

AFIT/GCE/ENG/94S-01

AD-A284 788



3D Stereo Data Visualization and Representation

THESIS

Kai-Heng Wang (Felix)
Captain, R.O.C.A.F.

AFIT/GCE/ENG/94S-01

1108 **94-30621**

Approved for public release; distribution unlimited

94 9 23 032

3D Stereo Data Visualization and Representation

THESIS

Presented to the Faculty of the Graduate School of Engineering
of the Air Force Institute of Technology
Air Education and Training Command

In Partial Fulfillment of the
Requirements for the Degree of
Master of Science in Computer Engineering

Kai-Heng Wang (Felix), B.S.
Captain, R.O.C.A.F.

Accession For	
NTIS CRA&I	☒
DTIC TAB	☒
Unannounced	☐
Justification _____	
By _____	
Distribution /	
Availability Codes	
Dist	Avail and/or Special
A-1	

September 1994

DTIC QUALITY INSPECTED 3

Approved for public release; distribution unlimited

Preface

People ask me why I call myself "Felix, The Man." I reply that my wife suggested I needed a lot of luck. The name Felix is an ancient Greek word meaning lucky. So we decided I should call myself Felix. Yes! I am a very lucky guy because I've met a lot of people who are so nice and friendly and make me feel at home.

The first person I'd like to thank is my advisor Dr. Steven K. Rogers. Dr. Rogers not only guided me in my research but also inspired me to continue searching for new knowledge. Thanks also to my committee members, Dr. Mark E. Oxley and Capt. Dennis W. Ruck for their help in reviewing my work. Many thanks to Lt. Col. Phillip Amburn for his valuable suggestions and assistance. Of course, I cannot forget to thank my dear friends, Capt. Laura Suzuki and Capt. Chih-H Young for their kindness help and generous assistance. Especially, Capt. John Armitstead, for his endless encouragement throughout the academic years, moral support and valuable help in developing my 'C' code program.

Special thanks to Capt. Armin Sayson, his wife, Malou, Lt. Hassan Dehmani, his wife, Donya for they didn't allow me to forget there is more to life than just AFIT. Our group, which we fondly dubbed "The WE Team," spent many long nights pouring over my thesis to ensure I graduated with my head still on my shoulders. Without their special help and encouragement, I could never have finished my work. Not only did they help me write my thesis, they also gave me a warm feeling of friendship.

I want to thank my family in Taiwan for their support. Especially, my mother, who provided endless prayers and financial assistance. Most of all, I'd like to express my love and appreciation to my gorgeous wife Feeling. You patiently endured the loneliness while I spent the past two and half years buried in my work. What made things more difficult was the separation of your being in Taiwan while I was at AFIT. Without your love and patience, I could never have survived AFIT.

Finally, I would like to thank myself for not letting AFIT get the better of me.
Well done Felix – you did a fantastic job!

Kai-Heng Wang (Felix)

Table of Contents

	Page
Preface	ii
List of Figures	vii
List of Tables	viii
Abstract	ix
I. Introduction	1
1.1 Background	2
1.2 Problem Statement	3
1.3 Scope	3
1.4 Methodology	4
1.5 Overview of Thesis	4
II. Current Knowledge and Literature Review	5
2.1 Introduction	5
2.2 Image Visualization	5
2.2.1 Visual Perception	6
2.2.2 Human Factor Considerations	12
2.3 Current Display Systems	16
2.3.1 Stereoscopic Display System	16
2.3.2 Autostereoscopic Display System	17
2.4 Data Representation	18
2.4.1 Kriging Algorithm	19
2.4.2 Neural Networks Algorithm	24
2.5 Summary	38

	Page
III. Methodology and Implementation	39
3.1 Introduction	39
3.2 AFIT Optical Display Device	39
3.2.1 Structure of the Device	39
3.2.2 Enhancement of the Device	42
3.3 Theoretical Foundation	43
3.3.1 Spherical Mirror	43
3.3.2 Polarized Monitor	50
3.4 Image Transform and Acquisition	53
3.4.1 Rayshade 4.0 Program	53
3.4.2 Utah Raster Toolkit	54
3.5 Image Representation	56
3.5.1 Data Interpolation	57
3.5.2 LNKmap	57
3.5.3 Pattern Classifier with Clusterer	57
3.6 Summary	61
IV. Results and Analysis	62
4.1 Introduction	62
4.2 Experimental Setup	62
4.3 Experimental Results	63
4.4 Performance of the Modified AFIT ODD	63
V. Conclusions and Recommendations	68
5.1 Introduction	68
5.2 Conclusions	68
5.3 Recommendations	70
Appendix A. Source code for get4d	72

	Page
Appendix B. Source code for stereo-on mode	90
Appendix C. Source code for stereo-off mode	91
Bibliography	92
Vita	98

List of Figures

Figure	Page
1. Linear perspective of an extended road	8
2. Geometric test caption	11
3. Diagram of three Models	23
4. Perceptron learning procedure	29
5. Three-layer multilayer perceptron	30
6. Single perceptron	31
7. Backward Error Propagation	32
8. RBF structure	34
9. Taxonomy	36
10. Device illustration	40
11. Focusing of rays via a spherical mirror	41
12. Optical Axis	42
13. Concave spherical mirror	45
14. Invert image	46
15. Spherical aberration	48
16. Distortion	50
17. Stereo Image	56
18. LNKnet Algorithms	58
19. Five different types of classifiers	59
20. Left eye image	66
21. Right eye image	66
22. Left eye image	67
23. Right eye image	67

List of Tables

Table		Page
1.	Sign Convention	44
2.	Image formation	44

Abstract

This thesis investigates a technique for improving the perception and visualization of three dimensional information. This improvement will eventually give physicians the ability to both visualize and understand biomedical image data. Raw data is used to generate a stereo image pair. The raw data must be put in run length encoding (RLE) format and then converted to a stereo pair. Once the stereo pair is generated, they must be displayed on the Silicon Graphics ONYX workstation in stereo mode. This research uses 3D images of planets and satellites. The display information is transmitted across cable and is projected onto a spherical mirror. Under appropriate arrangement of both the mirror and the monitor, a real 3D stereo image can be observed.

Biomedical data of interest to this research are slices of a volume of tissue. To convert the slice information into volume data a representation that allows meaningful interpolation between the slices for stereo image pair generation is necessary. A second thrust of this thesis is to investigate neural networks and Kriging for potential use in interpolation of the biomedical volume data.

3D Stereo Data Visualization and Representation

I. Introduction

Ever since the first visual display system was invented, scientists and engineers have not stopped improving display capabilities and applications. Today, display systems are not only widely used for entertainment but also as communication tools. In addition, through the power of computers, displays bring us into another world that we have never dreamed of. For example, through scientific visualization we can see molecular structures that lead us to an understanding of the nature of the world, helping us improve our life and welfare. According to Kantowitz et al.:

Visual displays are extensively used for many of these reasons. 1. illusions of the third physical dimension, 2. pictorials that are easily recognized, 3. relationship diagrams, and 4. information of a historical or predictive nature; all of which can be readily interpreted by most people with little or no instruction (1:20).

Three Dimensional (3D) medical imaging is one of the more commonly used techniques to visualize the interior structures of our bodies. With this powerful tool, physicians can find abnormal or irregular tissues. Localization of the position of a tumor is critical for surgical planing. Physicians can also use the tool to trace the distribution of metabolism in order to find abnormalities in the brain or judge the impact of illnesses (2). 3D imaging will improve the efficiency of medical diagnostics and lessen the possibility of incorrect diagnosis. It will also minimize the danger and unnecessary pain that patients may suffer. As Gabor T. Herman et al. said: "complete perception of the 3-D structure of internal organs and their surroundings is critical to both diagnosis and treatment" (2:39). Other uses for 3D imaging include radiation therapy planning, surgical simulation, virtual health, computer-aided surgery and medical education (2)(3)(4)(5).

1.1 Background

There are several ways we can visualize and represent biomedical image data. According to Jayaram K. Udupa et al., there are several approaches and techniques that are normally used (6). They are slice imaging, projective imaging, and volume imaging (6). Slice imaging represents images in a 2D state by directly projecting the 2D cross-sectional data from the 3D scene space onto the 2D view space. Its most distinguishing feature is its ability to provide high fidelity transverse imaging data. Projective imaging uses either surface rendering or volume rendering techniques based on its data acquisitions and different applications. By implementing computer graphic manipulation techniques, the images will appear with certain perception cues, creating a 3D like scene but not in a true 3D form (2)(5)(6)(7)(8)(9)(10)(11)(12). When those image data are complex or abstract like most biomedical data, a 2 and 1/2 D projective approach may not provide enough clues to the users to give them a picture of the objects. With a true 3D display, the users may more readily recognize and understand the data (2). A good display device should provide a clear and precise image. According to Kantowitz et al., an effective visual display should have the following properties:

1. *visibility* in that all critical elements of the displays are seen.
2. *distinguishability* exists between all parts and symbols of a visual display. and
3. *interpretability* of all display variations into appropriate actions (1:205).

Medical imaging involves applying one or more combinations of the different modalities (such as Computed Tomography (CT), Magnetic Resonance Imaging (MRI), Positron Emission Tomography (PET) or Ultrasound (US), etc.) to get the sampled digital data of a part of the body. The choice of modality depends on the body's characteristics and application. Using the appropriate rendering technique, for example, volumetric rendering, the sampled data is combined with certain signals and image processing procedures. The image is output to a video or optical device to visualize the volume of interest.

1.2 Problem Statement

This research focuses on the application of visualization techniques being tested at AFIT for military images to biomedical applications. In particular, the projection of data through the AFIT optical display device for a 3D stereo visualization is studied. Representation techniques such as conventional statistical techniques – kriging and neural network algorithms for interpolating and reconstructing the biomedical data – will also be investigated. Previous work on the AFIT optical display device will also be discussed and analyzed.

1.3 Scope

The AFIT optical display device is based on the research conducted by Capt James R. Brandt (13). However, it extends the application by projecting the biomedical data which is generated on the SGI workstation onto a spherical mirror, creating a floating true 3D image. The goals of this research are:

- Project data onto a spherical mirror to produce an image that enables operators to visualize the volume that they are interested in, in a natural form (floating in space).
- Test the improvement of projecting the data onto the spherical mirror by using a stereo monitor, in order to study the enhancement of stereopsis cues to improve visualization.
- Investigate techniques such as kriging estimation algorithms and neural networks for possible use in improving the quality of volumetric data to fulfill the need for interpolation for potential use in medical applications.
- Investigate possible improvements to the AFIT optical display device using current display technology.

1.4 Methodology

Images used in this research will be obtained from a data base containing various images. The first image will be of a planet with a satellite orbiting around it. The satellite appears to be closer to the viewer compared to the planet. The second image is of another planet with a satellite orbiting it. In this image, the relative distances of the various parts of the satellite are clearly distinguishable. The size of the planet is small giving the impression that it is at a much greater distance from the viewer than the satellite. A stereo image pair will be produced from the two image files. The stereo image pair will consist of left and right views. This images will be projected onto a spherical mirror where it will be reflected onto a glass viewing plane. A pair of circularly polarized glasses will be required to create a single image with depth information.

1.5 Overview of Thesis

The remainder of this thesis is in four chapters :

- Chapter II provides an introduction to and a review of current literature and knowledge in the area of 3D display systems. The criteria for an ideal 3D display system, including human visual perceptions, mental perception and human factor considerations are emphasized. In addition, Kriging and neural networks algorithms for the application of data representation will also be introduced.
- Chapter III presents the methodology, the theory, and the algorithms that are incorporated in this research.
- Chapter IV analyzes the results with the methodologies that have been adopted for this effort.
- Chapter V gives conclusions and suggests further research.

II. Current Knowledge and Literature Review

2.1 Introduction

In the real world, we see images in 3D stereo mode. Normally, we have no trouble distinguishing an object's geometric size and dynamic structure. We can usually judge the distances between us and the objects we see. The reason we can do this is because of the disparity of our eyes and several other depth cues. Whenever we look at an object, the light that is either reflected or refracted or both from the object and the environment is sensed by our eyes. The brain, or more precisely the visual cortex, interprets the sampled image data from these two separate channels as a single image. The difference in view points from the two channels with other perception cues causes the depth perception, allowing us to see in a stereo mode. Some species, such as fish, which have their eyes on each side of their head cannot see in stereo and 'feel' things differently.

In biomedical engineering, we see things through a visual display system in a virtual environment (VE). However, if we have had no previous experience with the object we are seeing and if the display system does not provide the depth cues and other perception cues allowing us to see in a 3D stereo mode, we will not have a complete understanding or natural perception of the object. We will misjudge the object's relationship because our brains do not have enough clues to perceive things correctly.

2.2 Image Visualization

How we see has long been of interest to scientists and engineers. However, not all functions of the visual system are fully understood. Most features and limitations of our visual perception can be explained using either physics, psychophysiological, or psychological phenomena. Under normal circumstances, we can clearly visualize a scene in its most real and true 3D form, though some argue that the image we see

is unreliable (14). However, when we try to visualize a scene from a visual display medium, if there is not enough peripheral information provided by the display system and if we have had no previous experience with the image we see on the display system, we may interpret the image to be something other than what it really is. Most of the time we have no idea what we want to see or the actual image is not what we thought it should be. As Jose et al. pointed out, a "better understanding of how we perceive and understand the information could improve the visualization process and results" (15:68). According to Mary K. Kaiser:

Rather than deride the utility of visual exploration because of nonveridicalities of perception, the challenge is to understand how the visual system extracts information from the dynamic visual array. Then, visualization tools can be designed to exploit our natural perceptual capabilities for exploration and discovery (14:211).

In order to visualize and understand a clear and high fidelity image of an object in a 3D space, the image must provide necessary perception cues and ideal viewing conditions. Display systems must be designed so that they take into account the visual abilities and limitations of the users. In the following section the basic concept and critical criteria in developing a visual display system will be introduced.

2.2.1 Visual Perception. To be able to see a high fidelity and quality image through a display system, the image must provide the necessary depth perception cues. These cues are either monocular cues or binocular cues. Each cue has its own function; monocular cues have less influence over stereo perception, while binocular cues are needed to see in stereo. In addition, the state of our minds, our psychological make-up, and human factors play a very important role in this process.

2.2.1.1 Ambient Mode and Focal Mode. There are two visual modes in our visual system. These are the ambient mode and the focal mode. As Peter Lennie et al. pointed out, "... ambient vision is sensitive to motion and is very dependent on the peripheral visual field. Focal vision is used for the detail exami-

nation and identification of objects..." (16:103). According to Robert E. Clapp, the ambient system is not only sensitive to motion and changes in visual field but it also stabilizes and orients the focal modes toward the scene of interest (17). The focal mode is directed by our conscious awareness and its function and is closely related to the physical parameters of the eye (i.e., neuron firing and lighting effects) (17). In addition, as Robert E. Clapp said, failure of the ambient mode to coordinate the perceptions from other senses may cause the feeling of motion sickness and related uncomfortableness. The failure of the focal mode will result in the inability to visualize the image details (17). For example, a person with prosopagnosia cannot distinguish people's faces. The following explanations are mainly based on (18)(4)(19) and other sources listed.

2.2.1.2 Depth Perception Cues. We can further subdivide the two visual modes into the following depth perception cues: monocular cues, binocular cues and mental perception. These are the real factors that effect what we perceive.

Monocular cues: Also known as the pictorial cues or static depth cues, monocular cues are perceptions that occur in a single eye or under a static state even though the scene is in a 2D form. It has the following features:

- **Brightness** - Also known as Aerial Perspective, it is the phenomenon where distant objects look blurred and less clear than nearer objects. This perception is a result of the influences of the natural atmosphere such as humidity in the air, heat transfer from the ground, the reflection and refraction of light in the environment, or air pollution. It also provides a strong depth cue. Hence, when we look at an object in an open space, portions of the object that are further away from us appear more blurred or hazy than closer portions.
- **Linear Perspective** - The closer side of an object appears larger than the farther side of the same object. The reason for this is due to both the brightness effect and the changes of the viewing angles thus causing the phenomenon where the

more distant objects project less light into our retina than the closer objects.
See figure 1.

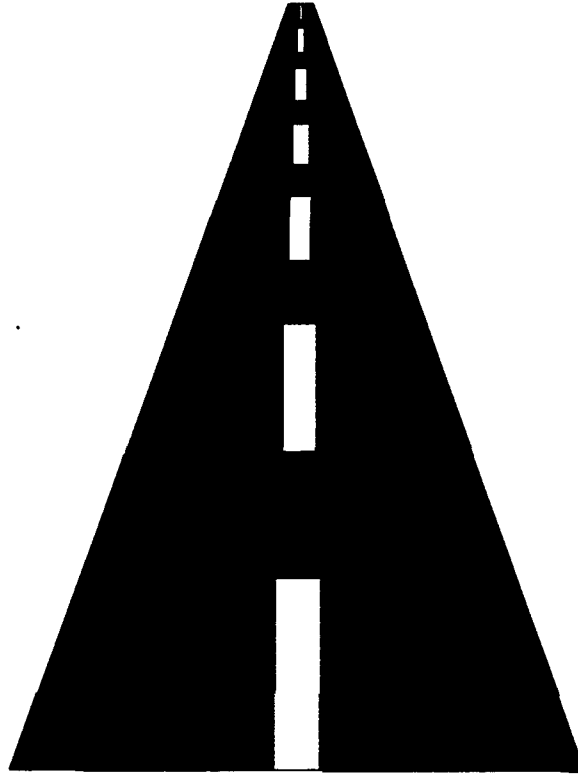


Figure 1. Linear perspective of an extended road

- **Size** - This is a strong depth cue wherein a distant object appears to have a small size compared with a close object even though the distant object may actually be much larger than the close object. The reason is related to the linear perspective cue wherein a farther object may have less light directly reflected into the eyes thus appearing smaller than the closer one, even though its real size is bigger. In addition, our experiences influence the judgement of size. For example, you would not think a puppy is the size of an elephant.
- **Shading** - The shadows of an object give a strong indication of the object's physical features or its geometric relationships with dynamic structures. According to Eberhart Zrenner et al.:

Shading is an important cue for the perception of depth and shape. Under most natural light sources, shading consists mainly of differences in brightness rather than hue. Therefore, depth information conveyed by shading is primarily luminance - contrast information, so the part of the visual system coding 'shape from shading' need not carry any color information (16:198).

- **Overlap or Interposition** - Information is also obtained by a partially observed object being more distant than the one blocking it. The object in front blocks light projected from the object behind it, causing the retina to receive only part of the projected light from the rear object. In computer graphics people use hidden line removal to achieve this effect.
- **Texture Gradient** - The texture of distant objects is less intense and is less detailed than closer objects. This is related to the linear perspective, brightness, and the size cues. Observers can see only an approximation of distant objects because their eyes receive less details. (Believe me, sometimes you would rather see an "attractive person" from far away instead of at a close distance or you might break your Ray Bans).
- **Motion Parallax** - The perception that a closer object moves faster than a more distant object is motion parallax. This is a result of the closer object passing through a smaller viewing divergent space faster than the distant object passing through a bigger viewing divergent angle. Although it is not a static cue, motion parallax is a monocular cue in the sense that we can get information from one or both eyes. It is one of the most important cues that provide information on objects in motion. According to Roy S. Kalawsky, "motion parallax contains information about the direction and the magnitude of the depth relative to the fixation point" (20:60).

Binocular Cue: A physiological depth cues or depth perception that provides the following features:

- Accommodation and Convergence - Also called oculomotor cue. These functions are a result of the six muscles of each eye working to adjust focus and bring our two eyes into a relative position such that we can concentrate on an object which we are interested in seeing. In accommodation, for example, if an object is suddenly moved in front of our eyes, the object might appear blurred until our eyes can properly focus on the object. Near-sightedness occurs when our eyes can not properly focus on an object and the image forms in front of the lens of our eyes. On the other hand, far-sightedness results when an image forms behind the lens of our eyes. In convergence, when we attempt to focus on a certain spot or area, our two eyes must converge toward the center of the viewing field in order to get a clear view. For instance, we are unable to see an object clearly if we just stare our two eyes toward the object. However, both accommodation and convergence have their limitations. If the object is either too close or too far, the image we see is unclear. In addition, if both of them can not coordinate with one another a headache, nausea and sickness etc. will result (16)(20)(21).
- Stereoscopic cue or Binocular Disparity - Since our eyes are located in different positions on our face, two slightly different images are interpreted in the visual cortex as a single image resulting in the depth perception known as stereopsis. Stereopsis not only helps people judge depth but also influences the judgement of an object's physical size and dynamic structure. As Yeh Yei-Yu pointed:

Stereoscopic cues can enhance visual perception and performance. Research in the basic visual sciences has revealed the extent and significance of the role that stereopsis plays in complex visual processing. (22:45-46).

