

**The Active Wave-front Sampling based 3D Endoscope**

by

Hemanth Prakash

B.Tech., Mechanical Engineering  
Indian Institute of Technology Madras, 2004

Submitted to the Department of Mechanical Engineering and the  
Department of Electrical Engineering and Computer Science  
In Partial Fulfillment of the Requirements for the Degrees of

Master of Science in Mechanical Engineering


and

Master of Science in Electrical Engineering and Computer Science


at the  
Massachusetts Institute of Technology  
February 2007

© 2007 Massachusetts Institute of Technology  
All rights reserved

Signature of Author.....

 Department of Mechanical Engineering  
January 14, 2007

Certified by.....

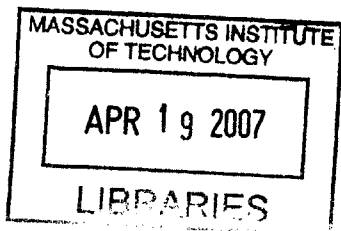
  
Professor Douglas P. Hart  
Professor of Mechanical Engineering  
Thesis Supervisor

Accepted by.....

Professor Lallit Anand  
Chairman, Department Committee on Graduate Students  
Department of Mechanical Engineering

Accepted by.....

  
Professor Arthur C Smith  
Chairman, Department Committee on Graduate Students  
Department of Electrical Engineering and Computer Science



**BARKER**



Room 14-0551  
77 Massachusetts Avenue  
Cambridge, MA 02139  
Ph: 617.253.2800  
Email: [docs@mit.edu](mailto:docs@mit.edu)  
<http://libraries.mit.edu/docs>

## **DISCLAIMER OF QUALITY**

Pagination issues by author, all content is accounted for.

# The Active Wave-front Sampling based 3D Endoscope

by

HEMANTH PRAKASH

Submitted to the  
Department of Mechanical Engineering and the  
Department of Electrical Engineering and Computer Science  
on January 14, 2007, in partial fulfillment of the  
requirements for the degrees of  
Master of Science in Mechanical Engineering  
and  
Master of Science in Electrical Engineering and Computer Science

## **ABSTRACT**

This thesis investigates the potential of Active Wave-front Sampling (AWS) for real time quantified 3D endoscopy.

AWS is a technique by which phase information from an aperture area of a lens is obtained by sampling sub-regions of the aperture area in time. The resulting phase information is then processed to obtain an accurate 3D measure of an imaged target object. While AWS has multitudinous applications and can take on many forms, this thesis examines the possibility of using an off-the-shelf monoscope and a single camera in conjunction with a rotating off-axis aperture to generate quantified real time three dimensional surface models of a surgical scenario. This AWS based stereo-endoscope (referred to as the AWS-scope) is shown to possess optical attributes that are superior to those of the current industry standard. Quantified 3D data generated using this AWS-scope is used to create enhanced synthetic stereo image pairs that exhibit superior depth perception characteristics comparable to the current industry standard.

## **Acknowledgements**

I extend my sincere thanks to Leslie Regan without whose wonderful assistance I would've never made it to MIT, Prof. Douglas Hart who imparted me with a skill or two besides helping me envision my research and of course my best friend, my Mother, without whom I would've neither entered nor endured this life.



## TABLE OF CONTENTS

<b>CHAPTER 1: INTRODUCTION.....</b>	<b>4</b>
1.1 Introduction.....	4
1.2 History of Endoscopy.....	4
1.3 Previous Research on Stereo-endoscopy.....	7
1.4 Motivation for Current Study.....	10
1.5 Roadmap for In-Depth Conceptualization of AWS based stereoscope.....	11
<b>CHAPTER 2: CONCEPTUALIZING THE AWS-SCOPE.....</b>	<b>13</b>
2.1 Active Wave-front Sampling (AWS).....	13
2.1.1 Fundamental Principle.....	13
2.2 Implementing AWS on a Monoscope to Obtain a Stereoscope.....	15
2.2.1 Paraxial Geometric Optical Model of the AWS-scope.....	16
2.3 Multiple Sample AWS.....	19
2.4 Performance of the AWS-scope.....	21
2.4.1 Effect of Aperture Radius.....	22
2.4.2 Depth of Target Object and Position of In-focus Plane.....	23
2.4.3 Field of View and Disparity Sensitivity.....	24
2.5 Comparison of AWS to Standard Stereo Imaging.....	27
<b>CHAPTER 3: REALIZING THE AWS-SCOPE.....</b>	<b>32</b>
3.1 Stages in the Realization Process.....	32
3.2 Components of the AWS-scope system implemented at MIT.....	40
3.3 Comparison of AWS-scope Disparity Levels: Lab Implementation vs. Parametric Model.....	44
<b>CHAPTER 4: COMPARISON OF THE AWS-SCOPE WITH THE INDUSTRY STANDARD.....</b>	<b>47</b>
4.1 Optical Characterization Tests.....	47
4.1.1 Field of View.....	47
4.1.2 Optical Resolution.....	49
4.1.3 Optical Transmission.....	57

4.1.4 Depth of Focus .....	60
4.2 Image Distortion.....	62
4.3 Vignetting in Images.....	65
<b>CHAPTER 5: QUANTIFIED 3D IMAGING AND THE PERCEPTION</b>	
<b>OF DEPTH BY THE HUMAN EYE.....</b>	<b>70</b>
6.1 Depth Perception.....	71
6.1.1 Pictorial Depth Cues .....	72
6.1.2 Oculomotor and Binocular Depth Cues.....	73
6.1.3 Interplay between depth cues .....	75
6.2 Convergence and Vision.....	77
6.3 Display of stereo images for effective binocular vision.....	83
6.3.1 Parallax.....	83
6.3.2 Zero Parallax .....	84
6.3.3 Positive Parallax.....	84
6.3.4 Negative Parallax.....	85
6.3.5 Divergent Parallax.....	86
6.4 Design Guidelines for Synthetic Stereo Generation.....	89
<b>CHAPTER 6: GENERATING SYNTHETIC STEREO WITH THE</b>	
<b>AWS-SCOPE.....</b>	<b>90</b>
6.1 Generating Stereo Views .....	91
6.2 Comparison of Synthetic Stereo and Raw Stereo.....	97
6.3 Mapping Color onto Synthetic Stereo .....	98
<b>CHAPTER 7: SUMMARY, CONCLUSION AND FUTURE WORK</b>	<b>102</b>
<b>REFERENCES.....</b>	<b>109</b>
<b>APPENDIX A: CAMERA CALIBRATION PARAMETERS.....</b>	<b>114</b>
<b>APPENDIX B: GENERATING SYNTHETIC STEREO FROM</b>	
<b>SURFACE MODEL .....</b>	<b>116</b>
<b>APPENDIX C: CORRELATING LEFT AND RIGHT IMAGES FOR</b>	
<b>MAPPING COLOR.....</b>	<b>119</b>

## LIST OF FIGURES

1.1 Precursors to the modern day endoscope .....	5
1.2 Modern day monoscope .....	6
1.3 Stereo-laparoscope .....	7
1.4 Split pupil stereo-endoscope .....	8
1.5 Commercial stereo-endoscope .....	9
2.1 Depth from defocus .....	14
2.2 Implementing AWS on a Monoscope to Obtain a Stereoscope .....	15
2.3 Paraxial ray trace for the AWS-scope .....	17
2.4 Plot of Normalized disparity vs. Normalized Target Distance .....	20
2.6 Plot of Disparity Sensitivity vs. Aperture Offset Radius .....	22
2.7 Plot of Disparity Sensitivity vs. Target Distance .....	23
2.8 Plot of Disparity Sensitivity vs. System Field of View .....	24
2.9 Disparity Sensitivity for the Constant Focal Distance and Constant FOV cases .....	26
2.10 Canonical Stereo Imaging System .....	27
2.11 Disparity Sensitivity for the Canonical and AWS systems .....	29
3.1 AWS-scope schematic in practice .....	34
3.2 AWS-scope Ray Trace .....	35
3.3 AWS-scope: Evaluation Points on CCD .....	36
3.4 AWS-scope: Polychromatic Diffraction MTF .....	37
3.5 AWS-scope: Spot Diagram .....	38
3.6 AWS-scope set-up at MIT .....	39
3.7 Olympus Monoscope .....	40
3.8 AWS module .....	41
3.9 AWS-scope Imaging Optics .....	42
3.10 Dalsa CCD Sensor .....	43
3.11 Comparing AWS-scope set-up to Theoretical Model .....	45
3.12 Comparing AWS-scope set-up to Theoretical Model .....	45
3.13 Comparing AWS-scope set-up to Theoretical Model .....	46
4.1 AWS-scope schematic for the Field of View Test .....	48
4.2 Field of View of the AWS-scope .....	49
4.3 AWS-scope schematic for the Optical Line Resolution Test .....	50

4.4 ISO 12233 Chart for Line Resolution Testing .....	50
4.5 Horizontal Line Resolution of AWS-scope .....	51
4.6 Vertical Line Resolution of AWS-scope .....	52
4.7 ISO 12233 Chart for MTF Testing .....	54
4.8 MTF profile for Schoelly in Horizontal Direction.....	55
4.9 MTF profile for AWS-scope in Horizontal Direction .....	55
4.10 MTF profile for Schoelly in Vertical Direction.....	56
4.11 MTF profile for AWS-scope in Vertical Direction.....	56
4.12 AWS-scope schematic for the Optical Transmission Test .....	58
4.13 Transmission Ratio curves for the AWS-scope and the Schoelly .....	59
4.14 AWS-scope schematic for the Depth of Focus Test .....	61
4.15 Optical Line Resolution vs. Target Depth for AWS-scope .....	61
4.16 Calibration Target imaged in various Orientations.....	63
4.17 Projecting calibrated points onto Target using Tsai Calibration Model .....	63
4.18 Re-projection Error from the Calibration Model .....	64
4.19 Re-producing Input Calibration Data for the Various Orientations.....	64
4.20 Raw image captured using the AWS-scope.....	65
4.21 Intensity Fall-off plot for the raw image captured using the AWS-scope .....	66
4.22 Improving the quality of the raw image's Intensity Profile.....	67
4.23 Correcting a Stereo pair for Intensity Variations.....	69
5.1 Stereo pairs captured using the Schoelly and the AWS-scope .....	70
5.2 The various Monocular Depth Cues .....	73
5.3 Convergence for the Human Eye.....	74
5.4 Accommodation for the Human Eye .....	75
5.5 Regimes of depth cues at various distances for the Human Eye .....	76
5.6 Schematic of the Geometrical Model of Depth Perception of the Human Eye .....	78
5.7 Plot of Convergence vs. Disparity Sensitivity to Convergence.....	79
5.8 Plot of Convergence vs. Disparity Sensitivity to Depth .....	82
5.8 Plot of Convergence vs. Disparity Sensitivity to Depth .....	82
5.9 Positive Convergent Parallax .....	83
5.10 Zero Parallax .....	84
5.11 Positive Parallel Parallax .....	84

5.12 Negative Parallax .....	85
5.13 Divergent Parallax .....	86
5.14 Effect of Scaling on Parallax .....	87
5.15 Parallax Changing with Baseline of Imaging .....	88
6.1 Coral Scene used for demonstrating Synthetic Stereo .....	91
6.2 Controlling Disparity Sensitivity .....	93
6.3 a-f Synthetic Stereo of the Coral Scenario at various Disparity Sensitivities....	94-96
6.4 Comparison of Synthetic Stereo and Raw Stereo imaged identically .....	97
6.5 Mapping Color onto the Synthetic Stereo pair .....	99
6.6 Comparison of the color mapped Synthetic Stereo and Raw Stereo .....	100
7.1 Super-resolution using the AWS technique .....	105
7.2 Comparison of a Monoscope and a Stereo-scope .....	107
7.3 An Artist Rendering of the AWS-scope .....	108

# CHAPTER 1: INTRODUCTION

## 1.1 Introduction

Endoscopy is the examination and inspection of the interior of body organs, joints or cavities through an endoscope; a device that uses fiber optics and powerful lens systems to provide lighting and visualization of the internal viscera. Endoscopy's greatest benefit is that it reduces the need to perform open-surgery. In addition, by providing a powerful channel for visualization, it allows for closer examination of internal abnormalities such as cysts, biliary (liver) cirrhosis, bleeding, bronchitis, cancer, degenerative disease, gallbladder stones, hernia, inflammation, metastatic cancer, polyps, tumors, ulcers, and other diseases and conditions. Endoscopy has also revolutionized surgery by opening the doors to the extraordinarily convenient concepts of out-patient and ambulatory surgery. Consequently, reductions in after-surgery care and observation, lowered risks of infection and shorter patient recovery times have been some of its indirect benefits [44].

## 1.2 History of Endoscopy

The history of endoscopy [44, 45] can be traced back to the early 1800s when attempts were first made to look inside the body using lighted telescopes. In 1805 Philip Bozzini made the first attempt to observe the living human body directly through a tube (figure 1.1) he created known as a Lichtleiter (light guiding instrument) to examine the urinary tract, rectum and pharynx. In 1853, Antoine Jean Desormeaux of France developed an

instrument (figure 1.1) specially designed to examine the urinary tract and the bladder. He named it "endoscope," and it was the first time this term was used in history.

Since those crude attempts, the field of endoscopy has matured tremendously aided in no small measure by developments such as the invention of the microprocessor, fibreoptics, photography, sufflation etc. Figure 1.2 shows a modern day monoscopic system complete with accessories such as HDTV display, fibre-optic light and imaging channels and other electronic displays.



*Figure 1.1: Pre-cursors to the modern day endoscope*

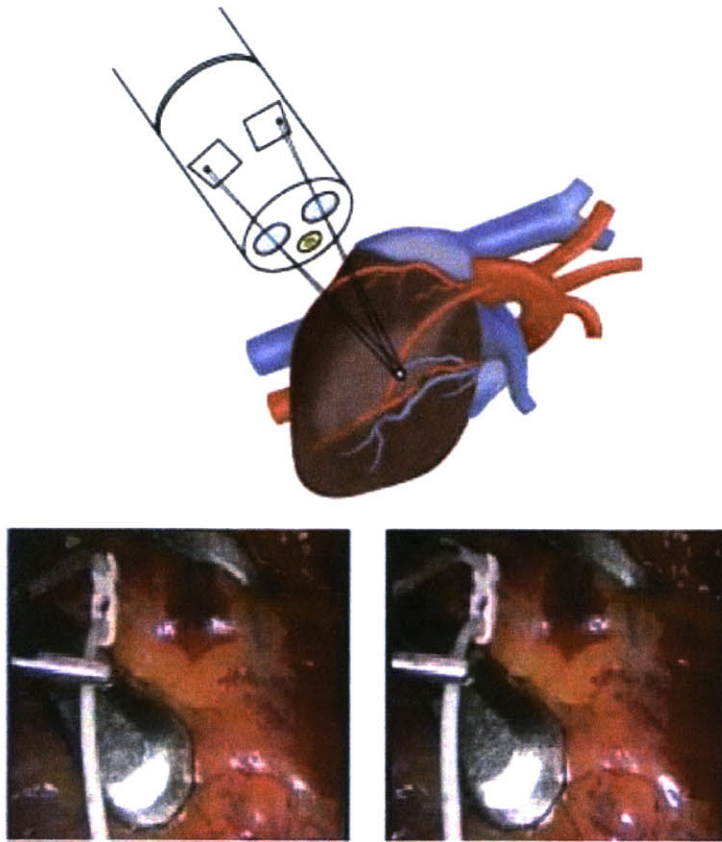
However, up until the late 1980s even the best endoscope systems (equipped with sophisticated optics and electronics such as high resolution HDTV display systems) had the great limitation of being monocular i.e. doctors had to perform surgery looking at images from a single channel. In doing so, they had to use subtle cues such as shading and their a-priori knowledge of the surgical scene to extract the sensation of depth from these two dimensional projections. The probability of error on the part of surgeons while performing delicate surgical procedures was invariably large in monoscopic surgery.



*Figure 1.2: A modern day monoscope based endoscopic system complete with HDTV display, fiber-optic light delivery mechanism and a fiber optic data transmission system for the captured images*

It was only in the 1990s that the concept of stereo-endoscopy took shape. Stereo-endoscopes provided surgeons with binocular vision inside the body. For the first time, surgeons actually perceived depth as they would in the real world while operating on the viscera (figure 1.3). A natural consequence of this technological advancement was a greater level of accuracy in incision, cauterization and other surgical techniques.



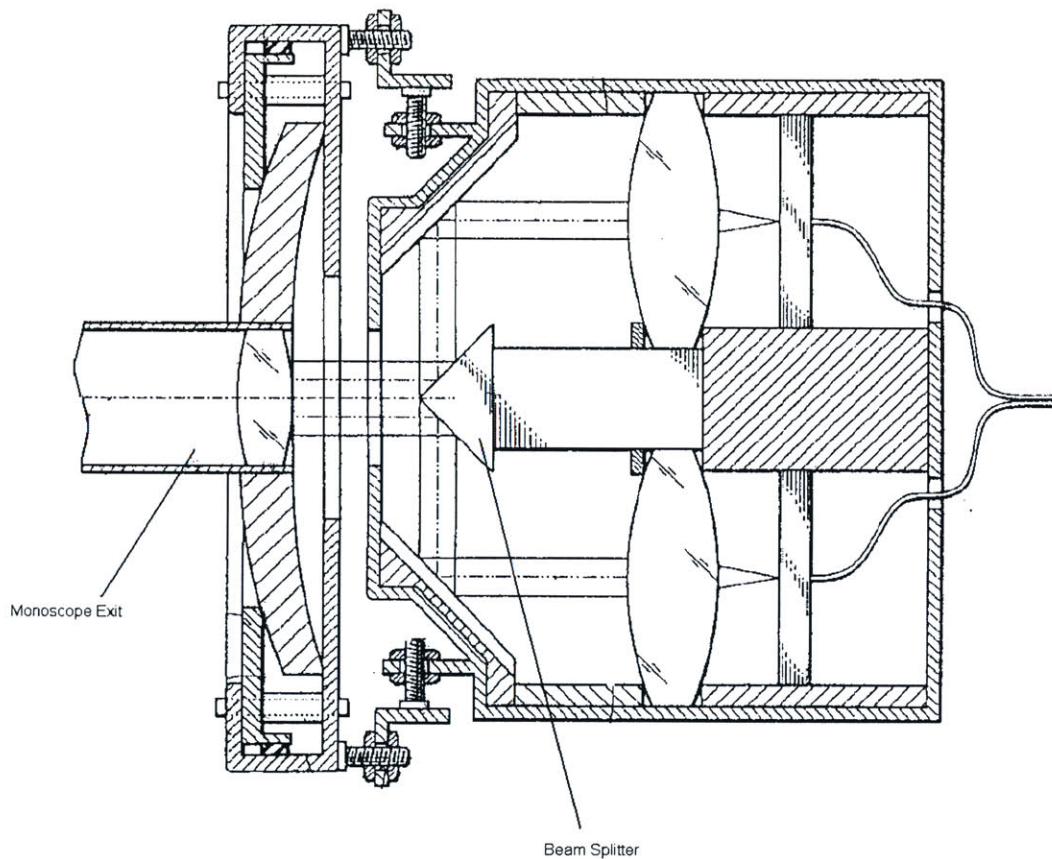


*Figure 1.3: Schematic of a stereo-laparoscope observing the heart and in vivo stereo pair obtained from such an observation. [46]*

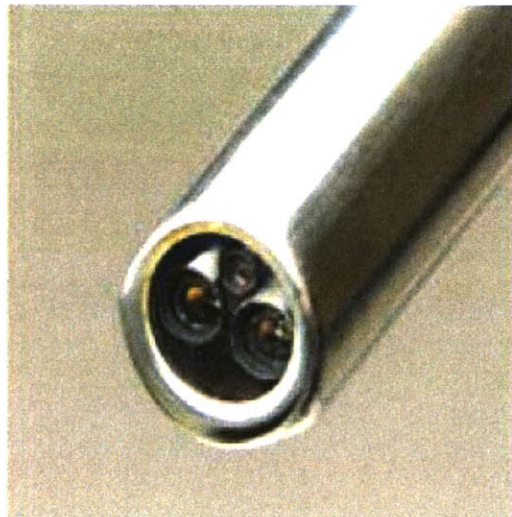
### **1.3 Previous Research on Stereo-endoscopy**

Stereo-endoscopes that have been developed so far have enabled three-dimensional perception either by using two adjacent optical channels fabricated into a single endoscope unit (figure 1.4) [1, 19] or by dividing the pupil of a single optical channel endoscope into left and right segments usually along a horizontal line [21, 34, 36, 23, 4, 33]. Depending on the implementation, the latter allowed some control over the magnitude of the created parallax displacement. More elaborate versions of pupil dividing stereoscopic endoscopes (figure 1.5) have allowed for changing the disparity direction as well [33, 26, 5]. The primary objective of these three-dimensional endoscopes has been the generation of stereo parallax for direct viewing, displaying or

endoscopes has been the generation of stereo parallax for direct viewing, displaying or recording of three-dimensional images. These stereoscopes have traditionally had problems of increased diameter of the endoscope so as to accommodate multiple optical channels and ensure adequate light levels. There have been cost issues associated with them as well, since they are based on custom made non-standard endoscope channel designs.



**Figure 1.4** Cross-section of a split pupil stereo-endoscope by Sekiya et al. The endoscope consists of a monoscope coupled to a module which splits the monoscope output into two optical trains in order to generate stereo.



*Figure 1.5* Cross-section of a typical stereo-endoscope. The endoscope consists of two optical channels represented by the left and the right cameras and a smaller channel in the center which represents the fiber optic light source

The few quantitative surface imaging methods applied in the field of three-dimensional endoscopy currently rely either on fixed baseline stereo imaging as in the case of a bisecting or dual optical channel endoscope [26, 19], modulation of a projected structured pattern by a scene that requires two endoscopes or one with two optical channels [17, 18, 3], or shape from shading methods [27, 42, 28] that also use two optical paths. These three-dimensional imaging methods are limited either due to their static nature (e.g. fixed baseline stereo) or because of the special illumination/projection requirement (e.g. space encoded laser scanning) that makes simultaneous 3-D and texture measurement difficult. Shape from shading techniques suffer from highly specular and non-Lambertian surface characteristics in addition to problems created by the closeness of the illumination source to the object.

## 1.4 Motivation for Current Study

There is a chasm between available optical hardware solutions to generate stereopsis in endoscopic systems and current quantitative three-dimensional imaging methods. The applied techniques usually do not allow simultaneous access to both the stereo image pairs and the quantitative surface models. Problems related with space encoding requirements (laser stripe projection), processing speed, resolution, and hardware robustness do not allow real time synthetic generation of stereo image pairs; a priority in minimally invasive surgery. Besides these specific technical drawbacks, the major failure of all currently available three-dimensional surface imaging technologies is their limited flexibility. Indeed, current three-dimensional systems are specifically designed to operate under a strictly defined set of parameters and are not able to perform in other applications defined by different conditions. These drawbacks have severely limited the widespread commercial use of three-dimensional measurement information.

This thesis proposes a unique system design for three-dimensional imaging in endoscopes which shows promise to overcome several shortcomings of the current three-dimensional endoscopes besides providing additional features. The design which is based on the Active Wave-front Sampling (AWS) technique uses an off-the-shelf monoscope and a single camera for real time three-dimensional imaging. This monoscope based stereo endoscope possesses the key advantages of lower bore diameter as well as lower cost of construction. But most importantly, the AWS technique enables the generation of real-time quantified three-dimensional data of the surgical scenario which creates powerful possibilities for endoscopy in general. For the first time it would be possible to obtain super-resolution images of a surgical scenario, digitally alter camera to obtain better

perspectives of a scene, perform intra-operative anatomy and finally attain practical implementations of tele-surgery.

