason for that is the lack of high frequencies and sharp edges in this volume, which are necessary to perceive blur. This cue could also be unsuitable for the visualization application in which blurring and therefore the loss of details is not permissible.

For all the remaining cues the results are statistically significant for all volumes, which means that there is evidence that these cues help in the depth ordering task. However, the success rate varies between the cues.

The convex cue yielded a success rate of 75.96%. This lower than expected success rate can be explained by the nature of the volumes, which resulted in slices of very similar colour. To show the effect of convex contour on the depth perception Bertamini & Lawson (2008), Bertamini (2006), Bertamini & Croucher (2003) and Vecera et al. (2002) used two contrasting colours to separate the two layers. There was no such colour difference in our stimuli.

Similar success rates observed for the shadow cue, through all volumes except the abstract, suggests that it is not affected by the content within the volume. The shadow produced success rate of 83.32%. The false ordering, in the case of shadow, was most likely caused by the ambiguity between shadow and shading; the observer may see the "shadow" as if it was the shading of a curved surface. Because of this ambiguity, the observer may assume that the volume is illuminated from the right, consisting of curved surfaces and the order of the slices is opposite to the intended order (refer to Figure 4.14(b)). Another weak point of adding shadow is that less detail may be visible in the shaded parts of an image.

The contrast, brightness and relative size cues have the highest ratio of correct judgement with the average of 91.68%, 92.86% and 92.62% respectively. Relative size is a very strong cue that could be used with any type of display device. However, it

requires that the presented phenomena contains the features of common or known size. In our case, it was the equal size of the volume slices. It also requires perspective rather than orthogonal projection, which may be less suitable in some applications. Brightness and contrast cues provide intuitive depth ordering independently of the content or shape of the visualized volume. This supports our hypothesis that an HDR display can be useful in visualisation applications. From these two cues, the brightness cue could be more universal as it does not affect the contrast of the displayed data.

Contrary to our expectations, some cues, such as transparency and blur, resulted in very poor performance in the depth ordering task. Also, convex overlap and shadow cues were not as strong as we expected. Such a result could be specific to our data set, which consists mostly of abstract CT images, often lacking strong edges, and 4 discrete depth layers. We also expected the relative size to result in the highest probability. However, surprisingly, the brightness cue resulted in very comparable performance.

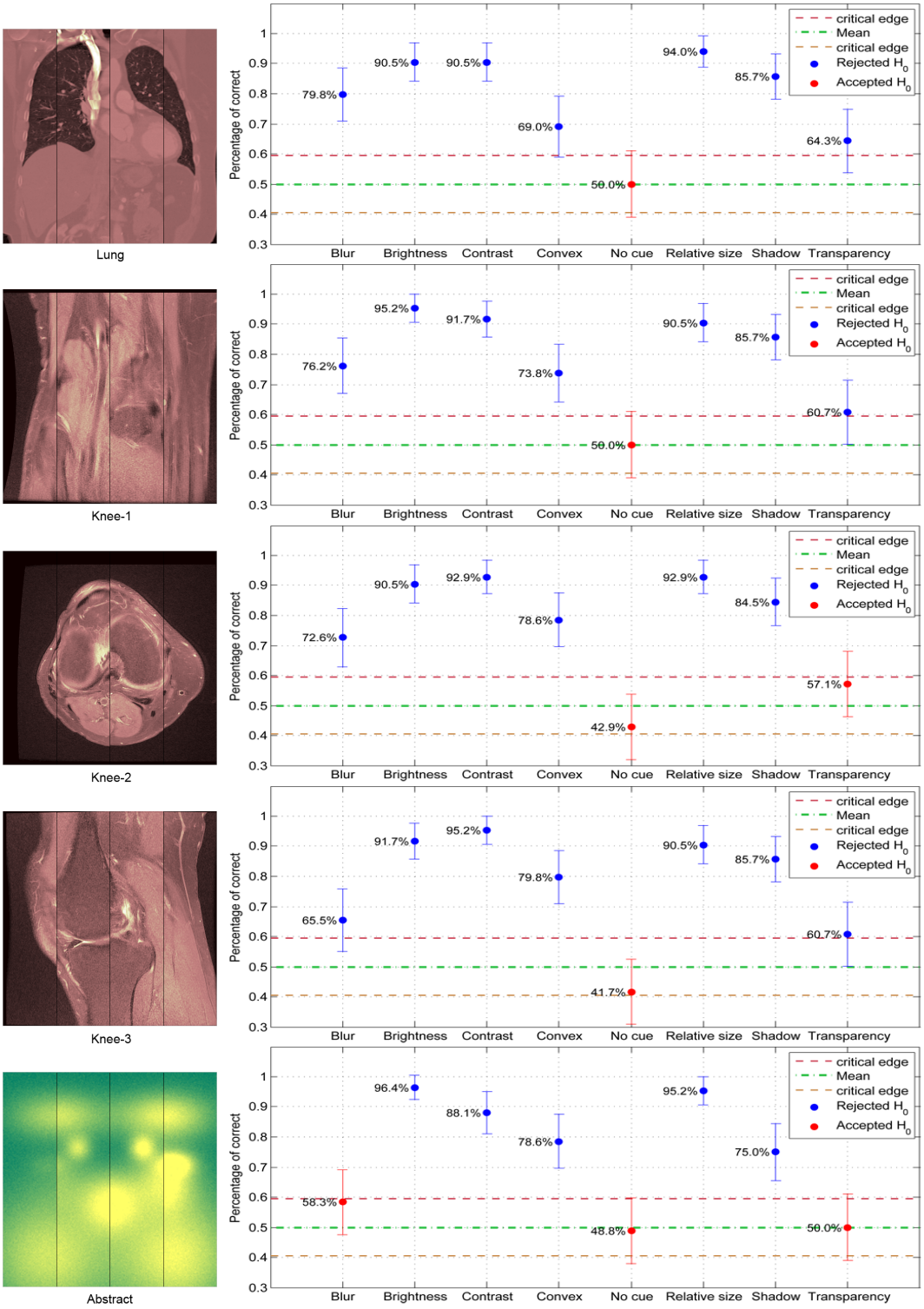
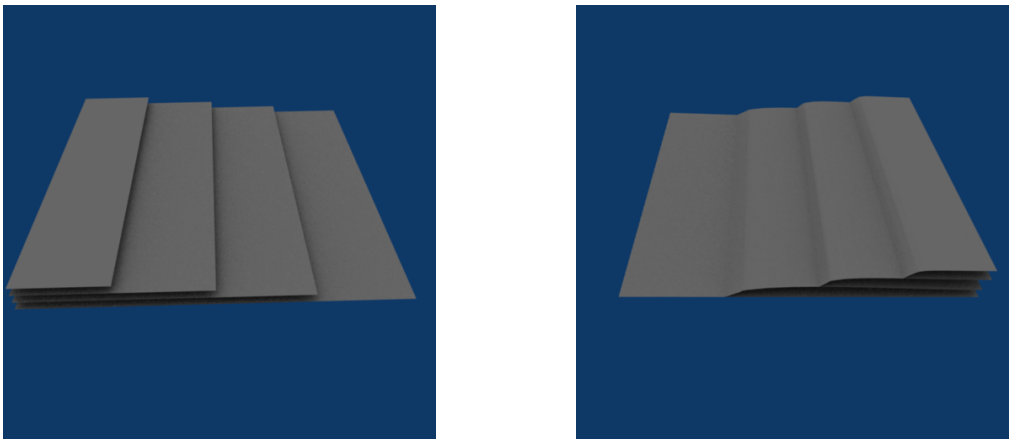


Figure 4.13: Percentage of correct depth ordering judgments for each monocular cue. The error bars show 95% confidence intervals. The data points marked in red indicate no statistically significant effect of a given cue on the depth ordering performance.



(a) (b)

Figure 4.14: An example of some possible ways to generate shadows. In (a) the shadow generated as a result of occluding lights by the layers on the top and the light source, in this case, will be in the left side, whilst in (b) the shadow is the result of a curved layers shade and the light source is in the right side.

Table 4.1: Percentage of correct depth ordering judgement for all cues. The highest ratio of each volume is indicated by a boldface font.

	Monocular depth cues							
CT volumes	**Blur**	**Brightness**	**Contrast**	**Convex**	**No cue**	**Relative size**	**Shadow**	**Transparency**
Lung	79.8	90.5	90.5	69.0	50.0	**94.0**	85.7	64.3
Knee-1	76.2	**95.2**	91.7	73.8	50.0	90.5	85.7	60.7
Knee-2	72.6	90.5	**92.9**	78.6	42.9	**92.9**	83.32	58.56
Knee-3	65.5	91.7	**95.2**	79.8	41.7	90.5	85.7	60.7
Abstract	58.3	**96.4**	88.1	78.6	48.8	95.2	75.0	50.0
Total average	70.48	92.86	91.68	75.96	46.68	92.56	75.0	50.0

4.6 Conclusions

The following summarizes the work that was done in this chapter and highlights its results. The work was based on an experiment conducted to find the most effective cue on depth perception when it is being used in visualisation tasks. CT and MRI scan images were used as the experimental stimuli. The reason was to prevent the test subjects from depending on their knowledge, when they were making a decision through the experiment. This would make them depend on the cues alone, when making their decision. Seven monocular cues were tested. These cues were blur, brightness, contrast, convex, shadow, relative size and transparency. The results showed that these cues differ from each other in the rate of their effect on depth perception. However, in general, each individual cue increased the possibility of correct depth ordering beyond chance. The most surprising cue was transparency. It was the least effective cue with the lowest percentage of correct judgements. Whereas, it is considered to be the most common cue, being used in many computer graphics applications. Both blur and convex cues observed a low success rate. The shadow cue provided a better success rate, compared to the blur and convex cues, but not as much as expected. The relative size cue predictably resulted a high success rate. However, it could be less effective in cases where the visualised data did not contain elements of a common size. The most interesting performance was the high success rate of the two cues that utilised the extended dynamic range of a HDR display; contrast and brightness. This promotes the importance of using HDR displays in visualization applications.

This work differs from that carried out by Surdick (1994), in both the procedure used and results. Regardless of the difference in the procedure in each experiment and the utilised cues, the results of the commonly used cues was contradictory. The

brightness and relative size cues where the only ones used by both groups. Surdick stated that the effectiveness of the brightness cue was vastly inferior, and the relative size cue was not one a superior cue. Whereas our results demonstrated the opposite.

4.7 Limitations and Future Work

The main limitation of this work is that the results apply to the case of parallel slices of CT and MRI scans. More experiments need to be done to confirm whether these findings generalize to more complex cases. For example, in a complex shape, altering the brightness might interpret the brightness change incorrectly as shading, because of the curvature of the surface, rather than discontinuity of depth. This ambiguity can be solved by changing the luminance levels on a HDR display, which would be interpreted as illumination discontinuity rather than shading. However, a significant luminance contrast in an image may result in glare, which reduces the visibility of details in darker regions. For future work, we will experiment with such complex cases and test whether the current results carry over to more natural settings. Furthermore, we wish to experiment with a combination of the monocular depth cues, as such a combination is expected to improve the accuracy and speed of depth detection (Mather & Smith, 2004).

4.8 Summary

This chapter has examined seven depth cues in order to find the most effective cues on depth perception. These cues were blur, brightness, contrast, overlap, relative size, shadow and transparency. With overlap cue, the edges of these layers were cut with

a curved contour rather than in a straight contour edge. Therefore, we called this cue throughout this study a “convex” cue. These cues were selected carefully based on previous studies, which have proven the impact of these cues on depth perception. These cues were tested using a HDR display device, because of the capability of such display devices. The hypothesis was that this type of display device will be helpful to the observer in making decisions throughout the experiment. The most important part of the results in this study was the high rate of correct depth judgement, of both the brightness and the contrast cues, which are considered the most important characteristics of a HDR device. This supported our hypothesis about the importance of using such kind of devices in visualisation applications or in depth ordering tasks.

Chapter 5

The Effect of Translucency on Decomposition of Semi-transparent Layers

Previous studies revealed that blurring the foremost layer in a multiple-layered image might be useful for separating it visually from the layers underneath when it is semi-transparent (Singh & Anderson, 2002a). Moreover, previous research also showed that the amount of blurring is helpful in estimating the distance between the focused and unfocused (blurred) layers, or between the parts of an image that have different amounts of blur (Held & Cooper, 2010). In this chapter, we attempted to determine which type of translucency (also known as *defocus blur*), uniform or non-uniform, yields more accurate depth perception. An experiment was conducted, and the results showed that both types cause an outcome on the impression when they are utilized. However, according to the statistical analysis, we were not able to perceive which type is better and has a stronger effect when it is used in depth-ordering tasks, and we conclude that there is no advantage to using uniform translucency over non-uniform translucency.

5.1 Introduction

In Chapter 2, we noted that transparency is just one of the cues that have been extensively employed in many computer-based applications. It plays an important role in direct volume rendering, which is used in many visualization applications (Strothotte & Schlechtweg, 2002; Zheng et al., 2013). Furthermore, dozens of psychology studies have been conducted to understand the result of transparency on human perceptual experience.

Despite the extensive use of the transparency cue, it sometimes gives a wrong impression that leads to making a wrong decision, especially with abstract or unknown images. In the previous chapter, it caused the lowest success rate of all the monocular cues used in the experiment. We concluded that one of the reasons for that small success rate was the ambiguity caused by X-junctions. The other reason was the visual ambiguity created via the sorting of the unknown or abstract semi-transparent object layers, which was discussed in detail in Chapter 2.

One of the difficulties of using transparency is illustrated in Figure 5.1(a). It is hard to know how many layers there are in the image, as well as the order of these layers. It might be perceived as either a single-layer image or a partially overlapped multiple-layer image. Moreover, if it considered a multiple-layered image, what is the order of these layers?

One of the solutions offered by Singh & Anderson (2002a) was blurring the foremost layer, as show in Figure 5.2. From the figure, it is obvious that the blur helps to separate the foremost layer, but there might be another cause that makes the separation operation trivial. The continuity of the striped lines in the background layer works as an extra cue to indicate that the image consists of multiple layers. Furthermore, all types of X-junctions found in the image are single-reversing. This type of X-junction gives a clearer view of transparency and a correct depth order (Watanabe & Cavanagh, 1993). In other words, the layer separation could not be so simple if the background is plain and has the same colour as the foreground, as in Figure 5.1(c). In this study, we hypothesize that using a non-uniform blurring instead of uniform blurring will

make visual layer discrimination faster, as shown in Figure 5.1(b). It seems obvious that both ambiguities, the number of layers and their order, have been removed. The figure shows clearly that there are two planes, and that the translucent plane is in front.

(a) (b) (c)

Figure 5.1: The ambiguity of depth order is one of the X-junction cases. (a) It is difficult to determine the correct order of the planes. (b) By making the first layer translucent, the ambiguity in (a) is overcome. (c) Uniform blurring is another way to remove the ambiguity, as suggested by Singh & Anderson (2002a).

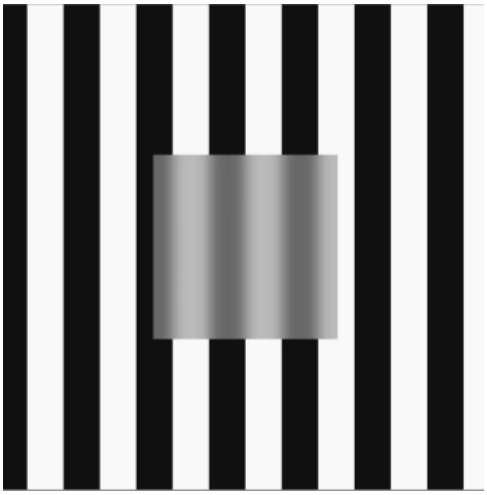


Figure 5.2: The role of image blurring in initiating a decomposition into multiple layers (one seen through the other). Image re-produced from (Singh & Anderson, 2002a).