Bruce M. Dow gave such a diagram as showed in figure 2:

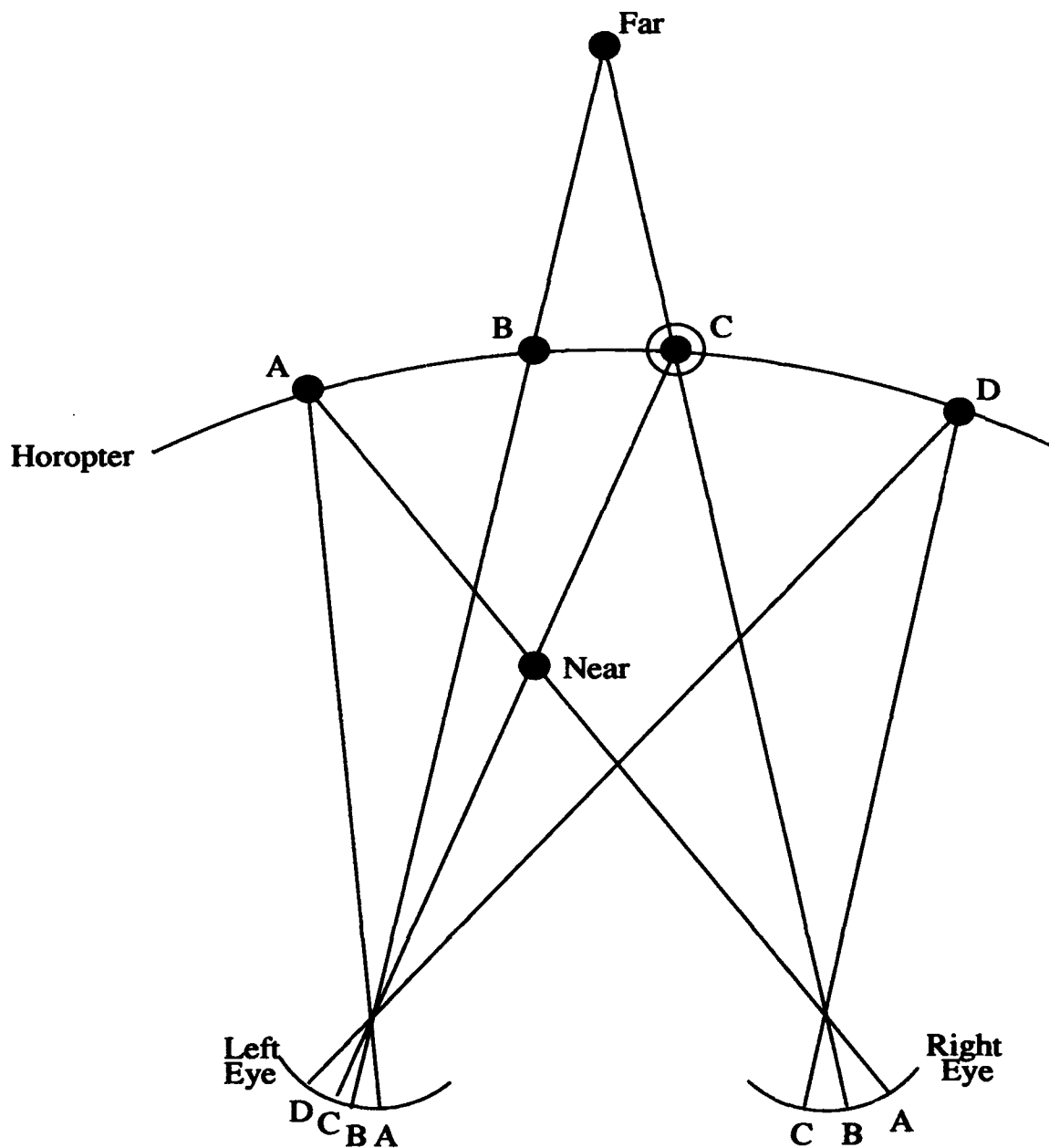


Figure 2. Geometric of binocular disparity in relation to depth vision. A,B,C,and D indicate objects located on the horopter and their images on the retinas of two eyes. "Far" and "Near" indicate objects located either behind (far) or in front of (near) the horopter. The "far" object is imaged at B (left eye) and C (right eye). The "near" object is imaged at A (right eye) and C (left eye). C (circled) is imaged at the fovea center of the two eyes (23:111).

Mental Perception: Mental perception is less understood than monocular cues and binocular cues, because it is totally different with the real world. However, no one can overlook this psychological phenomenon. For example, the attention can drive a person to overlook everything around him but the one thing he is interested (do not tell me you pay more attention to your commander than Cindy Crawford if they stand next to each other). In addition, imagination will also influence vision. For instance, you might imagine Sharon Stone sitting in front of you, instead of your wife. In addition, experience, expectancy, consciousness, and awareness all influence our perception of the image we see (14)(24)(25)(26).

2.2.2 Human Factor Considerations. Another important issue that we should not overlook while designing a 3D display system is human factors. If we don't understand a human's visual capabilities or limitations, we won't be able to implement useful and quality display systems. Most human factor considerations are also closely related to the visual perception cues discussed above. These are:

- **Field of View** - The static view of a normal person covers approximately 180 degrees horizontally for both eyes and 120 degrees vertically for a single eye. If head, neck, body, and eye movement were considered, the coverage would be 360 degrees horizontally and 180 degrees vertically. There is approximately 120 degrees of cross-sectional coverage between the views of each eye. A system designed with an ideal field of view should provide an operator with the most flexibility and freedom to see what he wants to see instead of limiting him to a certain view. An ideal field of view can give observers a complete view and clear understanding of an object. However, according to Aukstakalnis et al., field of view might also induce motion sickness. Some researchers have found that limiting the field of view under 60 degrees horizontally may prevent motion sickness (4). According to Roy S. Kalawsky, display field of view, display resolution and light have a close relationship. If field of view were increased,

resolution must also be increased properly, otherwise, jaggy or other artifacts may influence the visualization and recognition of the object (20).

- **Spatial Resolution** - The acuity of an image from a display system is influenced by the number of pixels that we can perceive, in other words, it is the minimum number of pixels that we can perceive. According to Roy S. Kalawsky, there are several factors that influence spatial resolution: display structure effects on perceived eye level, spatial frequency response of a display system, and spatial frequencies relationship to frame rate and bandwidth (20:61-62). In a human's eye, there are two kinds of photo-receptors: rods, which sense brightness, are distributed less in the center of the retina and are dense in the peripheral, therefore, allowing the periphery of the eye to sense even a little light; and cones which are dense in the middle of retina fovea but less in the peripheral area, allowing us to concentrate on a certain point in front of us and still see an object to the side without identifying what it is. For a more detailed explanation see (16)(21)(23).
- **Refresh/Latency and Update Rates** - In order to keep an image stable on a display system, we need to refresh or redraw the image within a certain time period to prevent the phenomenon of flicker. Normally a 60 hertz (Hz) rate is enough. However, some display systems may need higher rates (27). The update rate is the frequency at which an object in motion must be updated in order to avoid the appearance of erratic movement. An update rate above 10 - 15 Hz will eliminate such problems. According to Aukstakalnis et al., there are three basic factors related to refresh rate in virtual environments (VE) : polygons, display method, and display size (4:265-268). The more complicated and the bigger the size of image data, the more execution times are needed to reduce the refresh rate. Basically, it depends on the hardware and algorithms being used.

- **Luminance, Brightness, Lightness and Contrast** - These are the effects that not only will influence visual intensity but also view acuity, perception of color, etc. According to R. M. Boyton, "luminance is a measure of visually effective radiance, and it has exactly the same geometric properties" (23:231). Luminance is given as $L = K E(\lambda) R(\lambda) V(\lambda) d\lambda$, where $E(\lambda)$ is the spectral emittance of the light source, $R(\lambda)$ is the diffuse spectral reflectance of the surface, $V(\lambda)$ is the proportionality factor which is a function of wavelength, and K is a constant that depends on the units of measurement (21:232). Brightness is related to the aspect of vision and lightness is the property of object surface. According to Jan Walraren et al., "the range of contrasts produced by reflecting objects is relatively small, spanning a factor of about 20. However, for specular (mirror-like) reflections or direct viewing of the light source the range is limited only by the intensity range of the source in question" (16:57). Contrast also influences our perceptions. As Deborah Walters points out, "the results of perceived contrast experiments can be used to suggest one set of image features that appears to be of perceptual significance to humans" (23:379). See (16), (21), and (23) for more information.
- **Color** - Our ability to discriminate color results from the interactions of three different types of cones (red, green, blue) in the retina of our eyes. According to Roy S. Kalawsky:(20:64)

Colour is an important aspect of visual perception, which describes an observer's ability to discriminate objects or surfaces with respect to spectral composition. There attributes are applied to colour perception.

- **Hue** : The correct term to describe a colour by name.
- **Saturation** : The term used to describe the amount of purity or pure chromatic colour in the perception.
- **Brightness/Intensity** : The degree of intensity of a luminous light source.

There are many factors which influence the perception of color. Some examples include, luminance, brightness, lightness, contrast, retinal position, and size of object (23).

- Information Rate and Bandwidth - According to Michael Mckenna et al., "the information rate refers to what rate of data (i.e.bits/sec) is needed to drive a display, whereas the bandwidth of a display refers to the maximum rate at which the signal (pixel values) can change. In other words, the highest frequency signal that can be displayed" (18:425). As mentioned by Michael Starks, in graphics there is little difference but it has important influence in a standard video camera imagery. Because, in mono images, each eye receives only half the bandwidth received from a stereo image (27:217).
- Viewing Zone/Volume Extent - Also known as the display dynamic range, this is the distance range which provides to observers with the best resolution of an image of an object. For example; viewing an object at too close of a distance may result in an unclear image; viewing an object at too far of a distance may result in a loss of detail. This is related to the perception of accommodation and convergence as discussed above. In a normal situation, the distance and angles do not influence the visualization as directly watching a real object.
- Distortion - In general, a linear system transmits a signal or waveform above its threshold in a linear form. The output is exactly equal to the amplitude of the inputs. However, in a nonlinear system, distortion occurs when the output is the combination of a linear part and nonlinear part under the threshold (21). For example, in a spherical mirror reflection display system, if an image size projected onto the mirror is bigger than the ideal viewing field of the mirror, distortion will occur along the boundary of the object.
- Number of View Points - If the image displayed is created as a volumetric scene, there is no limit in the number of view points. Otherwise, the number of viewpoints may limit the viewing freedom.

2.3 *Current Display Systems*

We can divide display systems into two main categories. For users that wear glasses or goggles or when an image comes from two channels, the display system used is called stereoscopic. For systems that do not need any peripheral equipment, the display system is called autostereoscopic.

2.3.1 *Stereoscopic Display System.* We can subdivide stereoscopic display systems into the following groups based on their implementations:

- **Stereo Pair Display** - Also known as stereoscopes, this is the oldest stereoscopic display system. This display separates both eyes' viewpoints into different channels and combines the two 2D images into a single 3D image. Several applications that belong to this category include the Wheatstone mirror, Brewster stereoscope, Mirror stereoscope, and stereogram (29:20-21). For a discussion of performance, see (18).
- **Anaglyph Display** - It applies color separation techniques to achieve a 3D effect. Color separation is accomplished by projecting stereo image pairs through filters consisting of complementary colors (29)(27)(30). According to Ian Sexton, "a major disadvantage is that the use of differently coloured filters causes retinal rivalry. The main advantage is that its relative simplicity makes it fairly economic" (30:88).
- **Polarization Display** - Polarizing glasses filter an imaging projected by two (left and right) polarizers that separate the image into two different channels to produce a 3D effect. However, if polarization is done improperly, a crosstalk factor (ghosting) will result. See (31) for more details on ghosting. As Ian Sexton pointed out, the view presented through a polarization display system is independent of the viewport (30:89). In this research this technique will be used to incorporate with other device to create a 3D stereo image, as will be further discussed later.

- Sequential Stereo Pair Display - This system intermittently displays a sequence of left and right eye images to create the 3D effect. Earlier systems suffered from severe flicker problems. Flicker has been reduced in current displays by increasing the refresh rates to 50 Hz (29).

2.3.2 Autostereoscopic Display System. An autostereoscopic display has certain advantages over a stereoscopic display system. The most obvious advantage is that users do not need a viewing aid to see an image, allowing more than one user to view the image at the same time. According to Michael Starks (28), we can divide autostereoscopic display systems into the following:

- Parallax Barriers - After placing vertical slits in front of a display plate to cause part of the view to be blocked, we can then alternately project the images seen by our two eyes side by side to cause our eyes to see only the appropriate images and interpret it as a stereo scene (27)(30). See (18) for the performance of parallax barriers.
- Dynamic Parallax Barriers - As Michael Starks explained, in dynamic parallax barriers, "one or more vertical slits are rapidly scanned in the horizontal direction and the appearance of the image points on the screen behind are timed precisely so that a viewer at any position will see a stereo image" (29:219). See (28) and (32) for more detail.
- Lenticular Display - The idea is stimulated by the limitation of parallax barriers to replace the slits with a cylindrical lenses. According to Bruce Lane, "a particularly valuable feature of lenticular is the multiple viewing zones. It is also possible to widen the viewing zone by increasing the number of viewpoints encoded behind the lenticules" (29:26). For its performance, see (18).
- Stereoptilexer - Invented by Robert Collender in early 60's, the Stereoptilexer is a dynamic parallax barrier but with its physical slits replaced by virtual slits (27).

- **Integral Photography** - Also called "fly's eye lens photography", integral photography uses an array of many small lenses to take and display images (26)(28).
- **Large Mirrors, Lenses, and Retro-Reflective Screens** - By using the commonly known physical phenomenon of properly projecting an image onto a curved mirror or screen, a stereo image will appear in an appropriate viewing space.
- **Laser-based Hologram Display** - By projecting a laser beam onto either a solid volumetric display device or a multi-planar display device, a true 3D image is created. For more detail, see (32),(33), and (34).
- **Other Devices** - Several other devices are also available in autostereoscopic. They are swan's cube displays, fresnel displays, space-filling multiplanar displays, varifocal multiplanar displays, volumetric displays, holovision, stereoscopy, and stereosculpting. They use either laser, mirror, other optical methods, or computer techniques to generate the 3D effect. For more details, refer to (28), (29), (30), (18), (33), (34), (35), (36), and (37).

2.4 Data Representation

The data representation method used in a display system is very important. Even with an ideal display system, if the data representation method yields an image that cannot properly represent the image data, an unsatisfactory or ambiguous image is expected. As Edward J. et al. said, "the user may wish to visualize a nebulous 3D region, data without distinct surfaces or structures, or a regions buried inside a larger structure, or interrelate region in different data sets" (38:153). In order to represent 3D imaging from its 2D space into 3D view space, there are several steps required. They are deciding on the volume of interest (VOI), filtering, data interpolation, segmentation, and rendering, etc. (6).

One of the most important procedures is data interpolation. The reason data interpolation is necessary is due to the limitation of image modality abilities. For example, in CT or MRI methods we obtain image data in a sampled 2D cross-

sectional form each time the scanner is activated. In order to see a 3D volume, we have to interpolate those sampled 2D slices as a continuous volume. Since those slices are not in a continuous linear form, a good interpolation algorithm is required to estimate the volume that we are interested in and to get the highest fidelity and most accurate image data.

2.4.1 Kriging Algorithm. Kriging started out as a geostatistical method that was developed for locating ore distributions and surface estimation from sampled data. The method has been modified to meet the statistical analysis needs of various applications. It differs from the classical statistical method in that a statistic is a function of a random variable while kriging is a function of a regionalized function. In a regionalized area, all variables not only have close geometric relations but also are anisotropies related wherein the range of influences in certain area varies according to the different directions (11). The method was named after D.G. Krige, a South African mining engineer, for his contributions in developing the algorithm. Matheron, a French engineer, formalized and generalized Krige's theory into a more practical methodology (39). The algorithm is based on the assumption that sampled data in a neighboring region has close correlation. Therefore, we can predict unknown data in a spatial region by using a linear combination in an unbiased and most optimal trend to get the mean of weighted random sampled data (40)(41)(11). As Jesus Carrera et al. stated, "by using kriging, not only can the estimates of the natural phenomena be determined, but the estimation variances reflect the uncertainty of the estimation process" (43:189). Noel Cressie cited the following definition from an article written by Hemyari and Nofziger (1987):

kriging is a form of weighted averaging in which the weights are chosen such that the error associated with the [predictor] is less than for any other linear sum...The weights depend upon the location of the points used in the [prediction] process and upon the [covariation] ... reflected in the semivariogram (39:239-240).

As Cressie mentioned, Krige has his own perception as:

the name given ... by Matheron to : the multiple regression procedure for arriving at the best linear unbiased [predictor] or best linear weighted moving average [predictor] of the ore grade of an ore block (of any size) by assigning an optimum set of weights to all the available and relevant data inside and outside the ore block (39:240).

From the assumptions and the methodology, we can divide kriging into the following primary types:

- Simple Kriging - Also known as linear regression. Simple kriging uses the mean of n known sampled values through linear combination to find the unknown neighboring points or values. According to Andre G., if the known values do not have the same attributes with those of the unknowns, the method is called cokriging. The only difference between kriging and cokriging is in the inference of the covariance/variogram models in solving for the linear regression weight (43). As Andre G. Journel pointed out, "simple kriging requires that the mean of the variable over the field being estimated be constant and known" (43:2). We can write the general equation as:(43:11)

$$\sum_{\beta=1}^n \lambda_{\beta} C_{\alpha\beta} = C_{\alpha 0}, \alpha = 1, \dots, n$$

and the corresponding minimized error variance is:

$$\sigma_{SK}^2 = Var\{Z_0 - Z_0^*\} = C_{00} - \sum_{\alpha=1}^n \lambda_{\alpha} C_{\alpha 0} \geq 0$$

where $C_{\alpha\beta}$ is a double sum of the covariance and $C_{\alpha\beta} = C_{\alpha'\beta}$, for all β , if and only if $\alpha = \alpha'$ and λ_{β} is the shift parameter. In addition, Z_0 is unknown value and Z_0^* is estimator, under the condition of unbiasedness the offset (or shift parameters) must sum to 1 so that the expected value of the estimation error is zero. i.e. $\sum_{\alpha=1}^n \lambda_{\alpha} = 1$.

As mentioned by Parrott (11), simple kriging is not very popular since the means are actually unknown most of the time (11). For more detail, see (43).

- Ordinary Kriging - Unlike simple kriging, in ordinary Kriging we assume the mean of sampled values are unknown, but are assumed constant (11:3-9)(43:2). In other words, it is a constrained linear regression. As Cressie said, "ordinary kriging is a simple kriging that uses the Best Linear Unbiased Prediction(BLUP)" (39). The equation is given as:(43:16)

$$\begin{cases} \sum_{\beta=1}^n \lambda_{\beta} C(x_{\alpha} - x_{\beta}) + \mu = C(x_{\alpha} - x_0), \alpha = 1, \dots, n \\ \sum_{\beta=1}^n \lambda_{\beta} = 1 \end{cases}$$

and the corresponding minimized error variance is:

$$\sigma_{OK}^2 = E\{[Z(x_0) - Z^*(x_0)]^2\} = C(0) - \sum_{\alpha=1}^n \lambda_{\alpha} C(x_{\alpha} - x_0) - \mu \geq 0$$

where λ_{α} and μ are constrained normal.

- Universal Kriging - universal kriging is an expansion of ordinary kriging. Noel Cressie gives the following quote:

The spatial prediction method known as kriging exploits second-order spatial correlation structure to obtain minimum variance predictions of certain average values of the random function. But to do so, it must be assumed that either the mean function(the drift) is known up to a constant or the second-order structure (the variogram) is known exactly. Knowledge of the drift allows the (stationary) variogram to be estimated and leads to ordinary kriging. Knowledge of the variogram allows the drift to be estimated and leads to universal kriging (41:623).

The equation can be written as:(38:19)

$$\begin{cases} \sum_{\beta=1}^n \lambda_{\beta} C(x_{\alpha} - x_{\beta}) + \sum_{\ell=0}^L \mu_{\ell} f_{\ell}(x_{\alpha}) = C(x_{\alpha} - x_0), \alpha = 1, \dots, n \\ \sum_{\beta=1}^n \lambda_{\beta} f_{\ell}(x_{\beta}) = f_{\ell}(x_0), \ell = 0, \dots, L \end{cases}$$

and the corresponding minimized error variance is:

$$\sigma_{UK}^2 = E\{[Z(x_0) - Z^*(x_0)]^2\} = C(0) - \sum_{\alpha=1}^n \lambda_{\alpha} C(x_{\alpha} - x_0) - \sum_{l=0}^L \mu_l f_l(x_{\alpha})$$

2.4.1.1 Variogram. Variogram is the most important parameter in kriging. According to Lynda Kerbs, it can be thought of as a squared mean among data in the given direction and different distances (44). As Noel Cressie mentioned, "in Matheron's geostatistics and Gandin's objective analysis, the variogram is used to define coefficients in an optimal linear predictor" (40:199). Lynda defines variogram as:

The variogram provides the geologist with a quantitative value for the range of influence of a sample in any direction. This information can be used to design optimum sampling grids and drilling patterns. The behavior of the variogram near the origin provides additional information regarding the continuity and randomness of the data. The variogram is a tool that geologists can use in interpretation of geologic problems by combining the knowledge of spatial variability derived from the variogram with the geologic reasons for the occurrence of spatial variability (44:56,58).

The equation of the variogram is:

$$2\gamma(h) = E\{[Z(x) - Z(x+h)]^2\} = 2[C(0) - C(h)]$$

where $C(h) = C(0) - \gamma(h)$, $\gamma(h)$ is the mean of differences and where h is a vector distance between sampled data points. There are several models currently in use. Three of those models as Donald P. mentioned are:(45)

Linear Model: $\gamma(h) = ah + b$

De Wijsian Model: $\gamma(h) = a \ln(h) + b$

$$\text{Spherical Model: } \gamma(h) = \begin{cases} C(\frac{3}{2} \frac{h}{a} - \frac{1}{2} \frac{h^3}{a^3}) + C_0 & \text{if } h < a \\ C + C_0 & \text{if } h \geq a \\ 0 & \text{if } h = 0 \end{cases}$$

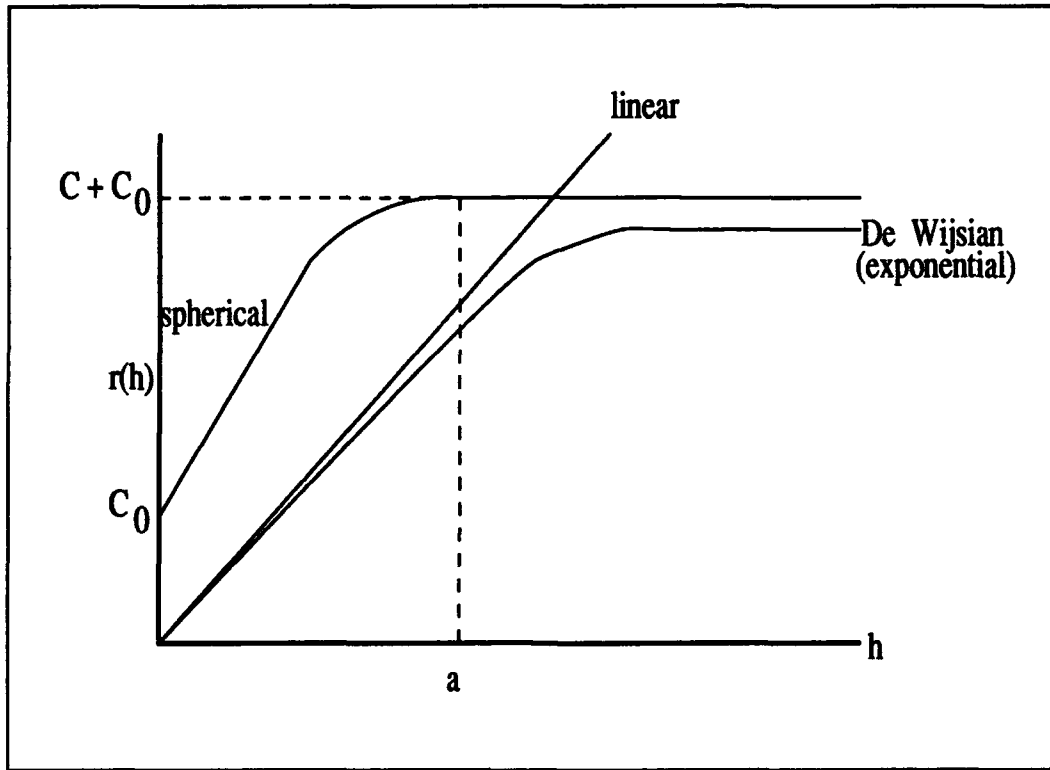


Figure 3. Diagram of three Models

The variogram is defined by three parameters: a , C_0 , and C . The parameter a stands for range, which is used to decide the rates of correlations of distance instead of the variables included in the predictor, i.e. defines the extent of partial dependency (46). C_0 is the nugget effect that consists of the combination of estimation error and a small variation. C is the partial sill which is combined with the nugget effect to form the sill function, which behaves as the variance of data. Its relative diagram is shown in figure 3. See (44), (45), and (46). As Lynda Kerbs said, "the variogram can be thought of as an 'average difference squared' between data a given distance apart in a given direction. This 'average difference' is calculated for several distances in a given

direction and a graph of distance versus average difference is constructed" (44:54). The variogram is a very important function because it provides a quantitative value for the influence extent of a regionalized area (44:56). As Ferenc said, "in applied sciences a great deal of data has to be collected and the structured features of the variable represented by the data must be analyzed" (47:331), a variogram functions as an analyzer.