## **1.5 Roadmap for In-Depth Conceptualization of AWS based stereoscope**

With the brief introduction to the history and evolution of endoscopy, and the need for a better three dimensional endoscope (stereoscope) in this first chapter, a detailed examination of the most important aspects of such an endoscope called the AWS-scope can be discussed in the subsequent parts of this thesis. In Chapter 2, a description will be given of how the AWS concept can be implemented on a monoscope to convert it to a stereoscope. Chapter 3 will concentrate on describing the implementation of an AWS-scope in practice and comparing the performance of such a realization to the previously developed parametric model. In Chapter 4, a comparison of the fundamental attributes of the AWS-scope such as resolution, field of view, optical transmission and depth of focus with those of the benchmark Schoelly stereoscope will be presented. Besides, the issues of vignetting correction and radiometric calibration of the images will also be addressed in this chapter. Chapter 5 will starts-off comparing the AWS-scope and the Schoelly on the key attribute of disparity sensitivity – the reasons behind AWS-scope’s lower sensitivity level will be briefly discussed and synthetic stereo generation presented as a potential solution to this hurdle. In order to generate effective synthetic stereo pairs, key drivers in the perception of a stereo pair by the human visual system will be identified and analyzed, and general design rules outlined. Chapter 6 will utilize the design rules of Chapter 5 to create flexible stereo views of a coral scenario captured using the AWS-scope. Finally, the proof-of-concept is completed for the AWS-scope by mapping color

onto the synthetic pair so as to replicate a ground truth real pair of equal disparity sensitivity. Chapter 7 finally summarizes the proof-of-concept AWS-scope. Future courses of action in terms of developing a prototype of the stereo-endoscope are recommended.

## **CHAPTER 2: CONCEPTUALIZING THE AWS-SCOPE**

### **2.1 Active Wave-front Sampling (AWS)**

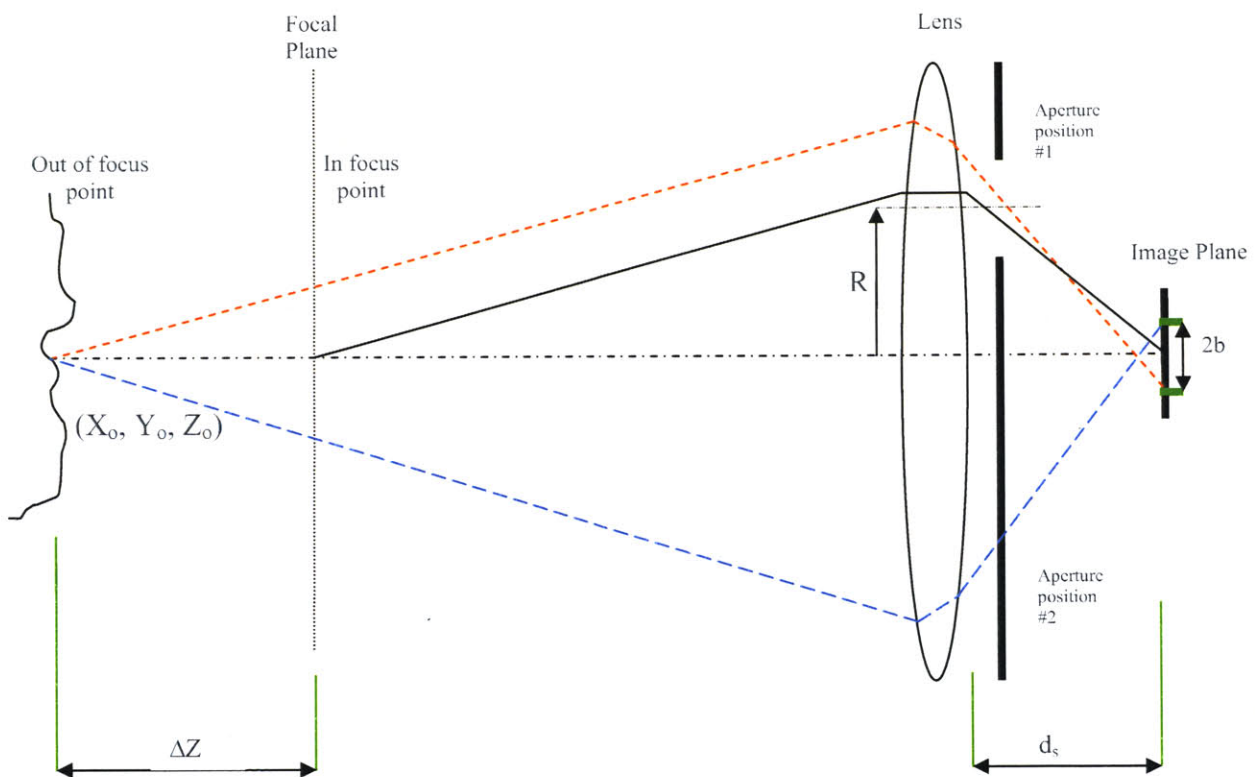
The active wave-front sampling (AWS) [6] approach to 3D surface imaging has several attractive attributes. Its most important characteristic is that, unlike stereoscopic systems, it requires only one optical path to capture depth information. This attribute reduces system costs by halving the cost of the optics, and also allows the technique to be applied to a wide range of currently available 2D imaging systems such as cameras, endoscopes, and microscopes. AWS also reduces system cost by dispensing with expensive laser based target illuminators, and minimizes target acquisition time by not having to slowly sweep an illumination sheet across the target. Instead, the AWS approach can calculate a 3D depth map with as little as two image frames (like a stereo system), thereby allowing real time (>30 Hz) operation. But perhaps the greatest benefit of AWS stems from its ability to generate high accuracy quantified three dimensional information and enable robustness in 3D processing when used in combination with a class of novel image processing algorithms developed at MIT.

#### **2.1.1 Fundamental Principle**

AWS involves the sampling of the wave-front of a light wave as it traverses the optical train so as to extract phase and intensity information from which full three dimensional information of the target being imaged is obtained. In its simplest form, AWS can be implemented using a single off-axis rotating aperture, a lens and a CCD sensor as shown



further away from the lens keeping the sensor at its original position, it is observed that the point is imaged off-axis. If the aperture were to be rotated by 180 degrees to position #2 and a similar ray trace performed, a second image is observed which is off-axis and symmetrically placed on the other side of the optical axis. The diameter of the circle formed by these two off-axis images  $d$  can be directly related to the change in depth  $\Delta z$  of the object point using simple geometry. In this manner, the AWS uses defocus to encode three-dimensional information into two dimensions.

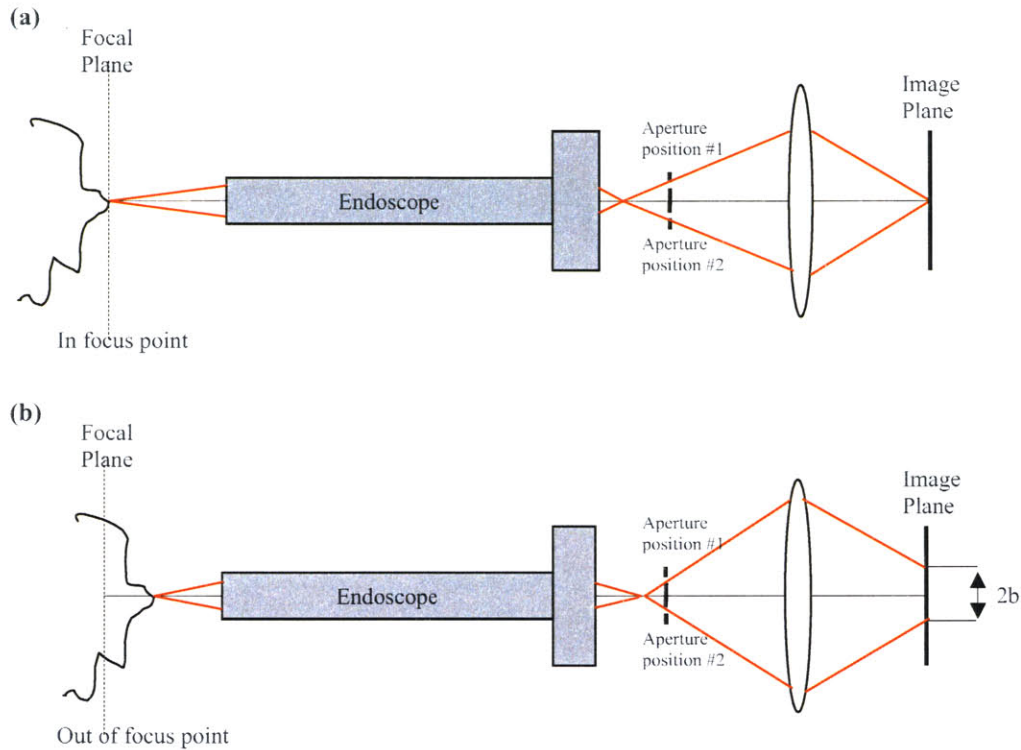


**Figure 2.1:** Schematic of depth-from-defocus approach. When the target feature of interest is located on the lens in-focus plane, the feature's image is perfectly in focus and the distance between the two off-axis images is zero. When the target feature of interest is located some distance from the lens in-focus plane the feature's image is out of focus, and the displacement between the two images relates directly to how far away the target is from the in-focus plane.



## 2.2 Implementing AWS on a Monoscope to Obtain a Stereoscope

Schematics of the proposed implementation of the AWS-scope are shown in figures 2.2a and 2.2b.



**Figure 2.2:** Off-axis rotating aperture codes three-dimensional depth information into two-dimensional disparity. Suppose the off-axis aperture position is rotated by  $180^\circ$  and two images at these two aperture positions are captured. **(a)** If the object point source is in the focal plane, the two image spots in the image plane completely overlap. **(b)** If the object is out-of-focus, there will be a disparity of ' $2b$ ' in the image plane between the two sampled spots. The depth position and disparity  $2b$  are uniquely related. If we take away the rotating aperture and leave the lens iris fully opened, there will be a defocused blur in the image plane with a diameter of  $2b$ . AWS uses this principle to acquire an entire field of three-dimensional measurement points with each image frame.

This optical train uses an additional focusing lens in order to focus the sampled light rays on to the sensor. The off-axis aperture is placed at the location of the exit pupil of the monoscope. The principle of operation is identical to that described in the section 2.1. Depth from the focal plane is determined by image disparity.

### 2.2.1 Paraxial Geometric Optical Model of the AWS-scope

A paraxial ray-trace model for the AWS-scope design proposed in the section 2.1 is shown in figure 2.3. Symmetry of the optical ray trace has been exploited to depict only the top half of the optical axis so that clarity is preserved in the schematic.

The constitutive relations for this model can be written from the Gaussian lens law [49].

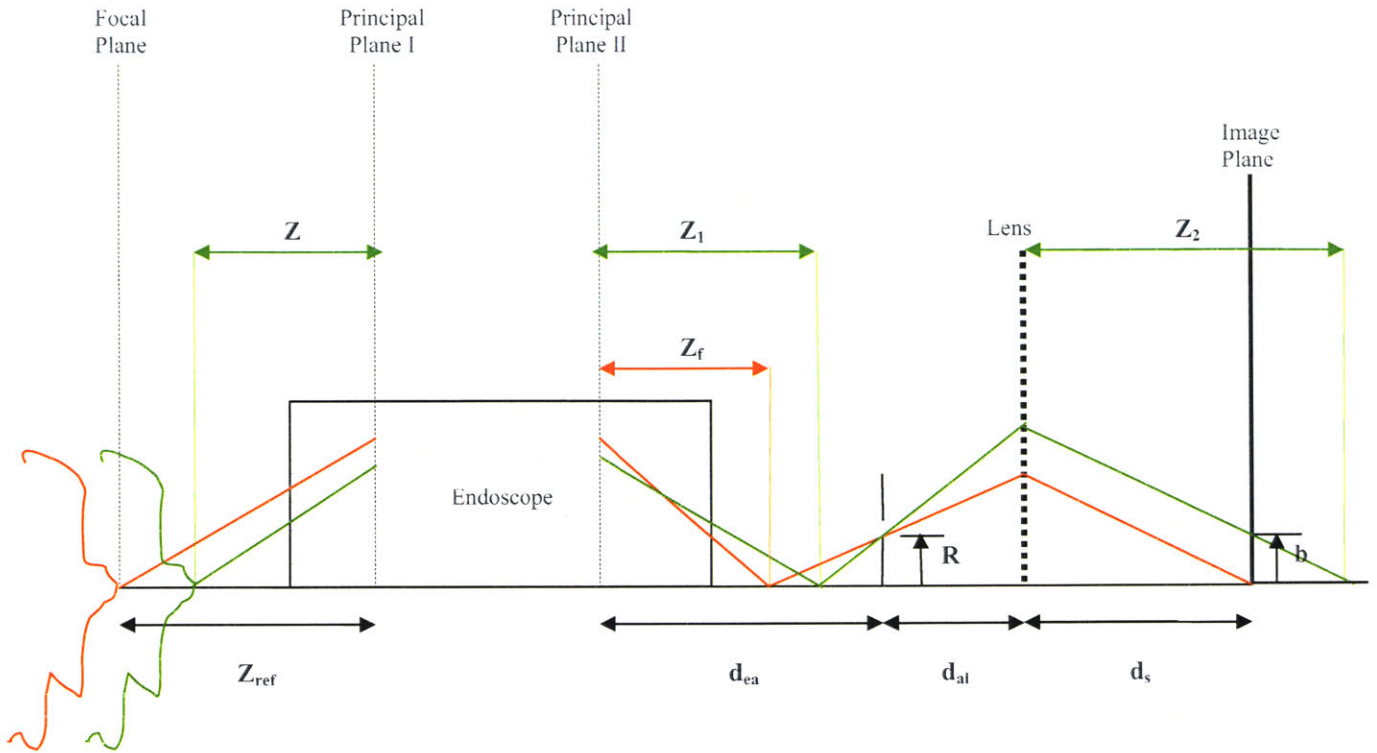
#### At the focal plane

$$\frac{1}{Z_{ref}} + \frac{1}{Z_f} = \frac{1}{EFL} \quad \frac{1}{(d_{ea}+d_{al}-Z_f)} + \frac{1}{d_s} = \frac{1}{f_L} \quad 2.1 \text{ (a) and (b)}$$

#### Away from the focal plane

$$\frac{1}{Z} + \frac{1}{Z_1} = \frac{1}{EFL} \quad \frac{1}{(d_{ea}+d_{al}-Z_1)} + \frac{1}{Z_2} = \frac{1}{f_L} \quad 2.2 \text{ (a) and (b)}$$

where EFL is the Effective focal length of the endoscope and  $f_L$  is the focal length of the focusing lens.



### 1. Endoscope related

i) In Focus condition (red):

Object distance:  $Z_{ref}$   
Image distance:  $Z_f$

ii) Out of Focus condition (green):

Object distance:  $Z$   
Image distance:  $Z_1$

All object distances measured from Principal Plane I of the endoscope and all Image distances measured from Principal Plane II of the endoscope.

### 2. Optical setup

$d_{ea}$ : Fixed distance from principal plane II of the endoscope to the aperture

$d_{al}$ : Fixed distance from aperture to focusing lens

$d_s$ : Fixed distance from focusing lens to the camera sensor

$R$ : Radius of offset of the aperture

**Figure 2. 3:** Paraxial ray trace for the AWS-scope system.

Note: Endoscope has been exaggerated for clarity

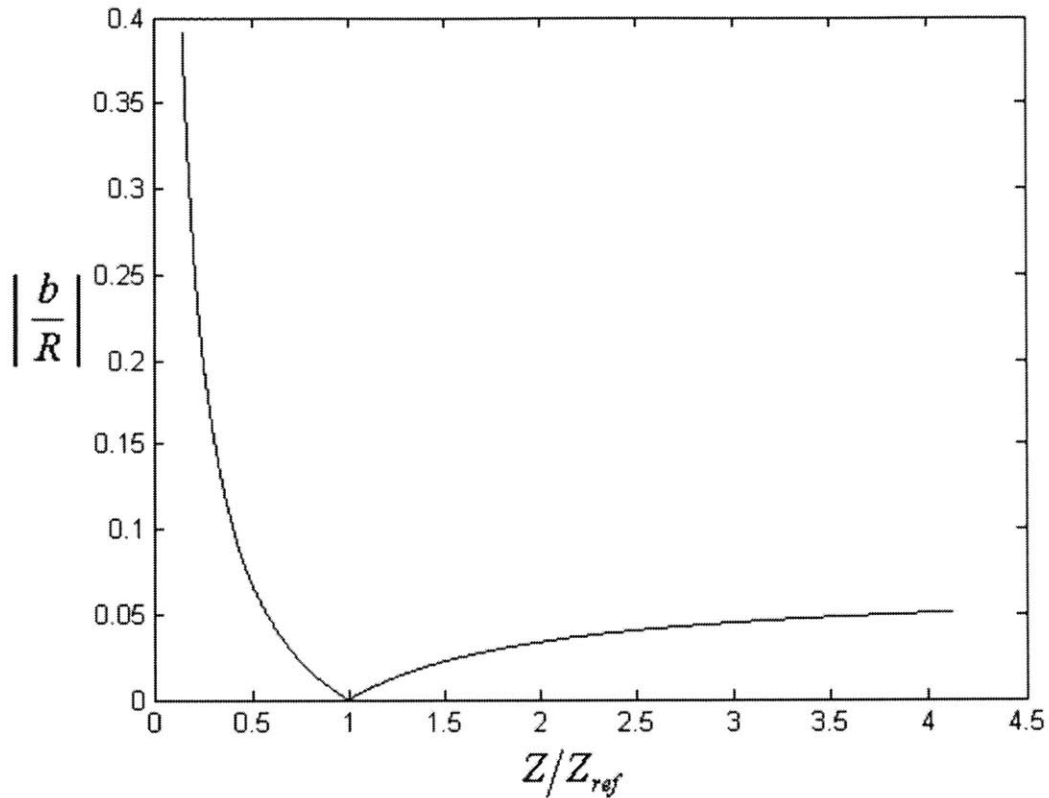
Using similar triangles, the disparity  $b$  can be written as

$$\frac{b}{R} = \left(1 - \frac{d_s}{Z_2}\right) \left(1 + \frac{d_{al}}{(d_{ea} - Z_1)}\right) \quad 2.3$$

$Z_1$  and  $Z_2$  depend on  $Z$  (Equations 2.1(a), 2.1(b), 2.2(a) and 2.2(b)).  $d_s$  can however vary depending on the distance of the in-focus plane  $Z_{ref}$ . Figure 2.4 shows a plot of the absolute value of normalized disparity ' $b/R$ ' vs. normalized target distance ' $Z/Z_{ref}$ '.

The following observations can be summarized from equation 2.3 and figure 2.4.

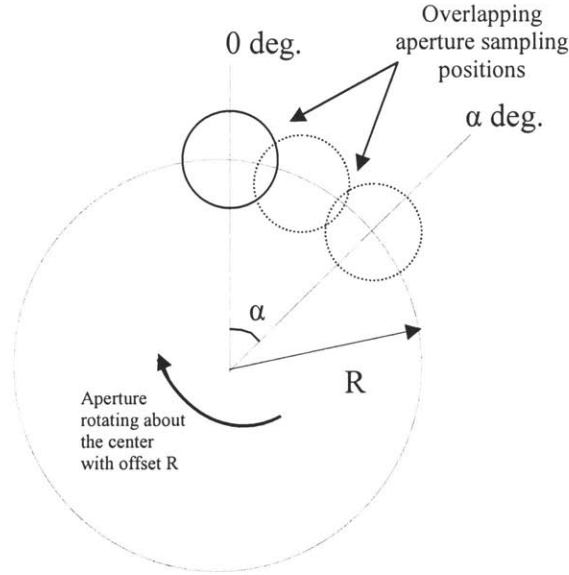
- For a specified value of  $Z_{ref}$ , disparity increases with decrease in target distance due to an increased sense of depth as objects are moved closer to the optics.
- For a specified target distance  $Z$  an increase in  $Z_{ref}$  increases the value of disparity. This is due to greater disparity for the same  $Z$  when the focal plane is farther away.
- The disparity ' $b$ ' can attain negative values. In fact disparity changes sign near the focal plane as the rays of light focus and cross-over to the opposite side of the optical axis at the focal plane. This phenomenon of sign change is a 180 degree change of phase for the disparity.



**Figure 2.4:** A typical plot of the absolute value of normalized disparity vs. normalized target distance for the AWS-scope system characterized by the parametric equations above. Normalized disparity expectedly goes to zero when the target distance is equal to the distance of the focal plane from the principal plane of the endoscope.

### 2.3 Multiple Sample AWS

One feature of the AWS technique is its ability to generate highly accurate three-dimensional information by allowing for super-sampling of the wave-front. Super-sampling is achieved by rotating the aperture through small and overlapping angular increments instead of just two positions 180 degrees apart. By sampling at several angular positions, a more accurate estimate of the diameter of the circle formed by the off-axis image disparity can be obtained which in turn results in highly improved estimates for the change in depth.



**Figure 2.5:** Front view of the AWS module showing the rotation of the aperture. Instead of obtaining two points on the disparity circle by sampling at points 180 degrees apart on the aperture circle, several sampling points are included at smaller angles so as to encode high accuracy and resolution depth information.

Figure 2.5 illustrates the concept of super-sampling. Consider two aperture sampling positions at an angle  $\alpha$  apart. Let the vertical position be the reference 0 degree position. If 'b' is the offset of the image when the aperture is at the zero degree position, then when the aperture is rotated by angle  $\alpha$ , the image offset would be 'b cos( $\alpha$ )'. Change in disparity as the image moves from zero to  $\alpha$  degree is shown in equation 2.4.

$$\Delta b = b(1 - \cos(\alpha)) = \left(1 - \frac{d_s}{Z_2}\right) \left(1 + \frac{d_{at}}{(d_{ea} - Z_1)}\right) (1 - \cos(\alpha)) \quad 2.4$$

Super-sampling enables robust data processing thereby making it possible to quantify low textured surfaces without projecting patterns or speckled light.

## 2.4 Performance of the AWS-scope

The idealized optical performance characteristics of the active wavefront sampling (AWS) approach are identical to those of the classic depth-from-defocus approach, as described quantitatively by Figure 2.4. Again, this figure shows that as a target feature moves farther away from the in-focus plane, the motion diameter of its image rotation caused by a circular aperture sampling pattern increases. This diameter increases more steeply if the target feature moves toward the lens than away from the lens. Also, it can be seen that if the target feature is moved away from the lens, the diameter asymptotically reaches a constant value. These characteristics suggest that the optimal operating regime for AWS based systems is defined by having the target located between the in-focus plane and the lens, where system sensitivity to depth is highest.