Another problem comes with transparency when two or more distinct layers are overlapped to produce a single image in such a way that the layer in the back is fully covered by the foremost layer. If the layer in the front is opaque, then the layers in the back will be fully

obscured. However, if the front layer is semi-transparent, it will be difficult to analyse such an image, especially when the content of the layers is unknown or abstract. This difficulty applies whether such an image consists of a single or multiple layers, as indicated in Figure 5.3.

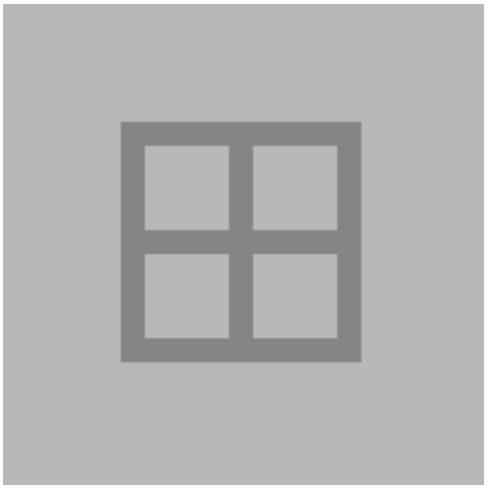


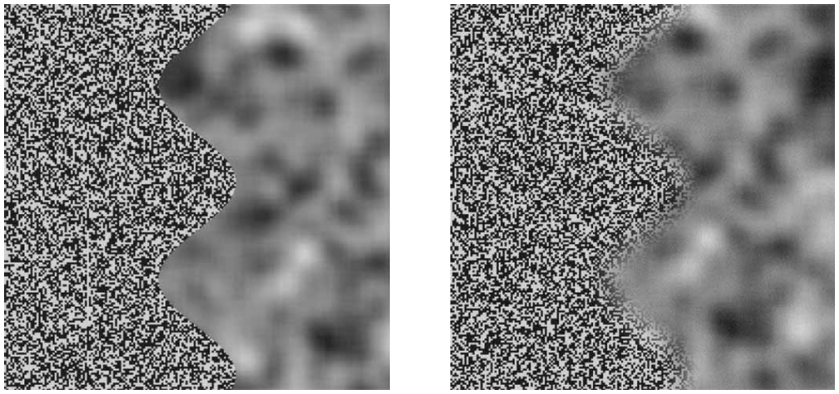
Figure 5.3: It is difficult to analyse the image and determine how many layers there are, and whether the window shape is in front of a grey background or at the back of the grey semi-transparent plane.

Singh & Anderson (2002a) considered how the presence of blurring modulates the perceived opacity of a partially transmissive surface given the precept of two distinct surfaces. They argued that increasing the degree of blurring in the region of transparency caused a decrease in perceived transmittance, as shown in Figure 5.2.

Mather (1996) claimed that the blur cue can also be utilized as a pictorial depth cue. Moreover, Mather & Smith (2002) discussed via an experiment how blurring can be useful in establishing depth order. As indicated in Figure 5.4, they used two regions, one sharp and one unfocused, adjacent along a sine-waved edge. They required viewers to select the nearest region to them; they found that most of the observers selected the sharp part as the closest region. When they made the separated line gradually blurred, they got the opposite results from the observers. In this case, most observers selected the blurred part as the closest.

Held & Cooper (2010) explained how the pattern of blurring, together with the relative depth cues, can be used to estimate the distances of the objects in a scene. The blurry part that

they used in their experiment was also transparent, allowing observers to see underneath it but not clearly so they argued that, depending on the amount of blur, the distance can be estimated.



(a) (b)

Figure 5.4: Stimuli used in an expeiment (Mather & Smith, 2002) asking the user to select the closest part of the two adjacent parts. (a) Most users selected the sharp part as being closer, whilst in (b) the blurred part was selected as the closest part. Image re-produced from (Mather & Smith, 2002).

We can conclude that blurring is an important cue in layer discriminating and in layer depth-ordering tasks in a multiple-layered image. The experiment described in this chapter aims to determine which type of blur can be more effective: uniform or non-uniform. In the former, all pixels have the same amount of blur applied, while each pixel in the latter has a randomly generated amount of blur applied. In this study, we tried to find out if there is a significant difference between the two types of blurring in a layer depth-ordering task.

5.2 Experiment

The main goal of this experiment was to compare the strengths of two different types of simulated translucency in a depth-ordering task. One type of simulated translucency simulates diffusion inside the material by blurring with a Gaussian filter. The other type of translucency employs non-uniform blurring, where the standard deviation of the Gaussian filter varies from pixel to pixel. Our hypothesis is that the latter type evokes a stronger impression of semi-

transparent material and thus results in better performance in a depth-ordering task. To produce a translucence like effect on a plane, we used the following procedure:

- For every pixel in the back layer, a different amount of Gaussian blur was applied. This was done by convolving the image with a Gaussian kernel. A different size was randomly assigned to the kernel, depending on the value of sigma (σ). The σ value (between 0.1 and 3) for each pixel was randomly selected. We called this type of blurring *non-uniform*, as shown in Figure 5.5. This type was compared with a normal type, which we called *uniform*, since the kernel size was fixed for all image pixels with σ=3.

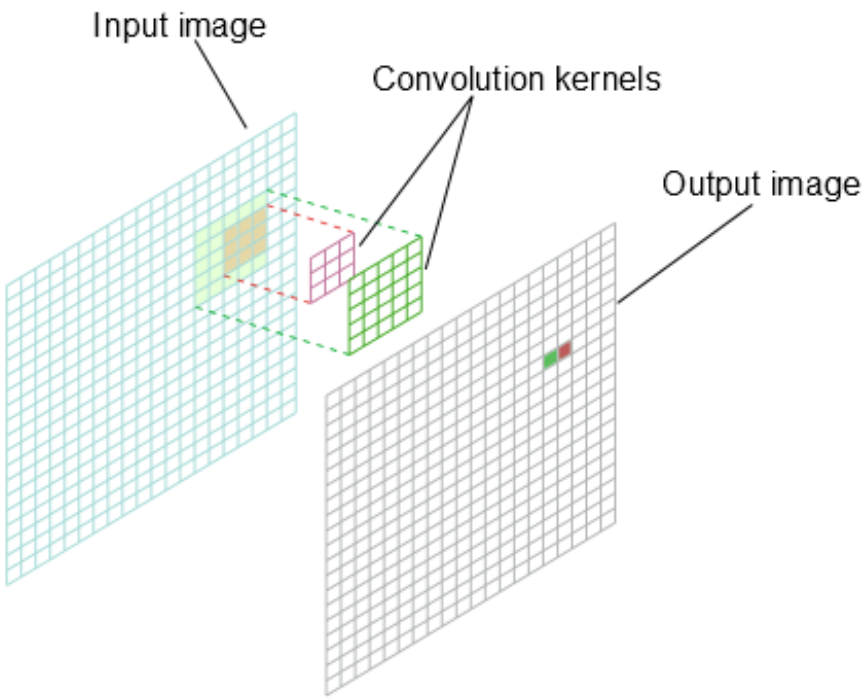


Figure 5.5: Producing a non-uniform blurring example. Every pixel in the output image has a different convolution kernel size, which has been selected randomly. The red pixel is the result of convolving the red convolution kernels, and the green pixel is the result of convolving the green convolution kernel.

- For both types of blurring, the kernel matrix size depended on the value of σ (NVIDIA,2014),

as in

$$s = round(3 \cdot \sigma), \tag{5.1}$$

where s is the size of the kernel matrix, and the following equation was used to generate the content of the kernel matrix:

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}, \tag{5.2}$$

where σ is used to control the amount of blurring and x & y are the distances from the centre of the kernel in the horizontal and vertical axis, respectively.

5.2.1 Stimuli

After generating the translucency layers, an alpha-blending based on Metelli (1974) model was applied to produced transparency, using Equation 5.3:

$$I = I_a\alpha + I_b(1-\alpha), \tag{5.3}$$

where I is the intensity of the result layer by compositing the intensity of layers a and b. α specifies the amount of intensity of each pixel in the layer a compositing with the 1-α intensity of the corresponding pixel in the layer b. The α value in this experiment was 0.5.

Each stimulus consists of three layers, as shown in Figure 5.6. Both the front and the back layers have the same size; they are bigger than the layer in the middle. The front layer is shifted to the right and bottom, which helps us clearly see the back layer (see Figure 5.7). The front layer is transparent, but with blurring added using either uniform or non-uniform blurring. The middle layer includes a shape of a window and is smaller than both the front and the back layers. The grey levels of both the front and the back layers were fixed throughout the experiment, while the grey level of the middle ranged from bright to dark. The minimum

brightness value used in this study was 0.3, increasing regularly by 0.1 to a maximum value of 1. We started with 0.3 to avoid reaching the same grey value of the back layer (0.2).

There were eight stimuli for each type of blurring. Every stimulus was repeated three times to make sure that the observer would not answer randomly. Thus, 48 total stimuli were shown to each observer.

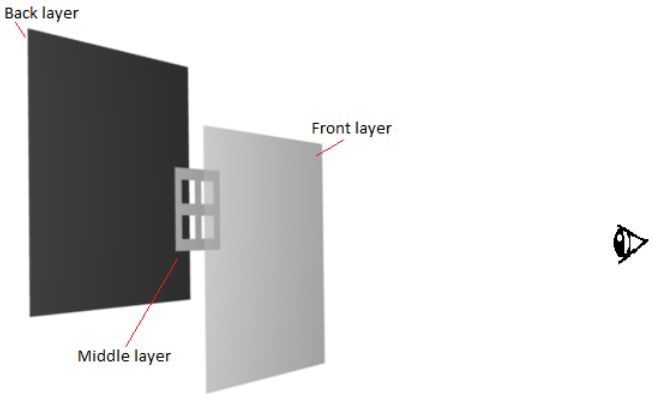


Figure 5.6: The stimulus' main parts. Each stimulus consists of three layers; the first is transparent.

(a) Non-uniform with high brightness

(b) Uniform with low brightness

(c) Uniform with high brightness

Figure 5.7: Experiment stimuli samples. The front layers of samples (a) and (c) have different types of translucency, but the brightness value for the window shape is the same in both samples, while the type of translucency of the front layer is the same in samples (b) and (c), though the brightness value of the window shape is different.

5.2.2 Apparatus

The images were assumed to be in the standard sRGB colour space, and were displayed using a normal TFT LCD monitor with a 17−inch screen and resolution of 1044×900. The viewing distance between the observer and the display screen was approximately 60 cm.

5.2.3 Observers

Fourteen volunteers, (4 female, 10 male), aged 23 to 40 years old, participated in this experiment. All observers had normal or corrected-to-normal vision. All observers were naive as to the objective of the experiment.

5.2.4 Experimental Procedure

A single image was displayed to the observer each time they were asked if the window shape was in front of or behind the transparent layer. The brightness of the window shape varied from trial to trial. Uniform and non-uniform blurring was presented randomly.

The strength of the translucency as a depth cue was measured by introducing a conflicting depth cue. For that conflicting cue, we chose to use object (a window shape in the middle layer) brightness. Several studies, including the work in Chapter 4, have shown that brightness is a strong monocular cue when presented in isolation. When the brightness of the middle layer is high, the observer might see the middle layer as being closer than the front layer because of the effect of the brightness cue. However, if the brightness of the middle layer is low, the object will appear behind the front layer. The difference in the brightness level that makes the object appear in front of the translucent layer for uniform and non-uniform blurring shows the effects of the strength of each type of translucency on depth perception.

5.3 Results

The best way to statistically analyse the results is to use a two-way analysis of variance (ANOVA) test. We used this test to determine if the blur type, brightness value, or blur and brightness value impacted the results. Thus, three null hypotheses were used. The first, H_{0A}, assumed that the blur type would have no significant effect on depth ordering. The second, H_{0B}, assumed that the brightness value would have no significant effect on depth ordering. The third,H_{0AB}, assumed that both factors together would have no significant effect on depth order.

After the two-way ANOVA test was applied, the first null hypothesis was rejected, with F_A=5.54 and p=0.0188. This means that the blur type had a significant effect on depth ordering for at least one brightness level. The second null hypothesis was also rejected, with F_B=77.74 and p=0, which means that the brightness value had a significant effect on depth ordering. However, the third null hypothesis was accepted, with F_{AB}=0.36 and p=0.9253, indicating that there was no significant effect on depth ordering by interactions of both factors. The level of significance used in the test was 0.05. To visualise the result, a cumulative Gaussian distribution curve was fitted, as shown in Figure 5.8.

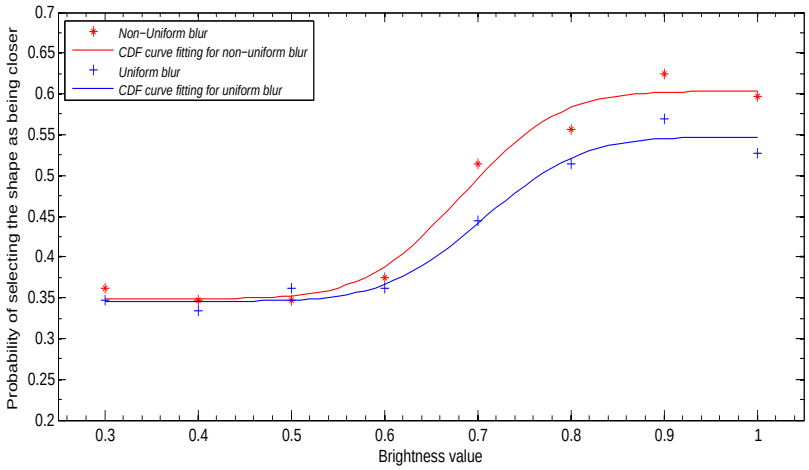


Figure 5.8: The result of the experiment. The non-uniform blur seems closer in all levels of brightness used in the experiment. No noticeable difference can be seen in the levels of brightness less than 0.7. A cumulative Gaussian distribution fitting was applied to visualise the result.

The figure shows that the probability values ranged between 0.3 and 0.7 for every brightness value, rather than between 0 and 1. This indicates that the observers were not in perfect agreement when they made their decisions. The figure demonstrates that there is little detectable difference between the two types of blurring at the lowest values of brightness, while the difference increased clearly at higher values of luminosity. It is also illustrates that the window shape was seen as being closer for the non-uniform type more than for the uniform type through all brightness values. This means that the non-uniform type appears more transparent than the uniform type, which was the opposite of our hypothesis.

The primary objective of the study was to determine if there a substantial difference between the two types of blurring in a depth-ordering task. Bootstrapping was used to determine the distribution of that data (Efron, 1979; Howell, 2013). The observer's outcomes were re-sampled 5,000 times using a bootstrapping function, and a cumulative Gaussian distribution was applied each time. To determine the distribution, the brightness value at specific points was selected for all 5,000 re-samples. The specific point selected was the middle of the probability of seeing the window shape as being closer. The histogram of the result for both types of blurring is shown in Figure 5.9.