By using the kriging method, we can estimate the surface or neighboring data between two adjacent points. In biomedical image representation, for example, either CT or MRI image data is obtained in a non-continuous slice form. In order to visualize a volumetric 3D image we can apply kriging to interpolate those slices into a whole volume data. Kriging has extended its power into fields other than mining. For example, at the Air Force Institute of Technology (AFIT) kriging has been applied in satellite image analysis, terrain modeling for flight simulators, and surface extraction in medicine (48).

2.4.2 Neural Networks Algorithm. The artificial neural network (ANN) is based on the general idea that a mathematical model may be used to emulate a creature's biological neural responses, especially in thinking, learning, vision, speaking, consciousness, and attention, in reacting to outer stimulus. Through the study of ANN, we may eventually understand a human's neural functions as they are related to mental/psychological or physical behaviors. The Defense Advanced Research Projects Agency (DARPA) gives the following definition for neural networks:

Neural networks are systems made of many simple processing elements operating in parallel whose function is determined primarily by the pattern of connectivity. These systems are capable of high-level functions, such as adaptation or learning, and/or lower level functions, such as data preprocessing for visual or auditory inputs. Neural networks are inspired both by biological nervous systems and mathematical theories of learning, information processing, and control (49:55).

As pointed out by Steven K. Rogers et al., "neurocomputers are apparently good at finding approximate solutions very quickly" (50:40). Steven K. Rogers et al. says that "one of the biggest potential advantages of artificial neural networks is the hope of eliminating the need for specific code development for a given discrimination task" (50:46). According to Richard P. Lippmann, "the potential benefits of neural nets extend beyond the high computation rates provided by massive parallelism" (51:13).

The neural net algorithm has been widely used in many applications. For example, the algorithm has been used in the military for target recognition and radar signals processing and commercially for speech, character, face, and voice recognition, financial analysis, pattern classification/segmentation, industrial machine vision, robotics, optimization/computation, and medical image analysis (49)(50)(52)(53)(54). Richard P. Lippmann explained:

Neural net models are specified by the net topology, node characteristics, and training or learning rules. These rules specify an initial set of weights and indicate how weights should be adapted during use to improve performance (51:13).

As Kunihiro Fukushima said:

Modeling neural networks is useful in explaining the brain and also in engineering applications. It brings the results of neurophysiological and psychological research to engineering applications in the most direct way possible. Once we complete a model, its simplification makes it easy to see the essential algorithm directly as a design principle for new information processors (55:65).

The neural net algorithm can also be used to model a stereo image (56)(57). In this research we will apply the neural net algorithm of biomedical image representation. The basic concept is the same as in target recognition in military applications as presented by Dennis W. Ruck. The algorithm includes three basic steps. First, data segmentation involves determining the potential target(s) which we are interested in. Second, feature extraction is the collection of the same or related features

from the raw data sets into a feature space for use in later computations. Third, image classification; by providing the neural net system with the appropriate training it can represent a high fidelity and accurate image (58). As Mehmed Ozlet et al. said, "their major advantage is that they do not rely on any assumption about underlying probability density functions, thus possibly improving the results when the data significantly depart from normality" (59:535). S.C. Amartur et al. also pointed out, "the new emerging field of artificial neural networks promises to provide unique solutions for the pattern classification of medical images" (60:215).

2.4.2.1 Theoretical Foundation. Artificial neural networks are inspired by biological activities. Therefore, in order to understand the ANN algorithms, it is best to start by trying to understand biological functions – how do neurons work? We need to understand the physiological and physical reactions as well as the psychological and mental processes of the brain. According to Richard P. Lippmann; "Designing artificial neural nets to solve problems and studying real biological nets may also change the way we think about problems and lead to new insights and algorithmic improvements" (51:13-14). As Teuvo Kohonen said:

Artificial neural networks' are massively parallel interconnected networks of simple (usually adaptive) elements and their hierarchical organizations which are intended to interact with the objects of the real world in the same way as biological nervous systems do (52:4).

DARPA gives the following definition:

Architectures and processing elements used in neural network models are simplified versions of those observed in biological nervous systems. Characteristics of biological neural networks that artificial neural network models hope to provide include:

- Fault tolerance to loss of a small number of computational elements,
- Insensitivity to small variations between computational elements,
- The need for primarily local connectivity and local learning rules,

- Realtime response, and
- Parallelism (49:60-61).

Since biological functions are not the scope of this research, the reader should refer to (50) and (60) for more information on how neurons function. However, with the present technology and the current knowledge, it is still impossible to develop a model as delicate and as complicated as real neurons. Therefore, we must assign an appropriate classifier according to the information's characteristics and the different applications to achieve the goal that we want. Richard P. Lippmann gave neural net models a clear metaphor:

Neural net models are specified by the net topology, node characteristics, and training or learning rules. These rules specify an initial set of weights and indicate how weights should be adapted during use to improve performance (51:13).

There are many methodologies that have been developed in order to achieve the best performance. No one can fulfill all the needs of all applications. However, neural net algorithms do promise to provide a better and efficient alternative solution to the traditional stochastic methods. Lippmann Richard gave the following reasons:

Neural nets typically provides a greater degree of robustness or fault tolerance than von Neumann sequential computers because there are many more processing nodes, each with primarily local connections. Damage to a few nodes or links thus need not repair overall performance significantly. Most neural net algorithms also adapt connection weights in time to improve performance based on current results. Neural net classifier are also non-parametric and make weaker assumptions concerning the shapes of underlying distributions than traditional statical classifiers (51:13).

2.4.2.2 Multilayer Perceptrons with Back Propagation. Since this research is trying to provide a sound solution to biomedical image recognition, the algorithms introduced here are based on the use of static pattern recognition. The multilayer perceptron algorithm with backward error propagation (backprop) learning rule is one of the best candidates. As Steven K. Rogers et al. said, "the key to

these multilayer networks is the learning algorithms that allow the training of the weights on the hidden layers, the layers of neurons that are not directly connected to the inputs or the outputs" (51:52). Craig M. Aendt also pointed out that "neural nets and feed-forward back-propagation neural nets in particular are good classifiers. They also exhibit the ability to derive complex relationships in data" (62:880). According to Yonghong Tang et al., "multilayer feedforward artificial neural networks possesses the ability of approximating an unknown mapping arbitrarily well. This is particularly useful for pattern recognition" (63:113). Kunihiko Fukushima also emphasizes the importance of the feed-forward procedure as:

Tolerating positional error a little at a time at each stage, rather than all in one step, plays an important role in endowing the network with an ability to recognize even distorted pattern. The network recognizes the "shape" of the pattern independent of its size and position (55:66-67).

George N. et al. had the following statement:

Most of the current approaches for solving a classification problem are implemented using a single network classifier. However, preliminary results have shown that as the number of classes involved in a classification task increase, the classification accuracy of the neural classifier is greatly affected. In addition, as the number of classes increases, the size of the network must be increases in order for it to be able to converge in a reasonable number of training epochs. Both of these effects are quite undesirable and restrictive, especially when real life classification problems are considered (64:221).

Steven K. Rogers et al. gave the following illustrations with a three-layer multilayer perceptrons and backward error propagation and the basic algorithms as showed in figure 4, 5, 6, 7:(50:57)

As Steven K. Rogers et al. further explained;

The algorithm is shown in box is characterized by inputting data into the bottom of a multilayer network; then propagating the information through the network in a feed-forward manner. When the input nodes are computed, their output is compared to the desired classification of the input data. The error is then used to compute a correction for the

1. Initialize weights and thresholds to small random numbers. We recommend using a uniform distribution from -0.5 to +0.5.
2. Present training vector and desired output.
3. Calculate the actual output:

$$y = f_h(\sum_{i=0}^{N-1} w_i x_i + \theta)$$

where N is the number of inputs.

4. Learning:

$$w_i^+ = w_i^- + \eta(d - y)x_i$$

where η is the gain, $0 \leq \eta \leq 1$, d is the desired output.

Note: only learn when in error.

Figure 4. Perceptron learning procedure (50:51)

weights. The correction is just a gradient descent down an error surface. The error is back-propagated through the network, correcting the weights (50:54).

2.4.2.3 Radial Basis Function. Another algorithm that has received attention is the radial basis function (RBF) proposed by Broomhead and Lowe (65). However, it has been improved and extended into a more practical and useful state (66). According to Cichocki A. et al.:

An RBF is a multidimensional function which depends on the distance $r = ||x - c||$ (where $||.||$ denotes a vector norm) between the input vector x and the centre c . Radial basis functions provide a powerful method for the multidimensional approximation or fitting which essentially does not suffer from the problem of the proliferation of adjustable parameters as the dimensionality of the problem increases (as multilayer perceptron does). Typical choices for radial basis functions $\varphi(x) = \Phi(||x - c||)$ are

- (i) $\Phi(r) = r$ piecewise linear approximations,

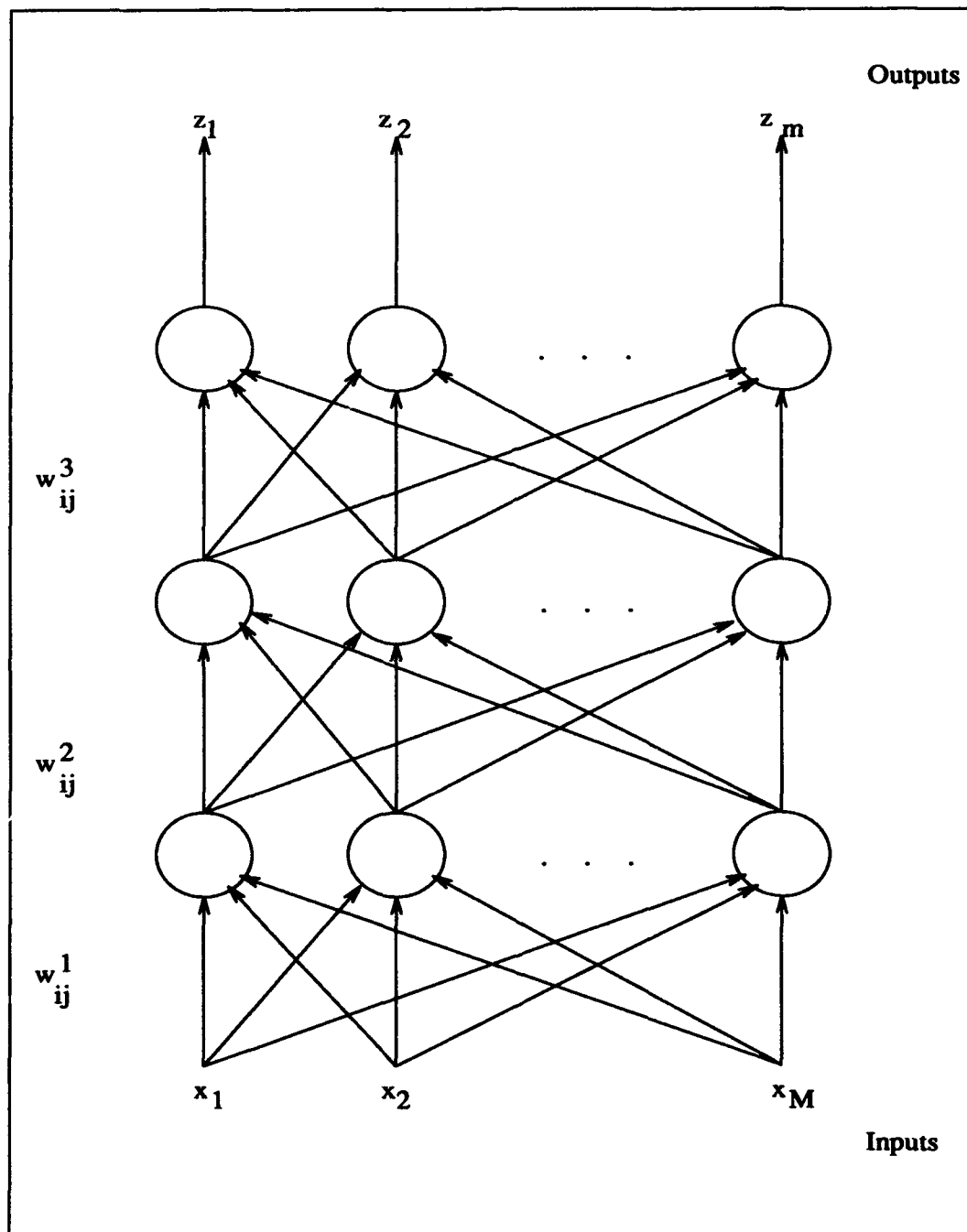


Figure 5. Three-layer multilayer perceptron (50:55)

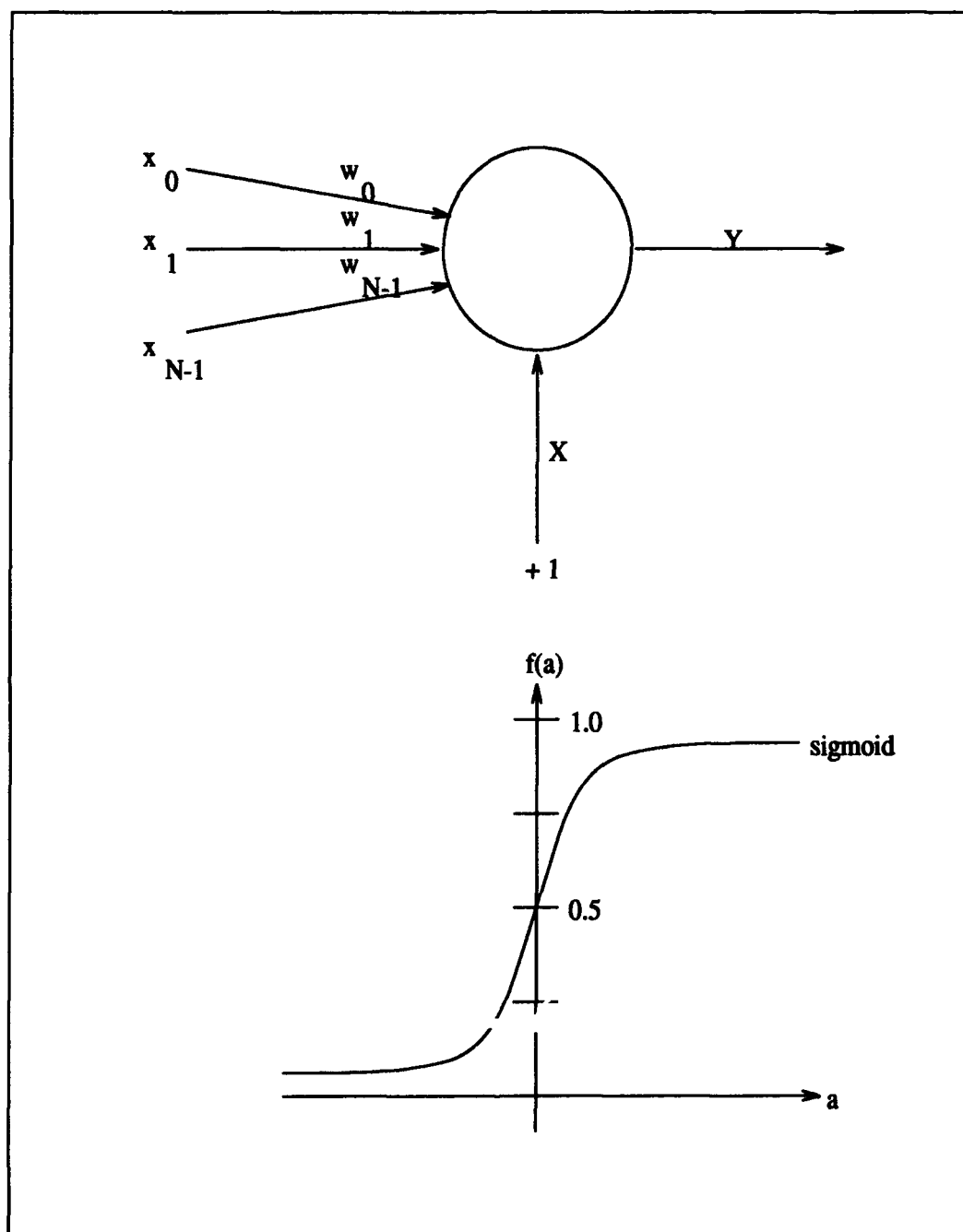


Figure 6. Single perceptron, where Y represents $y = f(\sum_{i=0}^{N-1} w_i x_i + \theta)$ and X is θ (50:48)

1. Initialize weights and thresholds to small random numbers. We recommend using a uniform distribution from -0.5 to +0.5.
2. Present training input and classification (desired output).
3. Calculate output
4. Learn (adapt weights and thresholds).

$$w_{ij}^+ = w_{ij}^- + \eta \delta_j x_i + \alpha (w_{ij}^- - w_{i-j-})$$

where w_{ij} is the weight from node i to node j in the next layer, x_i is the output of node i , and δ_j is the error associated with node j . η and α are learning rates (for example constants of .35 and .7 respectively). w_{ij}^+ is the new weight value and w_{ij}^- is the old weight value. w_{i-j-} is the value of the weight before the last update. Thresholds are adapted similarly where x_i is replaced by +1 if the threshold is added to the weighted sum and -1 if it is subtracted. The δ_j are defined as follows:

$$\delta_j = \begin{cases} y_j(1 - y_j)(d_j - y_j) & \text{for output node } j \\ x_j(1 - x_j) \sum_k \delta_k w_{jk} & \text{for hidden node } j \end{cases}$$

where d_j is the desired output for output node j and y_j is the actual output. For the hidden nodes the δ_k are the errors for the layers above, see appendix A.

Figure 7. Backward Error Propagation (50:56)

- (ii) $\Phi(r) = r^3$ cubic approximation,
- (iii) $\Phi(r) = \exp(-r^2/2\sigma^2)$ Gaussian function,
- (iv) $\Phi(r) = r^2 \log(r)$ thin plate splines,
- (v) $\Phi(r) = (r^2 + \sigma^2)^{1/2}$ multiquadratic function,
- (vi) $\Phi(r) = (r^2 + \sigma^2)^{-1/2}$ inverse multiquadratic function,

where σ is a real coefficient called the width or scaling parameter. Gaussian function has a peak at the centre c and decreases monotonically as the distance from the centre increases, therefore it is the most popular choice (66:452-453).

Basically, RBF has the same structure as MLP, however, only one hidden layer exists. In addition, unlike MLP, each of its hidden layer connections is weighted instead of each connection being a non-linear transformation of the basis function in the hidden layer of RBF. In other words, each unit of the hidden layer is a function of a non-linear decision making parameter base on the basis function which it uses. It computes the distance between the X input with its fixed centroid and depends on a radially symmetric function. Usually the nearest input node with the centroid is selected, thus the output layer give a weighted summation from those hidden layer (67)(68). As Chen S. et al. pointed:

An RBF network can be regarded as a special two-layer network which is linear in the parameters by fixing all RBF centers and nonlinearities in the hidden layer. Thus the hidden layer performs a fixed nonlinear transformation with no adjustable parameters and it maps the input space onto a new space. The output layer then implements a linear combiner on this new space and the only adjustable parameters are the weights of this linear combiner (69:302).

Chris Bishop also pointed out that "An important feature of radial basis function neural networks is the existence of a fast, linear learning algorithm in a network capable of representing complex nonlinear mapping (70:579). According to M.W. Mak et al.; the RBF approach not only pocesses the attribute of fast learning but also has a hyper-parabolic error surface, if a squared error measure of network performance is used and it has a firm analytical foundation that is paramount if this

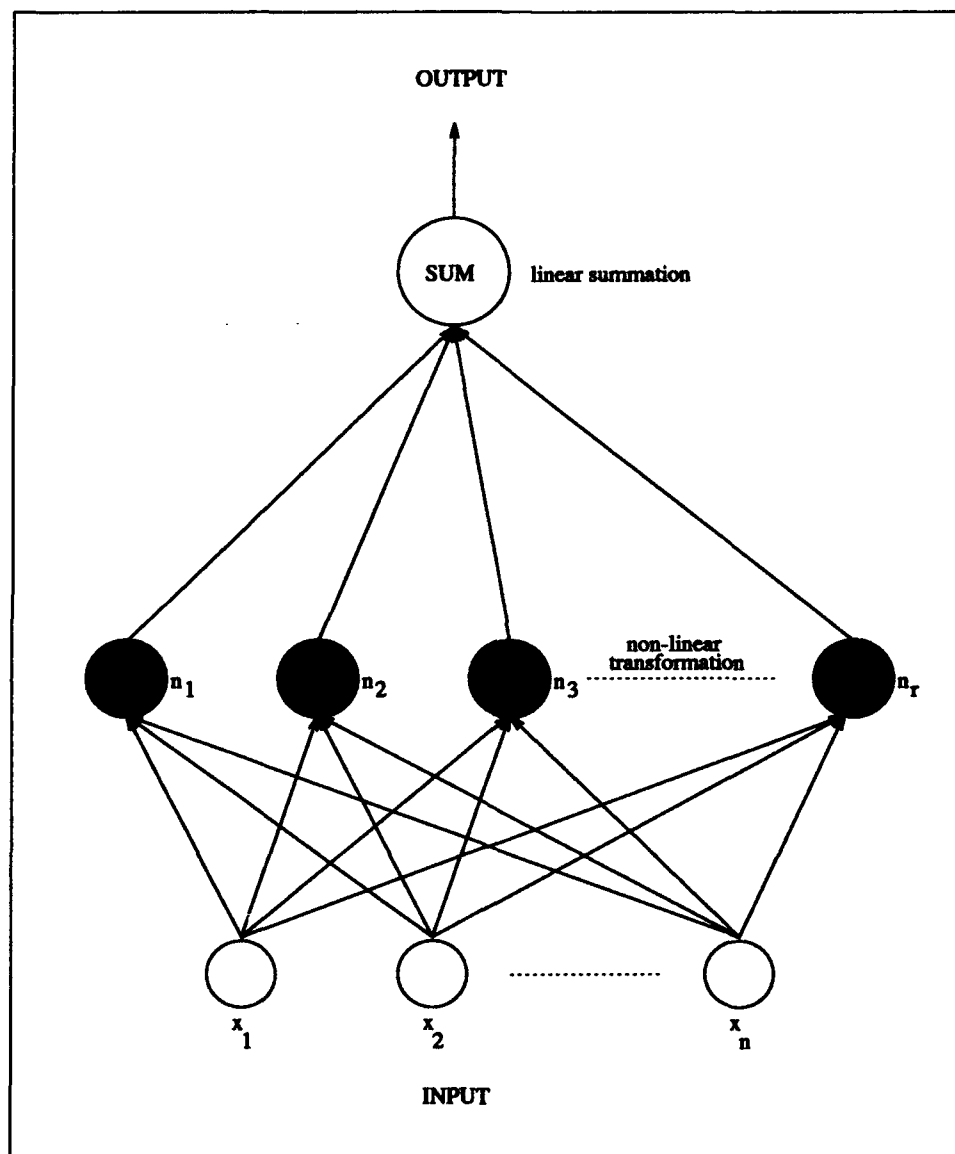


Figure 8. RBF structure

technique is used to solve real world problems (71:104). As A. Cichocki et al. say:, "The important advantage of the RBF network is that it offers training times one to three orders of magnitude faster than the standard back-propagation algorithm used for the multilayer perceptron of similar power and generality" (66:455). The reason that RBF are becoming important algorithms in neural networks is explained:

Neural networks have very general approximation capabilities and have proved to be highly successful in dealing with complex data. The structure of neural networks however is highly non-linear-in-the-parameters. Training times are usually very long and a theoretical analysis of the training algorithm is difficult when models are highly non-linear-in-the-parameters. By carefully pre-fixing centres, RBF networks can closely match the performance of two-layer perceptrons and yet have a linear-in-the-parameters formulation (72:84).