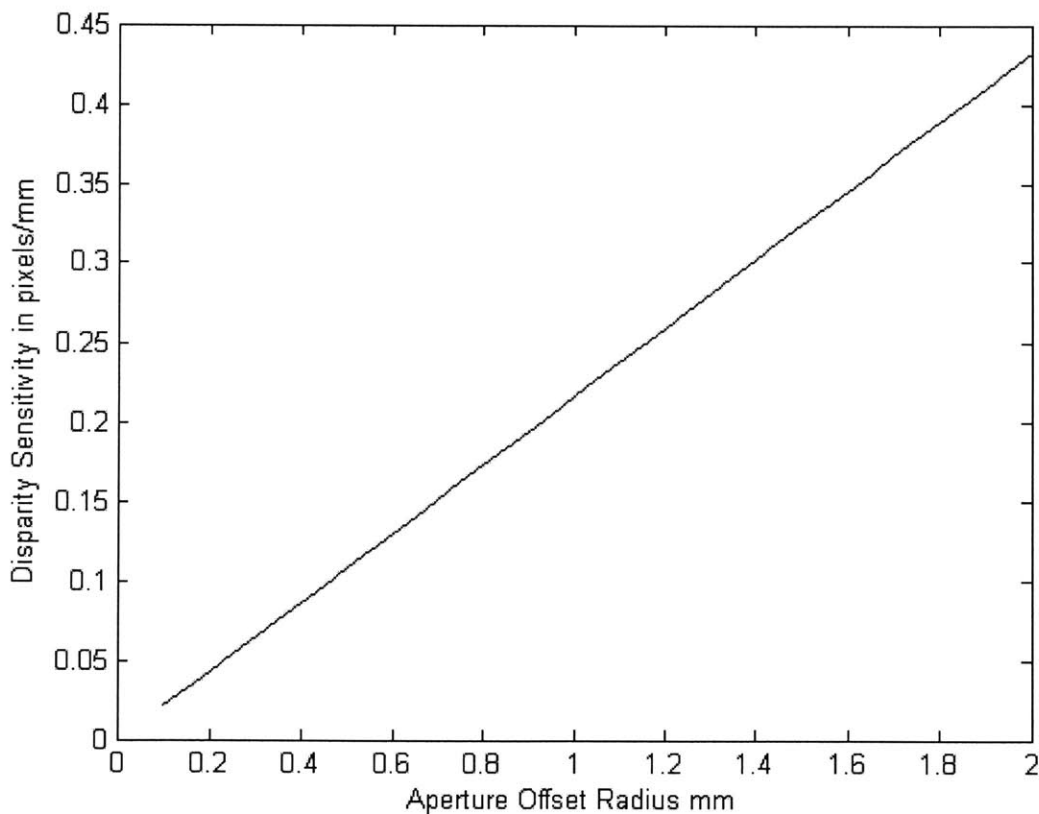
This ‘system sensitivity’ referred to here as ‘disparity sensitivity’ defines the disparity resolution of two closely spaced points separated by small  $\Delta Z$ . Mathematically, disparity sensitivity gives a measure of the change in disparity for an infinitesimal change in target depth. Equation 2.5 shows the mathematical expression for the disparity sensitivity of the system depicted by the ray-trace.

$$\frac{\partial \left( \frac{b}{R} \right)}{\partial Z} = \frac{\left( d_s \left( 1 + \left( \frac{d_{at}}{d_{ea} - \left( \sqrt{EFL} - \frac{1}{Z} \right)^{-1}} \right) \right) \right)}{\left( d_{at} + d_{ea} - \left( \sqrt{EFL} - \frac{1}{Z} \right)^{-1} \right)^2 \left( \sqrt{EFL} - \frac{1}{Z} \right)^2 Z^2} \left( \left( 1 - d_s \left( \frac{1}{f_L} - \frac{1}{\left( d_{at} + d_{ea} - \left( \sqrt{EFL} - \frac{1}{Z} \right)^{-1} \right)} \right) \right) d_{at} \right)}{\left( d_{ea} - \left( \sqrt{EFL} - \frac{1}{Z} \right)^{-1} \right)^2 \left( \sqrt{EFL} - \frac{1}{Z} \right)^2 Z^2} \quad 2.5$$

The effect of various imaging parameters on disparity sensitivity can be studied in greater detail as follows:

### 2.4.1 Effect of Aperture Radius

Increasing offset radius of the aperture 'R' linearly increases disparity sensitivity. Figure 2.6 illustrates this observation. However, there is a limit to increasing the aperture offset with limit being set by the size of the exit pupil of the endoscope as also the effect of vignetting and distortion when imaging from the lens periphery.

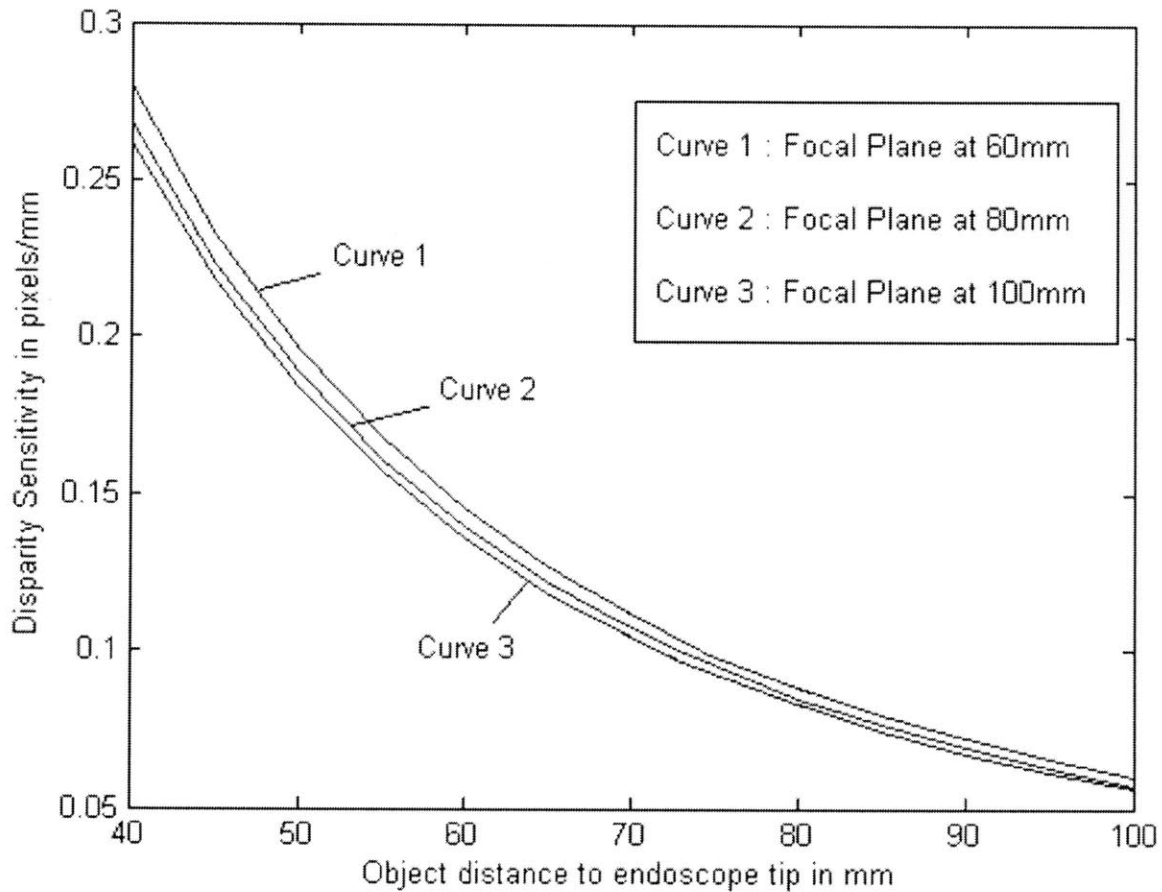


**Figure 2.6:** Simulated Disparity Sensitivity increases linearly with offset radius for the sampling aperture for an AWS coupled Olympus endoscope. For the purposes of simulation, the target object distance was 40mm and the in-focus plane was located at a distance of 100mm from the first principal plane of the endoscope. The field of view of the endoscope was about 80 deg.



## 2.4.2 Depth of Target Object and Position of In-focus Plane

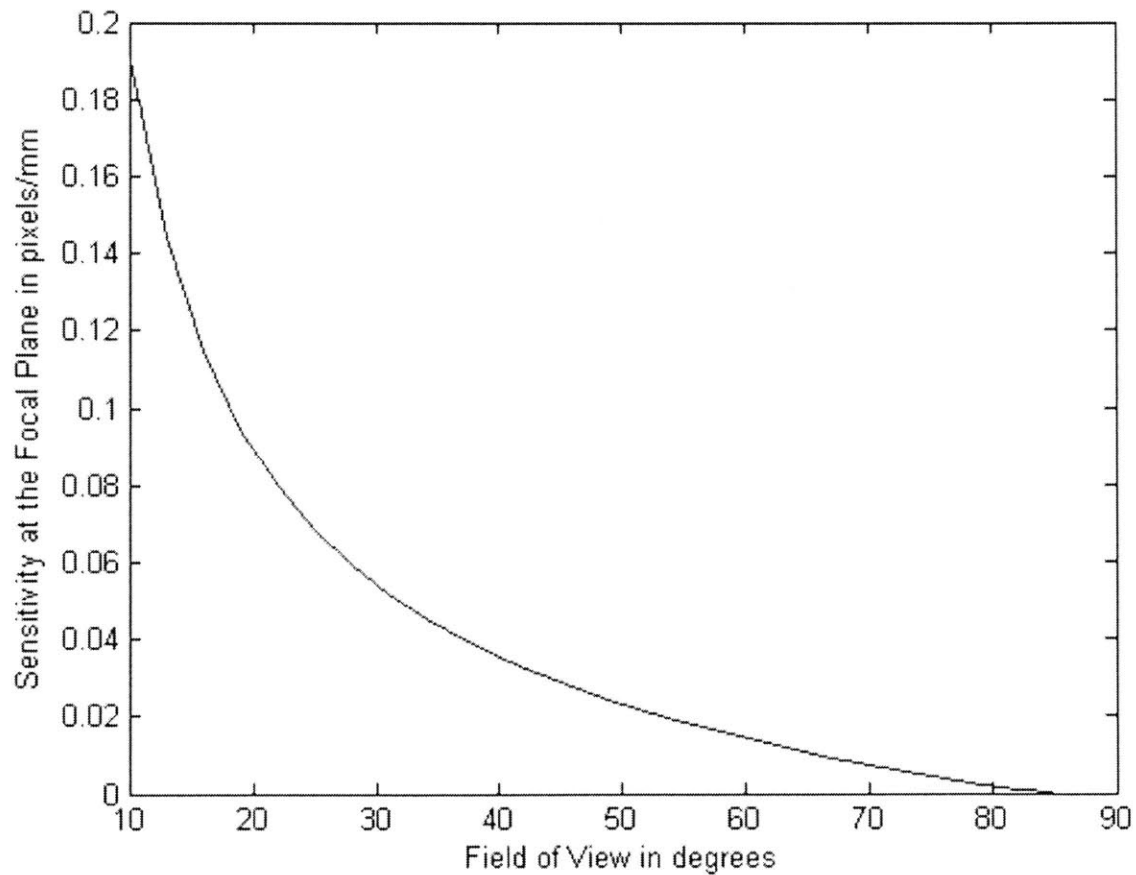
Decreasing the target object depth 'Z' to the first principal plane of the endoscope increases the sensitivity of the system in a highly non-linear fashion as depicted in figure 2.7. Further, the position of the in-focus plane also impacts disparity sensitivity as can be observed in figure 2.7.



**Figure 2.7:** Simulated Disparity Sensitivity decreases with increasing target object distance in a non linear fashion for AWS coupled Olympus endoscope. In addition, reduced distance between the first principal plane of the endoscope and the in-focus plane improves quality of disparity sensitivity for the same target object depth. The aperture offset was maintained at 1.25mm and the field of view at the sensor was about 80 deg.

### 2.4.3 Field of View and Disparity Sensitivity

A consequence of decreased target distance is decreased field of view of the target on the sensor. The effect of changing field of view on disparity sensitivity is illustrated in figure 2.8. As expected, disparity sensitivity drops off with increased field of view of the system. An implication of this observation is that the any proposed AWS-scope design would be limited in its capability to incorporate good disparity sensitivity without compromising on its field of view.



**Figure 2.8:** Disparity Sensitivity decreases with increasing system field of view in a non linear fashion for an AWS coupled Olympus endoscope. For the purposes of this simulation, the distance of the in-focus plane from the first principal plane of the endoscope is equal to 50mm. The aperture offset was maintained at 1.25mm and the field of at the sensor was about 80 deg.

Having analyzed the effect of changing field of view on disparity sensitivity, it is insightful to examine the effect of a constant field of view on disparity sensitivity. For the purposes of this exercise, the imaging train depicted in figure 2.3 is replaced with a lens of equivalent effective focal length at the position of the first principal plane of the endoscope [6]. Under this condition, if  $Z_{sensor}$  is the distance of the CCD sensor from this equivalent lens,

The expression for observed disparity is given by

$$\frac{b}{R} = Z_{sensor} \left( \frac{1}{Z} - \frac{1}{Z_{ref}} \right) \quad 2.7$$

and disparity sensitivity is given by

$$\frac{\partial \left( \frac{b}{R} \right)}{\partial Z} = -Z_{sensor} \left( \frac{1}{Z^2} \right) \quad 2.8$$

$Z_{sensor}$  can also be expressed in terms of the field of view of the system as

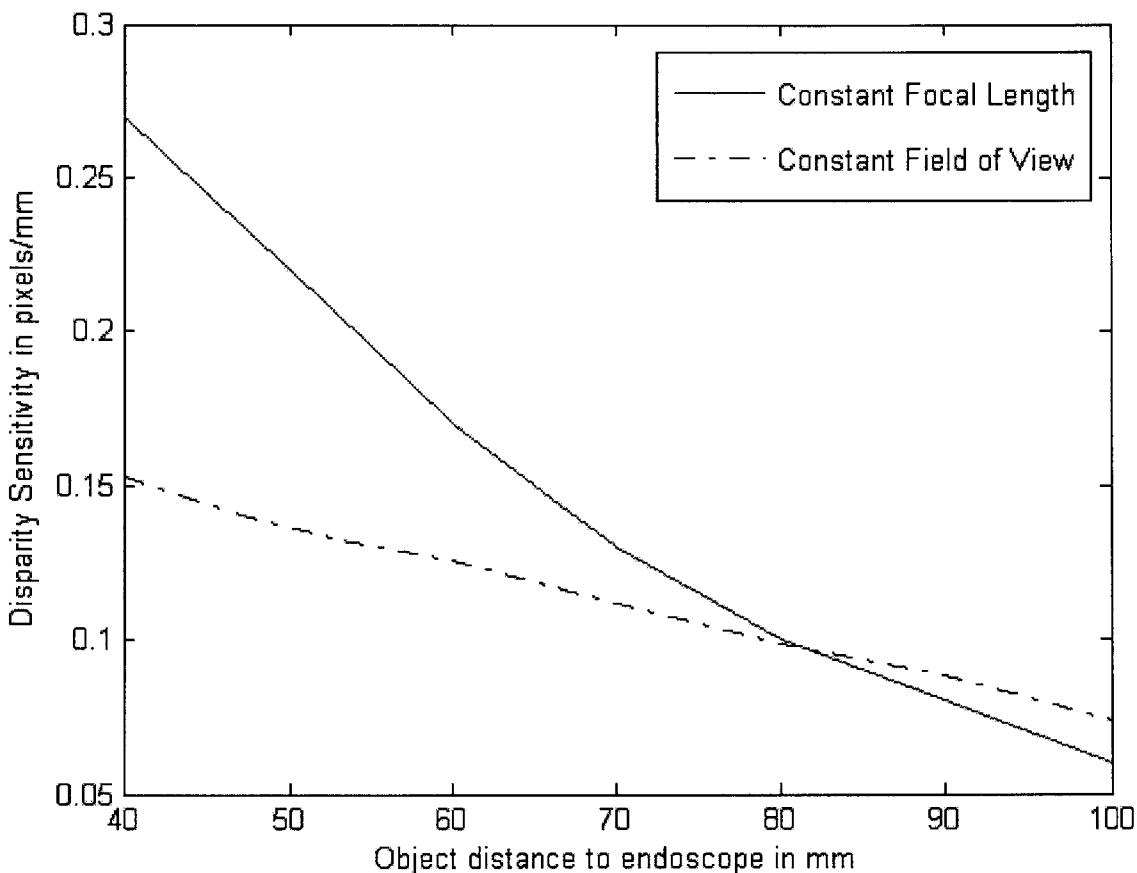
$$Z_{sensor} = S_{sensor} \left( \frac{Z}{S_{FOV}} \right) \quad 2.9$$

where  $S_{sensor}$  is the size of the sensor and  $S_{FOV}$  is the field of view of the system.

Replacing  $Z_{sensor}$  in equation 2.8 from the expression in equation 2.9, an expression for disparity sensitivity is obtained as

$$\frac{\partial \left( \frac{b}{R} \right)}{\partial Z} = -\frac{S_{sensor}}{S_{FOV}} \left( \frac{1}{Z} \right) \quad 2.10$$

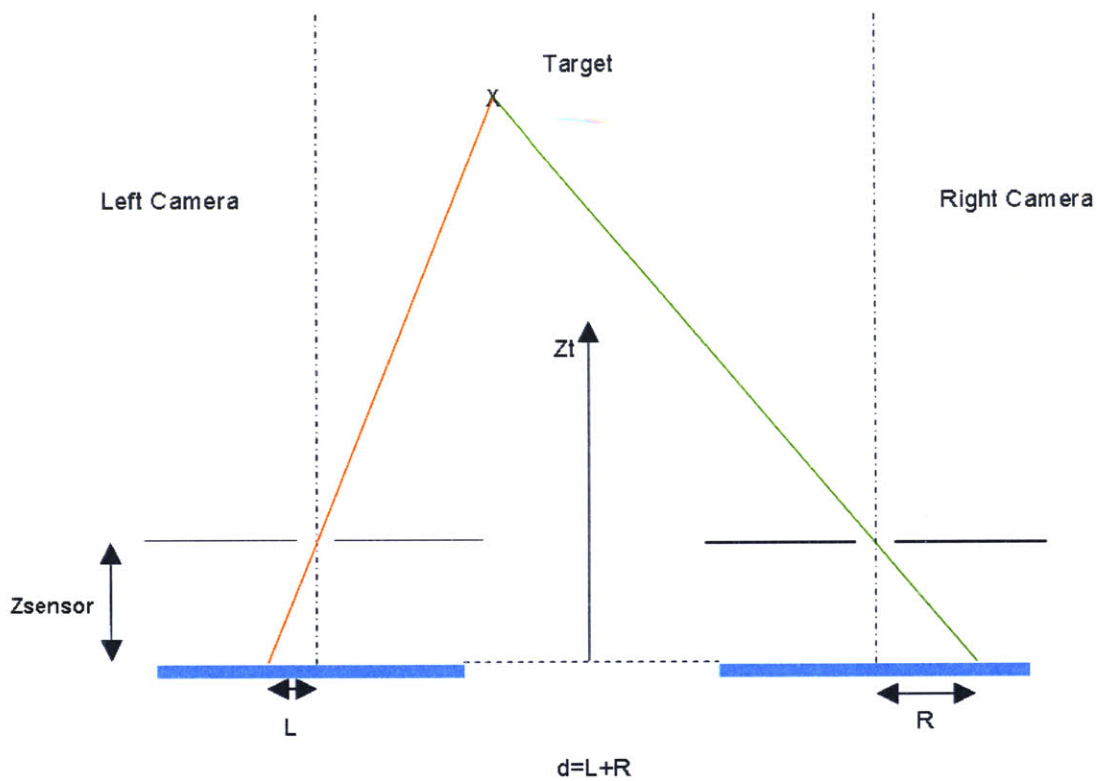
Clearly, for constant field of view, sensitivity decreases with increasing target distance. However, a comparison between equations 2.10 and 2.8 suggests that disparity sensitivity responds more rapidly to changes in target depth when the focal length of the system is held constant versus when the field of view of the system is maintained constant. Figure 2.9 illustrates this observation.



**Figure 2.9:** Comparison of simulated disparity sensitivities for the AWS-scope for the constant focal length condition and constant field of view condition. A CCD sensor of size 1 inch and a  $Z_{ref} = 100\text{mm}$  was used for the purposes of simulation. The constant focal length case had a focal length of 50mm while the constant field of view case had a field of view of about 80 deg.

## 2.5 Comparison of AWS to Standard Stereo Imaging

It is instructive to compare the performance of an AWS system to a standard depth-from-stereo imaging system [6]. A *canonical* stereoscopic system has two cameras whose sensors are perfectly aligned (ie: they have parallel optical axes and aligned x axes), as can be seen in Figure 1.10. This schematic shows that a target feature produces an image at a different location on each image sensor.



**Figure 2.10:** Canonical stereo imaging system composed by a left and right camera. The distance from the optical axis of the target image on the left and right cameras is denoted by “L” and “R” respectively. The total disparity  $d$  is just the sum of these two distances.

The disparity between these images,  $d$ , is clearly related to the distance of the target to the cameras in the  $z$  direction.

Indeed, a simple geometrical analysis reveals that the disparity between the two images is given by the following equation:

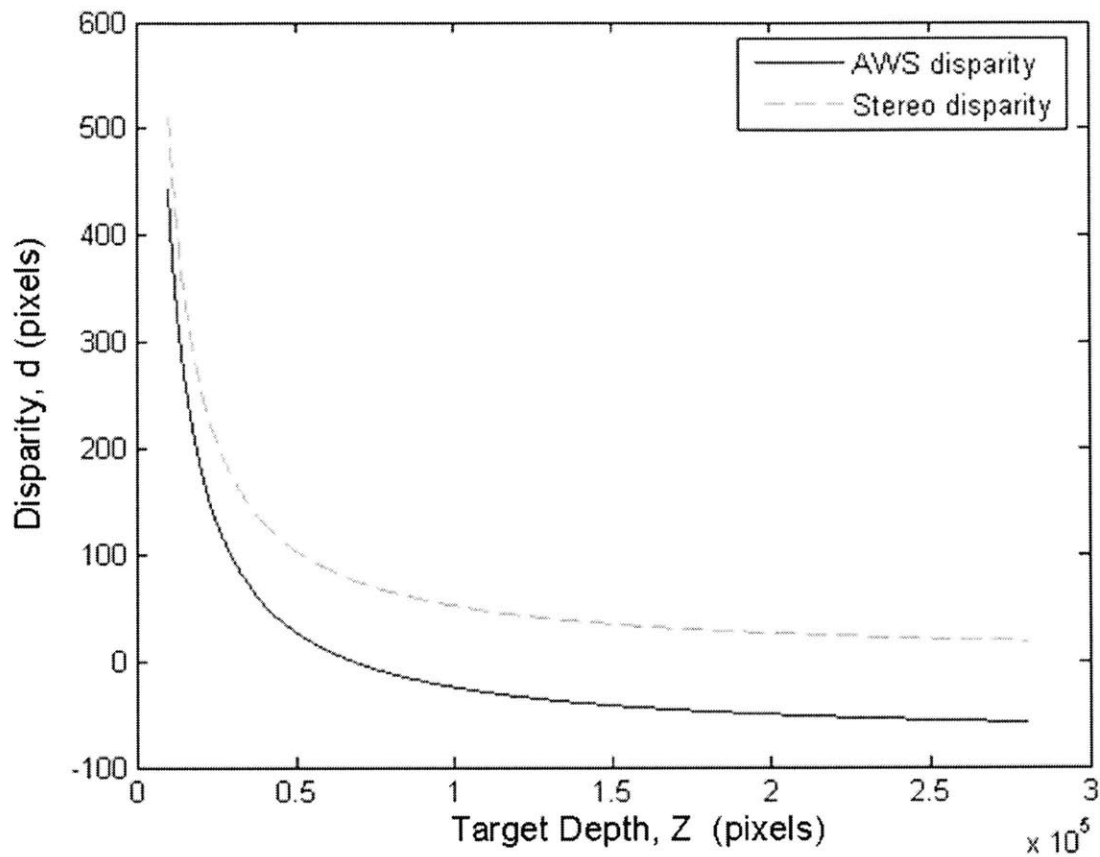
$$\frac{d}{b} = Z_{sensor} \times \frac{1}{Z} \quad 2.11$$

In this equation,  $Z_{sensor}$  is the distance from the lens' principal plane to the imaging sensor,  $b$  is the baseline distance between the optical axes of each camera, and  $Z$  is the distance of the target feature to the lens' principal plane. Comparing this expression to Equation 2.7, we can see that they only differ by a constant term involving the distance of the in-focus plane to the camera lens. This fact can be seen in Figure 2.11, which shows the disparity between two images of a target feature for both the nominal stereo imaging system as well as the AWS imaging system.