The figure shows that the distributions of both types are not the same and the non-uniform distribution is not Gaussian. To find out whether there is a significant difference between the two types of such distributions, bootstrapping was applied to both types, allowing us to find the difference between them, with cut-off values of 2.5 and 97.5 percentiles (Howell, 2013). If the limits included the 0.00 value, we could conclude that the difference was not statistically significant between the two types, as shown in Figure 5.10. The limits, shown in red lines, are -0.09 and 0.7, which means that there is no statistically significant difference between the two types of blurring. Therefore, the experiment could not determine which type of blurring was a stronger depth cue.

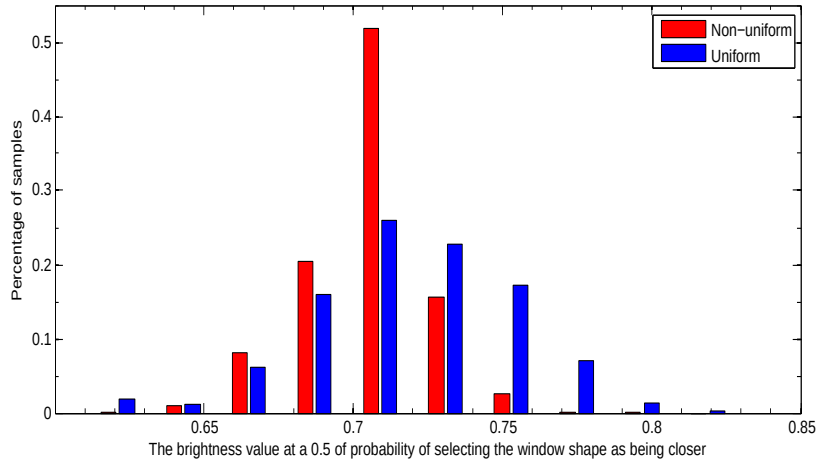


Figure 5.9: A grouped of histogram for both uniform and non-uniform blurring after applying the bootstrapping procedure 5000 times and selecting a specific brightness value. The figure shows that the distributions are not the same.

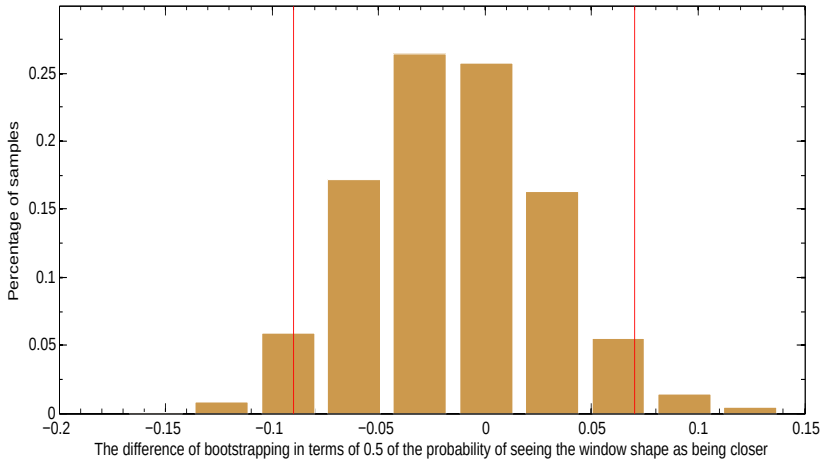


Figure 5.10: The differnce in the bootstrapped data for both uniform and non-uniform types, with the cut-off values of the 2.5 and 97.5 percentiles as the confidence limits. The red lines indicate the limits of the percentiles. The figure shows that there is no significant difference between the two types of bluring since the 0 value lies between the limits of the percentiles.

5.4 Conclusions

Based on the results of the experiment in this study, we are not able to decide which type of blurring (translucency) is better than the other in a depth-ordering task for multiple-layered

images. Although there was a small difference between the results of the two types of blurring that can be seen in Figure 5.8, it was not statistically significant. This difference was unexpected. We anticipated that the non-uniform type would be an effective depth cue because of the noise-like texture that resulted from randomly applying a Gaussian blur equation, but the results demonstrated the opposite. It seems that the layer blurred with the non-uniform type looks more transparent and less translucent than the layer blurred using the uniform type. One possibility is that not all pixels in this type of blurring have the same amount of blurring. Moreover, some have no blurring at all.

5.5 Limitations and Future Work

The stimuli used in this work were in the grey-scale level only, which might be regarded as a special case. To make the study more comprehensive in the future, the colour factor should be added. It would be better to see whether the translucency types have the same effect in colour stimuli as in the grey scale. In addition, the brightness value was the main changeable part in this experiment, while the style of displaying the stimuli was kept unchanged. By taking different styles for displaying the stimuli, the accuracy of the results should increase. For instance, we can use different grey-level values for all three layers and different shapes, while keeping the main idea unchanged, which is the shape layer lying between the other layers. Based on the outcomes of this work, we will focus more on the high values of brightness for the future experiments, since there was no detectable difference in the low brightness values. The σ value used to generate uniform blurring was kept the same for all stimuli, so checking different values for σ will provide more accurate results. In the end, the experiment was based on 2D flat planes for all layers, and it would be safer in the future if it includes 3D shape stimuli, such as curved surfaces.

5.6 Summary

This chapter has investigated the effect of translucency on depth perception. Two common types of translucence, uniform and non-uniform, were included in this study. The primary target was to find which type provides a stronger impression on depth perception. Our hypothesis was that the non-uniform translucency would provide a stronger impression. However, according to the outcomes of an experiment conducted for this study, we were unable to determine which type is stronger, since there was no statistically significant difference between the two types.

Chapter 6

The Impact of a Cutaway shape on Depth Perception in Three Dimensional images

Psychological studies revealed that a closed shape might be perceived either as a hole (cut) or as an object (figure). These studies used 2D images. Cutaway, is one of the techniques used in volume visualization to show the internal parts of 3D models. In this chapter, we tried to find out whether the shape of cutaway affects depth perception in 3D images. Three geometrical shapes were used. These shapes are rectangular, circular and convex. The results showed that the cutaway shape affects depth perception in 3D images, and the convex shape is preferred over other shapes to be cutaway shapes.

6.1 Introduction

The effect of cuts and hole shapes on the depth perception have been studied extensively by psychologist (Nelson et al., 2001; Kanizsa & Gerbino, 1976; Bertamini & Croucher, 2003;

Bertamini, 2006; Bertamini & Lawson, 2008). Their primary motivation behind these studies was because of the ambiguity that the shape of the hole provokes. They found that a hole can be perceived as a hole through a surface or sometime as an object in front of the surface. Their studies focused on finding the reasons that make our visual system depict the hole as an object or not. Figure 6.1 illustrates the ambiguity when trying to analyse the white disc in the centre which could be perceived as a hole in a black surface or as a white disc on the black background. They argued that there are clues which help the visual system to correctly perceive a hole and the most common cue in their study was the shape of a surrounded region of a hole. They agreed that the shape of the surrounded region has an affect whether or not it is perceived as a hole. The concave shape of the contour is more likely to be perceived as a hole whilst the convex shape perceived as an object. Their experiment were based on 2D surfaces.

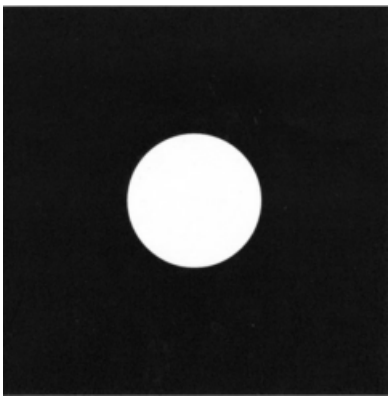
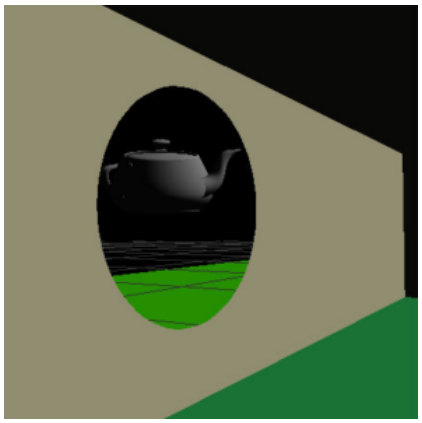


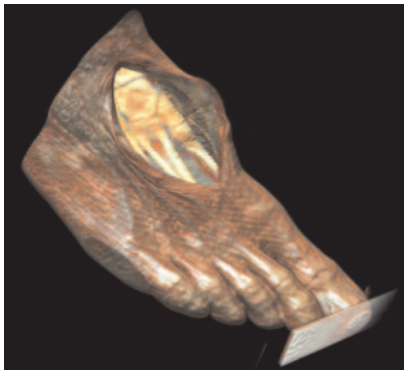
Figure 6.1: The ambiguity is in analysing the white centre which is could be perceived as a hole or as a disc on the black background. Image re-produced from (Nelson et al., 2001).

On the other hand, the effect of cutaway shape on a 3D surface was hardly studied. The cutaway is one of the techniques that is used in volume visualisation. Most researches dealt with techniques that were based on revealing the hidden parts inside a volume and they used random shapes for the cut. However, a few studies took into account the shape of the cut. Coffin & Hollerer (2006) used two geometric shapes; rectangular and circular, in addition to a random shape in their study, Figure 6.2(a). Correa et al. (2006) and Liang et al. (2005) tried

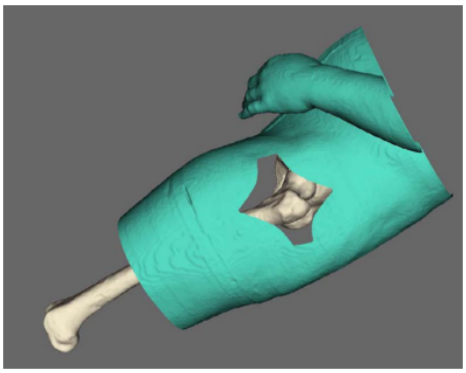
to use in their study an incision that reflects the shape expected by the surgeon when retracting the skin from the incision, as shown in the Figure 6.2 (b) and (c) respectively. Although they used a special cut shape in their research, they did not study the effect of the cutaway on depth perception.



(a)



(b)



(c)

Figure 6.2: Cutaway exmples. (a) A circular shape of cut is made to see behind the wall. Image re-produced from (Coffin & Hollerer, 2006), (b) and (c) the shape of cut is the same as a cut made by a surgeon. Images re-produced from (Correa et al., 2006) and (Liang et al., 2005) respectively.

The main objective of this chapter is to determine whether the cutaway shape influences the depth ordering in curved surfaces as in the planar surfaces. A geometrical shaped cutaway has been made in such a way that the internal parts can be seen through it. An ambient occlusion effect was added to the part under the region of the cut to add a 3D effect. Three geometric

shapes have been taken into account: circular, rectangular and convex. The reason for this choice is because firstly, these shapes were used in previous studies that cover the cutaway technique (Coffin & Hollerer, 2006; Liang et al., 2005) and secondly, using a geometric shape will help to automate the process of the cutaway technique.

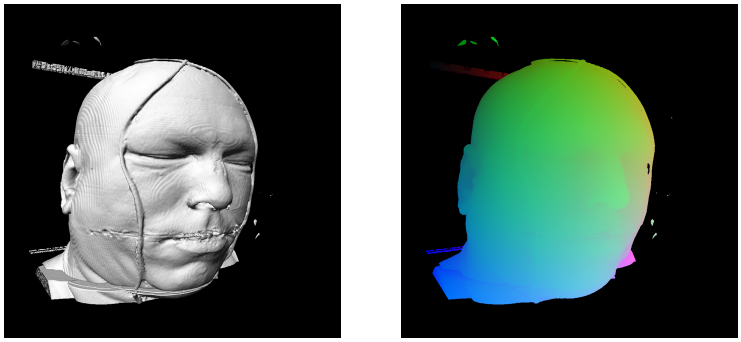
6.2 The Procedure

6.2.1 Simulating a Cut

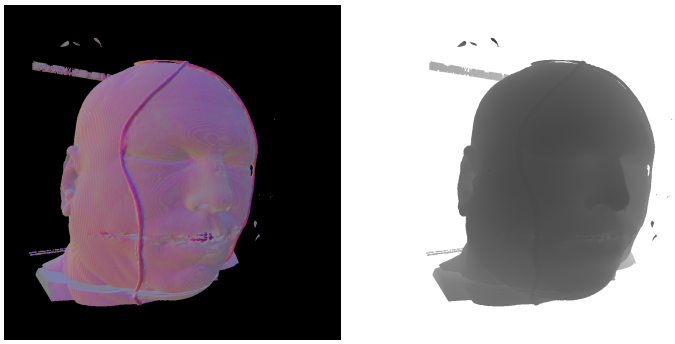
To asses the effect the shape of a cut has on the perception of depth in a 3D rendering, a multi-layered model is required. The outer layer of this model should encompass the internal layers. When cutting into the outer layer, it forms the only method of examining the internal layers. The shape of this cut impacts the intuitive perception of depth, in relation to the outer and inner layers. To examine this effect, a model of a head and skull were used. The model was taken from computerised tomography (CT) scans[1], and rendered using the *simple isosurface method* Hughes & Lim (2009). Each model was represented by four images, of size 1024×1017 pixels. These images are intensity, position, normal and depth, as shown in Figure 6.3. Images courtesy of (Hughes, David M.) at Hughes & Lim (2009).

The intensity image contains the intensity value of each pixel. Each of the RGB channels are rendered with the same intensity value. Therefore, the resulting image appears grey-scale, as shown in the Figure6.3(a). The position image holds the location values, in 3D, for each voxel in the image, see Figure6.3(b). The normal image, Figure6.3(c), includes the normal vector values for each pixel, which are used to determine the orientation of the surface. Finally, the depth image, contains distance information in 3D, from a specific viewpoint, see Figure6.3(d). The intensity image was used as the outer surface. The only change conducted was colouration of the intensity image. Normally, a portion of the skull is seen when the cor-

[1].The volume was downloaded from http://www.nlm.nih.gov/research/visible/



(a) Intensity (b) Position



(c) Normal (d) Depth

Figure 6.3: Isosurface volume rendering for the head CT scan[1] (Hughes & Lim, 2009). (a) *Intensity* image, includes the intensity value of each pixel, (b) *Position* image, contains information about the locations of each voxels in 3D space, (c) *Normal* image, contains the normal vector value of each pixel, and finally, (d) *Depth* image, which contains the depth information for each pixel to a specific viewpoint. Images re-produced from (Hughes & Lim, 2009).

responding part of the skin (outer surface) is removed. Therefore in order to distinguish the revealed skull from the skin image, the outer surface was coloured.

Based on these images, the following steps were followed to simulate a cut on the skin's surface:

1. On the *skin* layer, a point was selected to form the centre of the desired cut.

2. In the *position* image, a point $P(x_0,y_0,z_0)$ that corresponded to the selected point in 1 was determined. This point will be used as a first endpoint of a pseudo line L. This line will be used later to adjust a plane, and also as a direction to the projected shape on the

position image.

3. The direction of the line L was determined by taking a point in the *normal* image $N(n_x,n_y,n_z)$, which corresponds to the point P.

4. Calculating the other endpoint $P_1(x_1,y_1,z_1)$ of the line L, using the parametric equation for a line in 3D space, as follows:

$$\begin{aligned} x_1 &= x_0 + n_x.t, \\ y_1 &= y_0 + n_y.t, \\ z_1 &= z_0 + n_z.t, \end{aligned} \tag{6.1}$$

   where t is the distance from P to P_1. At this point, the pseudo line L was generated (the red line in the Figure 6.4).