When setting up the RBF network the most critical step is to decide the number of centers. As Chen S. et al. mentioned, if the number of centers is chosen among a well distributed input domain, then a much simpler RBF may possibly provide a reasonable representation for various functions (72). According to Chris Bishop:, "the most natural choice is to let each data point in the training set correspond to a basis function center" (70:579). However, Chris Bishop also pointed out this might cause the so-called overfitting problem, for more detail about the cause of the problem and its solution see (70). Therefore, as Bruce A. et al. said, in order to select appropriate centers and widths we need to choose certain clustering algorithm to the data points in the training set, for example, k-means clustering (73).

2.4.2.4 Other Algorithms. Although there are many different algorithms that have been developed, there are many new approaches being introduced. According to Richard P. Lippmann, there are six important neural nets that can be used for classification of static patterns.

As we see from the figure, those classifiers can be grouped into supervised and unsupervised training:

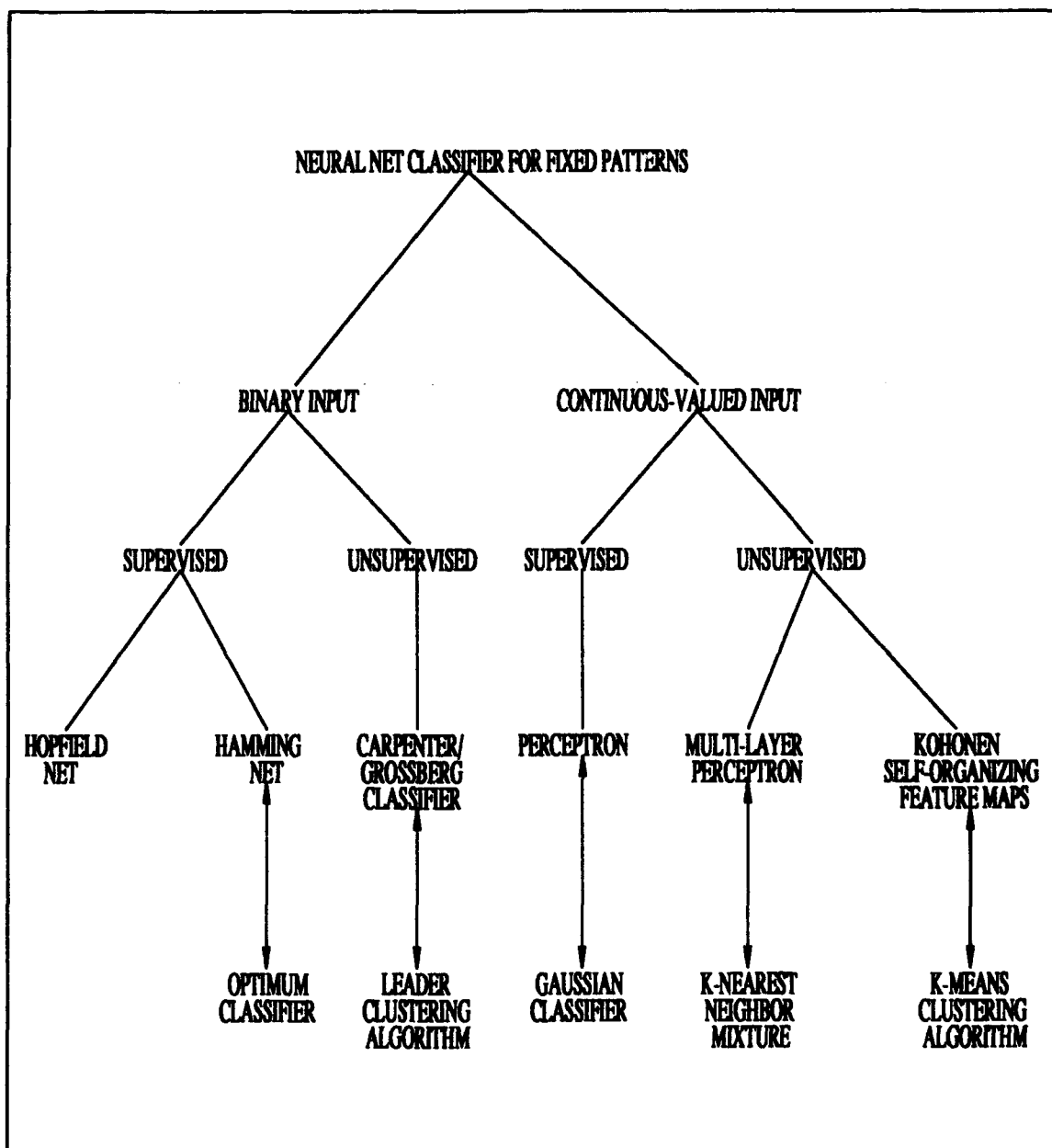


Figure 9. A taxonomy of six neural nets that can be used as classifiers. Classical algorithms which are most similar to the neural net models are listed along the bottom (51:6)

- Supervised training - When a system detects an incorrect response, it uses a cabled data set to send an error message to the network and train the network with the correct process. The system will remember never to make the same mistake.
- Unsupervised training - Also known as self-organization, this does not use any external labeled data set. Instead the system monitors its own performance by following certain function rules and uses the regularity from the input information to adjust itself to get the final result. As Kunihiro Fukushima pointed out that:

The self-organization of the network results from learning-without-teacher. The network by itself acquires the ability to classify and to recognize pattern correctly on the basis of similarity in shape (55:66).

Although its function is still not well known compared with the supervised training, it does have unlimited potential in different applications. For example, in military target recognition, sometimes we do not have suitable training data before we actually have contact with a target (61).

- hybrid network - By combining supervised training and self-organization into a hybrid network, a more efficient network that requires less overall training time may results in a faster learning process may be possible. According to Steven K. Rogers et al.,

The reason that a hybrid network may have an advantage over a strictly supervised network is that the lower self-organized layer hopefully organizes the data into an easier discrimination problem. This transformation of the input data provides a powerful learning process (50:73).

Although each type of training has its own features, learning rates determine the ability and efficiency of an algorithm. There are several trade-offs in learning rates, like the time for training, model for learning, network organization, theoretical

•
•
•
•

foundation, algorithm of transformation, etc. For more detail about the fundamental of neural networks see, (61), (74), and (75).

2.5 Summary

Visual perception is a very complicated issue, especially when we consider it along with our ability to visualize through a display system. It not only is influenced by our physical and physiological character but also our mental or psychological state. In addition, human factors also play an important role in the design of display systems.

In this chapter, a geostatistical algorithm called Kriging method has been introduced for image interpolation. Although it is originally used to estimate geological distribution, it received a lot of attention by researchers in the field of medical data interpolation and approximation. Two of the neural networks algorithms were also introduced for image representation. One is MLP which according to Cichocki et al. it is the most popular and widely used neural network for identification and prediction of the future (66:451). Another powerful algorithm is RBF. As Powell M.J.D. discussed, it has a very excellent performance in multivariable interpolation (76). Neural network algorithms have been widely used in pattern recognition, speech recognition, and in other fields. These algorithms simulate the neuron functions to achieve human-like abilities.

In the next chapter, this information will be applied and further discussed to show its capability in biomedical image data visualization and representation.

III. Methodology and Implementation

3.1 Introduction

The goal of this research is to provide biomedical engineers with the ability to represent biomedical image data in a 3D volumetric form thus providing better visualization for the physicians. This way, the physicians can get a better understanding of the interior anatomy structure and the relationships of the human body or other biomedical organizations. Therefore, a high fidelity and high quality image display system is necessary. In the following sections, an optical display system that can create a stereo image and a polarized monitor that enhances the stereo effects will be discussed. In addition, two neural network algorithms will also be discussed in order to provide the best interpolation of the medical raw slice data which is of interest to the physicians.

3.2 AFIT Optical Display Device

The AFIT Optical Display Device (ODD) is based on a video game by SEGA called TimeTravelerTM. By applying the physics of optics, when light rays projected onto a mirror, the mirror will reflect the light back to a relative place, depending on the characteristics of the mirror and several other parameters. Since images can be seen without any viewing equipment, the ODD is considered an autostereoscopic device. However, after the enhancement of the device it turns out as a stereoscopic device, to be discussed in the next section.

3.2.1 Structure of the Device. The device consists of a frame box that holds a spherical mirror and a TV monitor within it, as shown in the figure 10. When the TV monitor displays an image, the light rays of this image project onto the spherical mirror. When both the TV monitor and the spherical mirror are appropriately arranged, the reflected image will be focused above the reference viewing

plane. Therefore, the viewer can directly see a well formed floating image above the viewing plane.

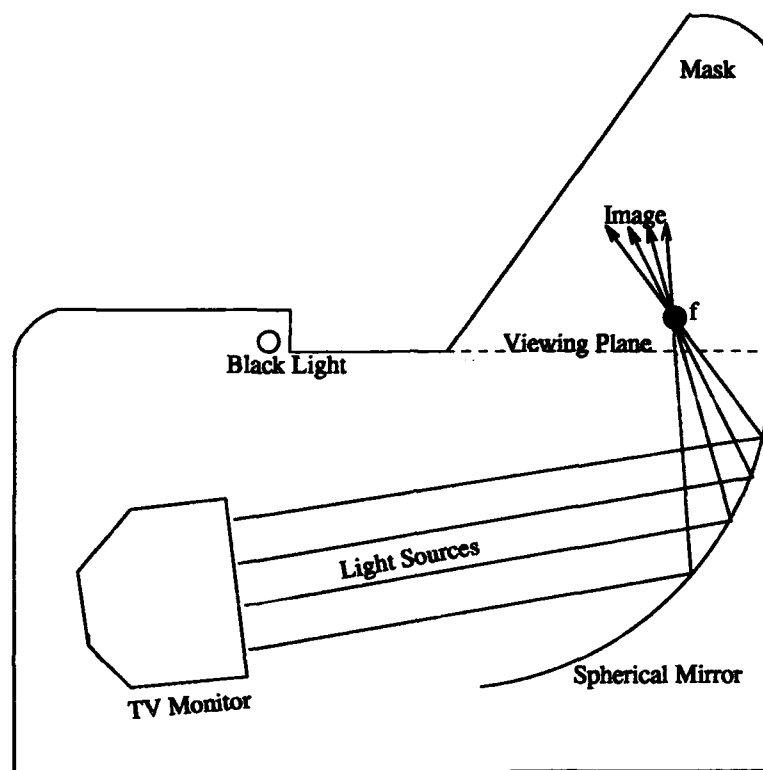


Figure 10. Device illustration

The spherical mirror used with this device is a plastic curved mirror with a radius of 16.93 inches. The focal point of the spherical mirror is given by the following equation:

$$\begin{aligned} f &= R/2 \\ &= 8.465 \text{ inches} \end{aligned}$$

where R is the radius of the mirror. The spherical mirror changes the size, type, and orientation of the image. These changes depend on the distance between the light source and the spherical mirror, as will be further discussed later. See figure 11.

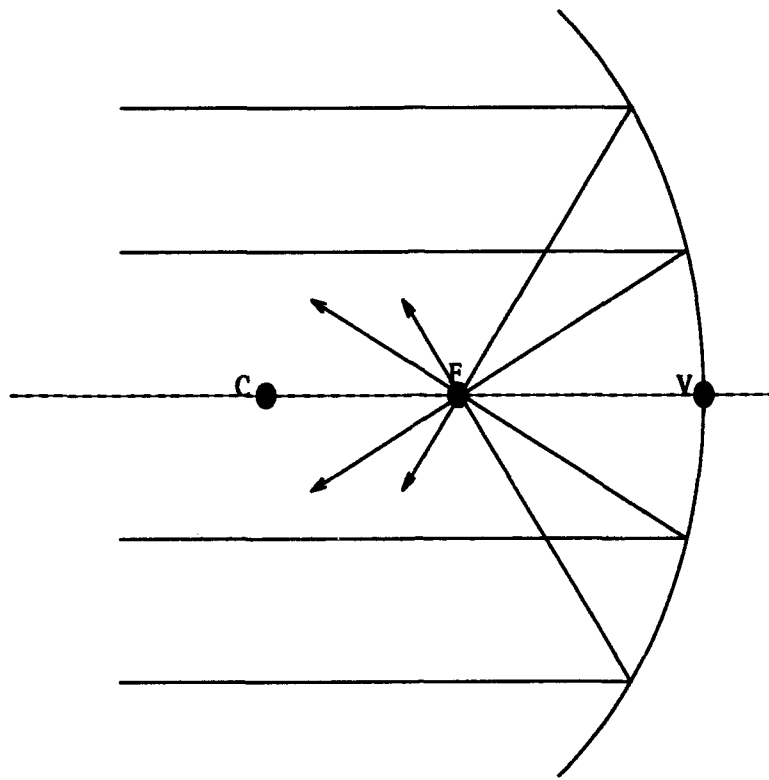


Figure 11. focusing of rays via a spherical mirror (77:160)

In addition, in order to avoid the high reflectivity from the back of the mirror which causes the glossing and smearing of the image, the back of the mirror is blackened. This technique largely decreases the reflection effect when compared with the normal silvered back mirror (a factor of 9.5) (13). The spherical mirror is arranged with an angle of -7.5 degrees with respect to the viewing plane and the optical axis while the TV monitor is placed at about 1.57 inches under the viewing plane, with an angle of 80.9 degrees with respect to the viewing plane and 88.4 degrees with respect to the optical axis, see figure 12. In addition, the back of the viewing space is masked by a frame which limits the aperture of the viewing space. Several other blocks are also arranged around the viewing space and are illuminated by a hidden black light in front of the viewing plane, thus producing a reflection off the viewing plane glass. This increases the effect of the illusion as objects are seen floating on the viewing space to the viewers.

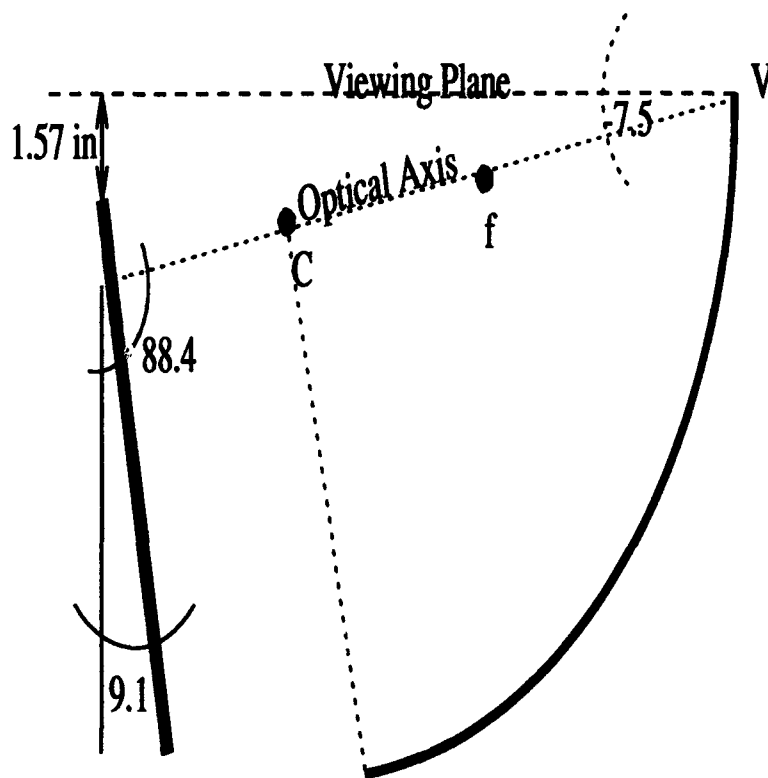


Figure 12. Optical Axis

3.2.2 Enhancement of the Device. Since the image formed by the spherical mirror is projected from a normal TV monitor, though it has the illusion of a stereo image, it is only a 2D flat image instead of a true 3D stereo scene. This is because of the fact that the normal TV monitor cannot provide the most important perception cue which is stereopsis that the spherical mirror can represent. For this reason, a polarized monitor from Tektronix Inc. was used.

The Tektronix monitor is a stereoscopic display monitor with a refresh rate of 63.25 Hz horizontally. In monoscopic mode, this monitor has a vertical refresh rate of 60 Hz, while it has 120 Hz refresh rate in stereo mode. In the monoscopic mode, it provides a horizontal and vertical pixel resolution of 1280 x 1024. In the stereoscopic mode, the vertical resolution is decreased to 512 pixels. When the stereoscopic mode is activated, an 18.2 inches diagonal liquid-crystal controller orients the polarization of any image that passes through it onto the image plane. Since it uses time mul-

tiplexing to display the stereo pairs, thus a stereoscopic modulator in front of the image plane will direct the image pairs switching the left eye and right eye image at 120 Hz. When the viewer is wearing a circularly polarized pair of glasses, he or she will see either a left or right eye image (78). As mentioned in the last chapter, when the visual cortex interpolates these two separate images into a single image, a stereoscopic illusion is produced. Also as mentioned before, the reason that the stereoscopic mode is produced in 120 Hz instead of 60 Hz vertically is because the frequency of the image received by each eye will be decreased by one half. This causes a flickering phenomenon. On the other hand, at 120 Hz, each eye will receive a 60 Hz vertical refresh rate, which is the minimum rate that a human being will not sense the flickering.

3.3 Theoretical Foundation

3.3.1 Spherical Mirror. A spherical mirror is a curvature mirror. Depending on the direction of the reflection of an incident light, this mirror can be considered as a concave mirror, if the light is reflected from a concave surface, or a convex mirror, if the light is reflected from a convex surface. The imaging equation of a spherical mirror is given by:

$$\begin{aligned} 1/s_o + 1/s_i &= 1/f \\ &= 2/R \end{aligned}$$

where s_o is the distance of an object from the center of the mirror and s_i is the focal length of the mirror. Using the characteristics of optic physics, a concave mirror and a convex mirror can produce different types of images. They are either real or virtual, magnified or minimized, erect or inverted, or an exact same size image, depending on its relation to the distance and the focal length of the mirror (see the table below and figure 13).

Quality	Sign	
	+	-
s_o	Left of V, real object	Right of V, virtual object
s_i	Left of V, real image	Right of V, virtual image
f	Concave mirror	Convex mirror
R	C right of V, convex	C left of V, concave
y_o	Above axis, erect object	Below axis, inverted object
y_i	Above axis, erect image	Below axis, inverted image

Table 1. Sign convention for spherical mirror (77:162)

CONCAVE				
Object	Images			
Location	Type	Location	Orientation	Relative Size
$\infty > s_o > 2f$	Real	$f < s_i < 2f$	Inverted	Minified
$s_o = 2f$	Real	$s_i = 2f$	Inverted	Same Size
$f < s_o < 2f$	Real	$\infty > s_i > 2f$	Inverted	Magnified
$s_o = f$		$\pm\infty$		
$s_o < f$	Virtual	$ s_i > s_o$	Erect	Magnified
CONVEX				
Object	Images			
Location	Type	Location	Orientation	Relative size
Anywhere	Virtual	$ s_i < f $, $s_o > s_i $	Erect	Minified

Table 2. Images of real objects formed by spherical mirrors (77:162)

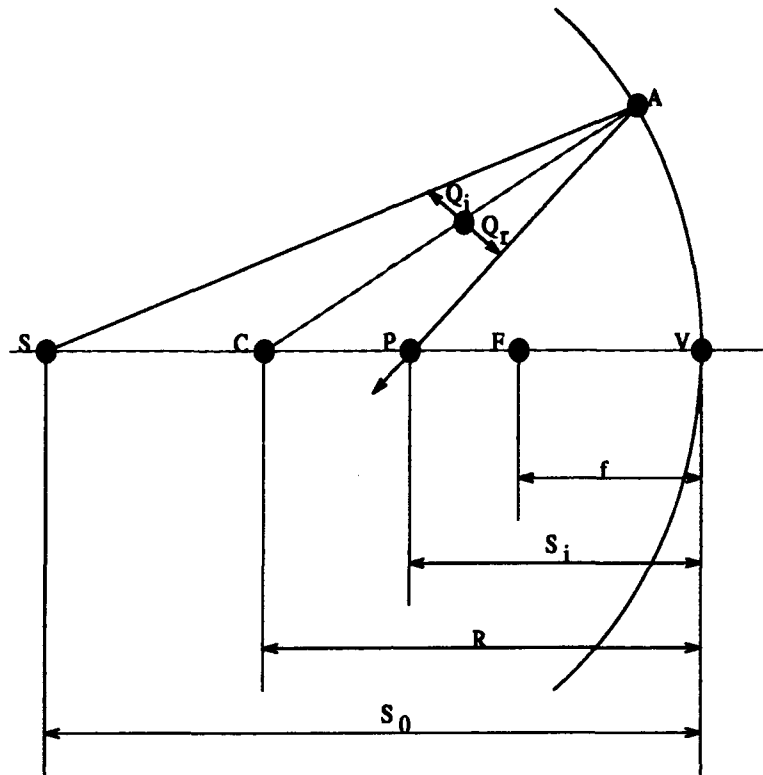


Figure 13. Concave spherical mirror (77:159)

Since the s_0 of the AFIT ODD is about 23.623 inches between the monitor and the vertex of the spherical mirror and the focal length is 8.465 inches, therefore from the table above we know $\infty > s_0 > 2f$, we get a real but inverted and minified image. By using the magnitude equation with $s_i = 13.189$ inches, and the magnification equation of an optical system:

$$M = -s_i/s_0$$

we get -0.22, which means that the image is 22 percent less than the original size. Therefore, in order to see a correct orientation of the object from the AFIT ODD, the image must be rendered in an inverted way or simply displayed it upside down on the monitor. See figure 14.