Just like in the AWS case, the sensitivity to depth of the canonical stereo system can be estimated by taking the derivative of equation 2.11 with respect to target distance. The following equation is the result of this operation:

$$\frac{\partial \left( \frac{d}{b} \right)}{\partial Z} = -Z_{sensor} \left( \frac{1}{Z^2} \right) \quad 2.12$$

Comparing the above expression to equation 2.8, we see that the sensitivity to depth of a canonical stereo system and of an AWS system is exactly the same. It can therefore be concluded that an AWS system with two diametrically opposed sampling positions responds exactly like a stereo system whose baseline is equal to the AWS sampling diameter. Therefore, the only physical system parameters that can be varied to increase depth sensitivity are the sampling diameter (equivalent to the baseline in a stereo system) and the distance of the principal plane of the lens to the imaging sensor ( $Z_{sensor}$ ).



**Figure 2.11:** Disparity as a function of target depth (both in pixels) for an AWS system and for a canonical stereo system. Using the same values baseline/sampling diameter and lens-to-sensor distance, it can be seen that the two curves are identical but displaced by a constant amount. This constant is introduced by the  $1/Z_{ref}$  term in equation 2.7.[6]

Comparing equation 2.12 to equation 2.8, we see that the sensitivity to depth of a canonical stereo system and of an AWS system is exactly the same. It can therefore be concluded that an AWS system with two diametrically opposed sampling positions responds exactly like a stereo system whose baseline is equal to the AWS sampling diameter. Therefore, the only physical system parameters that can be varied to increase depth sensitivity are the sampling diameter (equivalent to the baseline in a stereo system) and the distance of the principal plane of the lens to the imaging sensor ( $Z_{\text{sensor}}$ ). As discussed previously, disparity sensitivity increases with sampling diameter, but the latter is limited by the lens' exit pupil size; the bigger the lens, the larger the maximum sampling diameter. Disparity sensitivity also increases with  $Z_{\text{sensor}}$ , but the field of

view with a given lens also decreases as  $Z_{\text{sensor}}$  is increased. Compromises must therefore be made in increasing lens size (and therefore increasing cost), and increasing  $Z_{\text{sensor}}$  length (and decreasing the field of view).

A commercial stereo-endoscope system may be composed of two lenses and cameras separated by a baseline on the order of 6mm while an AWS system made with an identical lens and camera might only have a sampling diameter of 2mm. All other things being equal, therefore, an AWS-scope system using only two sampling positions would only have  $1/3^{\text{th}}$  the depth sensitivity of the stereo system (since disparity varies linearly with baseline). This difference in performance would at first seem to rule out the use of AWS-scope in favor of a standard stereo-endoscope system, but the comparison requires further examination. First of all, it must be remembered that the larger baseline present in a stereo-endoscope comes at the cost of greater occlusion (which is when a target feature is visible in one camera but not in the other), higher hardware cost, greater difficulty with calibration, and higher computational cost (since the matching problem needs to be solved over a larger disparity range). Additionally, in order to minimize patient trauma post surgery, the diameter of the endoscope needs to be as small as possible – a large baseline like the one found in standard stereo-endoscopes may therefore be undesirable. Finally, though a dual channel stereo-endoscope may be more sensitive to depth than a similarly built AWS-scope that uses two sampling positions, this performance differential can be significantly reduced by increasing the number of sampling positions used by the AWS system. Indeed the use of more than two sampling positions increases both the accuracy and the robustness of the algorithms that track target features from one image to another. The higher accuracy of these multi-sampling tracking algorithms results in a smaller uncertainty in the measurement of the target



feature's rotation diameter. The better rotation diameter estimate therefore compensates for the shallower sensitivity curve caused by the smaller baseline present in an AWS-scope.

## CHAPTER 3: REALIZING THE AWS-SCOPE

An exhaustive study of the optical feasibility and performance analysis of the proposed AWS-scope leads to the next logical step of realizing such a scope in practice. This chapter outlines the two step process used in implementing a laboratory version of the AWS-scope including a brief description of the various components used.

As explained in chapters 1 and 2, the four main components of an AWS-scope include a monoscope, the tag-on AWS module, an imaging sensor and some imaging optics. The laboratory implementation of the AWS-scope was performed in two stages:

### 3.1 Stages in the Realization Process

#### Stage 1: Simulation Stage

The first step of the realization process was a computer aided simulation of a potentially viable AWS-scope design. Zemax<sup>TM</sup> a ray tracing and optical simulation software package was utilized for the purpose of this exercise. Among other things, the objectives of this test included

- Appropriate positioning of the AWS sampling aperture with respect to the monoscope taking into account such imaging characteristics of the endoscope as the size of its exit pupil.
- Selection and positioning of imaging optics that minimized the effects of aberration/ image distortion besides complementing such imaging attributes of the monoscope as its field of view.
- Positioning of the sensor so as to obtain an ‘over-fill’ on the imaging sensor.

Figures 3.1, 3.2, 3.3, 3.4 and 3.5 illustrate the results of the Zemax™ simulation including the various performance attribute plots of the proposed implementation of the AWS-scope.

Figure 3.1

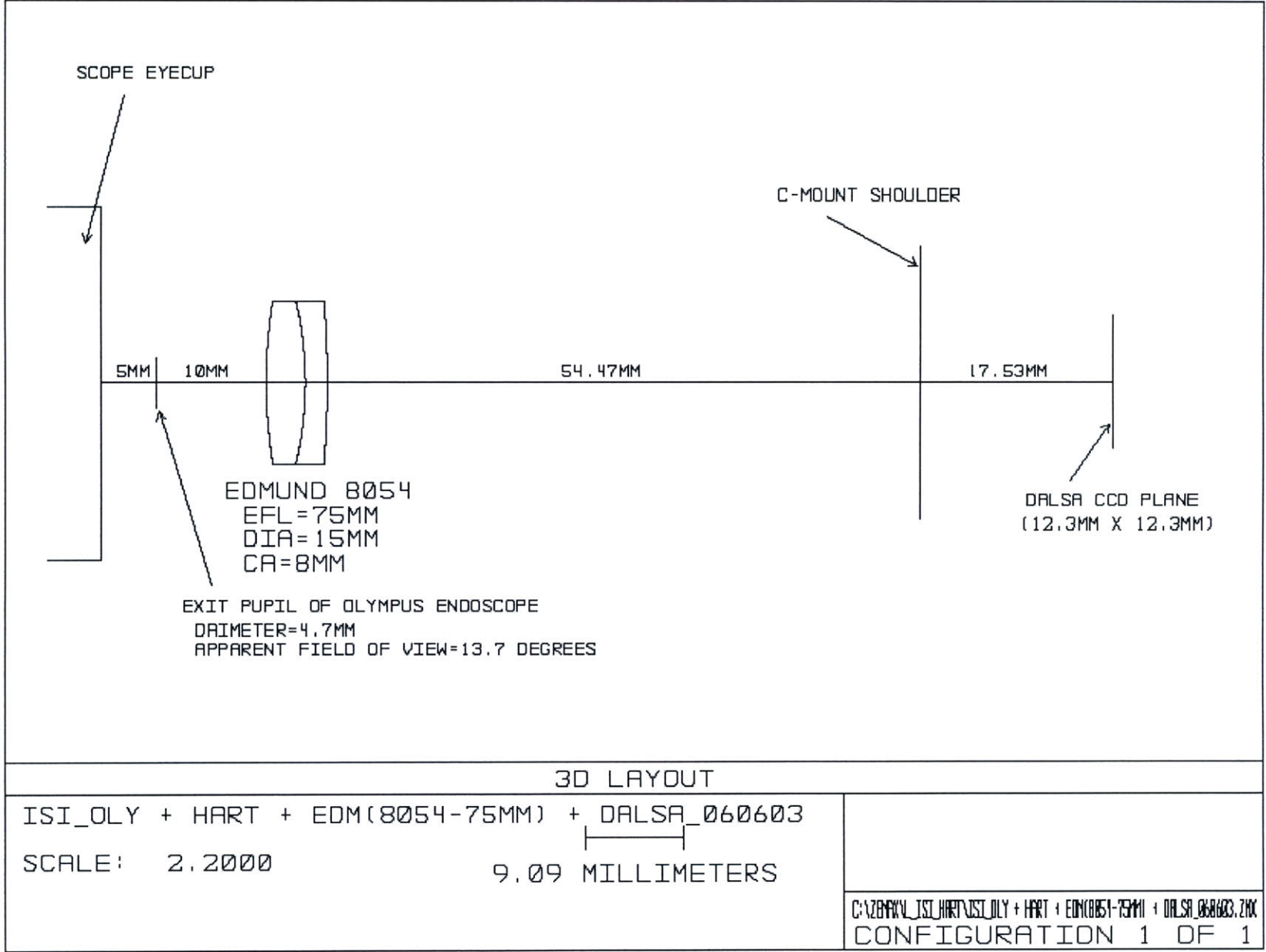


Figure 3.2

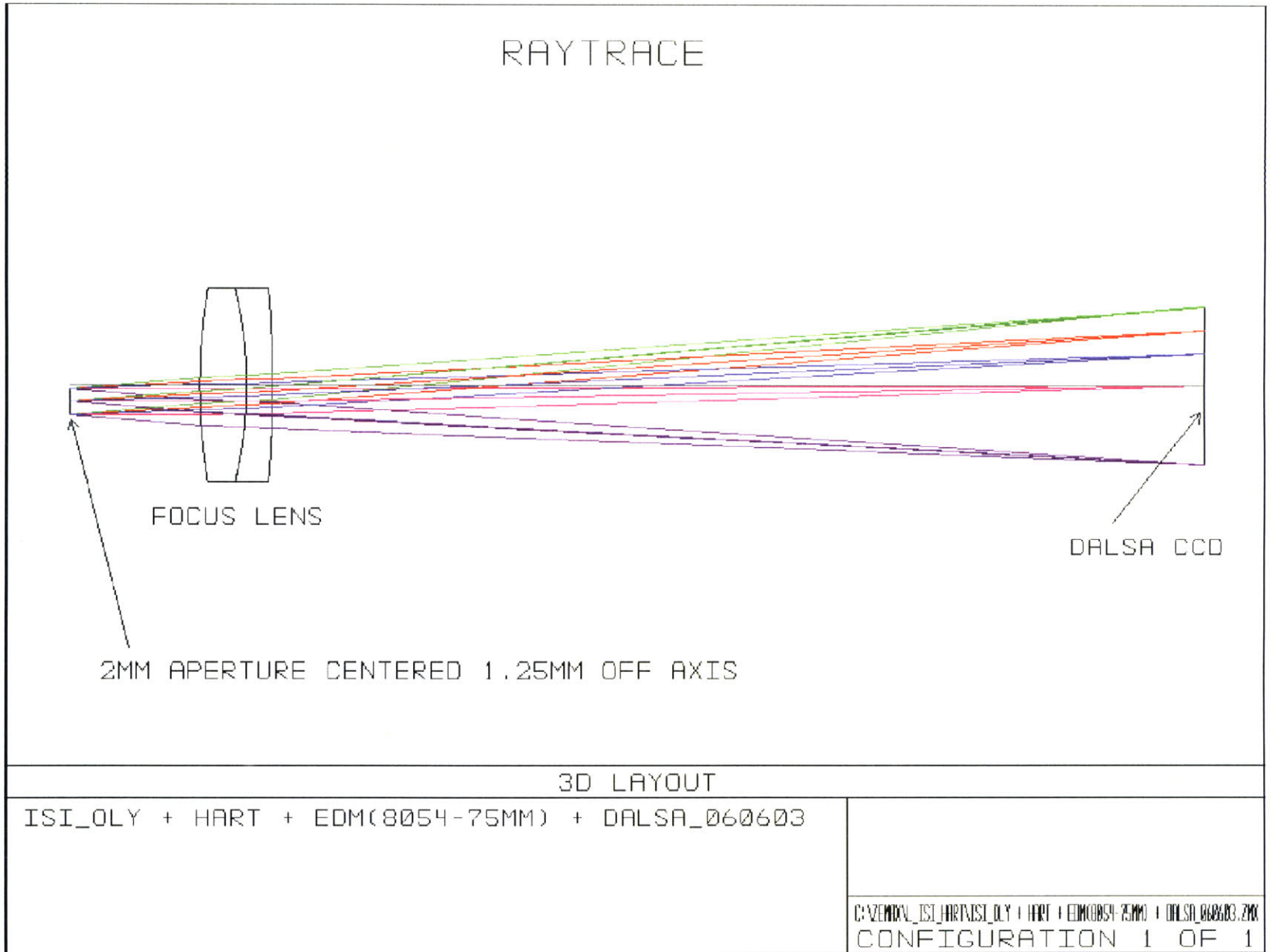


Figure 3.3

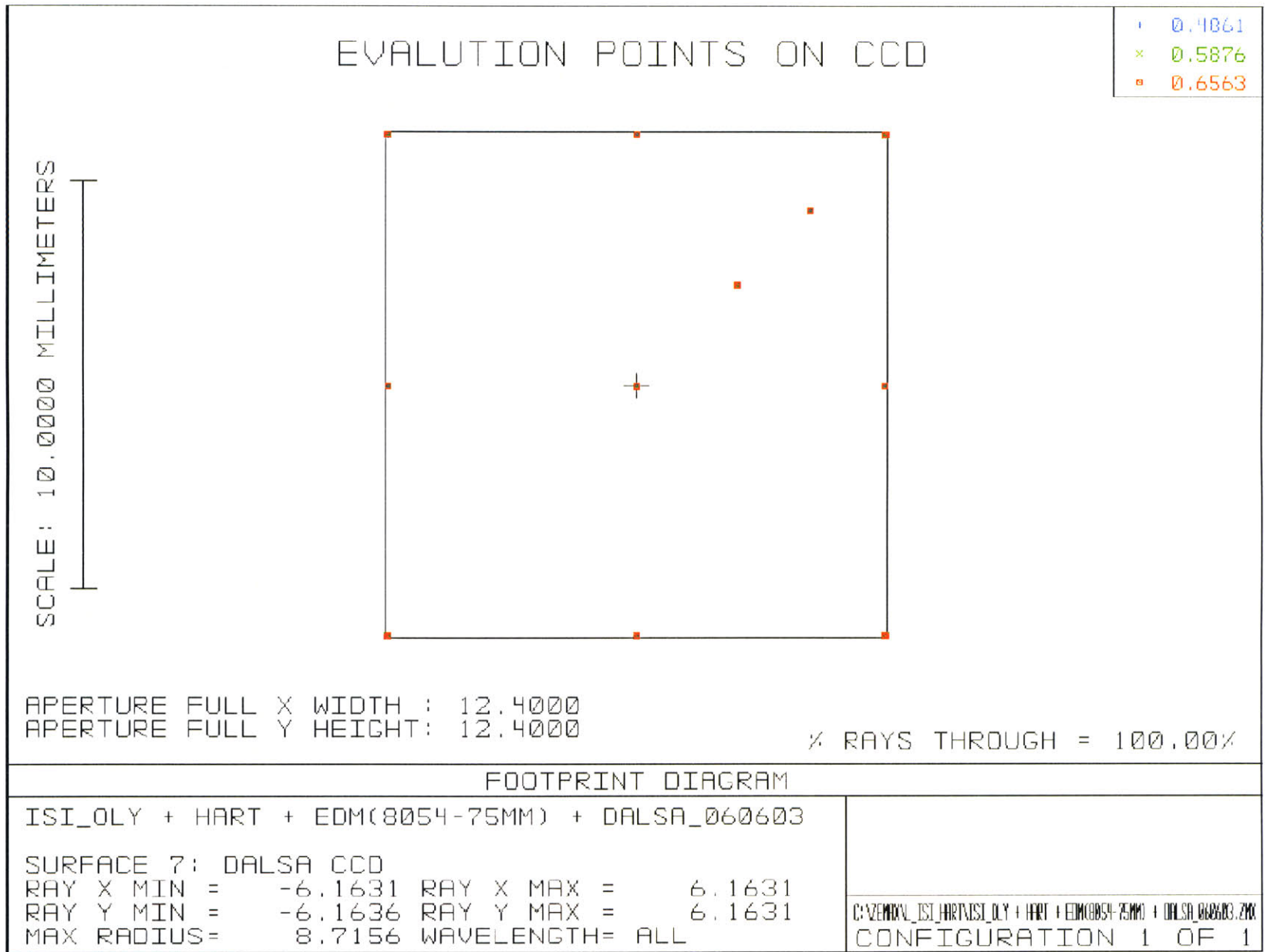


Figure 3.4

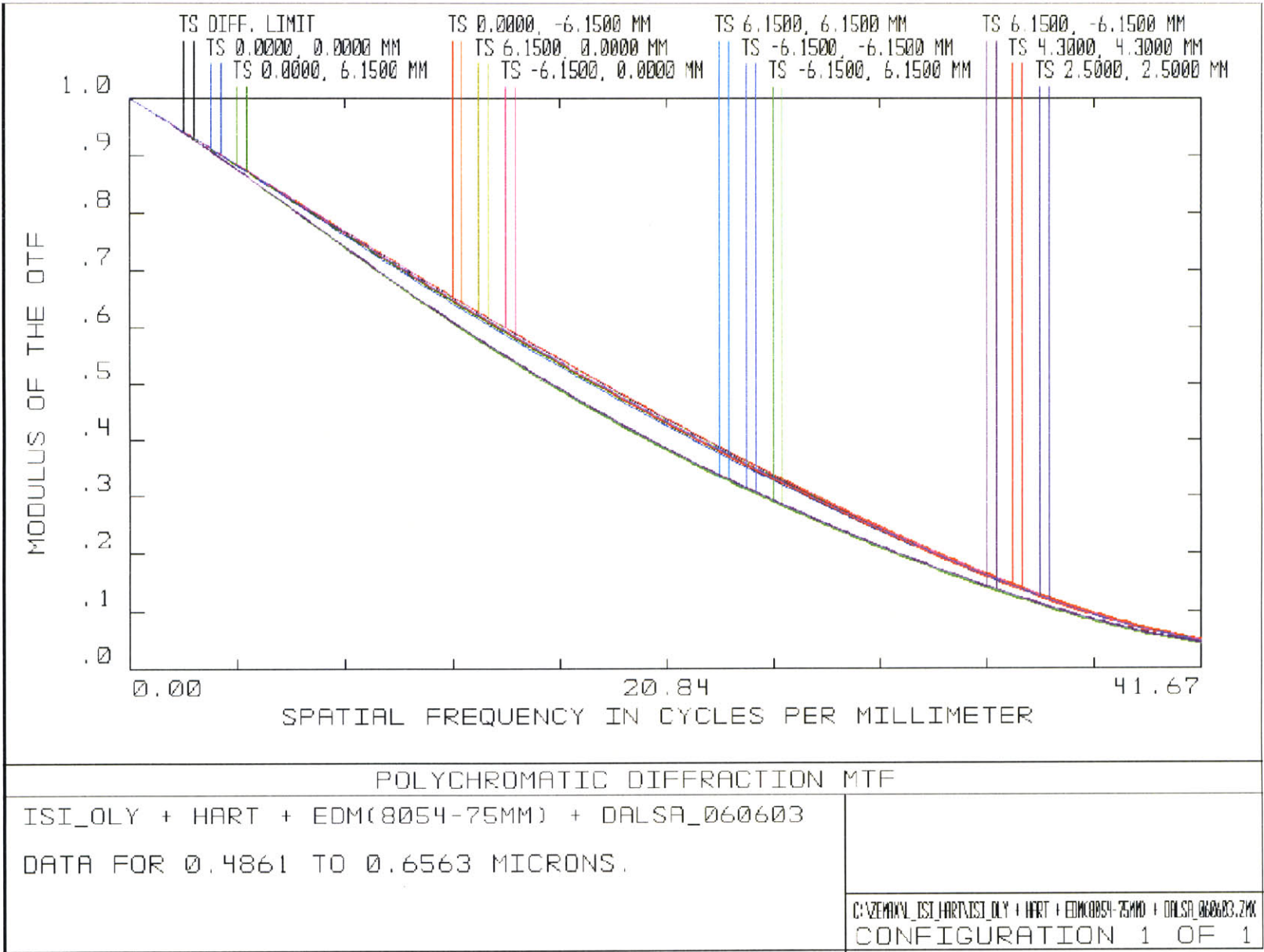
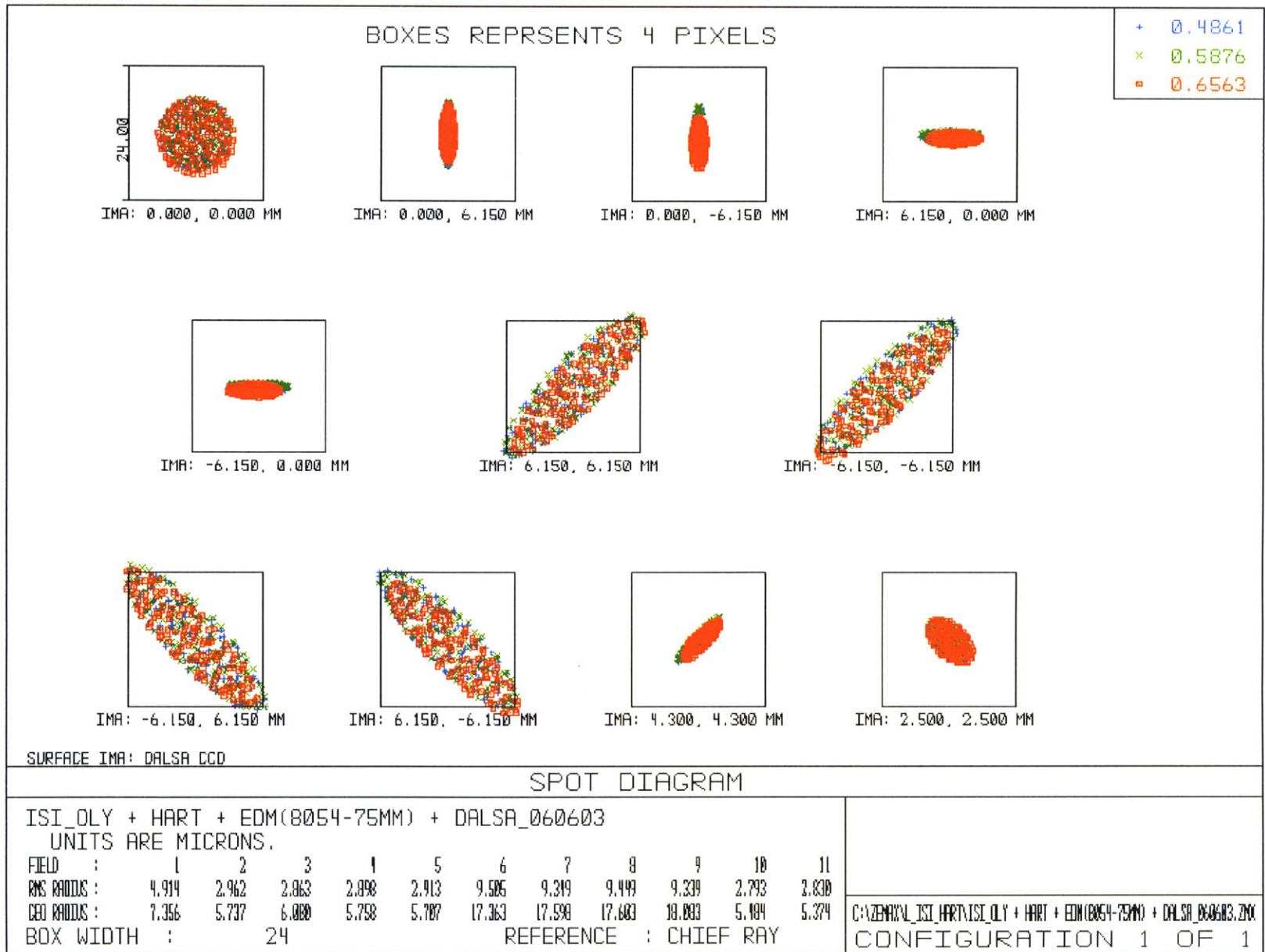