5. A 2D plane that contains the endpoint P_1 and perpendicular to the line L was prepared. The desired shape will be set in the plane and the point P_1 will be its centre.

6. Projecting all the plane pixels on the *position* image. This done by generating pseudo lines, which their start points were the pixels in the plane, and they were parallel to the line L, and have the same length of t. This process is similar to the ray-casting procedure, used in direct volume rendering. All pixels in the *position* image that intersect with these pseudo lines will be removed, as shown in the Figure 6.4.

7. Finally, on the *skin* layer, all pixels that correspond to those previously removed, are also removed. Figure 6.4 illustrates the procedure of producing the cut.

Appendix B contains a MATLAB function, which was used to generate the requested shape on a plane.

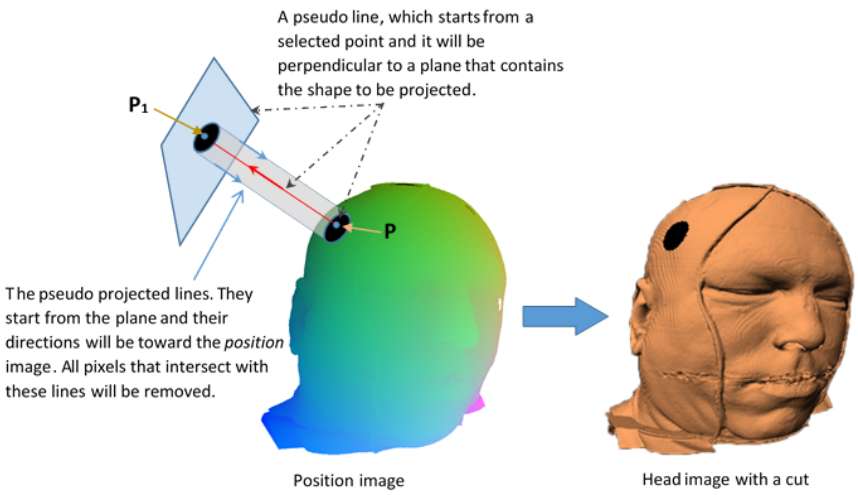


Figure 6.4: An illustration of producing a cut. Images re-produced from (Hughes & Lim, 2009).

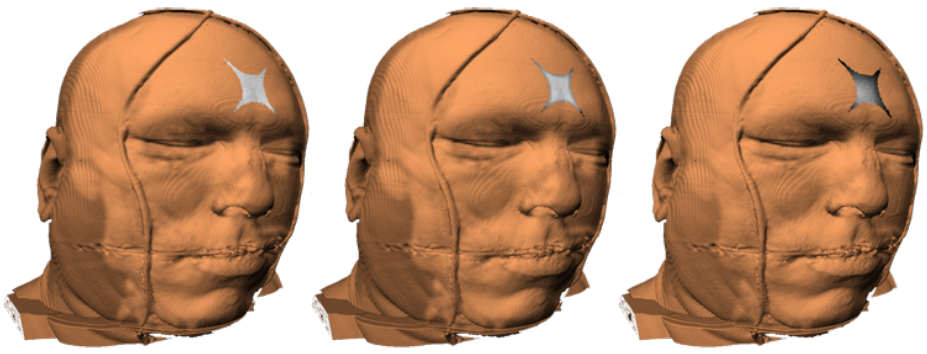
6.2.2 Preliminary Experiments

According to the composition of the human body, a part of the skull will appear if an incision is made to the skin on the head. Therefore, to simulate a cut, each pixel representing the skull that correspond to those removed will be displayed instead, as seen in Figure 6.5(a). The white convex shape appears as if it has been pasted on the surface more than it appears as a cut. In such a case, a monocular depth cue should be added to the image. The shadow depth cue was selected, since it is more appropriate cue in such situation. We first tried to generate a shadow effect, using the method described by Luft et al. (2006). Their method produces a flat shading between objects, that contain depth to separate them visually, which was the desired effect. Figure 6.5(b) shows the result of applying this method.

It was prudent to determine whether this type of shading influences depth perception before running the main experiment. The main experiment was to compare the effect of the shape of a cut on the depth perception. In order to check the effect of this method, a short empirical experiment was conducted to determine whether the observer perceived the convex shape as if it was located on or under the skin of the head, as depicted in Figure 6.5(b). To prevent the

participants from using their background knowledge about human anatomy, only a rectangular portion containing the cut shape was displayed. Five observers participated in this experiment. Nine images were displayed, three for each type of cut, a single image at a time. The shapes used were circular, rectangular and convex. Observers were asked to select the closest part of the rectangular portion to them, selecting between the white (skull) and the skin colour. The results showed that most of the observers (4 out of 5) chose the white part as being the closest. One of the reasons for this was because there was an equal amount of shade spreading along the whole contour of the shape, between the white and skin colours, which gave the impression of shading rather than shadowing. Thus, they selected the white portion as being the closest.

In our next attempt, ambient occlusion was used to generate a shadow-like effect, as shown in Figure 6.5(c). The same empirical experiment was performed again. The only different was the implementation of the shadow by using the ambient occlusion method. The outcome was the opposite of the previous trial. Almost all the observers (4 out of 5) chose the skin colour as being closer. This result gave us the motivation to run the main experiment using ambient occlusion, to find the effect of the shape of a cut on depth perception.



(a) (b) (c)

Figure 6.5: Enhanching the impression of a cut. The cut in (a), appears as if a shape was pasted on a head. In (b), after applying the algorithm found in Luft et al. (2006), the cut looks convincing compared to (a), but it also gave an ambiguous depth ordering impression. Finally, in (c), after applying ambient occlusion, the cut looked more convincing overall.

Stimuli were prepared using ambient occlusion, as shown in Figure 6.6. If the observers

were asked to select the shape that looks like a cut, all the shapes will have the same opportunity, The observers may depend on their knowledge about the human anatomy when they answer. In this case, all shapes seem as a cut. To prevent observers from using their knowledge, the part that contains the cut will be displayed only, as shown in Figure 6.7.

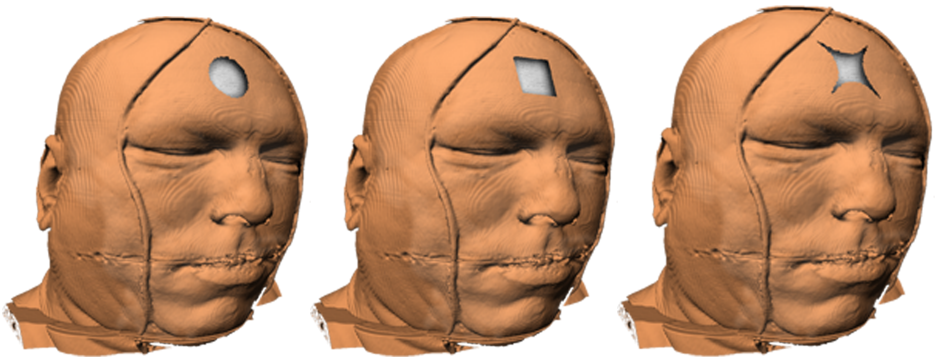
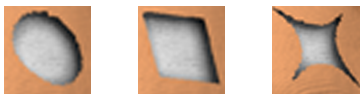


Figure 6.6: Three different cutaway shapes used in the experiment.

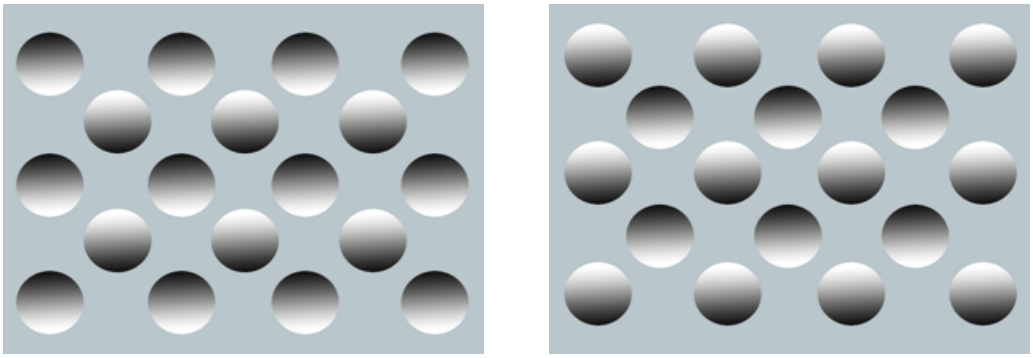


(a) (b) (c)

Figure 6.7: Example of stimuli used in the experiment. To prevent observers to use their knowledge about the human anatomy, only a part of the image, which contains the cut, will be shown through the experiment. (a) circular, (b) rectangular and (c) convex. All cuts are approximately the same size.

According to Kleffner & Ramachandran (1992); Ramachandran (1988), flipping an image containing shading might change the perception, as shown in the Figure 6.8. Figure 6.8(a) contains circles which are shaded to appear as spheres. The spheres in the second row are exactly the same spheres in the first row, the only different is the spheres in the second row are flipped by 180°. The first row appears as a set of concave holes in the plane while the second row appears as convex. Figure 6.8(b) is the same as Figure 6.8(a) but is flipped vertically. Based on this, we tried to find out which cut shape is not affected by this factor. The shapes

were rotated in the following angle degrees 90°, 180° and 270°. The other reason for flipping the image was to simulate different light directions, which display more natural cases in volume visualisation that use a cutaway technique.



(a) (b)

Figure 6.8: The effect of shading on depth perception. The image (b) is a copy of image (a) but is rotated by 180°. The result of this rotation is that all concave circular become convex and vice versa.

The main idea of the experiment is based on paired comparison. A pair of images was displayed to the observer. These images look the same but they only differ in the shape of the cut. Then the observer is asked to select the image in which its grey part looks deeper than the orange part. If both images were convincing or it was difficult to decide, the observer would select randomly.

6.3 Statistical Analysis

A subjective test was used to evaluate the effect of cut shape on depth ordering using paired comparison. A paired comparison is useful in cases when the objects to be compared can be only judged subjectively (David, 1963a). There are two types of paired comparison, *balanced* and *incomplete* (Setyawan & Lagendijk, 2004). The balanced type is used in this study because the test subject had to evaluate all possible comparison pairs and because the number of objects

in the test was not too large.

Let t be the number of cut shapes that we want to compare. Performing all comparisons of t objects will produce a $\binom{t}{2}$ paired arranged in $(t \times t)$ matrix. The matrix is called a *two-way preference matrix*. In balanced design, ties are not allowed which means that the observer has to vote one cut shape from a pair. Table 6.1 illustrates an example of such a matrix for t =4.

	O_1	O_2	O_3	O_4
O_1	-	1	0	1
O_2	0	-	1	0
O_3	1	0	-	1
O_4	0	1	0	-

Table 6.1: Two-way preference matrix example with four objects, the user has to vote on all comparisons. It is not allowed to compare the object with itself. Number 1 in the first row and third column indicates that O_1 is preferred to O_2

The number of objects in table 6.1 is 4 , this results in $\binom{4}{2} = 6$ comparison pairs. For each pair, the observer's choice was recorded. For example, the cell in row 1 and column 2 has the value of 1 which means that the observer considered that the image of cut shape of type O_1 gives a better order of depth than the image of cut shape type O_2. This can also be written as $O_1 \rightarrow O_2$ which is interpreted as O_1 *is preferred to* O_2.

After the preference matrix has been obtained, we have to analyse the test results. The main objective of this analysis is to see how large is the difference between the test objects (cut shapes in this study).

6.3.1 Coefficient of Agreement

Also known as Kendall Coefficient of Agreement(Ledda et al., 2005), which measures the difference of the preferences among all observers. When all observers make the same choice, we can say there is a complete agreement. To calculate the coefficient of agreement we have to apply the following equations:

$$\tau = \sum_{i=1}^{n} \sum_{j=1}^{n} \binom{O_{ij}}{2}, \mathrm{i} \neq \mathrm{j} \tag{6.2}$$

where n is the total number of observer and τ is the total number of agreements among n observers evaluating t objects. Then the coefficient of agreement can be found by using the following equation (Kendall & Smith, 1940):

$$u = \frac{2\tau}{\binom{n}{2}\binom{t}{2}} - 1 \tag{6.3}$$

u is equal to 1 indicates complete agreement, when all observers made identical choices. Smaller u means the less agreement between the observers. The minimum value of u is -1/(n-1) when n is even and -1/(n) when n is odd.

Having computed the coefficient of agreement, it is important to test the significance value. Details of these tests can be found in the Appendix A.

6.4 The Experiment

The experiment was conducted using an on-line SurveyGizmo[1] survey (See Figure 6.9). The survey started by giving the participant the instructions about the experiment and what is required. The survey consisted of several pages and each page contained a pair of images. The user was asked to select the image that gives him a better impression that the grey part in the image is underneath the red part. Three geometrical cut shapes were used in the experi-

[1] www.surveygizmo.com

ment *convex*, *rectangular* and *circular*. Each cut shape had three copies and they were rotated by a different angle. These angles are 90°, 180°,270°. In addition, there is another copy of each shape but without ambient occlusion. This will produce a 2D image, see Figure 6.11, which will help to compare the results of this experiment with the previous studies result.

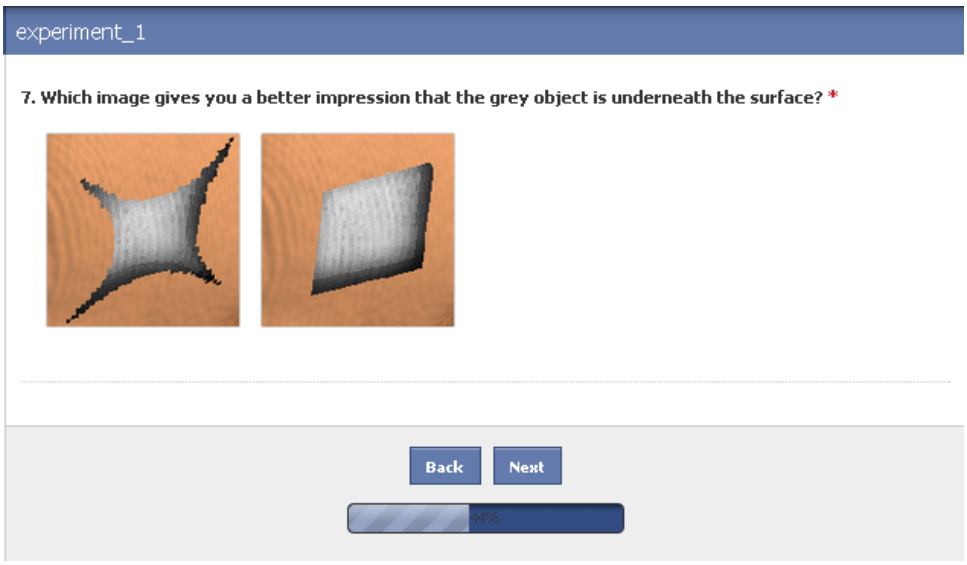


Figure 6.9: A screen shot of a user interface used in the experiment, showing convex and rectangular cut shapes. The observer has to select only one image by clicking on it

Thus, for every type of cut shape we have 5 images. We divided all images into groups based on the angles of each group. Then each group consists of 3 images. The images in the same group compared to each other to ascertain whether the rotation affects the observer's decision and to find out which shape is not affected by the rotation. After distributing the survey we received 92 outcomes. We eliminated 7 of them because they were not fully resolved. A total number of observers participated in the experiment was 85 and they were from 14 different countries around the world.