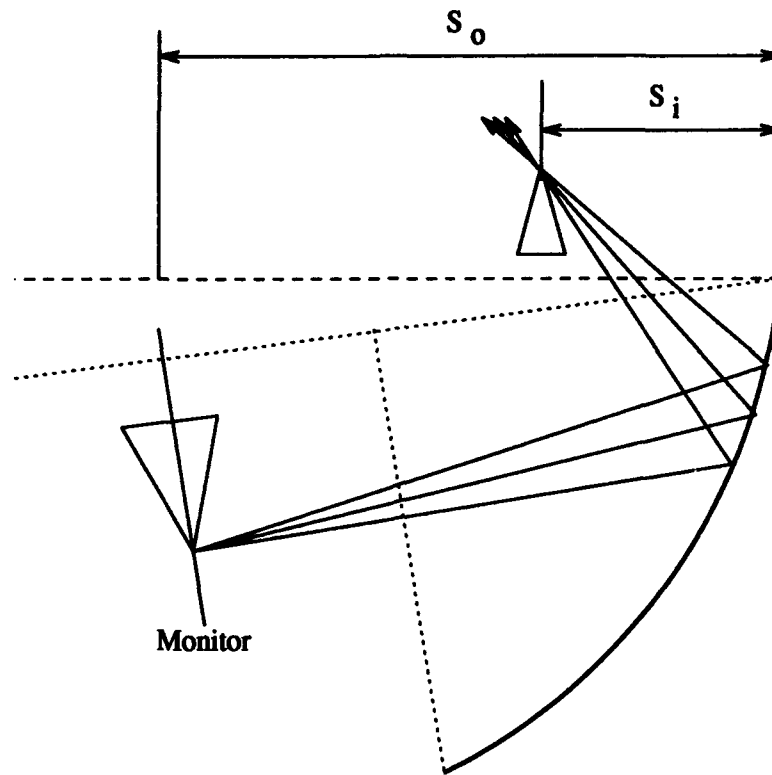


Figure 14. Invert image

3.3.1.1 Aberrations. In order to see a sharp and clear image from the reflection of a spherical mirror, correct focusing is the first thing that we have to be concerned with. This means keeping the image in the paraxial region of the mirror. In addition, if the mirror itself is covered with dust or the surface of the mirror itself is not homogeneous, an image degradation will happen due to the light being scattered out of the paraxial region (79).

However, even with a good focus and an ideal scatter free media, another phenomenon that causes the image degradation is the so-called aberrations. This is caused by the fact that an ideal paraxial condition can not actually hold (79).

There are two types of aberrations, one is monochromatic and the other one is chromatic. The chromatic aberration is caused by the variations of the wavelengths of the light source, thus the mirror cannot keep focusing all the spectral bandwidth. Since the chromatic aberration is not under consideration in the of AFIT ODD, no

further discussions will be included. For more information, see (80).

Monochromatic aberration or Seidel aberration, is a third order aberration. Theoretically, it is caused by the fact that in a spherical system, the wavefronts of the light rays fail to fall into the paraxial region and the wavefronts from the outer paraxial can not give a perfect spherical image. We can find an exact wavefront by using the Snell's law as:

$$n_1 \sin \theta_i = n_2 \sin \theta_r$$

where $n_1 \theta_i \sim n_2 \theta_r$, and $\sin \theta \sim \theta - \theta^3/3! + \theta^5/5! - \theta^7/7! + \dots$

then expand Snell's law to the third order form:

$$n_1(\theta_i - \theta_i^3/3!) = n_2(\theta_r - \theta_r^3/3!).$$

Finally, by summing the five different variations between the ideal spherical wavefront and the actual wavefront we can get the exact wavefront for the spherical system (80). Thus by correcting these five variations we can get a satisfactory image represented by the spherical mirror. We call these five different variations:

- Spherical Aberration - A spherical aberration occurs when either the light rays from the outer peripheral area focus in a paraxial focus of the spherical mirror creating a positive but blurred image or the light rays from the marginal zones focus off the paraxial focus point of the spherical mirror thus giving an inverse and poorer image as shown in the figure 15.

We can calculate the magnitude of the spherical aberration by simply measuring the transverse spherical aberration (TS) and lateral spherical aberration (LS). Since the spherical aberration for a spherical mirror is a function of the distances between an object and the optical axis of the spherical mirror and the radius of the mirror (79), we can use r^4 to represent a spherical aberration, where r is the distance between the axis of the mirror to the incident ray (13). There are several ways that can be used to correct the spherical aberration, like using an aspheric mirror instead of a spherical mirror or changing the focal

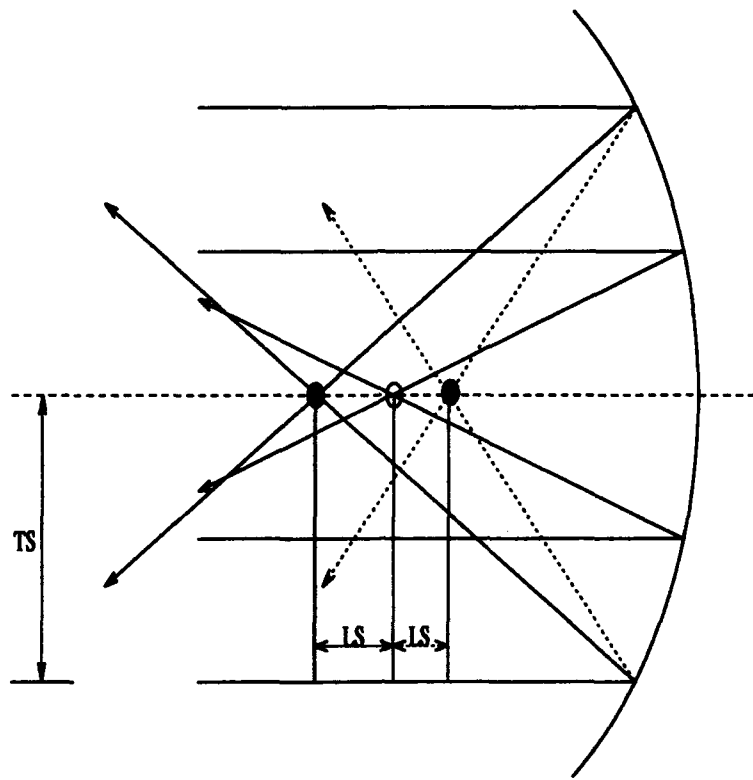


Figure 15. Spherical aberration

length, the distances of an object, or limiting a small amount of aperture into the mirror (77). Since in this research, a spherical mirror is used and the focal length and the distances of an object image are fixed, the spherical aberration is decreased by limiting the incident aperture of the object.

- Coma - Coma occurs when light rays focus off the paraxial point causing an inaccurate image in an asymmetrical distribution of the magnitude of the image. From its third order algebra term $hr^3\cos\theta$, where $\cos\theta$ is the angle of the optical axis with the ray which intersect, we know that when $h \neq 0$ and $\cos\theta$ is not a constant, the Coma aberrations appear to increase fast with the distance r . Therefore, in order to avoid this kind of ambiguous image on a spherical mirror, we need to limit the incident aperture angles and keep the image of the object close to the optical axis and place the object close to the spherical mirror's focus length.

- **Astigmatism** - Astigmatism is the phenomenon that occurs when the image's wavefront is asymmetric, thus giving a blurring image. It is a function of the image height and the mirror's shape and the focus length as shown in the third order algebra term $h^2r^2\cos 2\theta$. This happens when the image is located off the paraxial focus point, when the incident light rays are in distinguishable angles with the concave mirror axis that the image forms mutually perpendicular sagittal plane and a tangential plane image instead of a point image. Since the sagittal rays of the AFIT ODD focus the image plane, there is no sagittal astigmatism (13). However, if we try to eliminate the tangential plane, another aberration will come out as will be discussed next.
- **Field Curvature** - Also known as Petzval Surface, occurs when a focused plane image creates a curved surface. If we compare its algebra term h^2r^2 with the term of the astigmatism case, we can find a close relationship; the missing $\cos^2\theta$ part solves the astigmatism problem, but the field of the curvature is still there. Just like the spherical mirror, the convex mirror will generate an inward image, on the other hand a concave mirror will produce an outward image. As suggested by Brandt (13), this image defect can be reduced by placing an aperture in front of the mirror thus limiting the h and r terms.
- **Distortion** - Distortion occurs when the transversed light has a different displacement due to the paraxial theory. More clearly, it is caused by different parts of the mirror and gives a different focal length and magnification. Its algebra representation can be written as $h^2r\cos\theta$. As we can see h plays an important role in deciding the distortions. There are two kinds of distortions that may occur as illustrated in the figure 15. The barrel distortion is a negative distortion that occurs when the peripheral lateral magnification is smaller than the paraxial lateral magnification. This causes the outer center part of the image to be magnified more. On the other hand, the pincushion distortion has the reverse situation than the barrel distortion, therefore, it produces a

positive distortion (80). Fortunately, as Brandt mentioned, in the AFIT ODD, the aperture stop is the mirror instead of the lens, the ray bundles are not limited, therefore there is no evident distortion by the human vision (13).

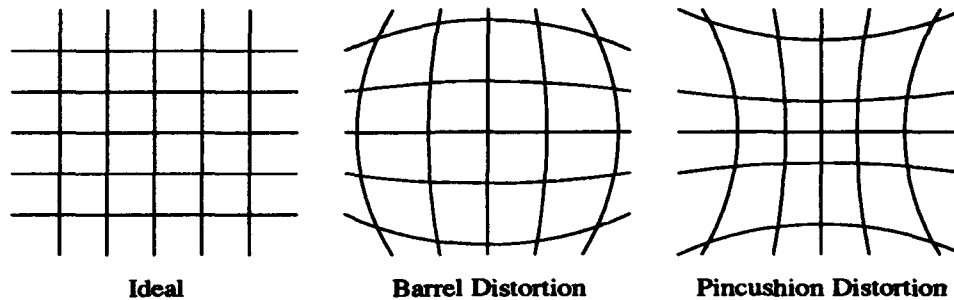


Figure 16. Distortion

3.3.2 Polarized Monitor. Light propagates in a transverse wave form. The oscillation of the wave is perpendicular to the direction of propagation; in other words, the phase of the light waves are the same in each polar of the electro-magnetic field or we can say it is a randomly and rapidly changed polarized light with equal amplitude. However, when those unpolarized light rays meet other environments, there will change due to the characteristics of the transverse of the electro-magnetic wave of the light ray. From the direction of the electro-magnetic wave fluctuating we can further categorize it as a vertical or horizontal polarization. If those horizontal and vertical waves change constantly in a certain form that becomes a polarized light. Depending on the phase and the direction changes of the waves we can distinguish it as elliptically polarized light, circular polarized light, or linear polarized light as the form of waves formed (81).

A polarized monitor is a monitor in which images are polarized horizontally and vertically by the polarization controller with a modulator in front of the image plane. Depending on the characteristics of the monitor, the viewers may see a stereo image directly without any additional viewing adds or with certain helping interface. In this research a passive circular polarized monitor is used. The observers have to

wear a polarized pair of glasses in order to see a stereo image respect to an active polarized monitor or an autostereoscopic monitor so that proper image is selected by the modulator on the screen lens. This way, no additional viewing device is needed.

In a circular polarized monitor, the superimposed image pair, first passes through a polarizer that generates a linear polarized beam. This beam must pass to the modulator in which it either varies the light wavelength, phase, polarization, or its direction. In this research, the polarized monitor used will polarize the light ray in a circular fashion before transmission to the screen following the modulator's switching of the image. If the observer wears a circular pair of glasses, which functions as a retarder or analyzer, the phase of the wavelength is retarded 90 degree in a circular way. Therefore, each time there is only one phase will be shown in one of the eyes. Since, the image keeps switching between the left eye and the right eye with a frequency of 120 Hz, each eye functions with a 60 Hz frequency. When the visual cortex interpret these two separate images as one image, the stereo effect is produced without flickering.

A polarizer is a device that can change the normal unpolarized light into a polarized light. Depending on the wave form it generates, we can further distinguish it as a linear polarizer, circular polarizer, or elliptical polarizer. There are four fundamental physical mechanism that govern the characteristics of the polarized light. These are dichroism, reflection, scattering, and birefringence. Although different polarized lights have different spectrums that are influenced by those four mechanisms, they do share the same property. This property is due to the asymmetry between them which is also the reason why we can polarize the natural light. The four mechanisms are:

- Dichroism - Or selective absorption, different interfaces have different absorption on a visible light depending on the characteristic of the interface. Normally, it is a function of the incident wave length. In addition, the power of

the absorption parameter and the length of the light travel will also decide the degree of dichroism (80).

- **Reflection** - There are two kinds of reflections. One is specular reflection and the other one is diffuse reflection. Specular reflection happens when light is reflected from a smooth mirror like surface and a diffuse reflection happens when light is reflected from an uneven or rough surface. A light polarizes when it reflects from the surface, the change of the interface also causes the change of the wavelength and the phase, thus generating a polarized light. However, a clear or perfect reflection can only happen in a certain angle which is called Brewster's angle, that is $\tan\theta_p = n_t/n_i$, where $\tan\theta_p$ is the angle of the polarization, and n_t and n_i are the transmit and the incident parameters of the interface respectively (77).
- **Scattering** - Scattering happens when a light ray passes through a non homogeneous interface thus the light wavelength will attenuate and cause the change of the wave and generate a polarized light. This phenomenon also depends on the molecules of the medium. Different mediums cause different degrees of polarization (77). In other words, if the polarized light is a visible light due to the change of the wavelength, different color of the light can be seen.
- **Birefringence** - Or simply double refraction, this happens when the characteristics of the medium are different in a different orientation thus the polarized light will show a different phase at the same time, usually a double image will be seen. We call this medium an anisotropic substance (77).

Although the above four incidences do affect the polarization of the light, however, none of them happen on the stereoscopic monitor. Therefore, in this research it is not under consideration and no more details will be discussed.

3.4 *Image Transform and Acquisition*

In order to represent the stereo image through the stereoscopic monitor, three steps are needed. First, a proper image pair is needed, second we need to properly put the relative image pair onto the correct coordinate of the monitor which provides the best image fit, and finally, a program to activate the stereo monitor in stereo mode, instead of the normal mode. A poor arrangement of the position of the images on the screen will affect the stereo features that we see and may cause the loss of the stereo ability. Therefore, in this section three application tools are introduced and their improvement will also be discussed.

3.4.1 *Rayshade 4.0 Program.* Rayshade is a graphic manipulation program. It uses ray tracing methods to create or refine a quality image. It was developed by Craig Kolb and Rod Bogart in 1991 (82). Whenever an appropriate command and parameters are given, the program will create a simple geometric object or generate some other graphical effects by the rayshading method to improve the image. Since each assigned parameter follows a proper procedure of calculations, therefore, depending on the image quality, the resolution, and the computation ability of the computer, time becomes important. Although rayshade support two image output formats, mult-image (MTV) and run length encoding (RLE) format files, however, RLE format file was adopted in this experiment. The reason is because the RLE file provide more flexible functions and give more benefits than MTV file, see (83). The rayshade program provides powerful functions that support the basic need for rayshading. For example, it can generate the shadow effect which is helpful in improving the stereo effect. Others, like atmosphere and surface effects, texturing functions, and the specifying of field of view and depth of field all will provide the necessary perception cues for a stereo image as discussed in chapter two.

One of its feature is to render an image into two different viewpoint images. With this, we can generate two view images one for each eye with a little disparity from each eye. Thus when properly put onto the polarized monitor with the stereo

mode on, than the viewer can see the stereo image as discussed above. However, as suggested by Graig E. Kolb, for the best result, the disparity should be equal to one tenth the distance from the view point to the look point. Because too large a separation will cause a hyper-stereo effect or so-called ghost image, if too close of the disparity no stereo effect will be seen at all. For more detail, read (80). In addition, in order to see the best result with a proper proportions, the rendering image should be anamorphically compressed in the vertical direction by a factor of two (84). In this research the stereo image pair will be generated by rayshade as a RLE format file.

3.4.2 Utah Raster Toolkit. Utah raster toolkit is a software package that was developed by John W. Peterson et al. from the University of Utah. The software was coded in the C language thus giving the possibility of extending or fixing for different situations or needs. It manipulates and composes the raster images by using the RLE format as the image file, for it provides the feature of saving image storage space which is very important for image processing or graphics need. It also provides a standard description of the image that can be supported by different machines thus solved the former restriction of the machine based image (85).

"get4d" is a function tool which is provided by Utah raster toolkit. When a RLE image file is assigned, it will read the file's header to get the information of the image. This includes the size of the file, number of channels, color map etc., for more detail see (85). As long as all the parameters are set up, a 1 x 1 inch window will show on the screen as the pointer of the mouse location. The user can then move the window to a proper location and enlarge the window to any size until the max rendering size is reached.

However, in order to allow two disparited rayshade images to be simultaneously present, and positioned at a proper relative places; it is modified to let two image files to be called in the same time. The program will check the size of the max resolution from the RLE file header, if the two image sizes are over the size of the screen, it

will cut the upper part of two images equally (we suggest only rendering the image with the resolution not over half of the size of the monitor, thus there would be any part of the image lost); on the other hand, if the two images are smaller than the screen, a proper distance between them will be automatically added according to the size of the images thus giving the best superimposing result when the stereo mode is activated. Because If the images are too close the stereo effect will be lost, if they are far apart, a ghost or double image will see.

In addition, the window is set at the same size as the screen instead of the max image size with a black background color. It has the following three immediate benefits: first, it hides other unrelated messages from the screen, thus only the image will be shown on the screen, second, the blackened background makes the image more specular or distinguishable, third, it confines the aperture output only from the present image pair. This minimizes the light rays projected onto the mirror which cut the light reflection from the peripheral parts of the monitor, this also avoids the annoying reflection from the mirror that would influence the visualization of the image. In addition, since the aperture is limited to only the present images, it also decreases the spherical aberration as discussed before. See figure 17.

In order to let the reflected image to represent a positive oriented image without change of the orientation of the monitor (or the observers), the images are rendered in an inverted way thus giving a positive representation of the image from the AFIT ODD. (There is another way that do not need to invert the image which is invert the observers, normally we do not suggest to do so unless the observers like to).

3.4.2.1 Stereo Activate Procedure. As long as the image pair is sent to the polarized monitor, with the circular pair of glasses we can see a superimposed image pair to be present on the screen. However, in order to see a stereo image, the monitor must be set at the stereo mode. The stereo mode changes the monitor from its normal 60Hz mode into 120 Hz. Thus we can see a stereo image. This procedure was provided by Capt Bob Caley. It was written in C++ language and use the SGI

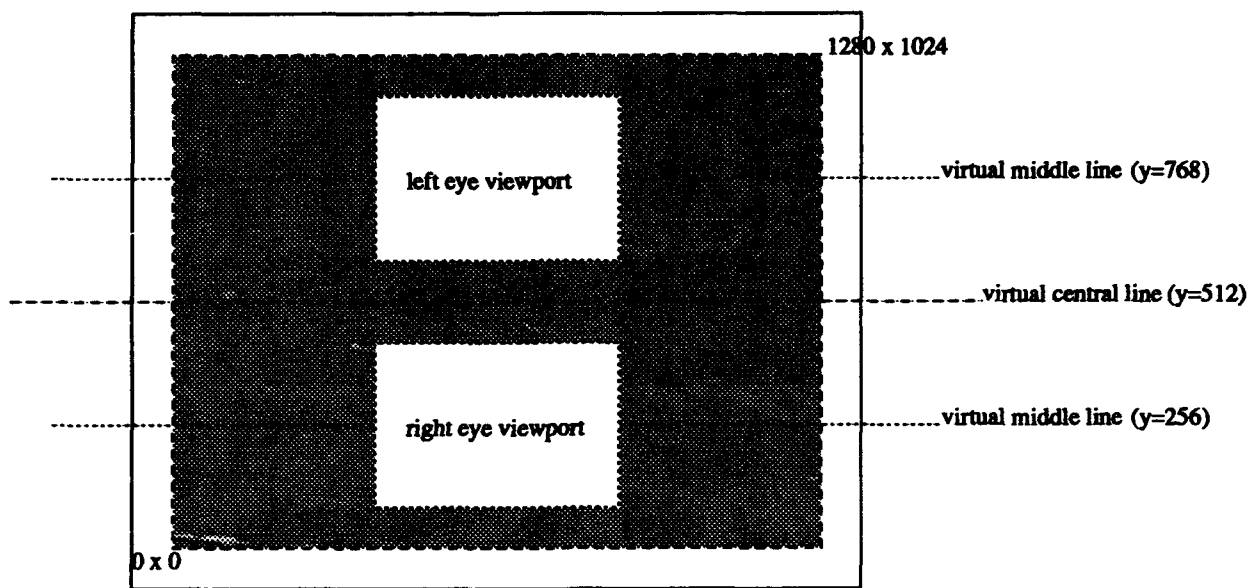


Figure 17. Stereo Image

GL call "setmode" to set the appropriate display mode for the monitor. Every time when a command is ordered, a procedure call will reacting by sending the parameter to graphic lab (GL) and activate the appropriate function thus active or deactivate the monitor mode between stereo or normal mode. When the stereo mode is on, the function will drive two 1280 x 492 image superimposed. If the stereo off mode is activated, the monitor will go back to the 60 Hz non-interlaced state, see (86) for more detail and the program is list in the Appendix 2 and 3.

3.5 Image Representation

The second experiment in this research is image representation. As discussed in the last chapter, biomedical image data is obtained as a sampled 2D sliced data. However, when we visualize and study this data in a 3D volumetric form, a precise and accurate representing image will help the physicians to better understand and judge the whole incidence of the data they are interested in. In order to integrate that nonlinear sampled data as a whole volume data, interpolation is needed.

3.5.1 Data Interpolation. Interpolation is the knowledge that finds the best point(s) or unbiased relations of neighboring non-linearly distributed point(s) or data, thus allowing the connection as a whole set. It is a very useful technique in many fields. For example, in finding the distribution of underground or undersea natural resources. In the field of medicine, interpolation has been extensively used to estimate the unknown region from the known region of sampled data. As mentioned before, the reason that only sampled data is taken instead a whole data set is due to both time and cost constraints.