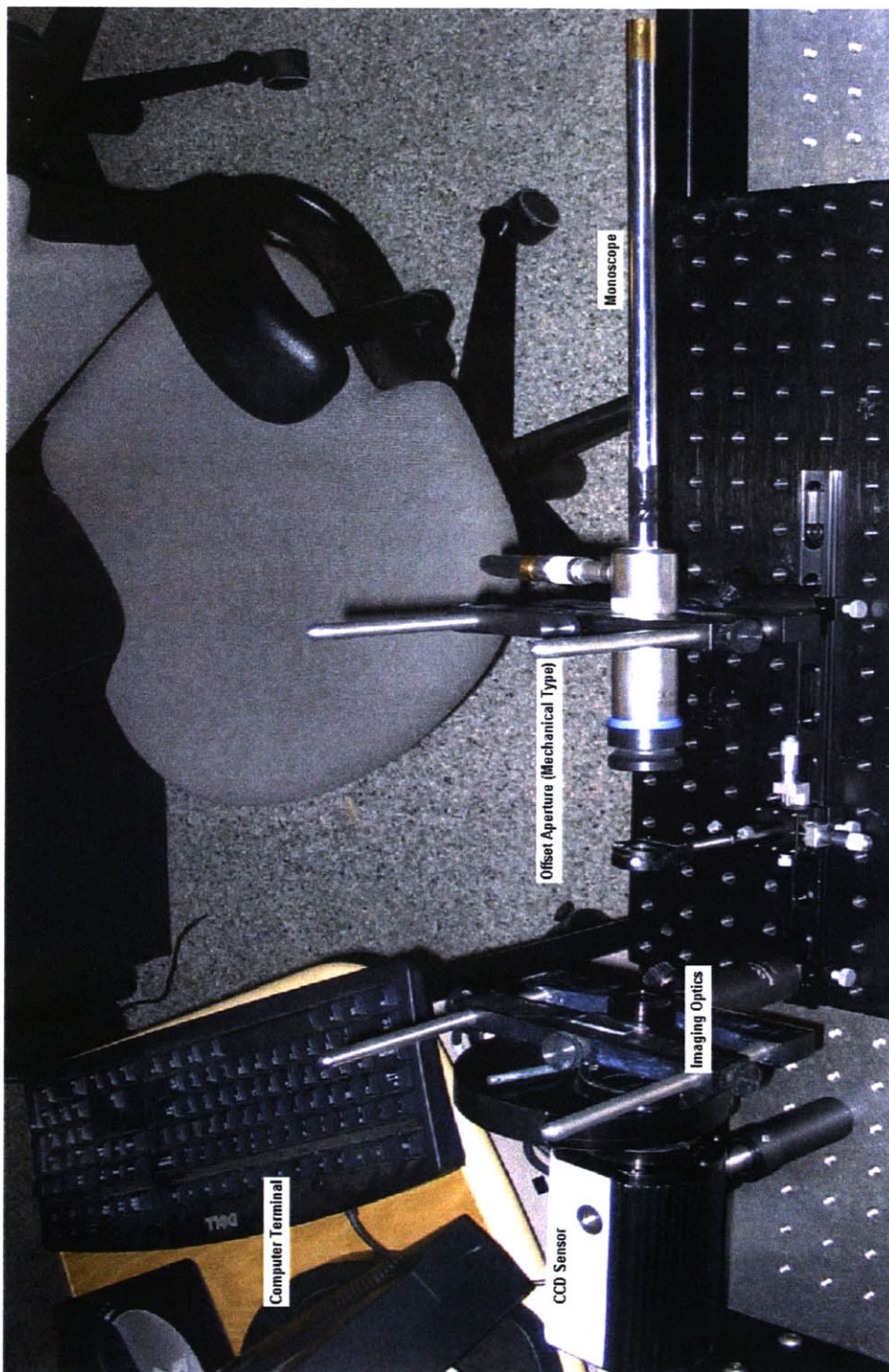
Figure 3.5





## Stage 2: Implementation Stage

The proposed design for the AWS-scope was implemented at the Hatsopoulos Microfluidics Lab at MIT – figure 3.6 illustrate the laboratory set-up.

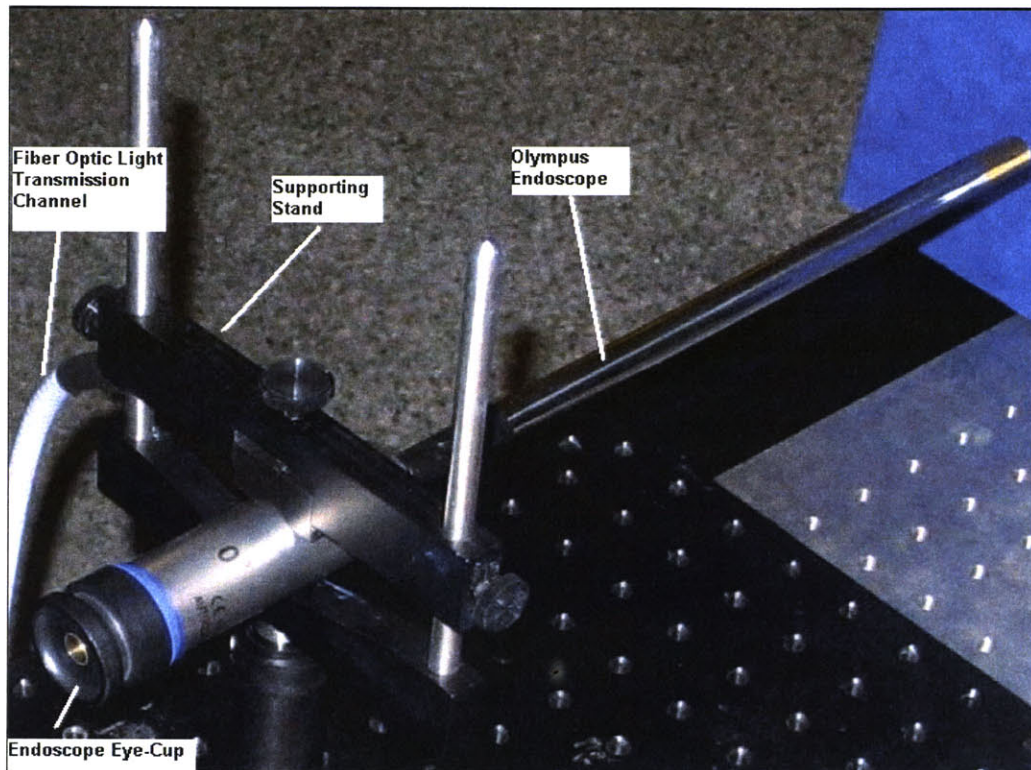


*Figure 3.6: A snap-shot of the AWS-scope set-up at MIT*

### 3.2 Components of the AWS-scope system implemented at MIT

The following section takes a closer look at the various constitutive components of the AWS-scope set-up. As can be observed, most of the components are relatively inexpensive and readily available.

#### Monoscope

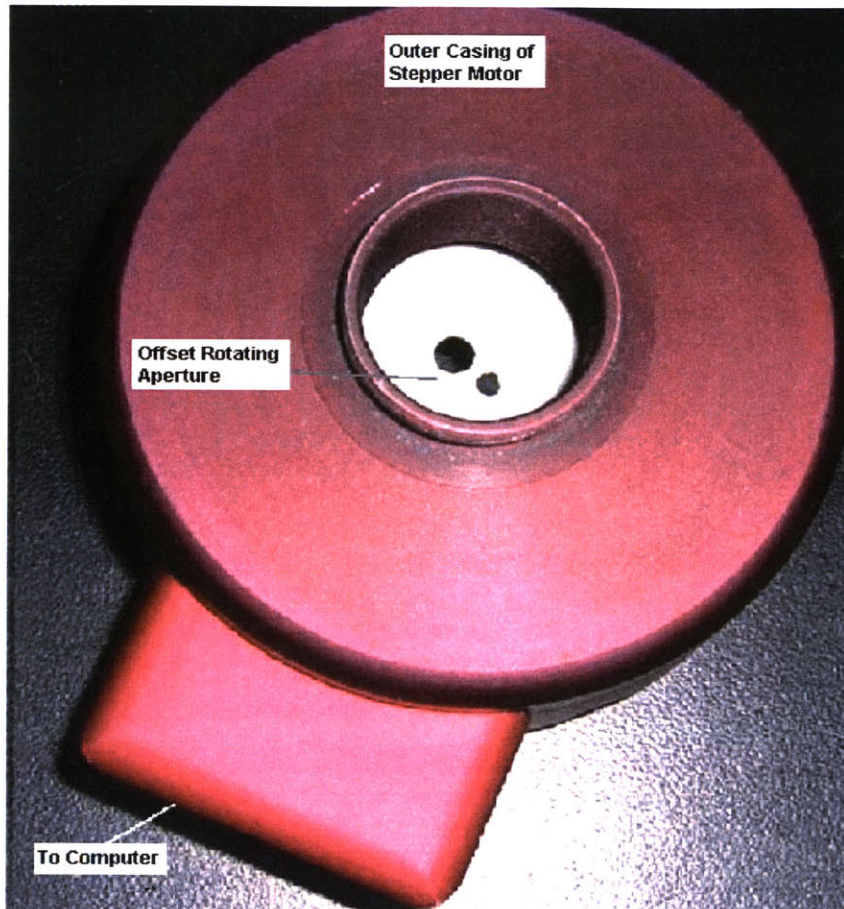


*Figure 3.7: A close-up of the Olympus Monoscope*

An Olympus monoscope with attached fiber optic illumination system was used for the purposes of implementation of the AWS-scope. The monoscope had an exit pupil with diameter 4.7mm located at a distance of 5mm from the eyecup. The scope's field of view was 85 degrees and the apparent field of view at the exit pupil was 13.6 degrees.



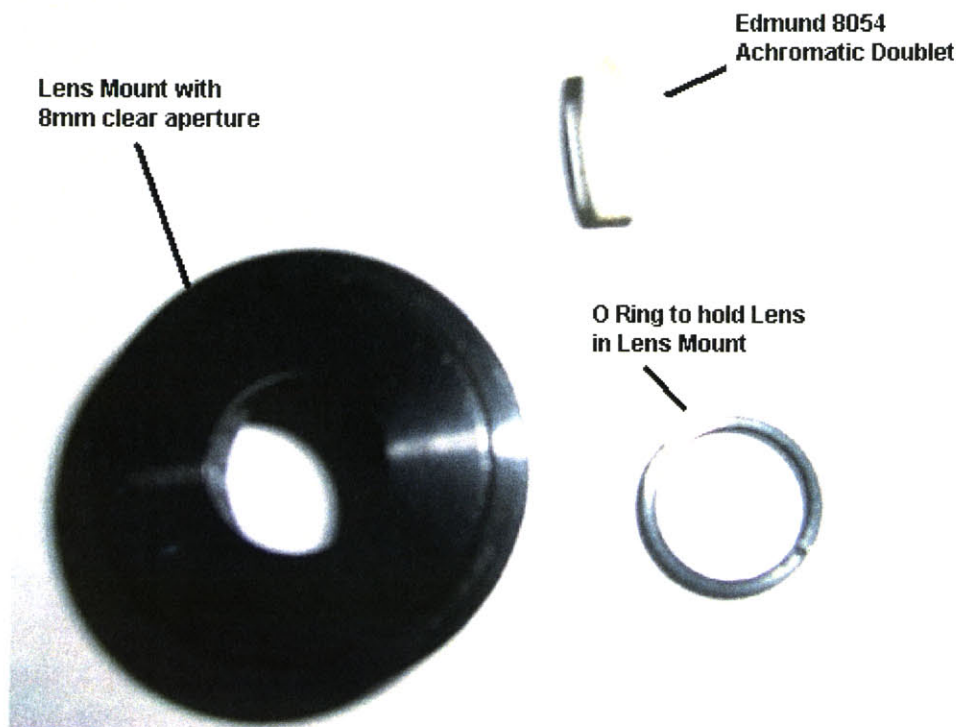
## AWS-module



*Figure 3.8: A close-up of the AWS Module*

A wave-front sampling mask which consisted of an aperture mask with offset radius of 1.25mm and an aperture diameter of 2mm was used for this set-up. The sampling mask was precisely driven by a computer controlled stepper motor. The choice of aperture diameter and offset radius was decided by the size of the exit pupil and illumination levels for the captured images. Alternatively, a mechanical, mounted mechanical aperture was used when data was collected from only two sampling points 180 degrees apart (shown in figure 3.6).

## Imaging Optics

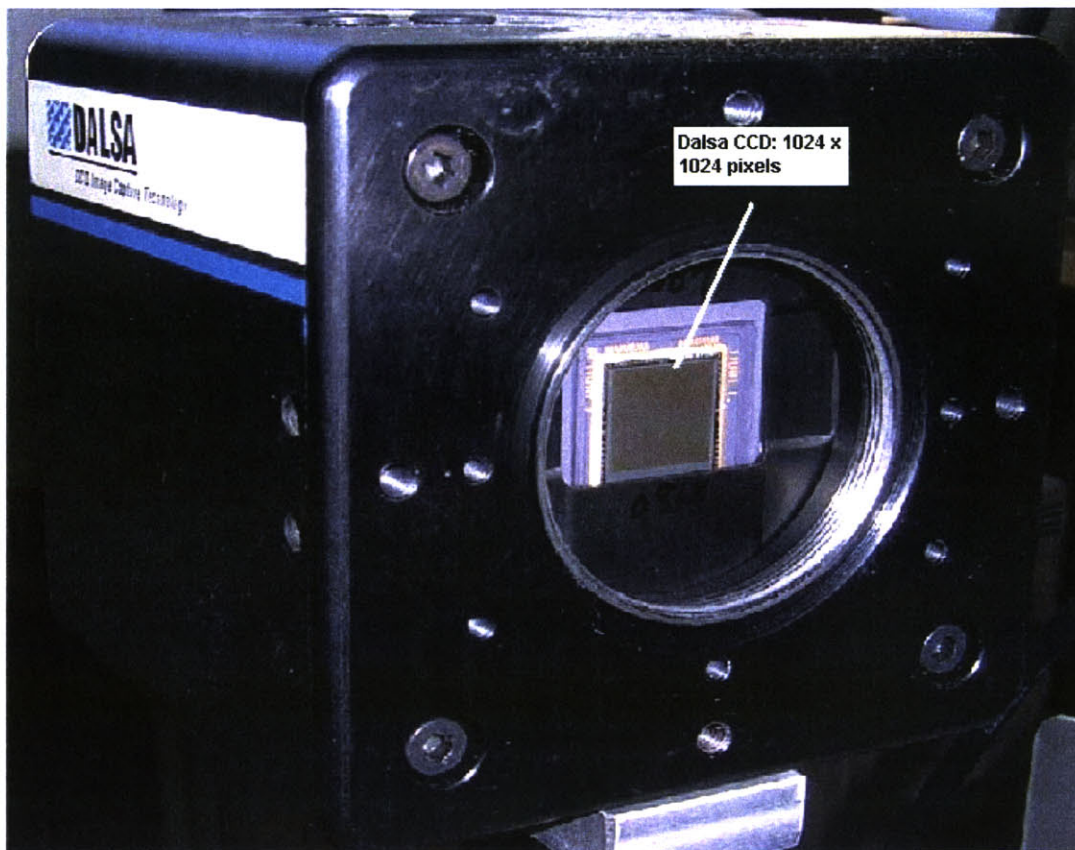


*Figure 3.9: A close-up of the Imaging Optics*

A diffraction limited monochromatic lens was used to fulfill the requirement of aberration free imaging at each aperture position and to give the highest possible resolutions (figures 3.1 to 3.5). The lens was an Edmund 8054 with an effective focal length of 75mm, a diameter of 15mm and a clear aperture of 8mm.

### **Imaging Sensor**

A Dalsa CCD sensor of size 12.3 mm x 12.3 mm with 1024 x 1024 pixels was used in overfill configuration for the set-up.



*Figure 3.10: A close-up of the CCD Sensor*

Alternatively, a Q-Imaging CCD sensor was used when capturing color image samples. The Q-Imaging camera was similarly computer controlled and is observed in figure 3.6.

### **Miscellaneous**

Other components used in the set-up include opto-mechanical parts such as rails and mounting stands for flexibly positioning the optical train components with respect to each other.

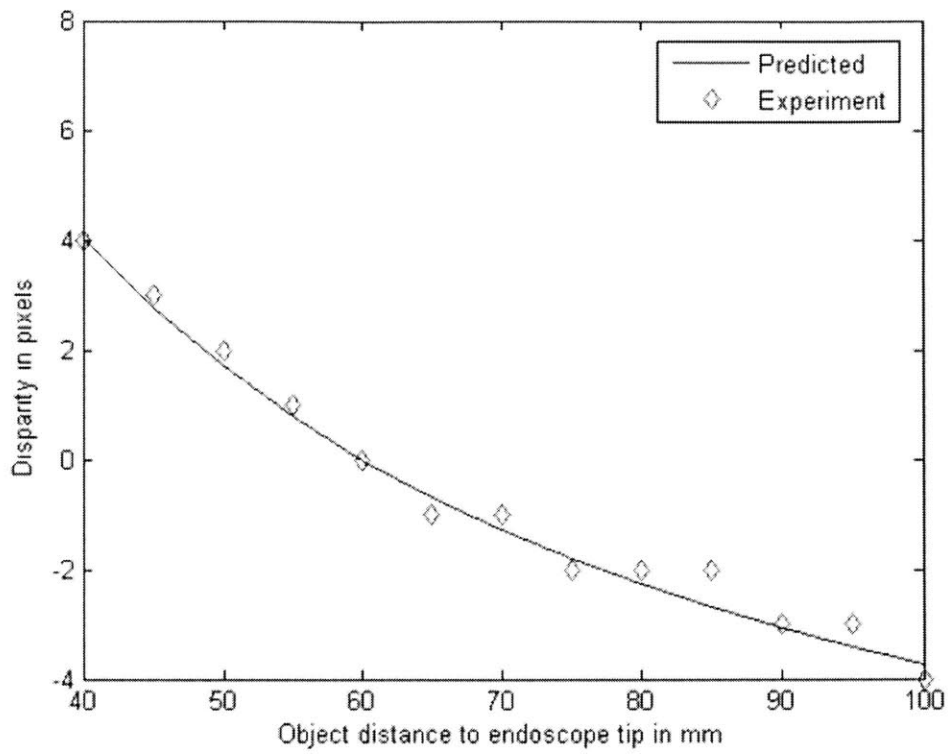
### **3.3 Comparison of AWS-scope Disparity Levels: Lab**

#### **Implementation vs. Parametric Model**

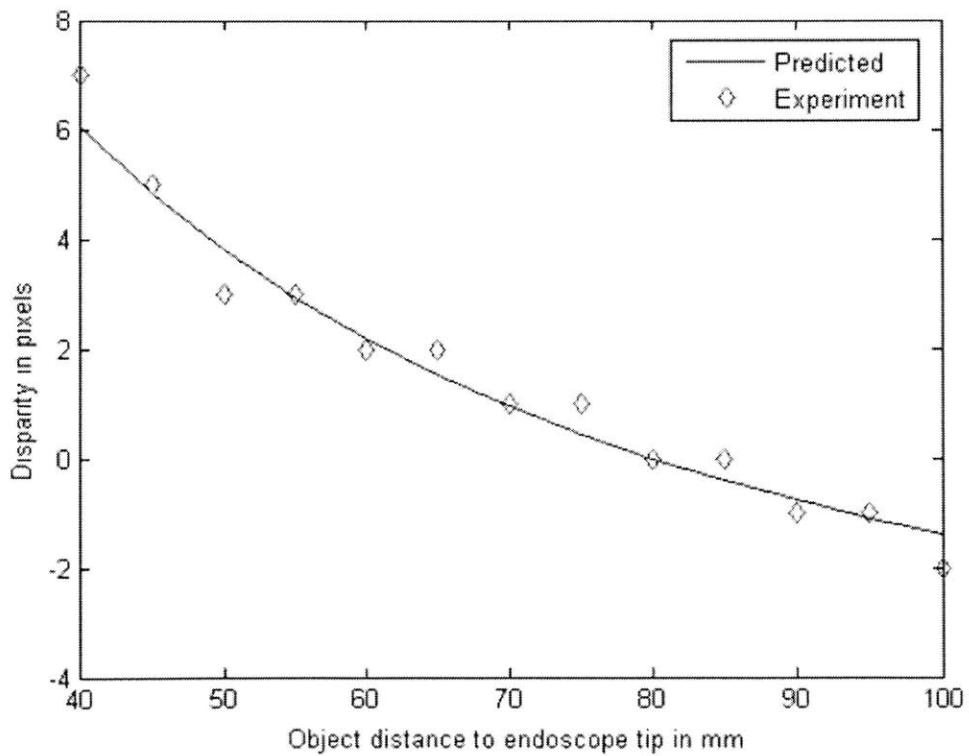
Having designed a feasible set-up to realize the AWS-scope in the laboratory, it is instructive to compare the disparity levels exhibited by this set-up versus those predicted by the parametric geometrical-optics model of the scope described in the previous chapter.

Figures 3.11, 3.12 and 3.13 present a comparison of the experimental and predicted values (using equation 2.3) of disparity for the AWS-scope system. The parameters of the simulation were identical to those of the implemented AWS-scope and the disparities exhibited by the AWS-scope in the lab were measured by sub-pixel detection of the movement of image features on a finely divided checker-board grid pattern. Three cases involving three different distances of 60mm, 80mm and 100mm of the in-focus plane to the principal plane were considered.