6.5 Results

The result of paired comparison data from the 85 subject tests, see Appendix C, was arranged in a preference matrix. The experiment resulted in five preference matrices for five groups. Each group consisted of all three cut shapes with the same light source direction. The direction of the light source was only different between the groups. Except the fifth group, which contained only the cut shapes without ambient occlusion 6.11. The highest coefficient of agreement is found in group 4 (0.0525). The preference matrix of group 4 is shown in Table 6.2. The number in each cell represents the total number of times that this cut shape was considered as having a better effect of depth ordering. For instance, the second cell in the first row is 54 representing that convex cut shape was judged as a better impression of depth 54 times out of 85 than a rectangular cut shape.

Table 6.2: The group 4 preference matrix which has got the highest coefficient of agreement . The table shows the result of paired comparison of group 4.

	Convex	**Rectangular**	**Circular**	**Total**
Convex	-	54	57	111
Rectangular	31	-	44	75
Circular	28	41	-	69

Table 6.3 shows the total values of overall results. The second column in Table 6.3 shows the coefficient of agreement amongst subjects through the test. The complete agreement is found when half of cells in the Table 6.2 is 85 and the other half is 0 which means that all participants have the same selection.

The third column of the table represents the results of *Kruskal-Wallis* test (H), explained in the Appendix A, is statistically significant ,which means that its value is equal to or larger than

the critical value of Chi-Square for a particular degree of freedom,in our case is 5.99, then we can say that there are differences between cut shapes, although we do not know where these differences lie. Table 6.3 shows that the H value of all groups are statistically significant.

Significance test of the score differences is performed in order to see whether the perceptual quality of any two cut shapes is perceived as different. Otherwise, we may have to conclude that the perceived quality of the two cut shapes is similar. In other words, we want to find $\acute{R}$ such that the probability P(R $\geq \acute{R}$) is less than or equal to the significance level α (usually α =0.05).

From Peaeson & Haetlet (1976) we have $W_{t,\alpha}$= 3.31. Substituting this value into Equation (A.2), we have R = 26.678 and thus we set R = 27. Therefore, only objects having a score difference of more than 27 are to be declared as significantly different.

Table 6.3: Overall results. The table shows that the convex cut shape has the higher rate of selection, shown in a green colour, than the other cut shapes in all groups when ambient occlusion was applied. Without ambient occlusion the circular cut shape has the higher selection.

Group	**u**	**H**	**Score**			**Rank**		
			Convex	**Rectangular**	**Circular**	1^{st}	2^{nd}	3^{rd}
1	0.0308	13.49	102	86	67	Convex	Rectangular	Circular
2	0.0446	14.09	105	68	82	Convex	Circular	Rectangular
3	0.0383	16.13	107	76	72	Convex	Rectangular	Circular
4	**0.0525**	**20.29**	111	75	69	Convex	Rectangular	Circular
5	0.0331	14.02	94	64	97	Circular	Convex	Rectangular

Figure 6.11 shows the result of applying the significance test of score differences for each group. The figure shows that the convex cut shape was highly preferred in all groups except the group without ambient occlusion. Figure 6.10 shows the significant difference of all groups together. It shows that convex is preferred among the other shapes. The reason could be because of the effect of ambient occlusion. Considering the image in Figure 6.7, the area covered by the shadow is greater than the other shapes area. The second possible reason is that the convex shape does not allow the observer to believe that the light being emitted sits on top of an elevated object. Such that, the lower-left tail of the convex adds extra information (regarding the light source) compared the other two shapes. The other possible evidence is that in circular and rectangular shapes we cannot determine the light direction in relation to the shadow. However, it is clear that the direction of light can be determined by the convex shape.

The results also showed that both the shapes rectangular and circular were affected by the rotation. Figure 6.11 shows that the circular cut shape in group 2 was preferred to the rectangular cut shape while in group 1 the rectangular cut shape was preferred. The convex cut shape was not affected by rotation.

The most surprising result was in the fifth group, in which the ambient occlusion was not used. According to the psychologists Nelson et al. (2001); Kanizsa & Gerbino (1976); Bertamini & Croucher (2003); Bertamini (2006); Bertamini & Lawson (2008), the convex shape should get the highest rate of selection among the other shapes and the circular shape the lowest rate. The results showed that the circular has got the highest rate.

Shadow was one of the factors suggested by Nelson et al. (2001) to increase distinguishing holes from the object. The results showed that this factor does not have a big effect when it is used with a specific shape, since although the rectangular and circular have a shadow but they still seem as an object not as hole. We expected the reason was because the observer looked at the shadow as a shade.

Table 6.3 shows the complete results of the overall experiment. For each group, the ta-

ble shows the coefficient of agreement(u), Kruskal-Wallis test (H), which is explained in the Appendix, and the rank of cut shapes from the highest preferred to the lowest preferred.

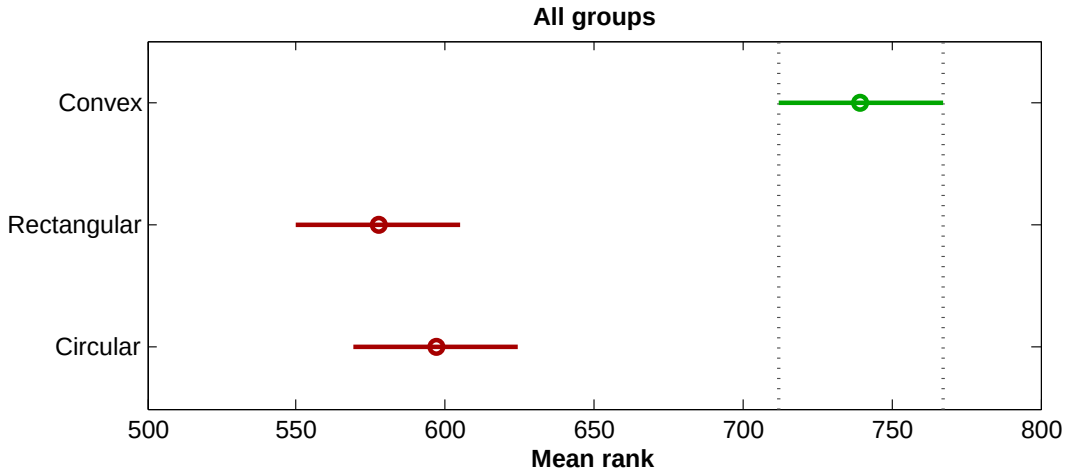


Figure 6.10: The result of all groups together. The convex cut shape is preferred to the other cut shapes in term of giving a better impression of depth order

6.6 Conclusions

This chapter focused on finding the effect of cutaway shape on depth perception, and the result showed that the cut shape influences the depth impression and the depth ordering on 3D images as well as in 2D images. The result showed that most observers preferred a specific shape, the convex shape in this study. Nelson et al. (2001) suggested that adding factors to a hole will remove ambiguity of seeing the hole or seeing it as an object, and the shadow was one of the suggested factors. However, this factor has not had a big contribution in this study since both rectangular and circular shapes selected as an object when they were compared with the convex shape even though they have a shadow. The reason could be because of both the rectangle and the circle shapes are very regular shapes that are rarely seen as a cut. Moreover, the observer may analyse the shadow in these shapes as shade. The other surprising result found in the images without shadow factor, or simply without ambient occlusion, the result showed that the

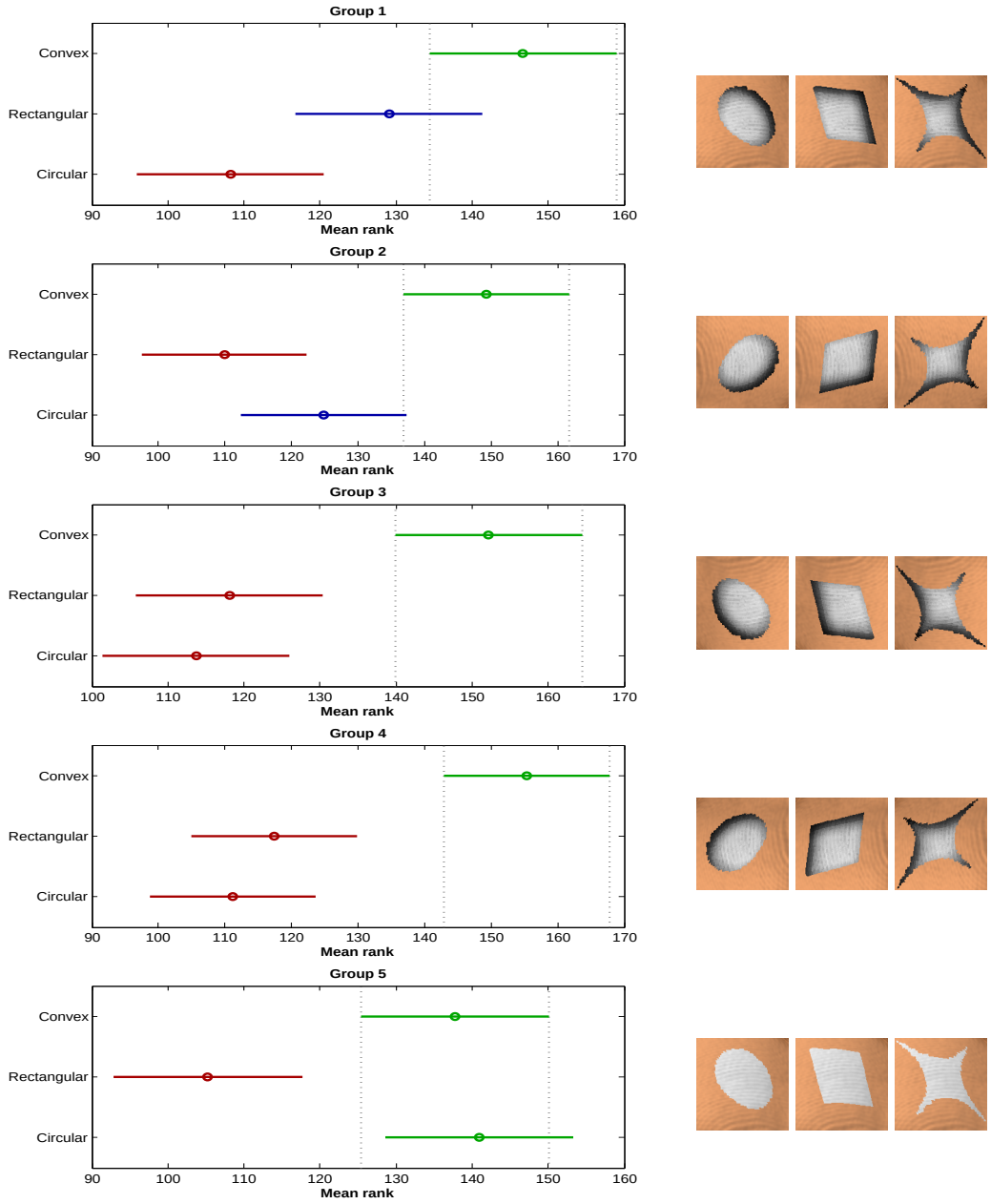


Figure 6.11: The result of applying significance test of score differences for all groups. The green line indicates this cut shape is significant difference from another cut shape and it has the higher rate and the other cut shape will be in red color. The blue color indicates that there is no significantly different of this cut shape with any other cut shapes

circular shape got the highest selection over the other cut shapes, which was different from what was expected.

The results also demonstrate that both rectangular and circular cut shapes were affected by the rotation which means that changing the direction of light may have an effect on the perception of the shape and the convex cut shape was not affected by rotation.

6.7 Limitations and Scope of the Work

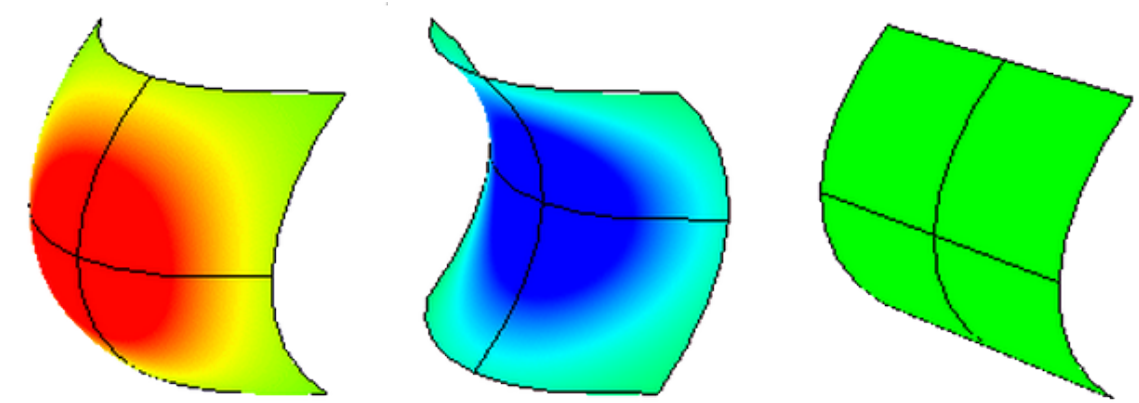
Our work has several limitations that will be addressed in the future work. According to the main types of curvatures that can be found with *Gaussian curvature* equation:

$$K = k_1.k_2, \tag{6.4}$$

where k_1 and k_2 represent the principle curvatures (Rhino5, 2014), the result of the product will be one of the three cases: positive, negative and zero. Each of which will result a different curvature surface. Figure 6.12 shows the types of curvature.

When the result of the equation 6.4 is positive the generated surface will be something like a bowl or a sphere surface as in Figure 6.12(a). Also, when the result is negative the saddle-like surface will be generated as shown in Figure 6.12(b) and finally when the result is zero, which means that at least one of the principle curvatures k_1 or k_2 is zero, the resulted surface will be flat at least in one direction as seen in geometric planes or cylinders as in Figure 6.12(c). Only one type of the Gaussian curvature has been used in this experiment which was the positive case. Trying all of the Gaussian curvature types will support the results more.

The experiment was also restricted with only one size for all cut shapes. To ensure that the impression of depth was not affected by the size of the cut, It would be advantageous to take different sizes for each cut shape into account. In addition, using different images in each group rather than a single image.



(a) (b) (c)

Figure 6.12: The result of Gaussian curvature equation 6.4. (a) A bowl-like curvature surface is the result when the Gaussian curvature is positive. (b) Saddle-like curvature is the result when Gaussian curvature is negative and finally, (c) when the result is zero the generated shape will be something like a planes, cylinders. Image re-produced from (Rhino5, 2014).

6.8 Summary

The main objective of this chapter was to check whether the shape of cut has an effect on depth perception or not. Three shapes were selected, based on previous research that dealt with volumetric cuts in visualisation applications. These shapes are circular, rectangular and convex. An experiment was conducted and the results proved our hypothesis that the depth perception was affected by the shape of a cut made on the outer layer of a multiple-layered image. They also showed that the convex shape cut had the highest correct rate than the other two shapes.

Chapter 7

Conclusions and Future Works

This chapter summarises all the works carried out in this thesis starting with an evaluation of the research hypotheses. It then discusses the main conclusions, and also outlines the novel contributions of this study. In addition, it highlights the possible direction, based on the outcomes of this study, for future work.