3.5.2 LNKmap. LNKmap is a software environment package which not only provides neural network algorithms but also conventional pattern recognition and statistical algorithms (see figure 18) that lets the user perform different tasks and fulfill different needs in an easy and simplified way without traditional tedious procedures. Basically, as long as the user selects a proper data set and a classifier algorithm with a chosen clusterer followed by setting several parameters, the experience can be done automatically. It is developed by Linda Kukolich and Richard Lippmann, and is licensed by MIT Technology Licensing Office. It runs under the SUN OpenWindows UNIX system. Since the central features of the neural network lies in its classifier and which can be trained to perform many tasks. Therefore, in the same manner, neural networks can be trained to interpolate a 2D data as a 3D volume data. In this research, two of the neural network algorithms, one is Multi-Layer Perceptron (MLP) with Back-Propagation (BP) supervised training algorithm and the other one is Radial Basis Function (RBF) combined unsupervised-supervised training algorithm, are adopted.

3.5.3 Pattern Classifier with Clusterer. In order to train a neural network, we need to find a classifier and a clusterer. A classifier is used to assign each input pattern to an unique class. The classification of the class is based on the features of the input data or patterns. Normally, features are discriminated according to a

LNKnet Algorithms			
	Supervised Training	Combined Unsupervised-Supervised Training	Unsupervised Training (Clustering)
Neural Network Algorithms	Multilayer perceptron (MLP) Adaptive step-size MLP Cross-entropy MLP Differential trained MLP Hypersphere classifier	Radial basis function (RBF) Incremental RBF (IRBF) Differential IRBF Learning vector quantizer (LVQ) Nearest-cluster classifier	Leader clustering
Conventional Pattern-Classification Algorithms	Gaussian linear discriminant Quadratic Gaussian K-nearest neighbor (KNN) Condensed KNN Binary decision tree	Gaussian-mixture classifier Diagonal/full covariance Tied/per-class centers	K-means clustering Estimate-Maximize (EM) clustering
Feature-Selection Algorithms	Canonical Linear Discriminant Analysis (LDA) KNN forward and backward search		Principal Components Analysis (PCA)

Figure 18. LNKnet Algorithms (87:2)

certain classification function. According to Lippmann there are three approaches to developing pattern classification. These are the probability-density-function (PDF), posterior-probability, and boundary-density-forming strategies (88:250). As Lippmann further explained:

These approaches differ in the statistical quantity that their outputs model and in the procedures they use for classifier training: PDF classifiers estimate class likelihoods or probability density functions, posterior-probabilities classifiers estimate Bayesian *a posteriori* probability, and boundary-forming classifiers form decision regions. ... The PDF functions formed by statistical classifiers are Gaussian shaped. These functions represent the distributions of the input feature for the classes. Posterior probabilities formed by many neural network classifiers have sigmoidal shapes. ... Finally, the binary indicator outputs of boundary-forming classifiers separate the input into two regions, one for class A and the other for class B. See figure 19 (88:250-251).

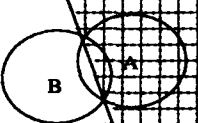
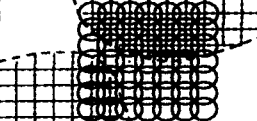
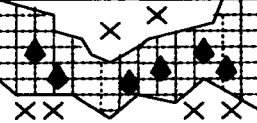
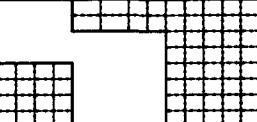
A Pattern-Classification Taxonomy			
Type of Classifier	Decision Region (shaded)	Computing Element	Representative Classifiers
PDF		Distribution dependent	Gaussian, Gaussian mixture
Global		Sigmoid	Multilayer perceptron, high-order polynomial network
Local		Kernel	Radial basis function, kernel discriminant
Nearest Neighbor		Euclidean norm	K-nearest neighbor, learning vector quantizer
Rule Forming		Threshold logic	Binary decision tree, hypersphere

Figure 19. Five different types of classifiers (88:252)

As Lippmann et al. pointed out: "the decision rule that most classifiers use is to assign the feature vector to the class corresponding to the discriminant function, or output, with the highest value" (88:250). A clusterer is the algorithm that finding the center of the cluster of the input data (patterns). The clusterer can be either trained on cabled or unlabeled training data. Which is, when trained, the clusterer(s) can be trained for each separate class or all the classes trained a single set of clusterers (88).

3.5.3.1 K-means clusterer. K-means clustering algorithm is one of the most important and popular clusterer. It is a binary splitting algorithm. As Linda Kukolich said: "When splitting a cluster, the initial means of new clusters are shifted a given percentage of the variance of that cluster away from the mean of

the original cluster. They use the variance of the old cluster as their initial variance (89).

3.5.3.2 MLP classifier. According to Lippmann et al. a MLP belongs to a global neural network classifier in which it: "form output discriminant functions from internal commutating elements or nodes that use sigmoid or polynomial functions having high nonzero outputs over a large region of the input space" (88:251). As further explained by Lippmann et al.:

They create decision regions by positioning ridge functions produced by sigmoids in the input space. In this limits where connections weights are "high" they create planes which define "half-spaces". These are combined to form class decision regions. The planes and their combinations are specified by weighted connections between the layers of the multi-layer perceptron. The weights are trained using a back propagation algorithm to perform a gradient which minimizes the error of the outputs according to some cost function (87:42).

According to Linda Kukolich:

In testing, a multi-layer perceptron computes several weighted sums of features in the input pattern. These sums are passed through sigmoid functions and the results are used as the input to the next layer in the network. Input values are fed forward this way until the the output layer of the network is reached. The class corresponding to the highest output node value is selected as the class of the pattern.

In training, the weights connecting each pair of adjacent layers are adjusted to minimize the error in the outputs of each input pattern. A gradient descent method is used for weight training, with the changes needed propagating backwards through the network (89).

Since MLP is the most popular algorithms, which provides excellent performance, in this software environment, MLP has many options in its application. For more details see (87), (88), (89), (90), and (91).

3.5.3.3 RBF classifier. The RBF classifier uses some output or results from the neural networks in order to minimize the mean squared error. As

Lippmann said: "Radial Basis Function classifiers calculate discriminant functions using local Gaussian functions instead of sigmoids of hidden node sum (87:44). As the description of user commands mentioned:

Hidden node mean locations and variances are determined using clustering. Output layer weights are then trained using a single pass through the data. Correlations between the output of all Gaussian hidden node outputs and between hidden node outputs and output nodes for each training pattern are summed and used to form a matrix which is inverted to find the weights from each center to each output node. These weights minimize squared error between desired and actual outputs of the training data. During testing, outputs are formed by forming weights sums of hidden node outputs (89).

3.6 Summary

This chapter focused on the different techniques of generating an image pair and how to put this image pair in a proper location. This image is transformed to a stereo mode and then projected onto the mirror. In addition, several other modifications are added to improve the viewing effects. Thus, when the viewer wears circular glasses, a real stereo image can be seen. Further more, the causes of the image defects from either the optical system or the display system and how to avoid or diminish those defects were also discussed. In addition, a software package was introduced in order to perform the interpolation of sampled data, thus providing a better image for the viewers.

In the next chapter, the results of the projected image data and its defects will be discussed. Furthermore, the analysis of those defects and their improvement will also be discussed.

IV. Results and Analysis

4.1 Introduction

This chapter discusses the experimental setup used to produce a 3D stereo image, the results obtained from the experiments, and the performance of the AFIT ODD.

4.2 Experimental Setup

The two images used in this research were obtained from a data base which contained various images. The first image was of a planet with a satellite orbiting around it. The satellite appeared to be closer to the viewer compared to the planet which appeared much farther away. The second image was of another planet with a satellite orbiting it. However, in this image, the relative distances of the various parts of the satellite were clearly distinguishable. The planet was small giving the impression that it was at a much greater distance from the viewer than the satellite. Although the images provided relative distances from the observer's point of view, in order to perceive a stereoscopic 3D image, they must be processed using a system which creates an illusion.

A stereo pair image was produced from each of the two image files using rayshading (82). The stereo image pair consisted of left and right views. In order to create a clear picture, the rayshading algorithm required a disparity parameter to be set.

The stereo image pair were fed into the modified "get4d" program to position the images such that they fit on the Tektronix stereoscopic monitor. The left and right images did not overlap and were each updated at 60 Hz. A procedure call to the SGI graphic library (setmode) created overlapping images and set the frequency of the composite image to 120 Hz. This image was projected onto a spherical mirror where it was reflected above a glass viewing plane. A pair of circular polarized glasses was required to create a single image with depth information.

In order to fool the brain into thinking it was viewing a true 3D image, references to the monitor's borders as well as other surrounding objects were removed from view using this modified AFIT ODD.

4.3 Experimental Results

Both of the images processed through the modified AFIT ODD appeared as real 3D stereo images. The 3D stereo images provided the missing depth cue information to the viewer. In addition, the relative distances between the different parts of both the satellite and the planet were enhanced. Therefore, the viewer obtained a better understanding of the objects in the image.

A 3D stereo image cannot be captured on film. Therefore, in order to show the results of the experiments, two pictures of each of the 3D stereo images were taken. The first picture was taken using a polarizing plate with no phase shift (see Figures 19 and 20). This picture represents the image that would have been perceived by the left eye of a viewer. The second picture was taken using the same polarizing plate set to provide a 90-degree phase shift (see Figures 21 and 22). This picture represents the image that would have been perceived by the right eye of a viewer. The difference between the two pictures was due to the disparity of the two viewpoints created by the modified AFIT ODD. The polarized glasses are based on the same concept as the polarizing plates. In effect, the polarized glasses fooled the brain into thinking the left and right images are one composite image represented in a 3D stereo mode.

4.4 Performance of the Modified AFIT ODD

Generally speaking, the AFIT ODD proved its potential ability to represent biomedical image data in a 3D stereo form. The following discussion describes the important features of the experimental setup using the AFIT ODD.

- **Field of View** - In an optical system, the field of view is defined as the extent of an object that can be seen. There are two elements which influence the field

of view: the aperture stop and field stop. The aperture stop limits the amount of light that is projected and the field stop controls the amount of light that passes through. In the AFIT ODD system, the aperture was confined by the size of the image. This restriction also limited the amount of light that hit the spherical mirror (i.e., entrance pupil). On the other hand, the field stop is the mask of the frame box where the final image was viewed. The reflected light from the spherical mirror to the viewing plane was also affected by the aperture of the projected image (i.e., exit pupil). Since the mask and the mirror were not symmetric, the reflection of the light rays was limited to the far left and right sides of the mirror. The glass viewing plane, located between the mirror and the frame mask, served two functions: as an entrance port when the reflected light rays struck the viewing plane and as an exit port when the light rays left the viewing plane.

- **Spatial Resolution** - The spatial resolution exhibited by the AFIT ODD was better than the resolution of the stereoscopic monitor. This is due to the fact that the spherical mirror minimized the size of the reflected image by 22 percent
- **Refresh/Latency and Update Rates** - In stereo mode, the monitor's frequency was run at 120 Hz. Since the polarized glasses divided this frequency in half, each eye received a 60 Hz image. This phenomenon eliminated flickering of the image. Also, since the image remained static, no update was required.
- **Luminance, Brightness, Lightness, and Contrast** - There were three factors that affected the image luminance, brightness, and lightness: the barrel of the stereoscopic monitor which limited the photons projected onto the spherical mirror, the low reflection mirror which decreased the photons reflected, and the pair of circular polarized glasses worn by the viewer which reduced the photons entering the viewer's eyes. In addition, the rayshading procedure used in generating the stereo image influenced the above three effects. If the shading

is dimmer, the output aperture will be much lower. Since the background of the window was black, only the object itself projected light onto the mirror. This was done so the edge of the display is not noticeable. Therefore, the object itself had a good contrast with respect to the surrounding environment. However, the contrast of the object itself depended on the rayshading between its parts.

- Color - Since the stereoscopic monitor supported a full color display, there was no limit on the colors of the projected image. Therefore, color depended totally on the shading parameters. However, due to the low reflection of the mirror, the colors with low intensity may have been lost during the light transmission.
- Information Rate and Bandwidth - The monitor alternately transmitted a top and a bottom image, each at a rate of 60 Hz. Therefore, no flickering phenomenon occurred. The information rate was doubled to 120 Hz and the bandwidth was one half (60 Hz).
- Viewing Zone/Volume Extent - The viewing zone of the AFIT ODD was limited by the stereoscopic monitor. This gave a vertical viewing angle of 27.2 degrees and a horizontal viewing angle of 36.86 degrees.
- Distortion - There was no distinguishable distortion caused by the spherical mirror.
- Number of Viewpoints - Since the image presented was a volumetric image, there was no limit in the number of view points.

Generally, the AFIT ODD provided a satisfactory 3D stereo image. However, in order to make it a more practical 3D stereo image display system, more research is needed. Specially to apply the device to biomedical visualization either neural networks or Kriging must be tested.

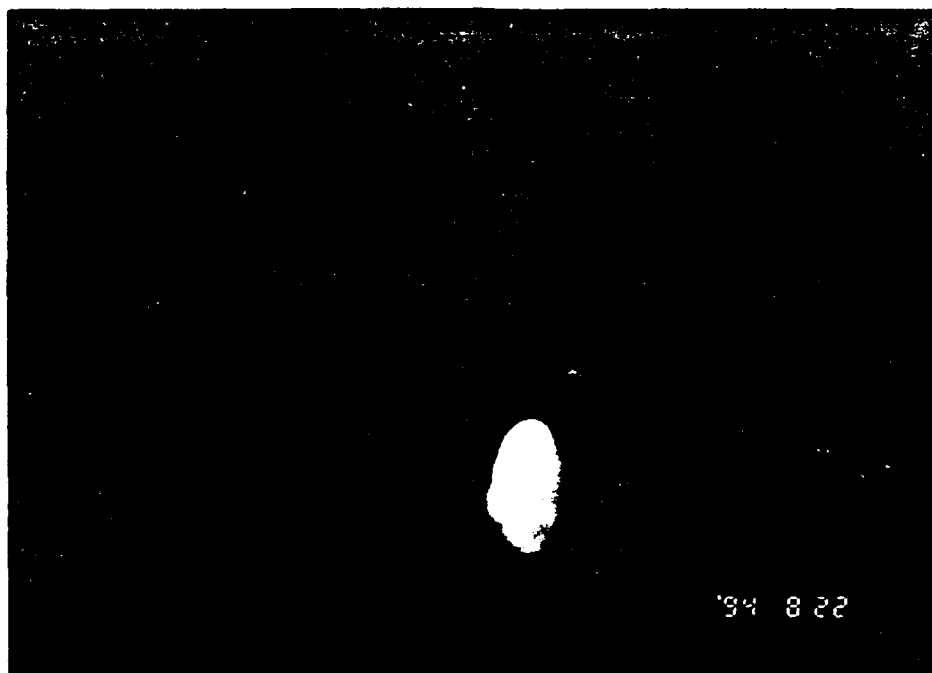


Figure 20. Left eye image

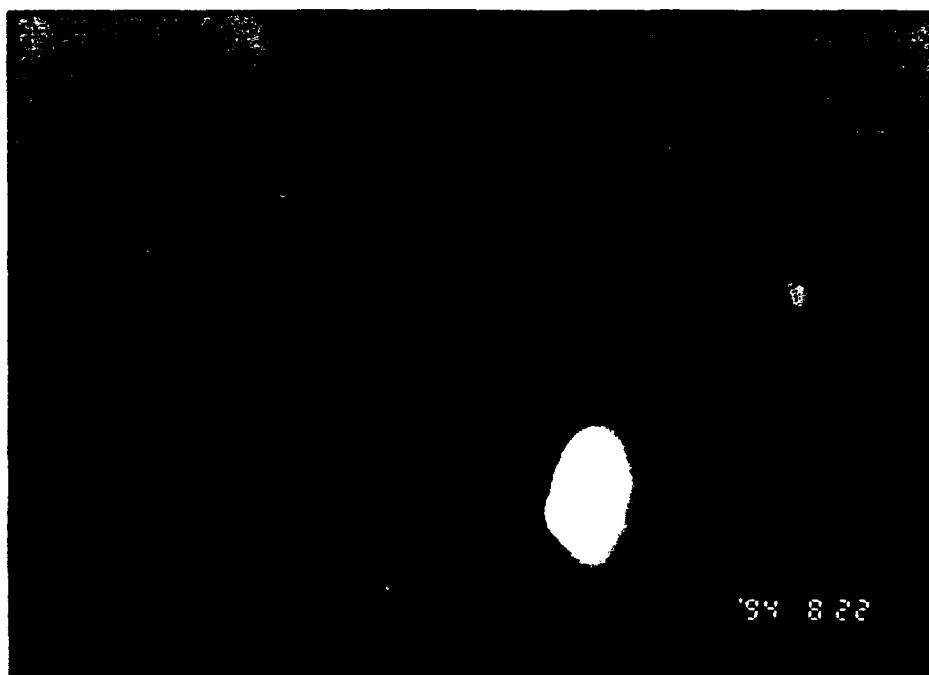


Figure 21. Right eye image (with 90 degrees phase shifting by polarizer)

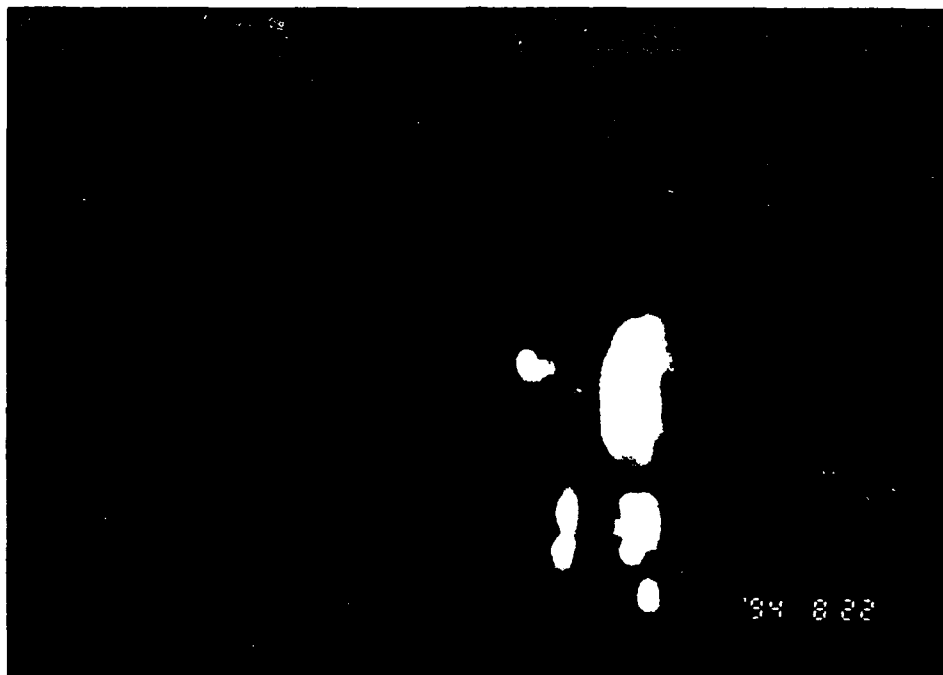


Figure 22. Left eye image

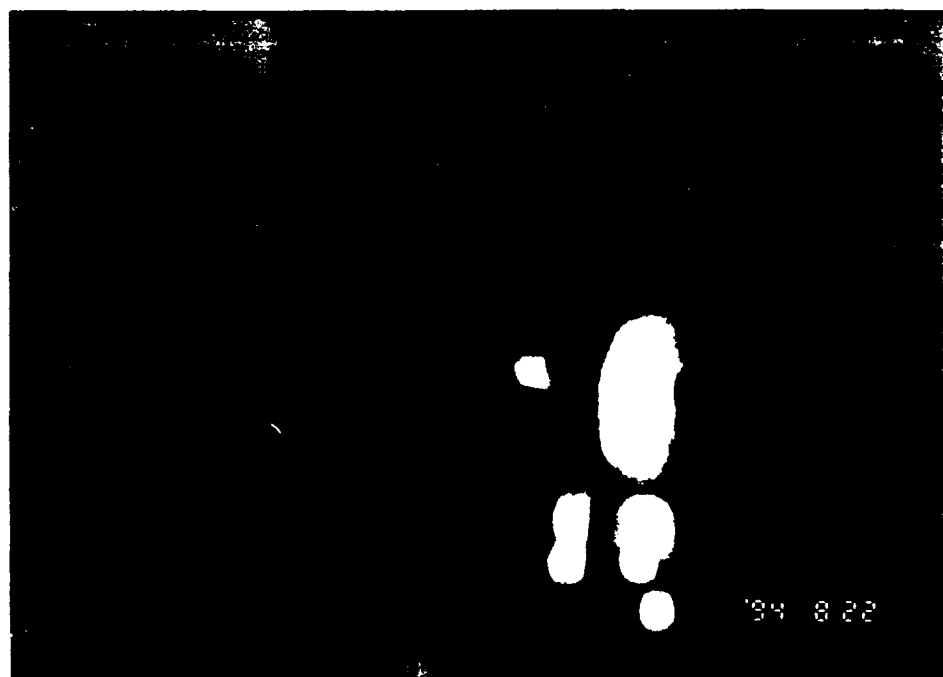


Figure 23. Right eye image (with 90 degrees phase shifting by polarizer)

V. Conclusions and Recommendations

5.1 Introduction

Advances in military technology often precedes advances in the private sector. Many commercial inventions have been based on military applications. Civilization has benefited from the discoveries made within the military. However, the military has also adopted the technology developed by the private industry. For example, a simple device that was developed for recreational use has become a useful tool for the military.

The goal of this research was to convert an image display device originally designed for entertainment into a device useful for biomedical image visualization. Images displayed by entertainment devices are usually not precise or of high quality. However, biomedical applications require a device that must be able to display both a high fidelity and high quality image. The first problem addressed in this research was to improve the quality and fidelity of an image. A geo-statistical algorithm and two neural networks algorithms were studied for potential application to the biomedical data.

In the following sections, conclusions and recommendations will be discussed.

5.2 Conclusions

In this research, optical phenomena, computer graphic techniques, and human factors were studied in order to develop a technique for converting an optical device from SEGA corporation into a real 3D stereo image display device. This technique projected images onto a spherical mirror and, under the appropriate arrangement of the image with respect to the mirror, a floating stereo image appeared. Precision and fidelity of the image was not considered because the SEGA device cannot provide the depth perception cue. However, if a biomedical image is to be visualized, a precise and high fidelity image becomes important. In this research the goal was to

apply this technique to create a 3D stereo biomedical image allowing physicians to study and analyze biomedical structures and their organization within the body.