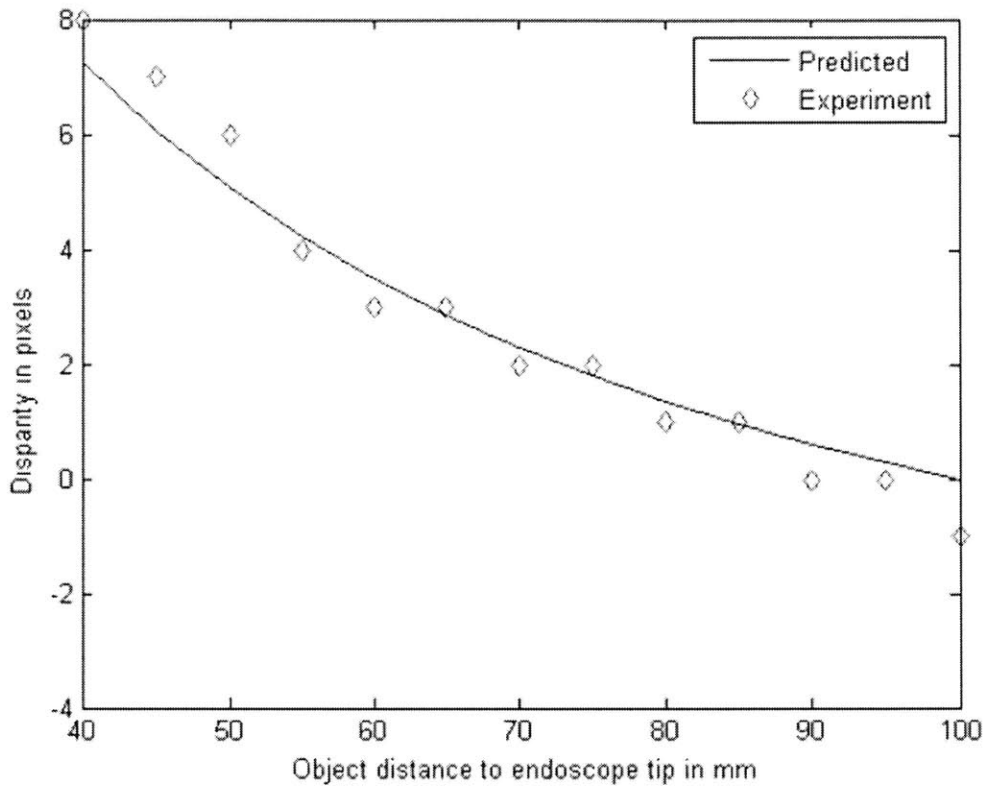
As can be observed from the figures, within the range of target object distances under consideration, the parametric model compared with the experimental results within a margin of about 15%. The difference in disparity estimates from the two methods arises primarily due to inherent assumptions in the parametric geometrical optics approach which ignores the effect of curvature of imaging surfaces.



**Figure 3.11:** Experimental vs. Predicted Disparity for distance of in-focus plane = 60mm



**Figure 3.12:** Experimental vs. Predicted Disparity for distance of in-focus plane = 80mm



**Figure 3.13:** *Experimental vs. Predicted Disparity for distance of in-focus plane = 100mm*



# CHAPTER 4: COMPARISON OF THE AWS-SCOPE WITH THE INDUSTRY STANDARD

Having conceptualized and realized the AWS-scope in chapters 2 and 3, it is instructive to study its optical performance attributes and contrast them to an existing stereoscope which is the industry standard. This chapter lays out the results of comparative tests that were carried out on the AWS-scope bench set-up and the Schoelly<sup>TM</sup> stereoscope. Test procedures, observations and inferences for a variety of optical attributes are described in greater detail in subsequent sections of this chapter. Besides, the issues of vignetting in captured images and image distortion for the AWS-scope are also addressed in the concluding sections of this chapter.

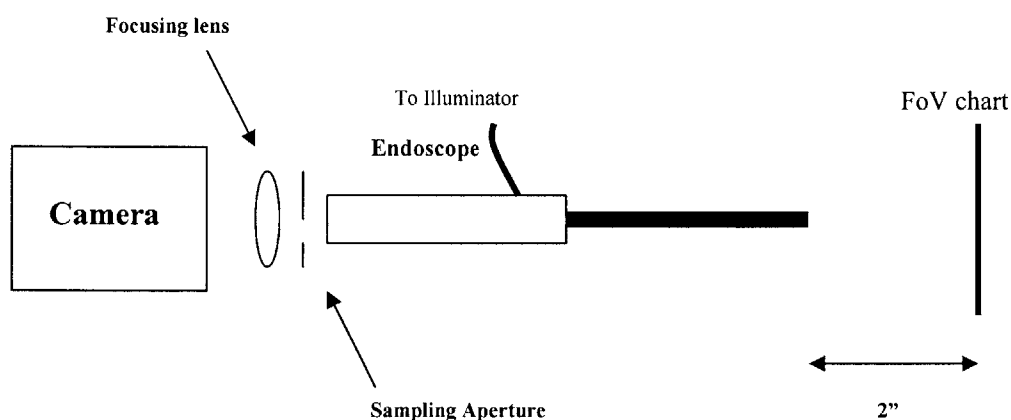
## 4.1 Optical Characterization Tests

Comparative Optical Characterization tests were performed on the AWS-scope on four main optical attributes: Field of View, Optical Resolution, Light transmission and Depth of Focus.

### 4.1.1 Field of View

Field of View (also called Angle of Coverage or Angle of view) is the amount of a given scene that can be seen with the aid of the optical system. The field stop of an optical system is the element which bounds the image plane and serves to limit the size or angular breadth of the object that can be imaged by the system. In a camera, the edge of the film itself (or the CCD plane as is the case in a digital camera) serves as the field stop.

The comparative test was conducted by viewing a standard Field of View chart placed at a standardized distance of 50.8 mm (2 inches) from the tip of the two endoscopes. The extent of angular coverage was read off of the portion of the chart captured on the sensor when viewed along the diagonal axis. Figure 4.1 illustrates the test set-up for the case of the AWS-scope and figure 4.2 shows the Field of View chart as viewed through the AWS-scope.

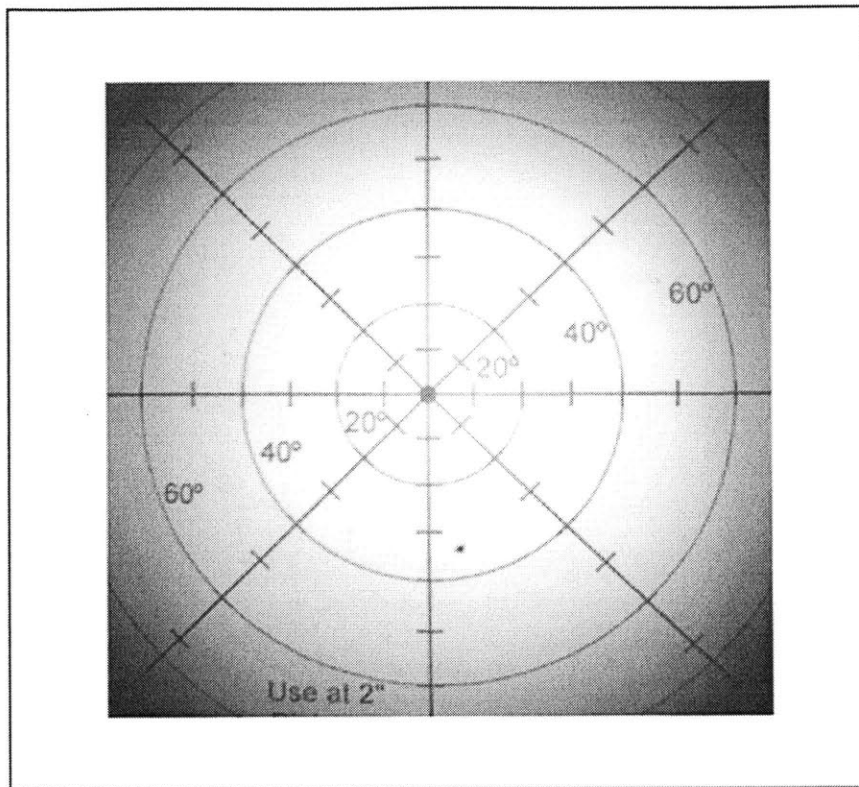


Note: Image not drawn to scale

**Figure 4.1:** Schematic of the experimental set-up used to estimate the field of view of the AWS-scope

The field of view as measured by the experiment was about 85 degrees for the AWS-scope while the benchmark Schoelly scope yielded a field of view of about 60 degrees under similar test conditions.

Clearly, it is preferable for a stereo-endoscope to possess a larger field of view so as to provide greater visibility of the surgical scenario to the surgeon. On that account, AWS-scope exhibits a superior field of view when compared to the benchmark.

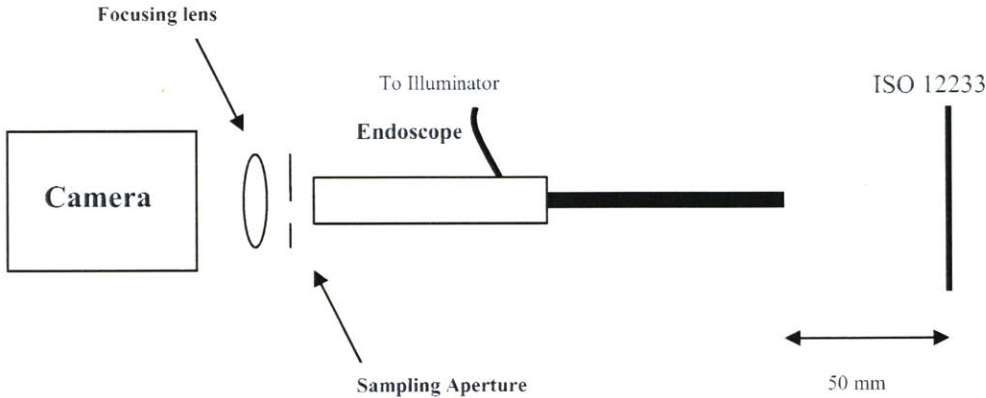


*Figure 4.2: Field of view chart as seen on the sensor under test conditions for the AWS-scope.*

#### **4.1.2 Optical Resolution**

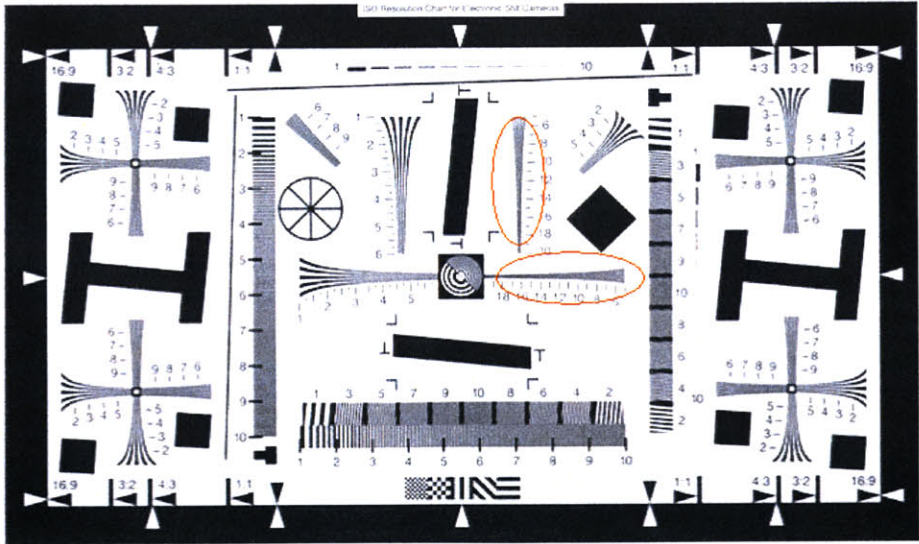
The resolution of an optical system is defined as the ability of the optical system to distinguish between two closely spaced points on an object (or in space) and is a function of the wavelength of light used as well as the numerical aperture of the system. Resolution, in the case of endoscopy, boils down to the fineness of detail that can be perceived in the captured surgical scenario. It is possible to quantify this detail by capturing an image which shows progressively finer detail and noting the finest detail that can be clearly distinguished. The ISO 12233 resolution testing chart was used for the purpose of comparing the AWS-scope and the Schoelly. Figure 4.3 illustrates a schematic

of the experimental set-up. Figure 4.4 shows the standardized ISO 12233 chart with the lines of interest encompassed by red bubbles.



Note: Image not drawn to scale

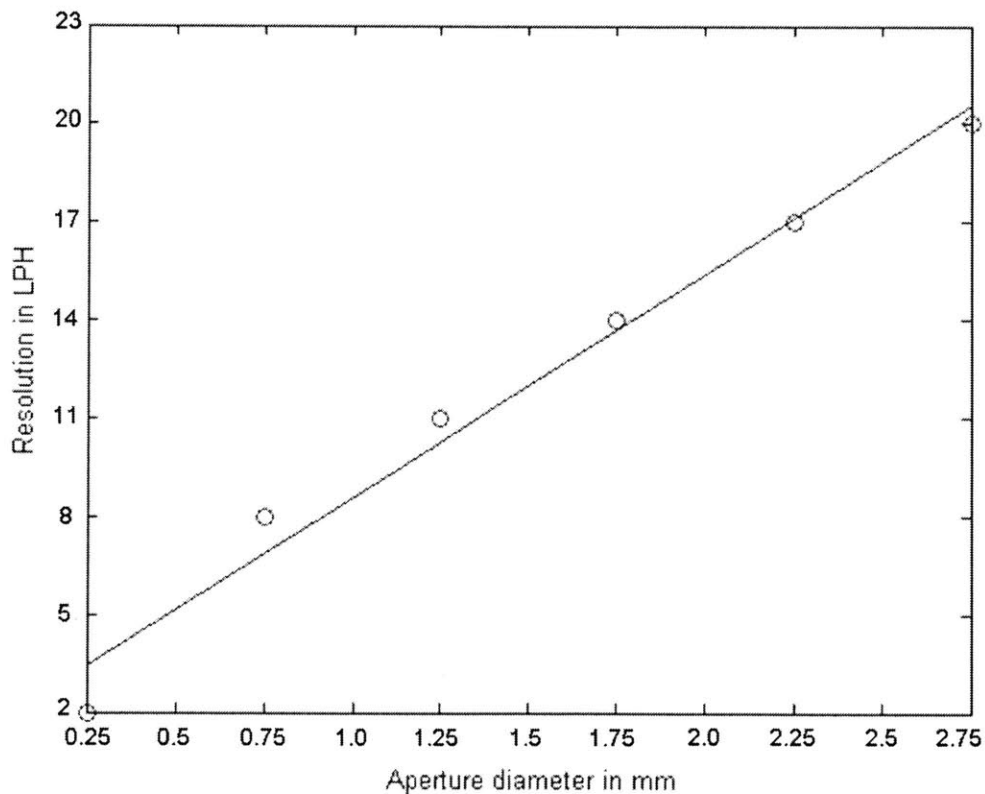
**Figure 4.3:** Schematic of the experimental set-up used to estimate the optical line resolution of the AWS-scope



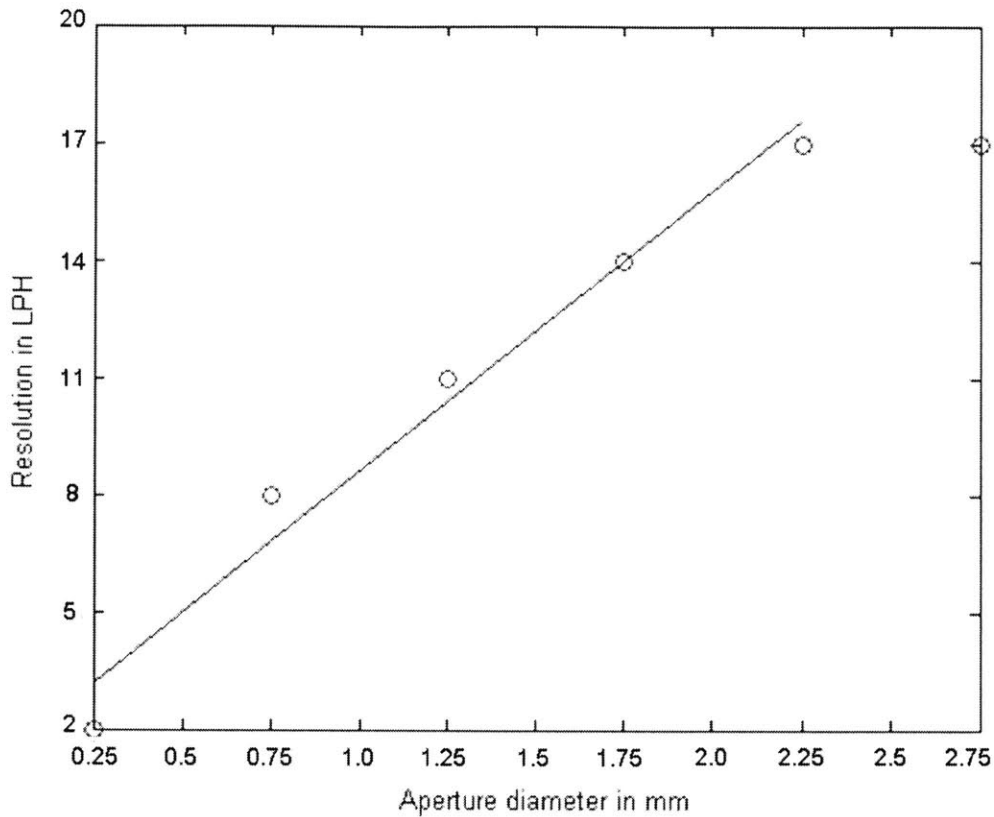
The ISO 12233 test chart for line resolution and MTF testing

**Figure 4.4:** ISO 12233 chart for testing line resolution of the optical system. The two sets of lines circled on the chart indicate the line sets used for comparative testing.

The chart was placed at a distance of 50 mm to the tip of the endoscopes – 50 mm is the typical operating distance of the endoscope tip from the surgical scenario during surgery and is hence an automatic choice for the imaging distance. The finest possible line separation (on the imaging sensor) that could be clearly distinguished was noted for both endoscopic systems in the vertical and the horizontal directions. Additional experiments were conducted on the AWS-scope wherein vertical and horizontal line resolutions were noted for changing sampling aperture diameters. Figures 4.5 and 4.6 depict the variation of line resolution with aperture diameter of the sampling aperture.



**Figure 4.5:** Horizontal line resolution with increasing aperture diameter for the AWS-scope.



**Figure 4.6:** Vertical line resolution with increasing aperture diameter for the AWS-scope.

The line resolution of the benchmark Schoelly in both vertical as well as horizontal directions was noted to be about 1200 LPH. At the design sampling aperture diameter of 2mm the AWS-scope yielded line resolutions of about 1500-1600 LPH in the horizontal direction and about 1600-1700 LPH in the vertical direction. Since higher line resolution is more preferable in precision tasks such as endoscopic surgery, the AWS-scope displays a superior optical characteristic in comparison to the benchmark.

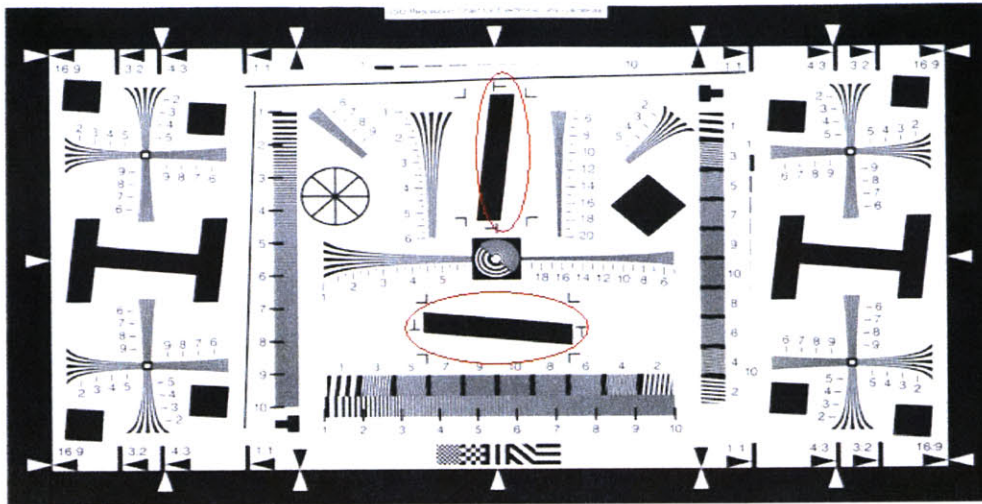
Figures 4.5 and 4.6 also indicate that increasing aperture diameter improves line resolution in a linear fashion – a trend that is consistent with the fact that increasing aperture size leads to higher values of numerical aperture thereby leading to better resolution for the imaging system.

Although measurement of optical resolution using the lines per height (LPH) metric is very convenient, it is based on perception and judgment (both of which are easily influenced by image contrast). It is therefore important that inferences from the subjective perception of line resolution be corroborated with other tests.

One method of corroboration involves use of the Modular Transfer Function (MTF) which represents the frequency response of an imaging system in the spatial domain. An MTF is a normalized curve on the Y axis with a 0-1 scale and a frequency scale on the X axis and it denotes the range of frequencies that the optical system passes through without contrast attenuation. The wider the MTF curve, the greater the number of frequencies that are passed through the optical system and therefore the better the contrast of the captured scenario due to the presence of higher frequency fine features.

MTF curves were generated for the AWS-scope system and the Schoelly by imaging standardized features (the vertical and horizontal blocks circled in red on figure 4.7) on the ISO 12233 chart. As in the case of optical line resolution testing, the chart was placed at a distance of 50 mm from the endoscope tips and the images captured. These images were subsequently processed in order to extract the contrast curve both in the horizontal and vertical directions.

The MTF profiles are shown in figures 4.8, 4.9, 4.10 and 4.11 for the Schoelly and the AWS-scope at its design condition of a 2mm sampling aperture.