7.1 Research Hypotheses

Comparing the work conducted in this study with the hypotheses (listed in chapter one), the results of this research affirms the following:

(H1) This thesis hypothesised that finding a depth cue, which gives a better impression of depth intuitively when used individually, will be useful in generating computerised imagery containing depth. It will help to find which depth cue should be included in a scene, to increase the impression of depth in 3D computer imagery. In addition, the study also assumed that HDR display devices will be more useful in visualisation applications than the LDR displays. The results proved these hypotheses to be true, since the study revealed that the brightness and contrast cues were the most effective. They are

considered as the basic features of HDR display devices.

(H2) The study also postulated that a non-uniform type of translucent texture provides a better result in depth-ordering tasks in multiple-layered images than a uniform type. This hypothesis was not proved. The reason was because we were unable to tell which type was better, since the results showed that there is no statistical difference between the two types of simulated translucencies.

(H3) The last hypothesis of this study was that the shape of the cutaway made on the outer surface effects the impression of depth. The results proved this hypothesis to be true. However, because the main objective was to see whether the hypothesis will succeed or not, the experiment was kept simple. Thus, further work is needed to strengthen these findings.

7.2 Summary

This study has presented an assessment to several monocular depth cues on the impression of depth perception. All the cues chosen to be assessed were based on previous studies, which proved the importance of these cues on depth perception. The study has shown that each cue has an individual effect on depth perception. Seven monocular depth cues were tested via an experiment. These cues were blur, brightness, contrast, overlap, relative size, shadow and transparency. The effectiveness of the brightness, contrast and relative size cues was superior to the other depth cues used. However, the effectiveness of the transparency cue was immensely inferior. This result reveals the potential error when the transparent cue is being used, since it considered to be one of the most common cues used in the field of computer graphics.

The study has confirmed the benefit of using a HDR display device, and showed that it can outperform typical LDR devices in visualisation applications. For instance, the brightness cue was considered the least efficient cue in depth perception by Surdick (1994). However,

the findings of our study were opposite to this. One of the reasons that prompted our studies to use a HDR device was it can reveal information that is hidden or unavailable when using a regular display devices (LDR) that may affect any decision-making processes. According to the findings of this study (as mentioned earlier), the brightness and the contrast cues had the highest rate of correct selection. These are considered the most important features that HDR devices provide, and also proved our hypotheses about the importance of using HDR devices in such applications. We consider that this study is the first study that compares several monocular depth cues using HDR display devices.

The two well-known types of the transparency, uniform and non-uniform, have been simulated in this study by using a Gaussian blurring function. The reason was to find which type provides a stronger effect on depth perception when it is used in depth-ordering task in multiple-layered images. The result will be utilised to remove or reduce the ambiguity caused by transparency, especially when it used in a multiple-layered image; in case that the first layer is semi-transparent and larger than the other layers. The result of the experiment has shown that there is no statistically significant difference between the two types of translucency in depth-ordering tasks. Moreover, in contrast to our expectation, the study showed that the non-uniform type appeared more transparent and less translucent than the uniform type. As a summary of this experiment, which was carried out on 2D layers, we were not able to decide which type provides a better result in a depth-ordering tasks for multiple-layered images.

The study has shown that the shape of a cutaway, being made on the outer surface of a multiple layered-image, effects depth perception. This proved our hypothesis. In spite of the simplicity of the experiment, this thesis has precedence in the study of the effect of the shape on a cutaway, on depth perception. Three shapes were used in an experiment that designed to find the most influential shape on the depth perception. These shapes were chosen because they have been used in earlier studies, as shapes for a cutaway (Coffin & Hollerer, 2006; Liang et al., 2005). The study revealed that the best impression of depth will be perceived when the

shape of the cut is convex. This study has not taken the random shaped cut because it is hard to decide which shape will be perfect among random shapes.

7.3 Future Works

Based on the results presented in this thesis, the following sections detail the possible methods for future work.

The main objective of the experiment conducted in this study was to find the monocular cue that provided the best impression of depth. However, due to the fact that many recent studies in the subject of visualisation move towards real-time applications. It would be useful to conduct further experiments that focus mainly on the speed of the observer's response, besides the correct answer. It is obvious that the speed of the response to each cue differs according to the nature of the cue. Moreover, in order to obtain the best results, it would be advantageous to hold the experiment in two phases. The first one, would be to find the monocular cues that have a fast response, and the second would be to make a combination of these fastest cues. Since, according to Mather & Smith (2004), a combination will improve the accuracy and speed of depth detection. For a wider study, the stimuli should contain more complex images besides the CT or MRI scans. The results of these experiments can be utilised as an external function, to be added to applications that deal with such images. Additionally, the experiments should be conducted using a HDR display device, to obtain the best results.

One possible reason, which hindered our ability to decide on which type of translucency was stronger, might be because of all the stimuli that was used in this study were based on grey-scale levels. Human vision is weak when finding the difference between grey levels that have close intensities, when compared with coloured intensities. Applying the same experiment using coloured stimuli would make the results more comprehensive. The main difficulty might be choosing the colour for the planes. Because human vision responds to each colour in a

different way. Although colour is not an independent depth cue, it is strongly affected by luminance contrast and to an object's location in a scene (Guibal & Dresp, 2004). Moreover, some people suffer colour-blindness (also known as *colour vision deficiency* CVD). In addition, the chromostereopsis phenomenon is another problem seen on a computer monitor, which red objects appear closer to the viewer than blue objects (Thompson et al., 1993). However in general, we think that by using coloured stimuli rather than grey-scale, the results will be more comprehensive. In addition, in order to make the work more integrated, curved layers should be used instead the flat surfaces. Unlike the curved layers, flat surfaces can be argued as a special case. Finally, to strengthen the efficiency of the work, it would be beneficial to use algorithmic translucency textures, such as used by Yan et al. (2012) or Hendrik & Goesele (2005). This will help to affirm whether there was a clear difference between the two types of translucency or not.

To improve the work done in this study further, future work should focus on applying the cut on different types of curvature surfaces. The positive Gaussian curvature was the only type that was used in this study. There are two other types of Gaussian curvature that this study did not cover. In addition, comparing different sizes of a cut and applying it on a different type of volume might increase the accuracy of the results. The next step would be focused on producing a real-time cutaway.

7.4 In Conclusion

Through experiments and analysing the results, we can affirm that monocular depth cues vary in their effect on depth perception. The brightness and the contrast cues are the most efficient cues and provide a better depth impression. Whilst the transparency cue has the lowest efficiency, and it often provides an ambiguous effect. Through this study we were not able to decide which type of translucency was more effective for depth ordering. The Gaussian blur function was

used to generate a simulation of both types of translucency used. This study also revealed the importance of using HDR displays in visualisation applications. Finally, the study has shown that the impression of depth is also affected by the shape of a cutaway, made on the outer layer of a multiple layered image.

Appendix A

Kruskal-Wallis Test

Kruskal-Wallis test is used to find out if there a statistical significant difference among cut shapes. A Kruskal-Wallis test normally is using to find out whether several samples considered as coming from the same population (Kruskal & Wallis, 1952). In most cases the samples differ; and the question arises is whether this difference is statistically significant or not. The first step to apply the Kruskal-Wallis test is to define the null hypothesis H_0 and the alternative hypothesis H_1. In our case we assume that:

H_0: there is no difference between cut shapes.

H_1: there is a difference between cut shapes.

The next step is rank the raw data and then applying the Kruskal-Wallis equation:

$$H = \frac{12}{N(N+1)} \sum_{i=1}^{c} \frac{R_i^2}{n_i} - 3(N+1) \tag{A.1}$$

Where c is the number of samples, n_i is the number of observations in the ith sample, N is the number of observations in all samples combined and R_i is the sum of ranks in the i the sample.

Then we have to find a critical value from Chi-squared table which H will be compared with. In our study we set $\alpha = 0.05$ with 2 degrees of freedom. The critical value that we have to compare the Kruskal-Wallis test result with it is 5.99147. We will reject H_0 if H is greater than critical value, which means that there is a significant difference in somewhere among the objects.

The strong agreements among the test subjects show that there is a significant differences, but we do not know between which pair these differences lie. In fact, this test is important since it supports making the final decision (Day & Quinn, 1989). In other words, we want to find $\acute{R}$ such that the probability P(R$\geq \acute{R}$) is less than or equal to the significance level α (and usually α =0.05). We declare the cut shapes within each group (scores difference<R) to be not significantly different. The distribution of the range R is asymptotically the same as the distribution of variance-normalized range, Wt, of a set of normal random variables with variance = 1 and t samples (David, 1963b). Therefore, we can use the following relation:

$$P\left(W_t \geq \frac{2R - 1/2}{\sqrt{nt}}\right) \tag{A.2}$$

where $W_{t,\alpha}$ is the value of the upper percentage point of W_t at significance point α. The values of W_t can be found in many statistics textbooks including Peaeson & Haetlet (1976). The equivalent function of the test can be achieved in MATLAB by using *multcompare* function.

Appendix B

Samples of Matlab and Mex-file Code

The following MATLAB function, called *plane_shape*, was used to generate the requested shape on a plane. Circular, rectangular and convex are the shapes used. However, with this procedure, any desired shape can be projected.

```
function [points]=plane_shape(area,shape)
global array_length ;
  % area  : contains the area size of the desired shape.
  % shape : represents the shape as 1=Circular ,2=Convex and 3=Rectangular
array_mid=(array_length/2.0);

points=zeros(array_length,array_length);
negative=ones(array_length,array_length);
%-------------------------------------------------------------
if shape==1 %Draw a circular shape on the plane (array)
    radius=sqrt(area/pi);
    radius=single(radius);
for m=1:array_length
```

```
    for n=1:array_length
        if ((array_mid-m)^2+(array_mid-n)^2)<radius^2
            points(m,n)=1;
        end
    end
end
%-------------------------------------------------------------------
elseif shape==3  %if the desired shape is rectangular
  rec_height=sqrt(area);
  rec_width= sqrt(area);
   for m=array_mid-int16(rec_height/2):array_mid+int16(rec_height/2)-1
    for n=array_mid-int16(rec_width/2):array_mid+int16(rec_width/2)-1
        points(m,n)=1;
    end
   end
%-------------------------------------------------------------------
elseif shape==2  %Draw a convex shape on the plane

h=sqrt(area);
w= sqrt(area);
ok=0;
    while ~ok
        xt=array_mid-(h/2);
        xb=array_mid+(h/2);
        yt=array_mid-(w/2);
        yb=array_mid+(w/2);
           if xt==0,xt=1; end
           if yt==0,yt=1; end
        w1=array_mid;w2=array_mid;
        [points]=bezier1(xt,yt,xt,yb,w1,w2,points);
```

```
        [points]=bezier1(xb,yt,xb,yb,w1,w2,points);
        [points]=bezier1(xt,yt,xb,yt,w1,w2,points);
        [points]=bezier1(xt,yb,xb,yb,w1,w2,points);
        % calculating the inner rectangle
        leng=0;
        raw1=array_mid;
        col1=array_mid;
            while points(raw1,col1)~=1
                leng=leng+1;
                raw1=raw1-1;
            end
        inner_height=leng*2;
        if inner_height<(sqrt(area)-7.5)
            ok=0;
            h=h+1;
            w=w+1;
            points=zeros(array_length,array_length);
        else
            ok=1;
        end
    end

[points,negative]=find_clear_points(points,negative,1,1,array_length,0);
[points,negative]=find_clear_points(points,negative,array_length,-1,1,0);
[points,negative]=find_clear_points(points,negative,1,1,array_length,1);
[points,negative]=find_clear_points(points,negative,array_length,-1,1,1);

points=negative;

end
```

```
%-----------------------------------
[r,c,v]=find(points);
points=[r,c];
points(all(points==0,2),:)=[];
```

Since the process of producing a cut takes time to calculate in MATLAB, a function called *MEX-files* was used to speed up the process. Each MEX-file contains only one function and can be written in C, C++ or Fortran. For example, the following C++ code is a MEX-file used to save the portion of an area from the *position* image, which is expected to contain the cut in order to reduce the calculation time.

```
#include <mex.h>
#include <math.h>
void mexFunction(int nlhs, mxArray *plhs[], int nrhs, const mxArray *prhs[])
{
    int    ind,dis,pontsrow,shape,i,j,x,y,area,distance,height,width;
    double *sp,*points,radius,pi,centrez,*rtnv;
    height=1017;
    width=1024;
    pontsrow=500000;
    shape = mxGetScalar(prhs[0]);  \\ the desired shape.
    x     = mxGetScalar(prhs[1]);  \\ x location in the head image.
    y     = mxGetScalar(prhs[2]);  \\ y location in the head image.
    sp    = mxGetData(prhs[3]);    \\ positon image
    area  = mxGetScalar(prhs[4]); \\ the expected area
    centrez=sp[(x-1)+(y-1)*height+2*(height*width)];
    plhs[0]=mxCreateDoubleMatrix(pontsrow,5,mxREAL);
    plhs[1]=mxCreateDoubleMatrix(1,1,mxREAL);
    points=mxGetPr(plhs[0]);
```

```
rtnv=mxGetPr(plhs[1]);
switch (shape)
{
    case 1:
    case 3:
        distance=(int)(sqrt(area)*2.5+0.5);
        break;
    case 2:
        distance=(int)(sqrt(area)*3.7+0.5);
        break;
}
dis=(int)(sqrt(area));
ind=-1;
for (i=x-distance;i<x+distance;i++)
 for (j=y-distance;j<y+distance;j++)
  {
  if((sp[(i-1)+(j-1)*height+2*(height*width)]>=centrez-dis)&& \
    (sp[(i-1)+(j-1)*height+2*(height*width)]<=centrez+dis))
   {
    ++ind;
    points[ind]=sp[(i-1)+(j-1)*height];
    points[ind+pontsrow]=sp[(i-1)+(j-1)*height+(height*width)];
    points[ind+2*pontsrow]=sp[(i-1)+(j-1)*height+2*(height*width)];
    points[ind+3*pontsrow]=i;
    points[ind+4*pontsrow]=j;
   }
  }

rtnv[0]=ind;
```

Appendix C

Survey Results

The following figure shows the results of the survey, which conducted to find the effect of cutaway shape on depth perception. The on-line survey software *Surveygizmo* was used. 85 observers from different countries around the world were participated.