This device is economical but provides satisfactory performance. For this research, performance was enhanced by replacing the projector. Instead of a normal TV monitor, a stereoscopic monitor was used as the projector. This provided the stereopsis cue that a normal TV monitor could not provide. Thus, a real 3D stereo image will be seen instead of the flat 2D stereo image. Human factor effects – field of view; spatial resolution; refresh/latency and update rates; luminance, brightness, lightness, and contrast; color; information rate and bandwidth; viewing zone/volume extent; distortion; and number of viewpoints – were also studied to improve the viewing environment.

This device represents images based on the principle of optical physical phenomenon as discussed in Chapter 3. Even with well formed images, they can be further enhanced by eliminating third order aberrations, providing larger images, and improving image visibility. In terms of computer graphic techniques, instead of using a biomedical image to test the device, a simple geometric object was used. Thus, the device may not represent biomedical images properly because these are more complex. Further testing of the device is necessary to validate its ability to provide a high fidelity visualization of biomedical data. A more efficient algorithm to deal with the stereo image generating procedures may give a better and less ambiguous image. In terms of human factors, improvements to the viewing condition are needed. For example, a wider viewing space and larger viewing angles allow more flexibility or more people to visualize a single image at a time. In addition, a well illuminated image is easier to view than one in a dimly lit environment.

In addition, a geo-statistical (Kriging) algorithm and two neural network (multilayer perceptron, and radial basis function) algorithms were studied and background information is provided for potential application to interpolate the medical image data. This data interpolation is necessary before further processing by the

AFIT ODD.

5.3 Recommendations

This research proved that the modified AFIT ODD has a greater potential for biomedical image visualization and representation than the original AFIT ODD. However, several enhancements may be incorporated and several problems must be addressed in order to further improve the performance of the modified AFIT ODD in displaying biomedical images:

- A major improvement can be obtained by replacing the stereoscopic monitor with an autostereoscopic monitor. This may eliminate the need for circular polarized glasses in order to visualize a stereo image. This improvement addresses two human factors problems: for people that don't wear glasses, they will not be required to put on glasses in order to see the 3D stereo image; for people that already wear glasses, they will not need to put on a second pair of glasses on top of their own glasses.
- An increase in the viewing field and the elimination of potential aberrations may result if the spherical mirror is replaced with a larger Cartesian oval mirror. Examples of Cartesian oval mirrors include aspherical mirrors and parabolic mirrors. Careful study of the replacement mirror is necessary in order to ensure new aberrations are not introduced.
- Increasing brightness, lightness, and luminance of the image may improve the viewing of a 3D stereo image. Currently, an image appears dim when viewed with the modified AFIT ODD. This is because the barrel of the stereoscopic monitor and the circular polarized glasses attenuate the transmission of light from the monitor to the viewer's eyes. In addition, a higher reflectivity mirror may be necessary to increase the aperture without introducing excessive noise.

- Investigate the possibility of integrating a data glove, digital pointer, or space ball in order to interactively move an image to the appropriate location in space. Also, a head movement or eye ball movement tracker may also be integrated with the device. With these enhancements, physicians may have more flexibility in observing the images of the biomedical structures they are interested in.
- Modify the Kriging and neural network algorithms to improve the interpolation of biomedical data to create better 3D stereo images.
- Integrate the image processing procedure with the stereo image generating procedure. This may cut data processing time by eliminating tedious data manipulation procedures.
- Add grid, scale, index, or other reference data with the image display system to allow an image to be readily located, measured, compared, or sized. In addition, zoom in and zoom out functions may be implemented to further increase the analysis of the 3D stereo image.

Appendix A. Source code for get4d

```
/*
 * get4d.c - Put RLE images on the Iris/4D display under the window manager.
 *
 * Author: Russell D. Fish (Based on getmex.c .)
 * Computer Science Dept.
 * University of Utah
 * Date: Thu May 26 20:49:11 1988
 * Copyright (c) 1988, University of Utah
 *
 * Date of Modified: Aug 14 1994
 *
 * Reason: Put a pair of RLE images on the Iris/4D display under the window
 * manager.
 */

#include <stdio.h>
#include <math.h>
#include "gl.h"
#include "device.h"
#include "rle.h"
#ifdef USE_STDLIB_H
#include <stdlib.h>
#else

#ifdef USE_STRING_H
#include <string.h>
#else
#include <strings.h>
#endif

extern char *getenv ();
#ifdef VOID_STAR
```

```

extern char *malloc ();
#else
extern void *malloc();
#endif
extern void free();

#endif /* USE_STDLIB_H */

#define ADJUST_2ND_IMAGE 18

#define MAX(i,j)  ( (i) > (j) ? (i) : (j) )
#define MIN(i,j)  ( (i) < (j) ? (i) : (j) )

/* Global variables. */
long window_number; /* Window number from MAX. */
int x_size, y_size; /* Size of image. */
int dbg = 0; /* Set if debug mode. */
int forkflg = 0; /* Set if not to run in background. */
int bwflag = 0; /* Set for greyscale output. */
int NoBorder = 0; /* Turn off border */
int gt; /* Whether we're using GT fast rectangle drawing. */
double disp_gamma = 1.0; /* Gamma of Iris display. */
double img_gamma = 0.0; /* Gamma of input image. */

/* Window preferences. */
int posflag = 0, w_xpos, w_ypos, sizeflag = 0, w_xsize, w_ysize;

long *rect_image[2]; /* GT Image data buffer pointer. */
char *rgb_image[2]; /* Non-GT Image data buffer pointer. */

long *rev_rect_image_buffer;
char *rev_rgb_image_buffer;

rle_pixel **cmap;

```

```

/*****
* TAG( main )
*
* Usage:
* get4d [-G][-S] [-p xpos ypos] [-f] [-D] [-w] [file]
* Inputs:
* -D: Debug mode: print input file as read.
* -f: Don't fork after putting image on screen.
* -G: GT mode: single fast rectangle panning disabled.
* (GT mode is the default on GT, GTX, and Personal Iris 4Ds.)
* -S: Slow mode: Allows resizing the window, and panning with the
* mouse. (Slow mode is the default on non-GT 4Ds.)
* -g disp_gamma: Gamma of the output display.
* -i image_gamma: Gamma the image was calculated for.
* -I image_gamma: Gamma of the image pixels (1.0 / (-i gamma)).
* -n: No border will be drawn.
* -p xpos ypos: Position of the lower left corner of the window.
* -s xsize ysize: Initial size of the window (slow mode only.)
*
* -w: Black & white: reduce color images to B&W before display.
* Advantage is that smoother shading can be achieved.
*
* file: Input Run Length Encoded file. Uses stdin if not
* specified.
* Outputs:
* Puts image in a window on the screen.
* Assumptions:
* Input file is in RLE format.
*/

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

```

```

char *iname[2];
FILE * infile[2];
char vbuff[50];
char *var;
int gtflag = 0, slowflag = 0, gflag = 0, iflag = 0;

iname[0] = NULL;
iname[1] = NULL;
infile[0] = stdin;
infile[1] = stdin;

/* Handle arguments. */
if ( scanargs( argc, argv,
"% D%- f%- GS%- g%-disp_gamma!F iI%-image_gamma!F n%- \n\
p%-xpos!dypos!d s%-xsize!ysize!d w%- file%s file%s",
&dbg, &forkflg,
&gtflag,
&gflag, &disp_gamma,
&iflag, &img_gamma,
&NoBorder,
&posflag, &w_xpos, &w_ypos,
&sizeflag, &w_xsize, &w_ysize,
&bwflag,
&iname[0],
&iname[1] ) == 0 )
exit( 1 );

if ( gtflag ) /* Mode args override if specified. */
gt = gtflag - 1; /* -G->2 means yes GT, -S->1 means no GT. */
else
{
#ifdef _IBMR2
/* See if we're on a GT. For this purpose, a Personal Iris is a GT. */
gversion( vbuff ); /* String like "GLADGTX-3.1". */

```

```

gt = strcmp( vbuff+4, "GT", 2 ) == 0 ||
    strcmp( vbuff+4, "PI", 2 ) == 0;
#else
gt = 1;
#endif
    }

    /* Diddle with gammas. */
    if ( iflag == 2 ) /* -i flag. */
if ( img_gamma != 0.0 ) /* Paranoid. */
    img_gamma = 1.0 / img_gamma;

#ifdef _IBMR2
    /* Look at home for .gamma file. */
    if ( !gflag )
    {
        char *var = getenv( "HOME" );
        char buf[BUFSIZ];
        FILE *gamfile;
        float flt_gam = 0; /* Just in case of 68000 iris botch. */

        if ( var != NULL && *var != '\0' )
        {
            sprintf( buf, "%s/.gamma", var );
            if ( (gamfile = fopen( buf, "r" )) != NULL )
            {
                fscanf( gamfile, "%f", &flt_gam );
                if ( flt_gam != 0 )
                    disp_gamma = 2.4/flt_gam; /* SGI displays have gamma 2.4. */
                fclose( gamfile );
            }
        }
    }
#endif

```

```

        infile[0] = rle_open_f("get4d", infname[0], "r");
        infile[1] = rle_open_f("get4d", infname[1], "r");
        get_pic( infile, infname );
/* Read image and make a window for it. */
        update_pic(); /* Keep drawing the window. */
    }

/*
    * Read an image from the input file and display it.
    */
get_pic( infile, infname )
FILE * infile[2];
char * infname[2];
{
    register int i, y;
    int loop_var;
    int ncolors;
    unsigned char *scan[3];

    for (loop_var = 0; loop_var < 2; loop_var ++ )
    {
        /*
            * Read setup info from file.
            */
        rle_dflt_hdr.rle_file = infile[loop_var];
        if ( rle_get_setup( &rle_dflt_hdr ) < 0 )
        {
            fprintf(stderr, "get4d: Error reading setup information from %s\n",
                infname[loop_var] ? infname[loop_var] : "stdin");
            exit(1);
        }

        if ( dbg )

```

```

rle_debug( 1 );

/* We're only interested in R, G, & B */
RLE_CLR_BIT(rle_dflt_hdr, RLE_ALPHA);
for (i = 3; i < rle_dflt_hdr.ncolors; i++)
RLE_CLR_BIT(rle_dflt_hdr, i);
ncolors = rle_dflt_hdr.ncolors > 3 ? 3 : rle_dflt_hdr.ncolors;

/* Do nicer b&w rendering if only one color channel in input. */
if ( ncolors == 1 && rle_dflt_hdr.ncmap <= 1 )
bwflag = 1;

/*
 * Compute image size and allocate storage for colormapped image.
 */
x_size = (rle_dflt_hdr.xmax - rle_dflt_hdr.xmin + 1);
y_size = (rle_dflt_hdr.ymax - rle_dflt_hdr.ymin + 1);
if ( gt )
{
rect_image[loop_var] = (long *) malloc(x_size * y_size * sizeof( long ));
    rev_rect_image_buffer = (long *) malloc(x_size * sizeof( long ));
}
else
{
rgb_image[loop_var] = (char *) malloc(x_size * y_size * 3 * sizeof( char ));
    rev_rgb_image_buffer = (char *) malloc(x_size * 3 * sizeof( char ));
}
/*
 * Set up for rle_getrow. Pretend image x origin is 0.
 */
for (i = 0; i < 3; i++)
scan[i] = (unsigned char *) malloc(x_size);
rle_dflt_hdr.xmax -= rle_dflt_hdr.xmin;
rle_dflt_hdr.xmin = 0;

```



```

cmap = buildmap( &rle_dflt_hdr, 3, img_gamma, disp_gamma );

/* For each scan line, pack RGBs into the image memory. */
while ((y = rle_getrow(&rle_dflt_hdr, scan)) <= rle_dflt_hdr.ymax)
{
switch ( ncolors )
{
case 1:
    /* Colormapped image */
    for (i = 0; i < x_size; i++) {
scan[2][i] = cmap[2][scan[0][i]];
scan[1][i] = cmap[1][scan[0][i]];
scan[0][i] = cmap[0][scan[0][i]];
    }
    break;

case 2:
    /* Weird image. */
    for (i = 0; i < x_size; i++) {
scan[2][i] = cmap[2][scan[1][i]];
scan[1][i] = cmap[1][scan[1][i]];
scan[0][i] = cmap[0][scan[0][i]];
    }
    break;

case 3:
    /* Normal image. */
    for (i = 0; i < x_size; i++) {
scan[2][i] = cmap[2][scan[2][i]];
scan[1][i] = cmap[1][scan[1][i]];
scan[0][i] = cmap[0][scan[0][i]];
    }
    break;
}
}

```

```

if ( bwflag )
{
    if (ncolors == 1)
    rgb_to_bw (scan[0], scan[1], scan[2], scan[0], x_size );
    else
        rgb_to_bw( scan[0], scan[1], scan[ncolors - 1],
            scan[0], x_size );
    /* Note: pack_scanline only uses channel 0 for B&W */
}

if ( gt )
    pack_rect_scanline( scan, x_size,
        &rect_image[loop_var][(y - rle_dflt_hdr.ymin) * x_size] );
else
    pack_rgb_scanline( scan, x_size,
        &rgb_image[loop_var][(y - rle_dflt_hdr.ymin) * x_size * 3] );
}
}

/*
 * Free temp storage
 */
for (i = 0; i < 3; i++)
free(scan[i]);

/* Size flag can cut down the window size, except in GT rectangle mode. */
/*
w_xsize = sizeflag && !gt ? MIN( w_xsize, x_size ) : x_size;
w_ysize = sizeflag && !gt ? MIN( w_ysize, y_size ) : y_size;

/* Window should be completely on the screen in GT rectangle mode. */
/*
if ( posflag && gt )
{
w_xpos = MIN( w_xpos, IMAISCREEN - w_xsize );

```

```

w_ypos = MIN( w_ypos, YMAXSCREEN - w_ysize);
}

w_xpos = MAX( 0, w_xpos ); /* Be positive. */
/*
w_ypos = MAX( 0, w_ypos );
*/

prefposition( 0, 1280, 0, 1024);

/* Register our preferences for the window size and location. */
if ( posflag ) {
    printf ("About to issue prefposition with %d %d %d %d\n",
        w_xpos, w_xpos + w_xsize - 1, w_ypos, w_ypos + 2*w_ysize - 1);

    prefposition( w_xpos, w_xpos + w_xsize - 1,
        w_ypos, w_ypos + 2*w_ysize - 1 );
} else if ( sizeflag || gt ) {
/* Have to keep the window full size to use the fast pixel commands. */
    printf ("About to issue prefsize with %d %d\n", w_xsize, 2*w_ysize );
    prefsize( w_xsize, 2*w_ysize );
} else {
    printf ("About to issue maxsize with %d %d\n", x_size, y_size);
    maxsize( x_size, y_size );
}

/*
* Get a window of the right size (user positions it with the mouse).
*/

#ifdef _IBMR2
    if ( forkflg ) foreground(); /* Don't fork. */
#endif

if ( NoBorder ) noborder();
window_number = winopen( "get4d" );

```

```

    if ( infname[0] ) wintitle( infname[0] );

    /* Loosen the constraints once the window is created. */
    /*
    if ( !gt )
maxsize( x_size, y_size );
    winconstraints();
    */
    RGBmode();
    gconfig();

    qdevice( ESCKEY );
    qdevice( REDRAW );
    unqdevice( INPUTCHANGE ); /* We don't pay attention to these. */

    if ( !gt ) /* Disable panning when in gt rectangle mode. */
    {
qdevice( LEFTMOUSE ); /* Pan the image under mouse control. */
qdevice( LEFTALTKEY ); /* Reset panning. */
qdevice( RIGHTALTKEY );
qdevice( F9KEY );
    }

    /* There was a redraw event sent when the window was created,
    * but we weren't listening for them yet.
    */
    qenter( REDRAW, window_number );
}

/*
* Track events & redraw image when necessary.
*/
update_pic()
{

```

```

short data;
long event;
int window_x_size, window_y_size, window_x_origin, window_y_origin;
int x_min, x_max, y_min, y_max, x_start, y_start, x_end, y_end, x_len;
int x_origin, y_origin, new_x_center, new_y_center;
int x_center, y_center, saved_x_center, saved_y_center;

register int y,x;
register char *y_ptr;

/* Looking at the center, at first. */
x_center = saved_x_center = x_size / 2;
y_center = saved_y_center = y_size / 2;

/* Redraw the window when necessary. */
while ( TRUE )
{
    event = qread( &data );
#ifdef DEBUG
    printf( "event %d, data %d\n", event, data );
#endif
    switch ( event )
    {
        case ESCKEY:
            gexit();
            exit(0);
            break;
        case REDRAW:
            winset( window_number );
            reshapeviewport();

            if ( gt )
            {
                /* On a GT, just blast out the whole rectangle.  If the

```

```

    * origin is off the screen to the left, it kills the
    * window server. Could duplicate the slow output logic
    * below, but avoid the complication for now...
    */
    getorigin( &x_origin, &y_origin );
    if ( x_origin < 0 )
    {
        RGBcolor( 0, 0, 0 ); /* Punt. */
        clear();
    }
    else
        lrectwrite( 0, 0, x_size-1, y_size-1, rect_image );
}
else
{
    /* Do panning in a resizable window. (Slow mode.) */
    RGBcolor( 0, 0, 0 );
    clear();

    /* Lower left corner of screen, in image coordinates.
     * (Keep the center of the image in the center of the
     * window.)
     */
    getsize( &window_x_size, &window_y_size );
    x_min = (window_x_size/2) - x_center ;
    x_max = x_min + (x_size-1);
    y_min = (window_y_size/4) - y_center;
    y_max = y_min + (y_size-1);

    /* Coordinate bounds have half a pixel added all around. */
    /* ortho2( x_min - .5, x_max + .5, y_min - .5, y_max + .5 ); */

    /* Draw just the part of the image in the window. */
    x_start = MAX( x_min, 0 );
    x_end = MIN( x_max, (window_x_size-1) );

```

```

    y_start = MAX( y_min, 0 );
    y_end = MIN( y_max, ((window_y_size/2)-1) );
    x_len = x_end - x_start + 1;

    /* Dump the scanlines. Check once in a while for another
     * redraw event queued up, and quit early if one is seen.
     */
    y_ptr = rgb_image[0]; /*debug + y_start*x_size*3; */

    for ( y = y_start;
    y <= y_end && (y%16 != 0 || qtest() != REDRAW);
    y++, y_ptr += x_size * 3 )
    {
        for (x = 0; x < x_size; x++)
        {
            /* reverse image left-to-right */
            rev_rgb_image_buffer[x] = y_ptr[x_size - x];
            rev_rgb_image_buffer[x + x_size] = y_ptr[(2 * x_size) - x];
            rev_rgb_image_buffer[x + (2 * x_size)] = y_ptr[(3 * x_size) - x];
        }
        cmov2i( x_start, y_end - y + y_start); /* using (y_end - y) so that picture is upside-down */
        writeRGB( x_len,
        rev_rgb_image_buffer, rev_rgb_image_buffer + x_size, rev_rgb_image_buffer + x_size * 2 );
    }

    y_ptr = rgb_image[1]; /*debug + y_start*x_size*3;*/

    y_start = y_start + ( window_y_size / 2 ) + ADJUST_2ND_IMAGE;
    y_end = y_end + ( window_y_size / 2 ) + ADJUST_2ND_IMAGE;

    for ( y = y_start;
    y <= y_end && (y%16 != 0 || qtest() != REDRAW);
    y++, y_ptr += x_size * 3 )
    {

```

```

        for (x = 0; x < x_size; x++)
        {
/* reverse image left-to-right */
            rev_rgb_image_buffer[x] = y_ptr[x_size - x];
            rev_rgb_image_buffer[x + x_size] = y_ptr[(2 * x_size) - x];
            rev_rgb_image_buffer[x + (2 * x_size)] = y_ptr[(3 * x_size) - x];
        }

cmov2i( x_start, y_end - y + y_start ); /* Using (y_end - y) so that picture is upside-down */
/* Must add y_start to remain in upper */
/* half of the (non-stereo) screen */
writeRGB( x_len,
    rev_rgb_image_buffer, rev_rgb_image_buffer + x_size, rev_rgb_image_buffer + x_size * 2 );
    }

}

break;

/* Alt key - Reset viewing to look at the center of the image.
 * Shift-Alt - Restores a saved view center.
 * Control-alt - Saves the current view center for Shift-Setup.
 * F9 is the same as the Alt keys.
 */
case LEFTALTKEY:
case RIGHTALTKEY:
case F9KEY:
if ( data == 1 ) /* Ignore button up events. */
{
    if ( getbutton(RIGHTSHIFTKEY) || getbutton(LEFTSHIFTKEY) )
    {
x_center = saved_x_center; /* Restore. */
y_center = saved_y_center;
qenter( REDRAW, window_number );
    }
    else if ( getbutton(CTRLKEY) )

```



```

    {
        saved_x_center = x_center; /* Save. */
        saved_y_center = y_center;
    }
    else
    {
        x_center = x_size / 2; /* Reset. */
        y_center = y_size / 2;
        qenter( REDRAW, window_number );
    }

}

break;

/* Pan a point picked with the left mouse button to the center
 * of attention. Beep if cursor is not on the image.
 */
case LEFTMOUSE:
if ( data == 1 ) /* Ignore button up events. */
{
    getorigin( &x_origin, &y_origin );
    new_x_center = getvaluator( MOUSEX ) - x_origin + x_min;
    new_y_center = getvaluator( MOUSEY ) - y_origin + y_min;
    if ( new_x_center >= x_start &&
        new_x_center <= x_end &&
        new_y_center >= y_start &&
        new_y_center <= y_end )
    {
        x_center = new_x_center;
        y_center = new_y_center;
        qenter( REDRAW, window_number );
    }
    else
        ringbell();
}

```

```

}
break;
}
    }
}

/*
 * Pack a scanline into a vector of RGB longs.
 *
 * Inputs:
 *  rgb: Pointers to buffers containing the red, green,
 *  and blue color rows.
 *  n: Length of row.
 *  line: Pointer to output buffer for packed color data.
 */
pack_rect_scanline( rgb, n, line )
unsigned char *rgb[3];
int n;
long *line;
{
    register int i;
    register long *dest = line;

    if ( !bwflag ) /* Color display. */
    {
register unsigned char *r, *g, *b;

for ( i = 0, r = rgb[0], g = rgb[1], b = rgb[2];
      i < n; i++, r++, g++, b++ )
    /* Combine 3 8-bit colors into a long. */
    *dest++ = *r + (*g<<8) + (*b<<16);
    }
    else /* Gray scale display. */
    {