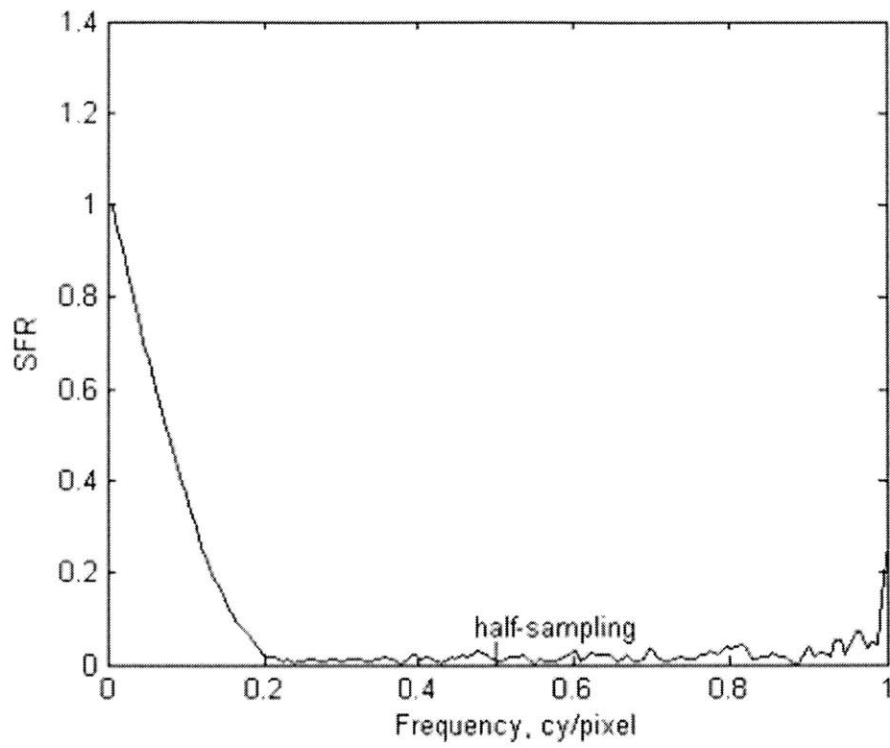


The ISO 12233 test chart for line resolution and MTF testing

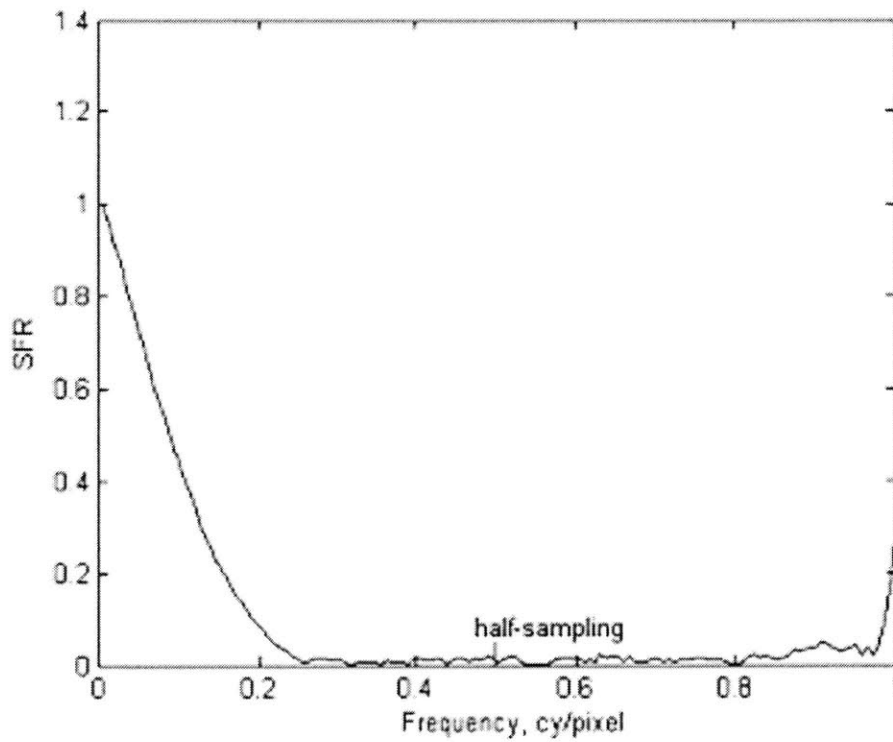
*Figure 4.7: ISO 12233 chart for MTF curve of the optical system. The two blocks circled on the chart indicate sets used for comparative testing in the vertical and horizontal directions.*

The MTF profiles indicate that the AWS-scope allows a greater portion of higher spatial frequencies to pass through when compared to the Schoelly thereby demonstrating the capability to capture finer details of a surgical scenario. This observation is also consistent with the previously observed data on optical line resolution testing thereby confirming that the AWS-scope is capable of superior resolution when stacked up against the industry benchmark.

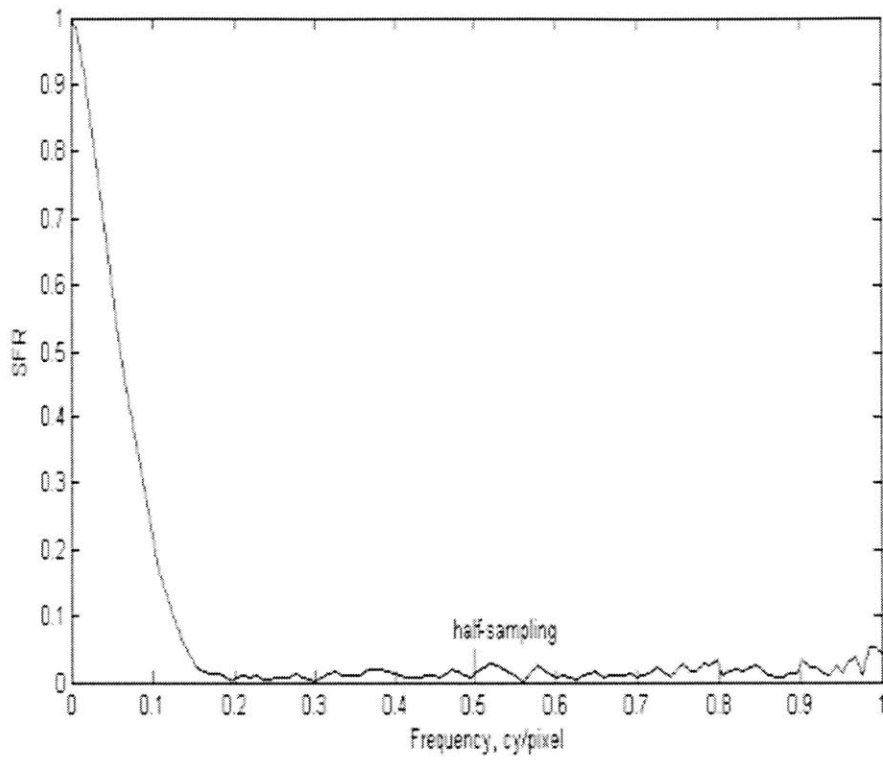




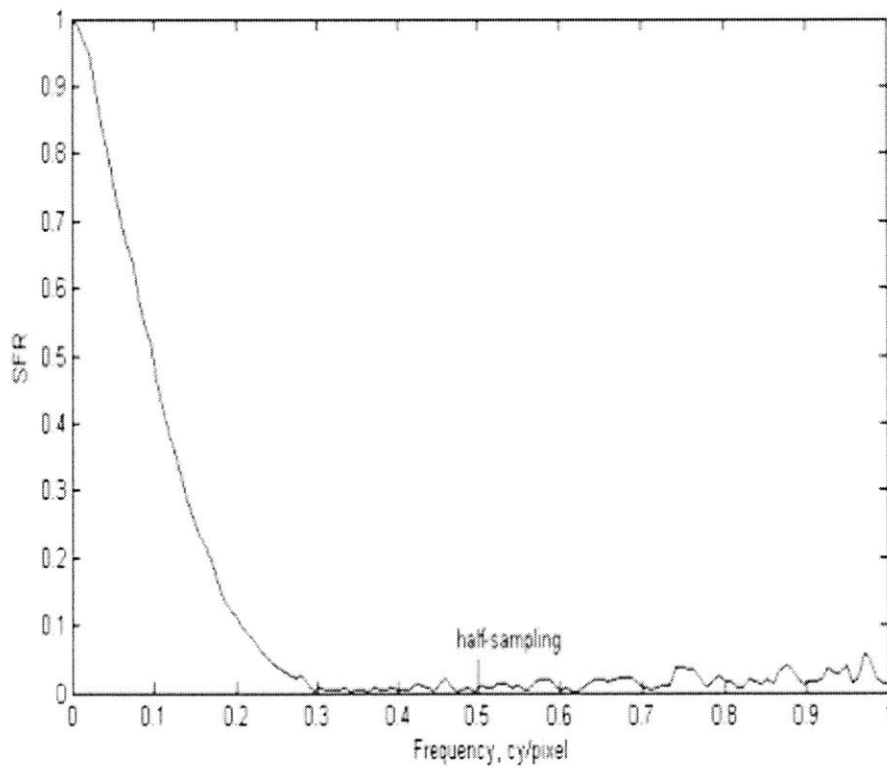
**Figure 4.8:** MTF profile for the Schoelly in the horizontal direction when the target is placed at a distance of 50mm from its tip.



**Figure 4.9:** MTF profile for the AWS-scope under design conditions in the horizontal direction when the target is placed at a distance of 50mm from its tip.



**Figure 4.10:** MTF profile for the Schoelly in the vertical direction when the target is placed at a distance of 50mm from its tip.



**Figure 4.11:** MTF profile for the AWS-scope under design conditions in the vertical direction when the target is placed at a distance of 50mm from its tip.

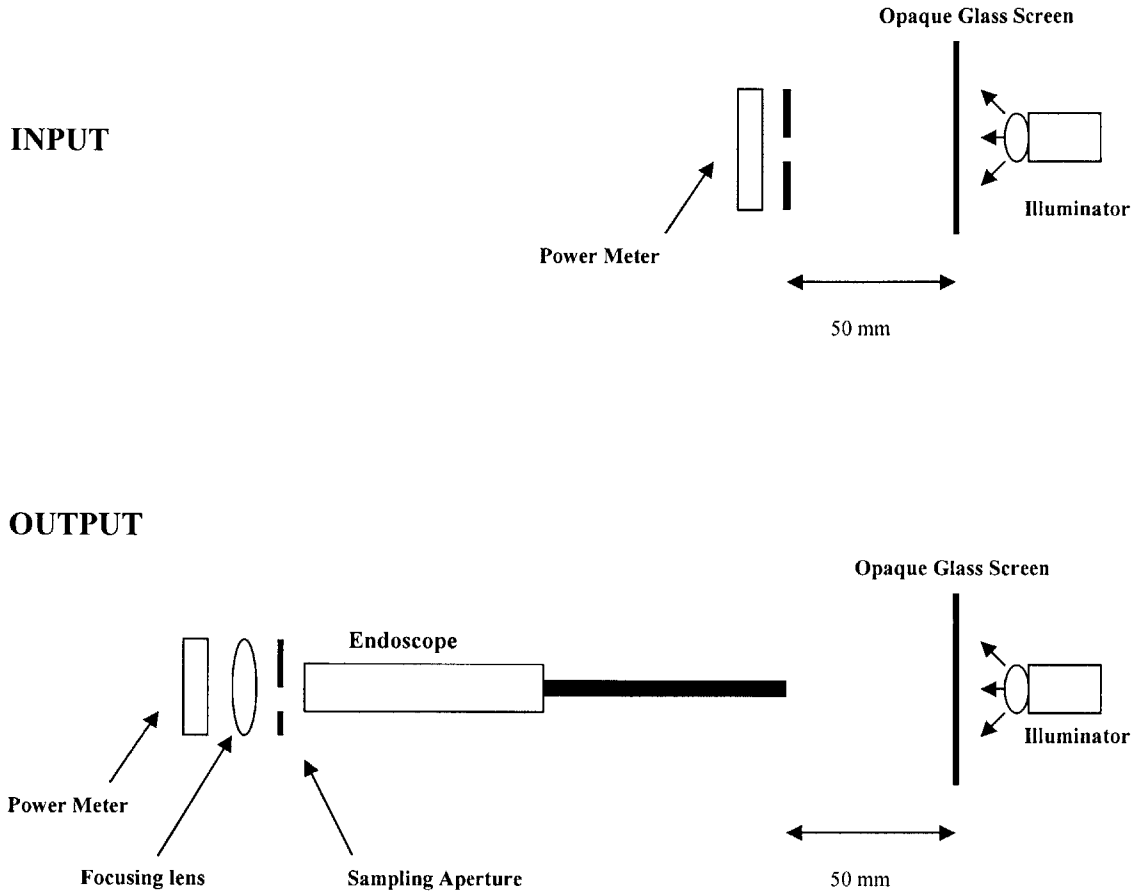
### 4.1.3 Optical Transmission

The optical transmission of an optical system is defined as the ratio of the amount of light exiting the system to the amount of light entering it. It serves as a test for a ‘Go/No go’ comparison between transmission ratios of two competing optical systems.

Since no standard procedures exist for optical transmission tests, the procedure employed for the present work was designed specifically for our configuration.

Figure 4.12 illustrates a schematic of the test set-up for measuring optical transmission ratios for the AWS-scope and the Schoelly. A standardized input light level-which consisted of light entering through a 3mm hole in a mask from a fiber optic light source - was provided to both systems. The output light levels were measured at the position of the imaging sensor using a power meter. The transmission ratio was then defined as the ratio of output light level to input light level. As in the optical resolution experiments, the light source was placed at a standard target distance of 50 mm from the endoscope tip.

The experiment was repeated on the AWS-scope for a variety of sampling aperture diameters. The power meter used for the purpose of measuring light-levels was initially calibrated to read a value of zero for the ambient light level of the lab where the experiment was performed.



Note: Image not drawn to scale

**Figure 4.12:** Schematic of the experimental set-up used to evaluate the Optical Transmission Efficiency of the AWS-scope.

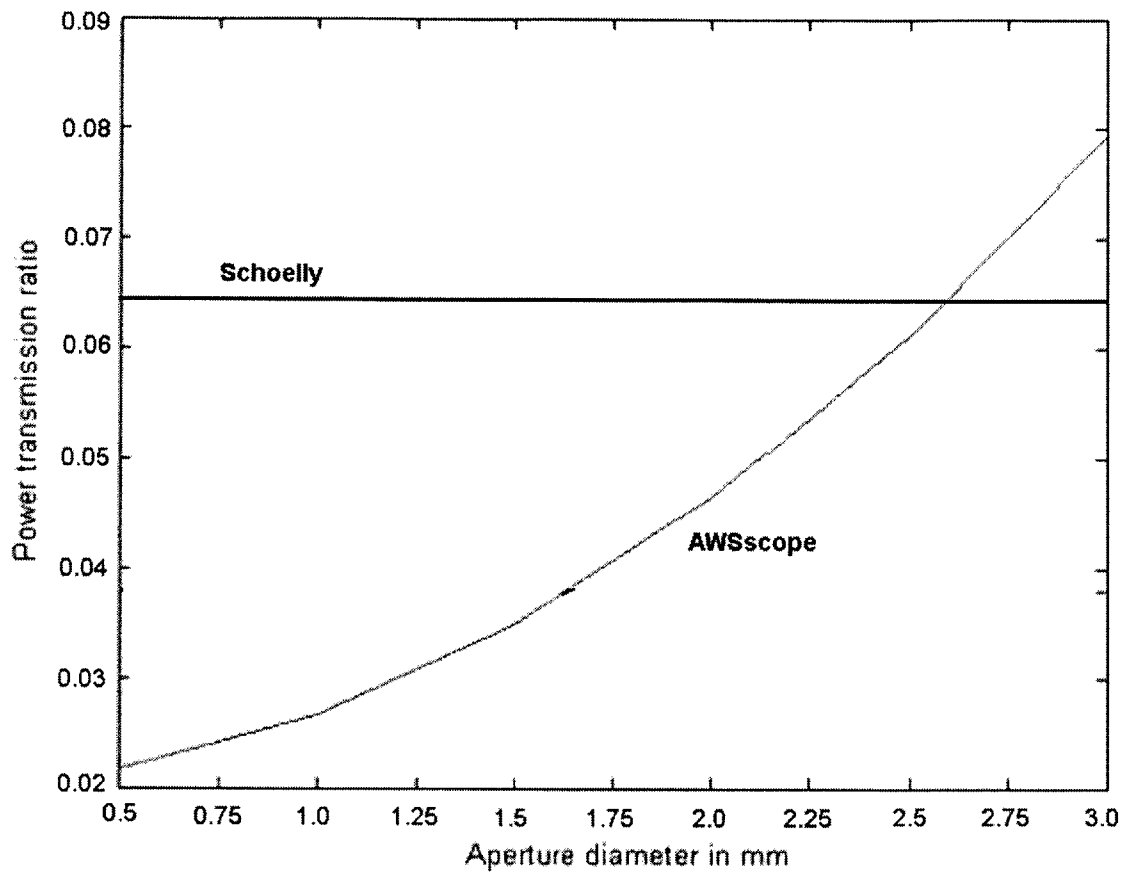
The following observations were made from the experiments:

Input Light Level = 16.8 microwatts

Schoelly scope output light level = 1.1 microwatts

AWS-scope output light level under design sampling aperture diameter of 2mm = 0.8 microwatts

Figure 4.13 summarizes the comparison of output levels of the Schoelly versus the AWS-scope at various aperture diameters.



*Figure 4.13: Transmission ratios for the AWS-scope at various sampling aperture diameters versus transmission ratio for the benchmark Schoelly.*

The comparative experiment indicates that the benchmark Schoelly is superior to the AWS-scope in its light transmission capabilities at the design condition of a 2 mm sampling aperture (figures 3.1 to 3.5). The difference in light transmission capabilities of the two systems could be partially due to poor alignment in the optical train of the AWS-scope. Finally, a slight increase in the aperture size substantially improves transmission capabilities of the AWS-scope – an observation that is consistent with the fact that power transmitted through an aperture mask increases in a manner proportional to the square of the diameter of the aperture.

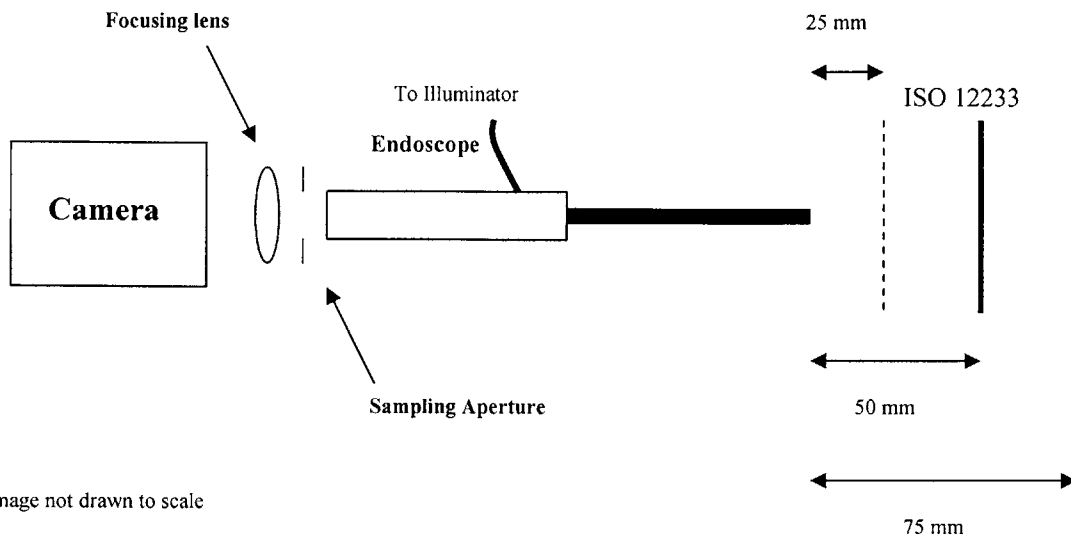
#### 4.1.4 Depth of Focus

The depth of focus of the endoscope and imaging system is the sum of the distance in front of the target object and behind it wherein the target image retains acceptable levels of sharpness. Mathematically, depth of focus depends on the numerical aperture of the imaging system and as a result, it is also related to the resolution of the imaging system. When a target object, which is in focus on the imaging sensor, is moved about its focal position without severely compromising optical line resolution, the imaging system could be characterized as possessing acceptable depth of focus in that range of movement.

For a typical endoscopic imaging system, the focal plane is positioned at a distance of 50 mm from the tip of the endoscope and the range of movement of the endoscope tip is about 25mm on either side of the focal position.

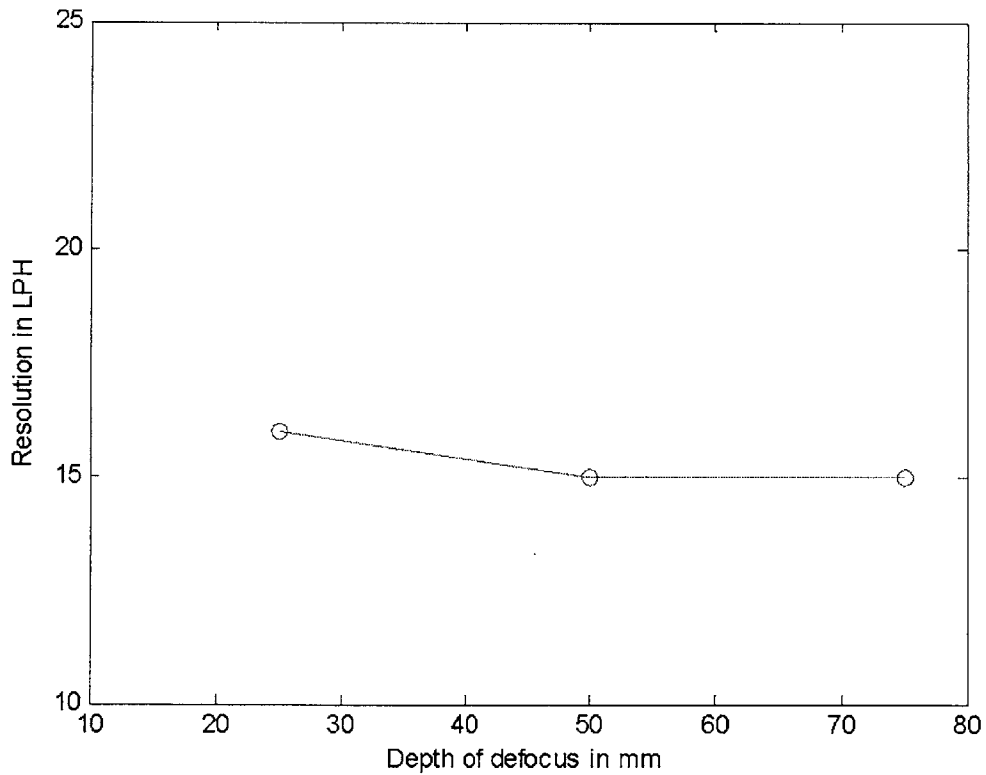
Figure 4.14 illustrates the experimental set-up used to evaluate the depth of focus characteristics of the AWS-scope. The AWS-scope was operated under design conditions with the optical line resolution target area on the ISO 12233 chart in focus at a distance of 50mm and an aperture diameter of 2mm. The chart was subsequently moved toward the endoscope tip by 25mm and imaged and then moved away from the endoscope tip by 25mm and imaged. Optical line resolutions were recorded for each of these target distances.

As can be observed in figure 4.15, optical line resolutions of about 1500 LPH remain constant over the depth of focus under consideration thereby indicating that the AWS-scope acceptable depth of focus characteristics.



Note: Image not drawn to scale

**Figure 4.14:** Schematic of the experimental set-up used to evaluate the depth of focus characteristics of the AWS-scope.



**Figure 4.15:** Optical line resolution versus target depth for a depth range of 25mm to 75mm with the focal plane positioned at 50mm from the tip of the endoscope.

## 4.2 Image Distortion

Image distortion in the AWS-scope system is a result of the interplay of a number of sources of error [48].

1. Imperfect alignment of various components of the optical train of the AWS-scope.
2. Aberrations in the imaging lens (the achromatic doublet).
3. Deformed pixels on the imaging sensor.

Distortion errors such as the ones listed above can be corrected by calibrating the right eye position and the left eye position for the AWS-scope using a three step procedure.