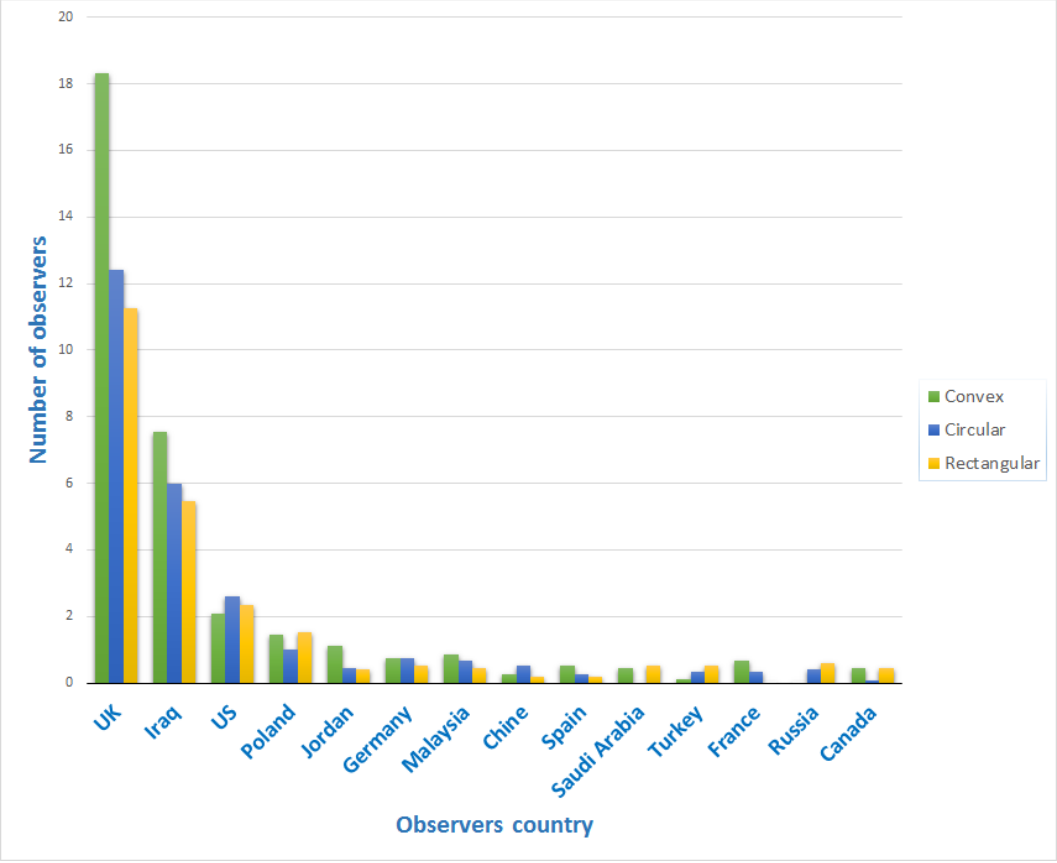


Figure C.1: Results of the survey.

References

Adam, F. (2005). Non-photorealistic rendering.
URL http://gfx.cs.princeton.edu/proj/sg05lines/ 31

Adelson, E. H., & Anandan, P. (1990). *Ordinal characteristics of transparency.*. Vision and Modeling Group, Media Laboratory, Massachusetts Institute of Technology. 4, 23, 39

Aks, J. T., Deborah J.; Enns (1996). Visual search for size is influenced by a background texture gradient. *Journal of Experimental Psychology: Human Perception and Performance*, *22:6*, 1467–1481. 18

Anderson, B. L. (1997). A theory of illusory lightness and transparency in monocular and binocular images: the role of contour junctions. *Perception*, *26*(4), 419–454. 24

Anderson, M., Motta, R., Chandrasekar, S., & Stokes, M. (1996). Proposal for a standard default color space for the internetsrgb. In *Color and Imaging Conference*, vol. 1996, (pp. 238–245). Society for Imaging Science and Technology. 63

Arditi, A. (1986). Binocular vision. *Handbook of perception and human performance.*, *1*, 23–1. 41

Arnheim, R. (1954). *Art and visual perception: A psychology of the creative eye*. Univ of California Press. xiv, 42, 43, 44

Arnheim, R. (1976). Art and visual perception: The new version. *The Journal of Aesthetics and Art Criticism*, *34*, 361–364. xiv, 44

Bavoil, L., & Myers, K. (2008). Order independent transparency with dual depth peeling. *NVIDIA Corporation*, *107*(February), 22020–5. 61

Beck, J., Prazdny, K., & Ivry, R. (1984). The perception of transparency with achromatic colors. *Perception & psychophysics*, *35*(5), 407–422. 60

Bertamini, M. (2000). Positional and symmetry information of concave and convex vertices. *Perception*, *29*. 43

Bertamini, M. (2006). Who owns the contour of a visual hole? *Perception*, *35*(7), 883–894. xiv, 42, 43, 44, 67, 90, 105

Bertamini, M., & Croucher, C. J. (2003). The shape of holes. *Cognition*, *87*, 33–54. 43, 67, 89, 105

Bertamini, M., & Lawson, R. (2008). Rapid figure-ground responses to stereograms reveal an advantage for a convex foreground. *Perception*, *37*(4), 483–494. 5, 37, 43, 56, 67, 90, 105

Bertamini, M., & Mosca, F. (2004). Early computation of contour curvature and part structure: Evidence from holes. *PERCEPTION-LONDON-*, *33*(1), 35–48. 43

Boucheny, C., Bonneau, G.-P., Droulez, J., Thibault, G., & Ploix, S. (2009). A perceptive evaluation of volume rendering techniques. *ACM Transactions on Applied Perception (TAP)*, *5*(4), 23. 29

Boynton, G. (2008). Psychology 333 sensation and perception.
URL http://courses.washington.edu/psy333/lecture_pdfs 11

Bruckner, S., Grimm, S., Kanitsar, A., & Gröller, M. E. (2005). Illustrative context-preserving volume rendering. In *EuroVis*, (pp. 69–76). 28, 44

Bruckner, S., & Groller, E. (2005). *Volumeshop: An interactive system for direct volume illustration*. IEEE. 22

Bruckner, S., & Gröller, M. E. (2007). Style transfer functions for illustrative volume rendering. In *Computer Graphics Forum*, vol. 26, (pp. 715–724). Wiley Online Library. 31

Burns, M., & Finkelstein, A. (2008). Adaptive cutaways for comprehensible rendering of polygonal scenes. In *ACM Transactions on Graphics (TOG)*, vol. 27, (p. 154). ACM. 4

Chan, M.-Y., Wu, Y., Mak, W.-H., Chen, W., & Qu, H. (2009). Perception-based transparency optimization for direct volume rendering. *Visualization and Computer Graphics, IEEE Transactions on*, *15*(6), 1283–1290. 21

Chen, C.-K., Yan, S., Yu, H., Max, N., & Ma, K.-L. (2011). An illustrative visualization framework for 3d vector fields. In *Computer Graphics Forum*, vol. 30, (pp. 1941–1951). Wiley Online Library. 29

Coconu, L., Deussen, O., & Hege, H.-C. (2006). Real-time pen-and-ink illustration of landscapes. In *Proceedings of the 4th international symposium on Non-photorealistic animation and rendering*, (pp. 27–35). ACM. xiii, 31, 32

Coffin, C., & Hollerer, T. (2006). Interactive perspective cut-away views for general 3d scenes. In *3D User Interfaces, 2006. 3DUI 2006. IEEE Symposium on*, (pp. 25–28). IEEE. xvii, 3, 5, 44, 46, 90, 91, 92, 112

Cognates (2013). Brightness.
URL http://art.nmu.edu/groups/cognates/wiki/71a63/Brightness.html 20

Correa, C. D., Silver, D., & Chen, M. (2006). Feature aligned volume manipulation for illustration and visualization. *Visualization and Computer Graphics, IEEE Transactions on*, *12*(5), 1069–1076. xvii, 3, 45, 46, 90, 91

Cunningham, D., & Wallraven, C. (2011). *Experimental Design: From User Studies to Psychophysics*. Natick, MA, USA: AK Peters, Ltd. 65

Cutting, J. E., & VIshoton, P. M. (1995). Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth. *(W. Epstein & S. Rogers, Eds.) Perception*, *5(3)*(3), 1–37. 13

Daly, S., Kunkel, T., Sun, X., Farrell, S., & Crum, P. (2013). Preference limits of the visual dynamic range for ultra high quality and aesthetic conveyance. In *IS&T/SPIE Human Vision and Electronic Imaging*, (pp. 86510J–86510J). Burlingame, California, USA. 26

David, H. A. (1963a). *The method of paired comparisons*, vol. 12. DTIC Document. 98

David, H. A. (1963b). *The method of paired comparisons*, vol. 12. DTIC Document. 117

Davis, H. (2008). *Practical Artistry: Light & Exposure for Digital Photographers*. O'Reilly Media, Inc. 14

Day, R., & Quinn, G. (1989). Comparisons of treatments after an analysis of variance in ecology. *Ecological monographs*, *59*(4), 433–463. 117

Delogu, F., Fedorov, G., Belardinelli, M. O., & van Leeuwen, C. (2010). Perceptual preferences in depth stratification of transparent layers: Photometric and non-photometric factors. *Journal of Vision*, *10*(2), 1–. 24

Deserno, T. M. (2011). *Biomedical image processing*. Springer. 27, 28

Deussen, O., Hanrahan, P., Lintermann, B., Měch, R., Pharr, M., & Prusinkiewicz, P. (1998). Realistic modeling and rendering of plant ecosystems. In *Proceedings of the 25th annual conference on Computer graphics and interactive techniques*, (pp. 275–286). ACM. 28, 31

Deussen, O., & Strothotte, T. (2000). Computer-generated pen-and-ink illustration of trees. In *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, (pp. 13–18). ACM Press/Addison-Wesley Publishing Co. 31

Diepstraten, J., Weiskopf, D., & Ertl, T. (2003). Interactive cutaway illustrations. In *Computer Graphics Forum*, vol. 22, (pp. 523–532). Wiley Online Library. 4, 44

Dinesha, V., Adabala, N., & Natarajan, V. (2012). Uncertainty visualization using hdr volume rendering. *The Visual Computer*, *28*(3), 265–278. 42

Efron, B. (1979). Bootstrap methods: another look at the jackknife. *The annals of Statistics*, *7*(1), 1–26. 85

Elder, J., & Zucker, S. (1993). The effect of contour closure on the rapid discrimination of two-dimensional shapes. *Vision research*, *33*(7), 981–991. 42

Elder, J. H., Trithart, S., Pintilie, G., & MacLean, D. (2004). Rapid processing of cast and attached shadows. *Perception*, *33*(11), 1319–1338. 37

Engel, K., Kraus, M., & Ertl, T. (2001). High-quality pre-integrated volume rendering using hardware-accelerated pixel shading. In *Proceedings of the ACM SIGGRAPH/EUROGRAPHICS workshop on Graphics hardware*, (pp. 9–16). ACM. 28

Epstein, W. (1961). The known-size-apparent-distance hypothesis. *The American journal of psychology*, (pp. 333–346). 39

Everitt, C. (2001). Interactive order-independent transparency. *White paper, nVIDIA*, *2*(6), 7. 39

Farnè, M. (1977). Brightness as an indicator to distance: relative brightness per se or contrast with the background?. *Perception*, *6*, 287–298. xiv, 38, 39

Ferwerda, J. A. (2001). Elements of early vision for computer graphics. *Computer Graphics and Applications, IEEE*, *21*(5), 22–33. 24

Fineman, M. B. (1981). Complexity of context and orientation of figure in the corridor illusion. *Perceptual and Motor Skills*, *53:1*, 11–14. 18

Fleming, R. W., & Ulthoff, H. H. B. (2005). Low-Level Image Cues in the Perception of Translucent Materials. *Psychology*, *2*(3), 346–382. 22

Fowlkes, C. C., Martin, D. R., & Malik, J. (2007). Local figure–ground cues are valid for natural images. *Journal of Vision*, *7*(8), 2. 37

GEFORCE (2014). Enabling ambient occlusion in games.
URL `http://www.geforce.co.uk/whats-new/guides/ambient-occlusion` xii, 17

Gibson, J. J. (1950). The perception of visual surfaces. *The American Journal of Psychology*, *63.3*, 367–384. 62

Gilchrist, A. L. (2013). *Lightness, brightness and transparency*. Psychology Press. 21

Gooch, B., & Gooch, A. (2001). *Non-photorealistic rendering*, vol. 2. AK Peters Wellesley. 29, 31

Guibal, C. R., & Dresp, B. (2004). Interaction of color and geometric cues in depth perception: When does red mean near? *Psychological Research*, *69*(1-2), 30–40. 114

Hadwiger, M., Laura, F., Rezk-Salama, C., Hollt, T., Geier, G., & Pabel, T. (2008). Interactive volume exploration for feature detection and quantification in industrial ct data. *Visualization and Computer Graphics, IEEE Transactions on*, *14*(6), 1507–1514. 28

Hansen, C. D., & Johnson, C. R. (2005). *The visualization handbook*. Academic Press. 27, 28

Harris, L. R., & Jenkin, M. (2011). *Vision in 3D environments*, vol. 116. Cambridge University Press Cambridge. 11

Held, R., & Cooper, E. (2010). Using blur to affect perceived distance and size. *ACM Transactions on Graphics*, *29*(2), 1–33. 13, 34, 75, 78

Hendrik, P. A. L., & Goesele, M. (2005). Realistic materials in computer graphics. URL http://resources.mpi-inf.mpg.de/siggraph05-course-realistic-materials/ 114

Hochberg, J. E., & McAlister, E. (1955). Relative size vs. familiar size in the perception of represented depth. *The American journal of psychology*, (pp. 294–296). 18

Howell, D. C. (2013). *Statistical methods for psychology*. Wadsworth, Cengage Learning. 85

Hughes, D. M., & Lim, I. S. (2009). Kd-jump: a path-preserving stackless traversal for faster isosurface raytracing on gpus. *Visualization and Computer Graphics, IEEE Transactions on*, *15*(6), 1555–1562. xvii, 27, 92, 93, 95

Hund, M., & Mertsching, B. (2009). Occlusion as a monocular depth cue derived from illusory contour perception. In *KI 2009: Advances in Artificial Intelligence*, (pp. 97–105). Springer. 19

Ichihara, S., Kitagawa, N., & Akutsu, H. (2007). Contrast and depth perception:An Effects of texture contrast and area contrast. *Perception*, *36*(5), 686–695. xii, xiv, 16, 35, 36

Ittelson, W. H. (1951). Size as a cue to distance: Static localization. *The American journal of psychology*, (pp. 54–67). 18, 39

Kajiya, J. T., & Von Herzen, B. P. (1984). Ray tracing volume densities. In *ACM SIGGRAPH Computer Graphics*, vol. 18, (pp. 165–174). ACM. 2

Kanizsa, G., & Gerbino, W. (1976). Convexity and symmetry in figure-ground organization. *Vision and artifact*, (pp. 25–32). 19, 37, 89, 105

Kanizsa, G., & Kanizsa, G. (1979). *Organization in vision: Essays on Gestalt perception*. Praeger New York. 19

Kasrai, R. (2004). *On the Perception of Transparency: Psychophysics and Applications to Medical Image Visualisation*. Ph.D. thesis, McGill University, Canada. 16, 21, 61

Kendall, M. G., & Smith, B. B. (1940). On the method of paired comparisons. *Biometrika*, *31*(3/4), 324–345. 100

Kingdom, F. A. (2008). Perceiving light versus material. *Vision Research*, *48*(20), 2090–2105. 16

Kleffner, D. A., & Ramachandran, V. S. (1992). On the perception of shape from shading. *Perception & Psychophysics*, *52*(1), 18–36. 97

Knödel, S., Hachet, M., & Guitton, P. (2009). Interactive generation and modification of cutaway illustrations for polygonal models. In *Smart Graphics*, (pp. 140–151). Springer. 44

Koenderink, J. J., et al. (1984). What does the occluding contour tell us about solid shape. *Perception*, *13*(3), 321–330. 19

Krüger, J., Schneider, J., & Westermann, R. (2006). ClearView: An interactive context preserving hotspot visualization technique. *IEEE Transactions on Visualization and Computer Graphics (Proceedings Visualization / Information Visualization 2006)*, *12*(5), 941–948. xiv, 3, 4, 48

Kruskal, W. H., & Wallis, W. A. (1952). Use of Ranks in One-Criterion Variance Analysis. *American Statistical Association*, *47*(260), 583–621. 116

Landis, H. (2002). Production-Ready Global Illumination. *SIGGRAPH course notes*, *16*(11), 87–101. 58