```

```

register unsigned char *bw;

for ( i = 0, bw = rgb[0]; i < n; i++, bw++ )
    *dest++ = *bw + (*bw<<8) + (*bw<<16);
}

}

/*
 * Pack a scanline into an RGB trio of vectors of bytes.
 *
 * Inputs:
 *  rgb: Pointers to buffers containing the red, green,
 *  and blue color rows.
 *  n: Length of row.
 *  line: Pointer to output buffer for packed color data.
 */
pack_rgb_scanline( rgb, n, lines )
unsigned char *rgb[3];
int n;
char *lines;
{
    int chnl;
    register int i;
    register unsigned char *src, *dest;

    for ( chnl = 0, dest = lines; chnl < 3; chnl++ )
    {
src = rgb[ bwflag ? 0 : chnl ]; /* Use just channel 0 for greyscale. */
        for ( i = 0; i < n; i++ )
            *dest++ = *src++;
    }
}

```

Appendix B. Source code for stereo-on mode

```
#include <gl.h>
#include "display.h"

/*****
/* Program:  stereo-on                                */
/* Author:   Capt Bob Caley                           */
/* Date:     23 November 1993                          */
/* System:   SGI VGXT/ONYX                            */
/*
/* Purpose:  This programs turns on SGI VGXT/ONYX stereo mode */
/*
*****/
main()
{
Display display;

display.stereo();
gexit();
}
```

Appendix C. Source code for stereo-off mode

```
#include <gl.h>
#include "display.h"

/*****
/* Program:  stereo-off                                */
/* Author:   Capt Bob Caley                            */
/* Date:     23 November 1993                          */
/* System:   SGI VGXT/ONYX                             */
/*                                                  */
/* Purpose:  This programs turns on SGI VGXT/ONYX standard 60HZ mode */
/*                                                  */
*****/
main()
{
    Display display;

    display.standard();
    gexit();
}
```

Bibliography

1. Kantowitz Barry H. and Sorkin Robert D. "Human Factors : Understanding People-System Relationships", *John Wiley & Sons Inc.*, 1983.
2. Gabor T. Herman et al, "Display of 3D Digital Images : Computational Foundations and Medical Applications", *IEEE Computer Graphics & Applications*, 39-46, Vol. 3, No. 5, Aug 1983.
3. Ken Pimentel and Kevin Teixeira "Virtual Reality : Through the new looking glass", *Windcrest Books*, 194-207, 1993.
4. Steve Aukstakalnis and David Blatner, "Silicon Mirage : The Art and Science of Virtual Reality", *Peachpit Press Inc.*, 1992.
5. Cover Steven A. et al. "Interactively Deformable Models foe Surgery Simulation", *IEEE Computer Graphic & Applications*, 68-75, Nov. 1993.
6. Udupa Jayaram K. and Herman Gabor T. (editors) "3D Imaging in Medicine", *CRC Press Inc.*, 1991.
7. T. Todd Elvins "A Survey of Algorithms for Volume Visualization", *Computer Graphics*, 194-201, Vol. 26, No. 3, Aug 1992.
8. Drebin Robert A., Carpenter Loren, Hanrahan Pat "Volume Rendering", *Computer Graphics*, 65-74, Vol. 22, No. 4, Aug 1988.
9. Ulf Tiede et al. "Investigation of Medical 3D-Rendering Algorithms", *IEEE Computer Graphics & Applications*, 41-53, Mar 1990.
10. Christian Barillot "Surface and Volume Rendering Techniques To Display 3-D Data", *IEEE Engineering in Medicine and Biology*, 111-119, Mar 1993.
11. Parrott Rob W. "Evaluation of Scalar Value Estimation techniques For 3D Medical Imaging", MS thesis, AFIT/GCS/ENG/91D-17, School of Engineering, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, Dec 1991.
12. Patrick Joseph Rizzuto Jr. "Three-Dimensional Medical Imaging Registration Using A Patient Space Correlation technique", MS thesis, AFIT/GCS/ENG/91D-18, School of Engineering, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, Dec 1991.
13. Brandt James R. Capt., "Real Image Visual Display System", MS thesis, AFIT/GEO/ENG/92D-03, School of Engineering, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, Dec 1992.
14. Kaiser Mary K. "Perceptual issues in scientific visualization", *SPIE Three-Dimensional Visualization and Display Technologies*, 205-211, Vol. 1083, 1989.
15. Jose Encarnacao et al. "Research Issues in Perception and User Interfaces", *IEEE Computer Graphics and Applications*, 67-69, Mar 1994.

16. Spillmann Lothar and Werner John S. (editors) "Visual Perception : The Neu-rophysiological Foundations", *Academic Press Inc.*, 1990.
17. Clapp Robert E. "Stereoscopic displays and the human dual visual system", *SPIE Advances in Display Technology VI*, 41-52, Vol. 624, 1986.
18. Michael Mckenna & David Zeltzer, "Three Dimensional Visual Display Systems for Virtual Environments", *Presence*, 421-458, Vol. 1, No. 4, Fal 1992.
19. Wickens Christopher D. "Three-dimensional stereoscopic display implementa-tion : Guidelines derived from human visual capabilities", *SPIE Stereoscopic Displays and Applications*, 2-11, Vol. 1256, 1990.
20. Kalawsky Roy S. "The Science of Virtual Reality and Virtual Environments", *Addison-Wesley Publishing Company Inc.*, 1993.
21. Cornsweet Tom N. "Visual Perception", *Academic Press Inc.*, 1970.
22. Yeh Yei-Yu and Silverstein Louis D. "Limits of Fusion and Depth Judgement in Stereoscopic Color Displays", *Human Factors*, 45-60, 32(1), 1990.
23. Leibovic K. N. (editor) "Science of Vision", *Springer-Verlag New York Inc.*, 1990.
24. Clapp Robert E. "Stereoscopic Perception", *SPIE True 3D Imaging Techniques and Display Technologies*, 79-84, Vol. 761, 1987.
25. Clapp Robert E. "Observing Eye and Perceiving Brain Organization in Vision", *SPIE Display System Optics*, 2-8, Vol. 778, 1987.
26. Francis Crick and Christof Koch "The Problem of Consciousness", *Scientific American*, 153-159, sept 1992.
27. E. A. Hibbard, M. E. Bauer, M. S. Bradshaw, D. G. Deardorff, K. C. Hu, and D. J. Whitney "On the Theory and Application of Stereogramphics in Scientific Visualization", *NASA ARC and Sterling Software*
28. Michael Starks "Stereoscopic video and quest for virtual reality An Annotated Bibliography of Selected Topics-Part II", *SPIE Stereoscopic Displays and Ap-plications III*, 216-227, Vol. 1669, 1992.
29. Bruce Lane "Stereoscopic displays", *SPIE Processing and Display of Three-Dimensional Data*, 20-32, Vol. 367, 1982.
30. Ian Sexton and David Crawford "Parallax Barrier 3DTV", *SPIE Three-Dimensional Visualization and Display Technologies*, 84-94, Vol. 1083, 1989.
31. Lenny Lipton "Factors affecting "ghosting" in time-multiplexed plano-stereoscopic CRT display systems", *SPIE True 3D Imaging Techniques and Display Technologies*, 75-78, Vol. 761, 1987.

32. Eichenlaub Jesse B. "Progress in Autostereoscopic Display Technology at Dimension Technologies Inc.", *IEEE Stereoscopic Displays and Applications II*, 290-299, Vol. 1457, 1991.
33. Rodney Don Williams and Daniel Donohoo "Image Quality Metrics for Volumetric Laser Displays", *IEEE Stereoscopic Displays and Applications II*, 210-220, Vol. 1457, 1991.
34. Clifton III T.E. and Wefer Fred L. "Direct Volume Display Devices", *IEEE Computer Graphics & Applications*, 57-65, July 1993. Edited by Daniel Thalmann "Scientific Visualization and Graphics Simulation", *John Wiley & Sons Ltd.*, 1990.
35. Hobbs Bruce A. "A User Interface To A True 3-D Display Device", MS thesis, AFIT/GCE/ENG/92D-06, School of Engineering, Air Force Institute of Technology(AU), Wright-Patterson AFB OH, Dec 1992.
36. Michael Starks "Stereoscopic video and quest for virtual reality: An Annotated Bibliography of Selected Topics-Part II", *SPIE Stereoscopic Displays and Applications II*, 327-342, Vol. 1457, 1991.
37. McAllister David F. (editor) "Stereo Computer Graphics and Other True 3D Technologies", *Princeton University Press*, 1993.
38. Farrel Edward J. and Christidis Zaphiris D. "Visualization of Complex Data", *SPIE Three-Dimensional Visualization and Display Technologies*, 153-160, Vol. 1083, 1993.
39. Noel Cressie "The origins of Kriging", *Mathematical Geology*, Vol. 22, No. 3, 1990.
40. Noel Cressie "Geostatistics", *The American Statistician*, 197-202, Vol. 43, No. 4, Nov 1989.
41. Noel Cressie "Kriging Nonstationary Data", *Journal of the American Statistical Association*, 626-634, Sept 1986.
42. Jesus Carrera and Ferenc Szidarovszky "Numerical Comparison of Network Design Algorithms for Regionalized Variables", *Applied Mathematics and Computation*, 16:189-202, 1985.
43. Journel Andre G. "Fundamentals of Geostatistics in Five Lessons", *American Geophysical Union*, Washington D.C., 1989.
44. Lynda Kerbs "GEO-Statistics: The Variogram", *COGS Computer Contributions*, 54-59, Vol. 12, No. 2, Aug 1986.
45. Donald P. Duckett, Jr. "The Application of Statistical Estimation Techniques to Terrain Modeling", MS thesis, AFIT/GCE/ENG/91D-02, School of Engineering, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, Dec 1991.

46. Noel Cressie "Spatial Prediction and Ordinary Kriging", *Mathematical Geology*, 405-421, Vol. 20, No. 4, 1988.
47. Ferenc Szidarovszky "Multiobjective observation network design for regionalized variables", *International Journal of Mining Engineering*, 331-342, 1 1983.
48. Yfantis E.A. and AU M. "An Image Compression Algorithm Based on Kriging", *SPIE Image and Video Processing*, 215-227, Vol. 1903, 1993.
49. Defense Advanced Research Projects Agency (DARPA) "DARPA Neural Network Study", *AFCEA International Press*, Nov 1988.
50. Rogers Steven K., Matthew Kabrisky, Tarr Gregory L. and Ruck Dennis W. "An Introduction to Biomedical and Artificial Neural Networks",
51. Lippmann Richard P. "An Introduction to Computing with Neural Nets", *IEEE Acoustics, Speech, and Signal Processing*, 4-22, Apr 1987.
52. Teuvo Kohonen "An Introduction to Neural Computing", *Neural Networks*, 3-16, Vol. 1, 1988.
53. Rogers Steven K. "Introduction to Biological and Artificial Networks for Industrial and commercial Applications", *SPIE - The International Society for Optical Engineering*, Sept 1993.
54. Lehr Michael A. and Widrow Bernard "Commercial and Industrial Applications of Neural Networks", *WNN'93*, 165-170, Nov 7-10 1993.
55. Kunihiro Fukushima "A Neural Network for Visual pattern Recognition", *IEEE COMPUTER*, 65-75, VOL. 21, NO. 3, March 1988.
56. Thomas Y. P. Lee and Clark C. Guest "A Binocular Interaction Neural network Model Using Parallel Channel Architecture", *SPIE Stereoscopic Displays and Applications IV*, 107-111, Vol. 1915, 1993.
57. Pepe Siy and Joe-E Hu "Edge-pixel-based stereo correspondence through ordering oriented neural networks", *SPIE Stereoscopic Displays and Applications IV*, 123-127, Vol. 1915, 1993.
58. Dennis William Ruck "Characterization of Multilayer Perceptrons and their Application to Multisensor Automatic Target Detection", PhD Dissertation, AFIT/DS/ENG/90-2, School of Engineering, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, Feb 1990.
59. Ozkan Mehmed, Dawant Benoit M. and Maciunas Robert J. "Neural-Network-Based Segmentation of Multi-Modal Medical Images : A Comparative and Prospective Study", *IEEE Transactions on Medical Imaging*, 534-544, Vol. 12, No. 3, Sep 1993.
60. Amartur S.C., Piraino D. and Takefuji Y. "Optimization Neural Networks for the Segmentation of Megnetic Resonance", *IEEE Transactions on Medical Imaging*, 215-220, Vol. 11, No. 2, June 1992.

61. Marilyn McCord Nelson and W. T. Illingworth "A Practical Guide to Neural Nets", *Addison-Wesley Publishing Company*, 1991.
62. Arndt Graig M. "Applications of Neural Networks to Landmark Detection in 3-Dimensional Surface Data", *SPIE Applications of Artificial Neural Networks III*, 880-886, Vol. 1709, 1992.
63. Tang Yonghong, Wee William G. and Han Chia Yung "Application of A Multi-layer Network in Image Object Classification", *SPIE Applications of Artificial Neural Networks II*, 113-120, Vol. 1469, 1991.
64. Bebis George N. et al. "Increasing Classification Accuracy Using Multiple Neural Network Schemes", *SPIE Applications of Artificial Neural Networks III*, 221-231, Vol. 1709, 1992.
65. Broomhead D.S. and Lowe D. "Multivariable functional interpolation and adaptive networks", *Complex System*, 321-355, Vol. 2, 1988.
66. Cichocki A. and Unbehauen R. "Neural Networks for Optimization and Signal Processing", *John Wiley & Sons Ltd.*, 1993.
67. Park J. and Sandberg I.W. "Universal Approximation Using Radial-Basis-Function Networks", *Neural Computation*, 246-257, Vol. 3, 1991.
68. Park Jooyoung and Sandberg Irwin W. "Approximation and Radial-Basis-Function Networks", *Neural Computation*, 305-316, Vol. 5, 1993.
69. Chen S. Cowan C.F.N. and Grant P.M. "Orthogonal Least Squares Learning Algorithm for Radial Basis Function Networks", *IEEE Transaction on Neural Networks*, 302-309, Vol. 2, No. 2, March 1991.
70. Bishop Chris "Improving the Generalization Properties of Radial Basis Function Neural Networks", *Neural Computation*, 579-588, Vol. 3, 1991.
71. Mak M.W., Allen W.G., and Sexton G.G. "Speaker Identification Using Multilayer Perceptrons and Radial Basis Function Networks", *Neurocomputing*, 6, 99-117, 1994.
72. Chen S. Gibson G.J. Cowan C.F.N. and Grant P.M. "Reconstruction of binary signals using an adaptive radial-basis function equalizer", *Signal Processing*, 77-93, Vol. 22, 1991.
73. Whitehead Bruce A. and Choate Timothy D. "Evolving Space-Filling Curves to Distribute Radial Basis Functions Over an Input Space", *IEEE Transaction on Neural Networks*, Vol. 5, No. 1, January 1994.
74. Lau Clifford (editor) "Neural Networks Theoretical Foundations and Analysis", *IEEE Press*, 1992.
75. Muller B. and Reinhardt J. "Neural Networks: An Introduction", *Springer-Verlag Berlin Heidelberg*, 1990.

76. Mason J.C. and M.G. Cos (editors) "Algorithms for Approximation", *Oxford*, 143-167, 1987.
77. "Stereoscopic 3-Dimensional Graphics Displays SGS 620/625 User's Manual", *Liquid Crystal Imaging Display Products Tektronix, Inc.*, 1987.
78. Eugene Hecht "Optics", *Addison-Wesley Publishing Company, Inc.*, 1987.
79. Sirohi Rajpal S. and Kothiyal Mahendra P. "Optical Components, Systems, and Measurement Techniques", *Marcel Dekker, Inc.*, 1991.
80. Keating Michael P. "Geometric, Physical, and Visual Optics", *Butterworths Publisher*, 1988.
81. Kliger David S. et al. "Polarized light in Optics and Spectroscopy", *Academic Press, Inc.*, 1990.
82. Kolb Graig E. "Rayshade User's Guide and Reference Manual", July 12, 1991.
83. Thomas Spender W. "Design of the Utah RLE Format", *University of Utah, Department of Computer Science*.
84. Lenny Lipton et al. "Programmer's Guide: The Creation of Software for Stereoscopic Computer Graphics", *StereoGraphics Corporation*, 1988.
85. Peterson John W., Bogart Rod G., and Thomas Spender W. "The Utah Raster Toolkit", *University of Utah, Department of Computer Science, Salt Lake City, Utah*.
86. "Graphics library programming guide", *Graphics Silicon Inc.*, 1991.
87. Linda Kukolich and Richard Lippmann "LNKnet User's Guide", *MIT Lincoln Laboratory*, July 9, 1993.
88. Lippmann Richard P., Kukolich Linda, and Singer Elliot "LNKnet: Neural Network, Machine-Learning, and Statistical Software for Pattern Classification, *The Lincoln Laboratory Journal*, 249-268, Vol. 6, No. 2, 1993.
89. Linda Kukolich, LNKmap User Commands.
90. Lee Yuchun "Handwritten Digit Recognition Using K Nearest-Neighbor, Radial-Basis Function, and BackPropagation Neural Networks, *Neural Computation*, 440-449, Vol. 3, 1991.
91. Steve Renals and Richard Rohwer "Phoneme Classification Experiments Using Radial Basis Functions", *Proc. Internet, joint conf. on neural networks, IEEE Washington DC*, 461-467, June 1989.

Vita

Captain Kai-Heng Wang (Felix) was born on 14 August 1963 in Hualian, Taiwan, Republic of China. In 1979, he joined the Chinese military to maintain Taiwan as a free democratic China. After graduating from Chung Cheng Preparatory School (CCPS), he attended the Chinese Air Force Academy. After earning a commission as a First Lieutenant, he was assigned as a computer systems officer. In May 1992, he was selected to come to the United States of America to attend graduate school at the Air Force Institute of Technology, Wright-Patterson Air Force Base, Ohio. Captain Wang earned a Master of Science degree in Computer Engineering. After graduating in September 1994, he was assigned to the Chinese Air Force Academy.

Permanent address: 5F 86-1 Kuling st.
Taipei Taiwan, R.O.C.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)		2. REPORT DATE September 1994	3. REPORT TYPE AND DATES COVERED Master's Thesis	
4. TITLE AND SUBTITLE 3D Stereo Data Visualization and Representation			5. FUNDING NUMBERS	
6. AUTHOR(S) Kai-Heng Wang (Felix)				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Air Force Institute of Technology, WPAFB OH 45433-6583			8. PERFORMING ORGANIZATION REPORT NUMBER AFIT/GCE/ENG/94S-01	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)			10. SPONSORING / MONITORING AGENCY REPORT NUMBER	
11. SUPPLEMENTARY NOTES				
12a. DISTRIBUTION / AVAILABILITY STATEMENT Distribution Unlimited			12b. DISTRIBUTION CODE	
13. ABSTRACT (Maximum 200 words) This thesis investigates a technique for improving the perception and visualization of three dimensional information. This improvement will eventually give physicians the ability to both visualize and understand biomedical image data. Raw data is used to generate a stereo image pair. The raw data must be put in run length encoding (RLE) format and then converted to a stereo pair. Once the stereo pair is generated, they must be displayed on the Silicon Graphics ONYX workstation in stereo mode. This research uses 3D images of planets and satellites. The display information is transmitted across cable and is projected onto a spherical mirror, a real 3D stereo image can be observed. Biomedical data of interest to this research are slices of a volume of tissue. To convert the slice information into volume data a representation that allows meaningful interpolation between the slices for stereo image pair generation is necessary. A second thrust of this thesis is to investigate neural networks and Kriging algorithm for potential use in interpolation of the biomedical volume data.				
14. SUBJECT TERMS			15. NUMBER OF PAGES	
			16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT UNCLASSIFIED	18. SECURITY CLASSIFICATION OF THIS PAGE UNCLASSIFIED	19. SECURITY CLASSIFICATION OF ABSTRACT UNCLASSIFIED	20. LIMITATION OF ABSTRACT UL	

GENERAL INSTRUCTIONS FOR COMPLETING SF 298

The Report Documentation Page (RDP) is used in announcing and cataloging reports. It is important that this information be consistent with the rest of the report, particularly the cover and title page. Instructions for filling in each block of the form follow. It is important to *stay within the lines* to meet optical scanning requirements.

Block 1. Agency Use Only (Leave blank).

Block 2. Report Date. Full publication date including day, month, and year, if available (e.g. 1 Jan 88). Must cite at least the year.

Block 3. Type of Report and Dates Covered. State whether report is interim, final, etc. If applicable, enter inclusive report dates (e.g. 10 Jun 87 - 30 Jun 88).

Block 4. Title and Subtitle. A title is taken from the part of the report that provides the most meaningful and complete information. When a report is prepared in more than one volume, repeat the primary title, add volume number, and include subtitle for the specific volume. On classified documents enter the title classification in parentheses.

Block 5. Funding Numbers. To include contract and grant numbers; may include program element number(s), project number(s), task number(s), and work unit number(s). Use the following labels:

C - Contract	PR - Project
G - Grant	TA - Task
PE - Program Element	WU - Work Unit Accession No.

Block 6. Author(s). Name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. If editor or compiler, this should follow the name(s).

Block 7. Performing Organization Name(s) and Address(es). Self-explanatory.

Block 8. Performing Organization Report Number. Enter the unique alphanumeric report number(s) assigned by the organization performing the report.

Block 9. Sponsoring/Monitoring Agency Name(s) and Address(es). Self-explanatory.

Block 10. Sponsoring/Monitoring Agency Report Number. (If known)

Block 11. Supplementary Notes. Enter information not included elsewhere such as: Prepared in cooperation with...; Trans. of...; To be published in.... When a report is revised, include a statement whether the new report supersedes or supplements the older report.

Block 12a. Distribution/Availability Statement. Denotes public availability or limitations. Cite any availability to the public. Enter additional limitations or special markings in all capitals (e.g. NOFORN, REL, ITAR).

DOD - See DoDD 5230.24, "Distribution Statements on Technical Documents."

DOE - See authorities.

NASA - See Handbook NHB 2200.2.

NTIS - Leave blank.

Block 12b. Distribution Code.

DOD - Leave blank.

DOE - Enter DOE distribution categories from the Standard Distribution for Unclassified Scientific and Technical Reports.

NASA - Leave blank.

NTIS - Leave blank.

Block 13. Abstract. Include a brief (Maximum 200 words) factual summary of the most significant information contained in the report.

Block 14. Subject Terms. Keywords or phrases identifying major subjects in the report.

Block 15. Number of Pages. Enter the total number of pages.

Block 16. Price Code. Enter appropriate price code (NTIS only).

Blocks 17. - 19. Security Classifications. Self-explanatory. Enter U.S. Security Classification in accordance with U.S. Security Regulations (i.e., UNCLASSIFIED). If form contains classified information, stamp classification on the top and bottom of the page.

Block 20. Limitation of Abstract. This block must be completed to assign a limitation to the abstract. Enter either UL (unlimited) or SAR (same as report). An entry in this block is necessary if the abstract is to be limited. If blank, the abstract is assumed to be unlimited.