Landy, M. S., Maloney, L. T., Johnston, E. B., & Young, M. (1995). Measurement and modeling of depth cue combination: In defense of weak fusion. *Vision research*, *35*(3), 389–412. 40

Ledda, P., Chalmers, A., Troscianko, T., & Seetzen, H. (2005). Evaluation of tone mapping operators using a high dynamic range display. *ACM Transactions on Graphics (TOG)*, (pp. 640–648).
URL http://portal.acm.org/citation.cfm?doid=1186822.1073242 25, 100

Legge, G. E. (1981). A power law for contrast discrimination. *Vision research*, *21*(4), 457–467. 15

Levoy, M. (1988). Display of surfaces from volume data. *Computer Graphics and Applications, IEEE*, *8*(3), 29–37. 4, 28

Levoy, M., & Whitted, T. (1985). *The use of points as a display primitive*. University of North Carolina, Department of Computer Science. 28

Li, W., Ritter, L., Agrawala, M., Curless, B., & Salesin, D. (2007). Interactive cutaway illustrations of complex 3d models. In *ACM Transactions on Graphics (TOG)*, vol. 26, (p. 31). ACM. 44

Liang, R., Clapworthy, G., Krokos, M., & Mayoral, R. (2005). Real-time predefined shape cutaway with parametric boundaries. In *Computer Graphics, Imaging and Vision: New Trends, 2005. International Conference on*, (pp. 227–231). IEEE. xiv, xvii, 3, 5, 43, 45, 46, 90, 91, 92, 112

Liu, R., Li, Z., & Jia, J. (2008). Image partial blur detection and classification. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, (pp. 1–8). IEEE. 14

Lorensen, W. E., & Cline, H. E. (1987). Marching cubes: A high resolution 3d surface construction algorithm. In *ACM Siggraph Computer Graphics*, vol. 21, (pp. 163–169). ACM. 28

Luft, T., Colditz, C., & Deussen, O. (2006). Image enhancement by unsharp masking the depth buffer. *ACM Transactions on Graphics*, *25*, 1206–1213. xvii, 41, 95, 96

Mamassian, P., Knill, D. C., & Kersten, D. (1998). The perception of cast shadows. *Trends in Cognitive Sciences*, *2*(8), 288–95. xiv, 37, 38

Mann, S., & Picard, R. (1994). *Being undigital with digital cameras*. MIT Media Lab Perceptual. 24

Mantiuk, R., Mantiuk, R. K., Tomaszewska, A., & Heidrich, W. (2009). Color correction for tone mapping. *Computer Graphics Forum*, *28*(2), 193–202. 58

Mantiuk, R., Myszkowski, K., & Seidel, H.-P. (2006a). Lossy compression of high dynamic range images and video. In *Proceedings of SPIE*, vol. 6057, (pp. 311–320). 62

Mantiuk, R., Myszkowski, K., & Seidel, H.-P. (2006b). A perceptual framework for contrast processing of high dynamic range images. *ACM Transactions on Applied Perception (TAP)*, *3*(3), 286–308. 25

Marshall, J., Burbeck, C., Ariely, D., Rolland, J., & Martin, K. (1996). Occlusion edge blur: a cue to relative visual depth. *Journal of the Optical Society of America*, *13*(4), 681–688. 34, 56

Mather, G. (1996). Image blur as a pictorial depth cue. *Proceedings B:Biological Sciences*, *263*(1367), 169–72. 13, 34, 78

Mather, G. (2006). *Foundations of perception*. Taylor & Francis. 12

Mather, G., & Smith, D. R. (2000). Depth cue integration: stereopsis and image blur. *Vision research*, *40*(25), 3501–3506. 34

Mather, G., & Smith, D. R. (2004). Combining depth cues: effects upon accuracy and speed of performance in a depth-ordering task. *Vision Research*, *44*(6), 557–562. 40, 73, 113

Mather, G., & Smith, D. R. R. (2002). Blur discrimination and its relation to blur-mediated depth perception. *Perception*, *31*(10), 1211–1219. xiii, xvi, 34, 35, 78, 79

McCormick, B. H. (1988). Visualization in scientific computing. *SIGBIO Newsl.*, *10*(1), 15–21.
URL http://doi.acm.org/10.1145/43965.43966 26

McGuffin, M. J., Tancau, L., & Balakrishnan, R. (2003). Using deformations for browsing volumetric data. In *Visualization, 2003. VIS 2003. IEEE*, (pp. 401–408). IEEE. 4

McReynolds, T., Blythe, D., Grantham, B., & Nelson, S. (1998). Advanced graphics programming techniques using opengl. *Siggraph 1998 Course Notes*. 44

Meißner, M., Pfister, H., Westermann, R., & Wittenbrink, C. (2000). Volume visualization and volume rendering techniques. *Eurographics tutorial*. 27, 28

Mendez, J. M. (2010). A Simple and Practical Approach to SSAO. In: Graphics Programming and Theory, gamedev.net.
URL http://www.gamedev.net/page/resources/_/technical/graphics-programming-and-theory/a-simple-and-practical-approach-to-ssao-r275359

Metelli, F. (1974). The perception of transparency. *Scientific American*, *230*(4). 81

Metelli, F., Da Pos, O., & Cavedon, A. (1985). Balanced and unbalanced, complete and partial transparency. *Perception & psychophysics*, *38*(4), 354–366. 21

Michelson, A. A. (1995). *Studies in optics*. Courier Dover Publications. 15

Moreland, K. (2013). A survey of visualization pipelines. *Visualization and Computer Graphics, IEEE Transactions on*, *19*(3), 367–378. 27

Nelson, R., Palmer, S. E., et al. (2001). Of holes and wholes: The perception of surrounded regions. *PERCEPTION-LONDON-*, *30*(10), 1213–1226. xiv, xvii, 42, 43, 89, 90, 105, 106

Newman, T. S., & Yi, H. (2006). A survey of the marching cubes algorithm. *Computers & Graphics*, *30*(5), 854–879. 28

NVIDIA (2014). Incremental Computation of the Gaussian.
URL http://http.developer.nvidia.com/GPUGems3/gpugems3_ch40.html 80

Okoshi, T. (2012). *Three-dimensional imaging techniques*. Elsevier. 11

O'Shea, J. P., Banks, M. S., & Agrawala, M. (2008). The assumed light direction for perceiving shape from shading. In *Proceedings of the 5th symposium on Applied perception in graphics and visualization*, (pp. 135–142). ACM. 17

O'Shea, R. P., Blackburn, S. G., & Ono, H. (1994). Contrast as a depth cue. *Vision research*, *34*(12), 1595–1604. 13, 35

O'Shea, R. P., & Govan, D. G. (1997). Blur and contrast as pictorial depth cues1. *Perception*, *26*, 599–612. 34, 35

Owada, S., Nielsen, F., Okabe, M., & Igarashi, T. (2004). Volumetric illustration: designing 3d models with internal textures. *ACM Transactions on Graphics (TOG)*, *23*(3), 322–328. 4

Peaeson, E., & Haetlet, H. (1976). Biometrika tables for statisticians. *Biometrika Trust*. 103, 117

Pflesser, B., Petersik, A., Tiede, U., Höhne, K. H., & Leuwer, R. (2002). Volume cutting for virtual petrous bone surgery. *Computer Aided Surgery*, *7*(2), 74–83. 4

Ramachandran, V. S. (1988). Perception of shape from shading. *Nature*, *331*, 163–166. 97

Reinhard, E., Heidrich, W., Debevec, P., Pattanaik, S., Ward, G., & Myszkowski, K. (2010). *High dynamic range imaging: acquisition, display, and image-based lighting*. Morgan Kaufmann. xiii, 26

Rempel, A. G., Heidrich, W., & Mantiuk, R. (2011). The role of contrast in the perceived depth of monocular imagery. In *Proc. of the ACM SIGGRAPH Symposium on Applied Perception in Graphics and Visualization*, (pp. 115–115). ACM. 26, 36

Rensink, R. A., & Cavanagh, P. (2004). The influence of cast shadows on visual search. *Perception-London*, *33*(11), 1339–1358. 37

Rheingans, P., & Ebert, D. (2001). Volume illustration: Nonphotorealistic rendering of volume models. *Visualization and Computer Graphics, IEEE Transactions on*, *7*(3), 253–264. 29

Rhino5 (2014). Curvatureanalysis.
URL http://docs.mcneel.com/rhino/5/help/en-us/commands/curvatureanalysis.htm xviii, 108, 109

Richards, W., & Miller, J. F. (1969). Convergence as a cue to depth. *Perception & Psychophysics*, *5*(5), 317–320. 11

Saxena, A., Chung, S. H., & Ng, A. Y. (2005). Learning depth from single monocular images. In *Advances in Neural Information Processing Systems*, (pp. 1161–1168). 41

Saxena, A., Chung, S. H., & Ng, A. Y. (2008). 3-d depth reconstruction from a single still image. *International Journal of Computer Vision*, *76*(1), 53–69. 41

Saxena, A., Schulte, J., & Ng, A. Y. (2007). Depth estimation using monocular and stereo cues. In *IJCAI*, vol. 7. 41

Sayim, B., & Cavanagh, P. (2011). The art of transparency. *i-Perception*, *2*(7), 679. 23

Scheidegger, C. E. (2007). Scientific visualization - project 4.
URL http://www.sci.utah.edu/~cscheid/fal05/scivis/project4/ xiii, 30

Seetzen, H., Heidrich, W., Stuerzlinger, W., Ward, G., Whitehead, L., Trentacoste, M., Ghosh, A., & Vorozcovs, A. (2004). High dynamic range display systems. *ACM Transactions on Graphics*, *23*(3), 760. 26, 65

Setyawan, I., & Lagendijk, R. L. (2004). Human perception of geometric distortions in images. In *Electronic Imaging 2004*, vol. 5306, (pp. 256–267). International Society for Optics and Photonics. 98

Shirley, P., & Morley, R. K. (2008). *Realistic ray tracing*. AK Peters, Ltd. 28

Sigg, S., Fuchs, R., Carnecky, R., & Peikert, R. (2012). Intelligent cutaway illustrations. In *Pacific Visualization Symposium (PacificVis), 2012 IEEE*, (pp. 185–192). IEEE. 44

Sima, J. F. (2014). *A Computational Theory of Visuo-Spatial Mental Imagery*. Ph.D. thesis, University of Bermen. 11

Singh, M., & Anderson, B. L. (2002a). Perceptual assignment of opacity to translucent surfaces: The role of image blur. *Perception*, *31*(5), 531–552. xv, xvi, 5, 22, 75, 76, 77, 78

Singh, M., & Anderson, B. L. (2002b). Toward a perceptual theory of transparency. *Psychological Review*, *109*(3), 492–519. 15, 21, 24

Solso, R. L. (1996). *Cognition and the visual arts*. MIT press. 12

Strothotte, T., & Schlechtweg, S. (2002). *Non-photorealistic Computer Graphics: Modeling, Rendering, and Animation*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc. 2, 29, 76

Sugano, N., Kato, H., & Tachibana, K. (2003). The effects of shadow representation of virtual objects in augmented reality. In *Mixed and Augmented Reality, 2003. Proceedings. The Second IEEE and ACM International Symposium on*, (pp. 76–83). IEEE. 37

Surdick, R. T. (1994). Relevant cues for the visual perception of depth: is where you see it where it is? *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *38 -19*, 1305–1309. 41, 46, 72, 111

Swain, C. T. (1997). Integration of monocular cues to create depth effect. In *Acoustics, Speech, and Signal Processing, 1997. ICASSP-97., 1997 IEEE International Conference on*, vol. 4, (pp. 2745–2748). IEEE. 40

Swain, C. T. (2000). Integration of monocular cues to improve depth perception. US patent 6,157,733. 38

Thompson, P., May, K., & Stone, R. (1993). Chromostereopsis: a multicomponent depth effect? *Displays*, *14*(4), 227–234. 114

Thompson, W. B., Fleming, R. W., Creem-Regehr, S. H., & Stefanucci, J. K. (2011). *Visual perception from a computer graphics perspective*. CRC Press. xiii, 22, 23

Vecera, S. P., Vogel, E. K., & Woodman, G. F. (2002). Lower region: A new cue for figure-ground assignment. *Journal of Experimental Psychology: General*, *131*(2), 194–205. 67

Viola, I. (2008). Illustrative visualization new technology or useless tautology? URL https://www.cs.purdue.edu/homes/aliaga/cs197-10/papers/illustrative-visualization.pdf 26

Viola, I., & Gröller, E. (2005). Smart visibility in visualization. In *Computational Aesthetics*, (pp. 209–216). xiv, 44, 45

Viola, I., Kanitsar, A., & Gröller, M. E. (2005). Importance-driven feature enhancement in volume visualization. *IEEE Transactions on Visualization and Computer Graphics*, *11*(4), 408–418. xiv, 3, 48

Vladusich, T., Lucassen, M. P., & Cornelissen, F. W. (2007). Brightness and darkness as perceptual dimensions. *PLoS computational biology*, *3*(10), e179. 20

Wanat, R., Petit, J., & Mantiuk, R. (2012). Physical and perceptual limitations of a projector-based high dynamic range display. In *Proc. of Theory and Practice of Computer Graphics*, (pp. 9–16). Goslar, Germany: Eurographics Association. 65

Wandell, B. A. (1995). *Foundations of vision.*. Sinauer Associates. 24

Ware, C. (2013). *Information visualization: perception for design*. Elsevier. 17

Watanabe, T., & Cavanagh, P. (1993). Transparent surfaces defined by implicit x junctions. *Vision Research*, *33*(16), 2339–2346. 23, 76

Weiskopf, D., Engel, K., & Ertl, T. (2002). Volume clipping via per-fragment operations in texture-based volume visualization. In *Proceedings of the conference on Visualization'02*, (pp. 93–100). IEEE Computer Society. 43

Weiskopf, D., Engel, K., & Ertl, T. (2003). Interactive clipping techniques for texture-based volume visualization and volume shading. *Visualization and Computer Graphics, IEEE Transactions on*, *9*(3), 298–312. 44

Wenger, R. (2013). *Isosurfaces*. CRC Press. 28

Yagel, R., Kaufman, A., & Zhang, Q. (1991). Realistic volume imaging. In *Proceedings of the 2nd conference on Visualization'91*, (pp. 226–231). IEEE Computer Society Press. 28

Yan, L.-Q., Zhou, Y., Xu, K., & Wang, R. (2012). Accurate translucent material rendering under spherical gaussian lights. In *Computer Graphics Forum*, vol. 31, (pp. 2267–2276). Wiley Online Library. 114

Yang, C.-K., Hung, C.-H., et al. (2008). A new seed-set finding approach for iso-surface extraction. *J. Inf. Sci. Eng.*, *24*(6), 1771–1785. 28

Yuan, X., Liu, Y., Chen, B., Yuen, D. A., & Pergler, T. (2007). Visualization of high dynamic range data in geosciences. *Physics of the Earth and Planetary Interiors*, *163*(1), 312–320. 42

Yuan, X., Nguyen, M., Chen, B., & Porter, D. H. (2006). Hdr volvis: High dynamic range volume visualization. *Visualization and Computer Graphics, IEEE Transactions on*, *12*(4), 433–445. 42

Zheng, L., Wu, Y., & Ma, K.-L. (2013). Perceptually-based depth-ordering enhancement for direct volume rendering. *IEEE Transactions on Visualization and Computer Graphics*, *19*(3), 446–459.
URL http://dx.doi.org/10.1109/TVCG.2012.144 37, 39, 76