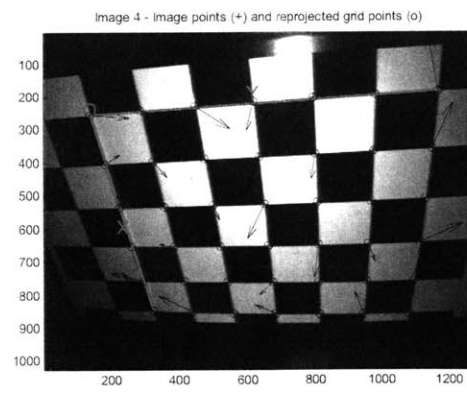
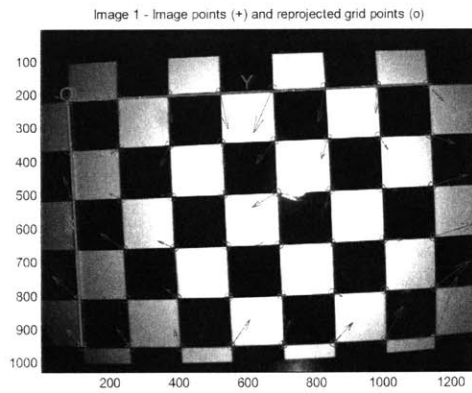
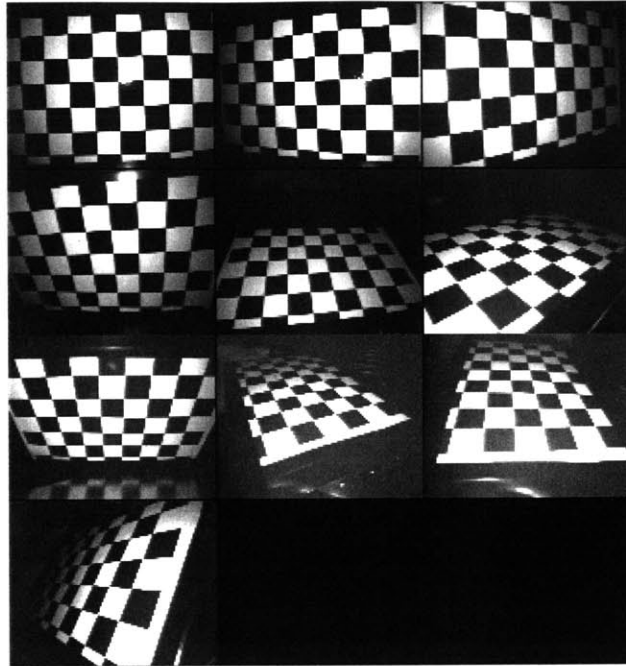
1. Calibration is performed by imaging a known target (a checkerboard pattern in this case) which is placed at a distance of 50mm from the tip of the endoscope.
2. Deviations from this known target geometry are then digitally evaluated and compensated for through image processing software.
3. The distortion correction coefficients are then applied to captured raw images so as to obtain rectified images.

Figures 4.16 to 4.19 illustrate the procedure of calibration of the AWS-scope for one particular position of the sampling aperture. As can be observed, calibration can potentially be performed on the AWS-scope so as to correct distortions up to sub-pixel accuracies.

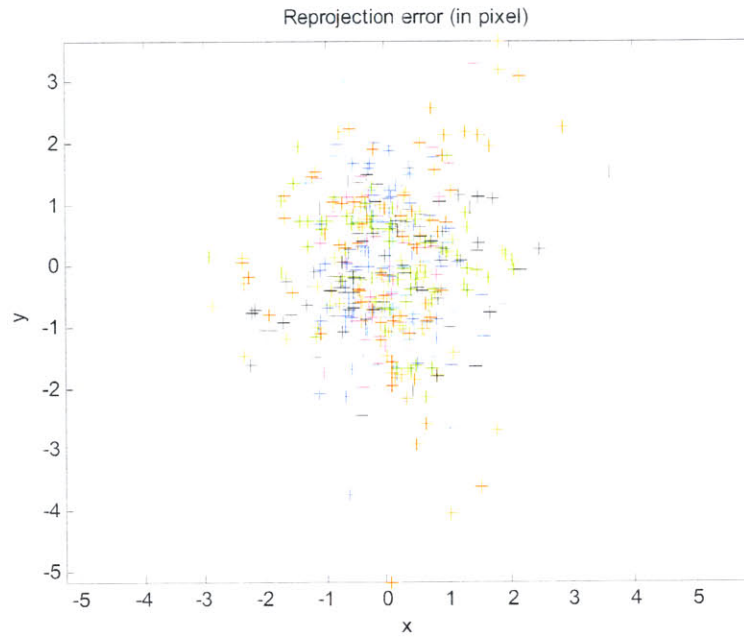
Appendix A provides greater detail on the distortion coefficients and other mathematical parameters associated with this particular calibration exercise.



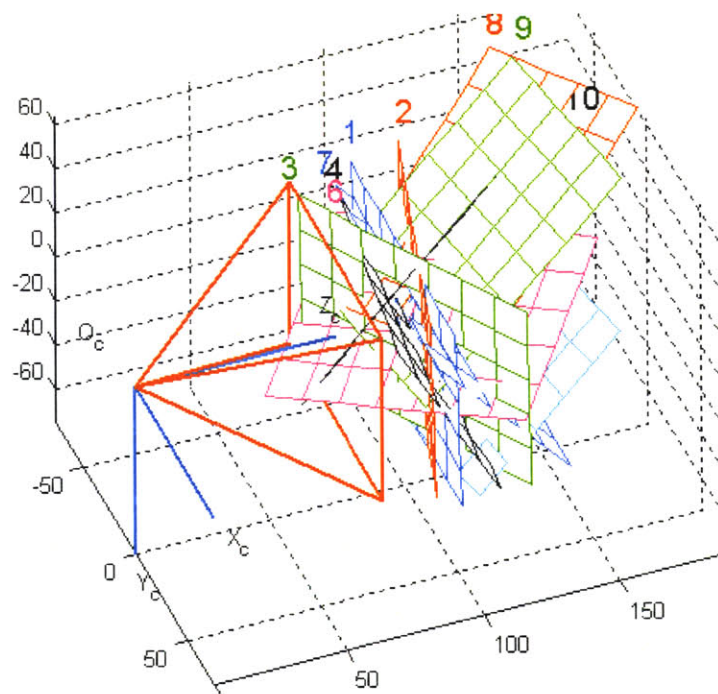
Calibration images



*Figure 4.17: The Tsai Camera Calibration Software [43] was used to digitally reproject points onto the known calibration target. Any mismatch in tracking the checkerboard pattern is compensated for as shown.*



**Figure 4.18:** A consolidated plot that shows the re-projection error for all captured orientations of the calibration target. It can be observed the most re-projections have occurred to sub-pixel accuracies in this particular calibration exercise.

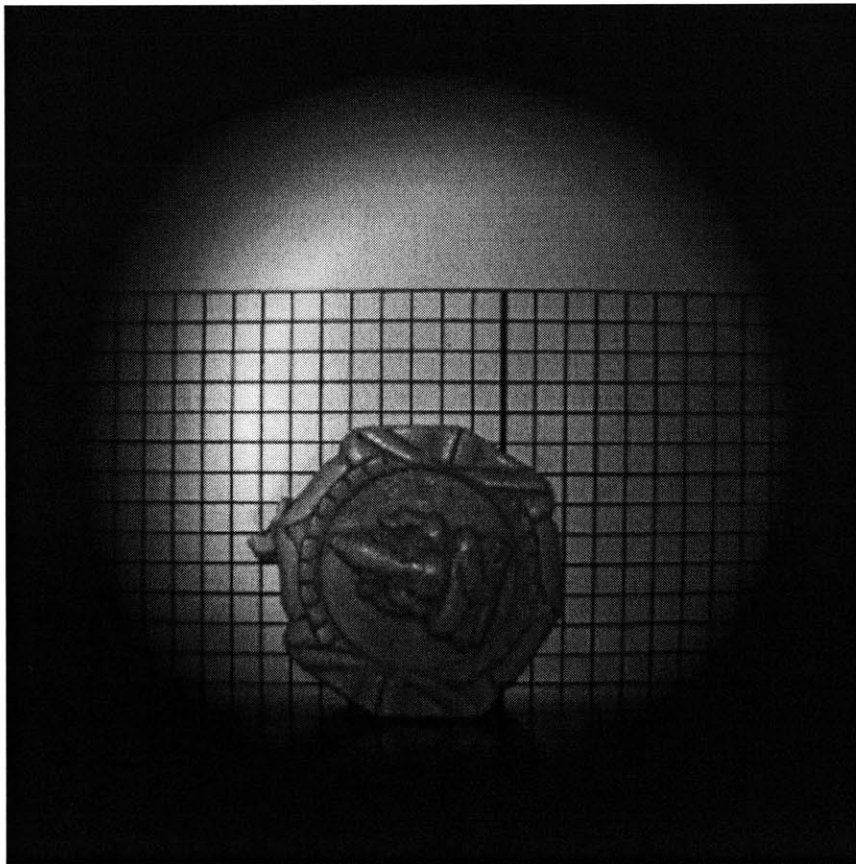


**Figure 4.19:** Using the calibration data to reproduce the position and orientation of the input calibration target data. The various planes numbered 1 to 10 are the calibrated orientations of the calibration target images that were captured.

### 4.3 Vignetting in Images

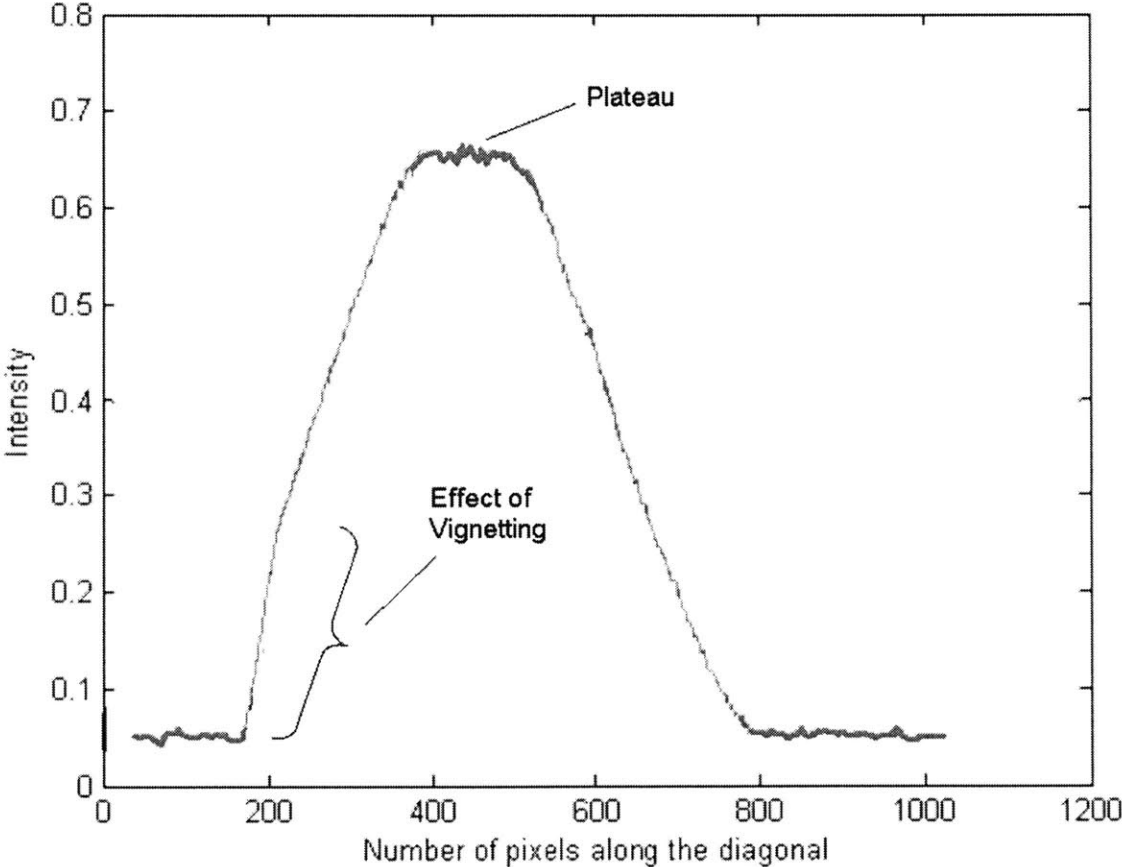
Vignetting is the phenomenon due to which a captured image does not exhibit a smooth fall-off in its intensity levels. Vignetting results primarily due to light rays being blocked from reaching the imaging sensor by one or more obstructing elements in the optical path of the imaging system. Vignetting is a highly undesirable effect in stereo-endoscopy because of its potentially distracting consequences when fusing the stereo pairs.

Figure 4.20 shows an image captured using the AWS-scope and the Dalsa CCD camera in the underfill configuration. As can be observed, the image exhibits moderate vignetting and intensity variations.



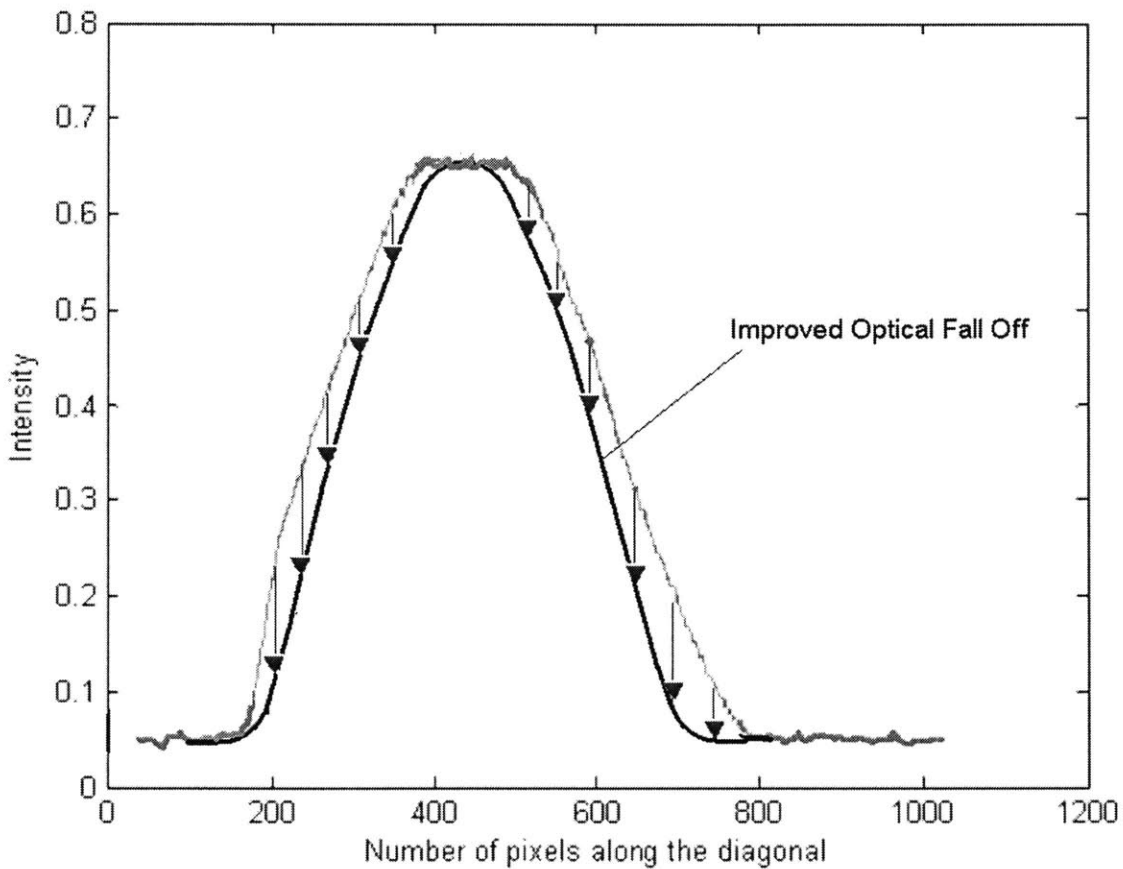
*Figure 4.20: A raw image captured using the AWS-scope. The image exhibits moderate to severe intensity variations besides vignetting at its periphery.*

It is instructive to plot the intensity variation plot for image in 4.20 along its diagonal so as to understand the effect of vignetting on optical fall-off. Figure 4.21 below illustrates the optical fall-off for image 4.20. As can be noted, vignetting at the periphery distorts the curve from having a smooth fall-off. The image is also ‘washed out’ around its central spot as can be noticed by the small ‘plateau’ of maximum intensity.



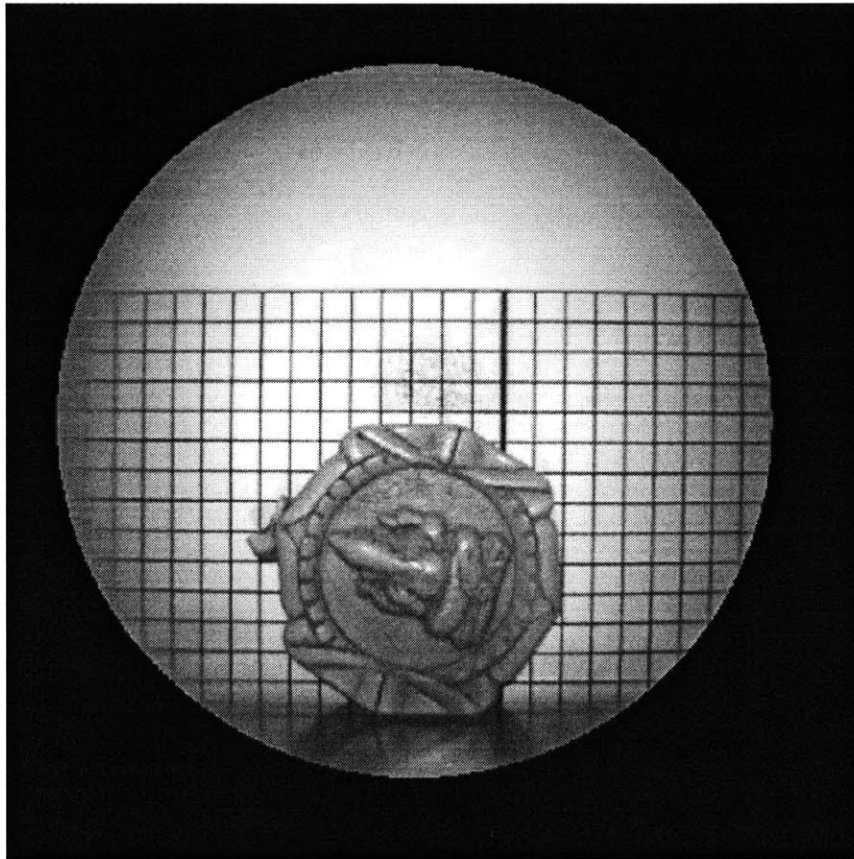
**Figure 4.21:** Intensity fall-off plot for the image in figure 4.20. Notice the effect of vignetting at the left hand corner of the image and the effect of ‘washout’ in its center.

The vignetting observed in figure 4.20 can be substantially corrected by improving the quality of optical fall-off. Appendix B demonstrates one such approach using software for digitally modifying optical fall-off from an irregular one to a smooth one. The technique fits a Gaussian curve to the average maximum and average minimum levels of the image under consideration and centers the curve at the mid-point of the plateau of maximum intensity. Figure 4.22 illustrates the effect of this technique on the fall-off represented in 4.21. The blue curve represents the new fall off curve which possess significantly superior intensity characteristics. The arrows indicate the adjustments effected on the pre-existing fall-off curve in order to obtain the new one.



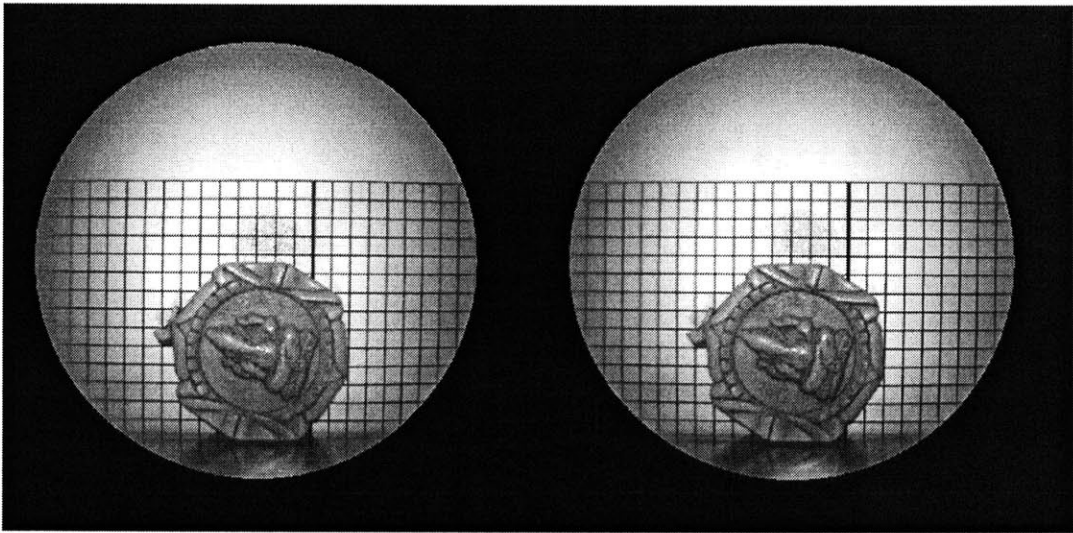
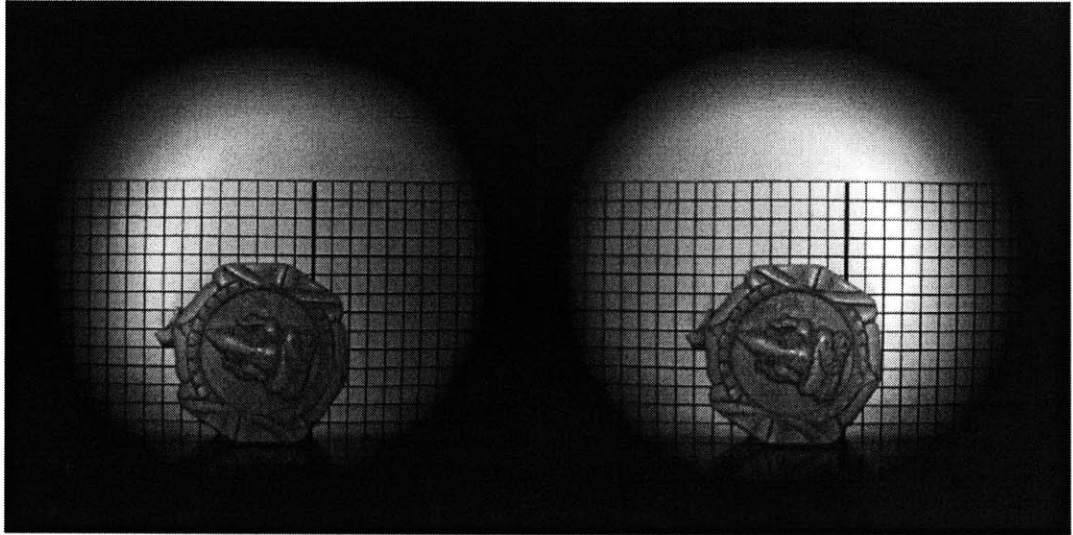
**Figure 4.22:** Correcting the irregular intensity fall-off to a smooth one digitally.

Figure 4.23 shows the effect of implementing the aforementioned intensity fall-off transformation on the image in figure 4.20. As can be observed, the quality of image has improved significantly with several portions of the image previously invisible now clear.



*Figure 4.23: Processed image corrected out for intensity variations. Notice the marked improvement in visual quality of the image.*

Finally, it would be useful to observe the marked improvement in quality of stereo when the vignetting correction technique is implemented on a stereo pair. Figure 4.24 illustrates this technique on a stereo pair.

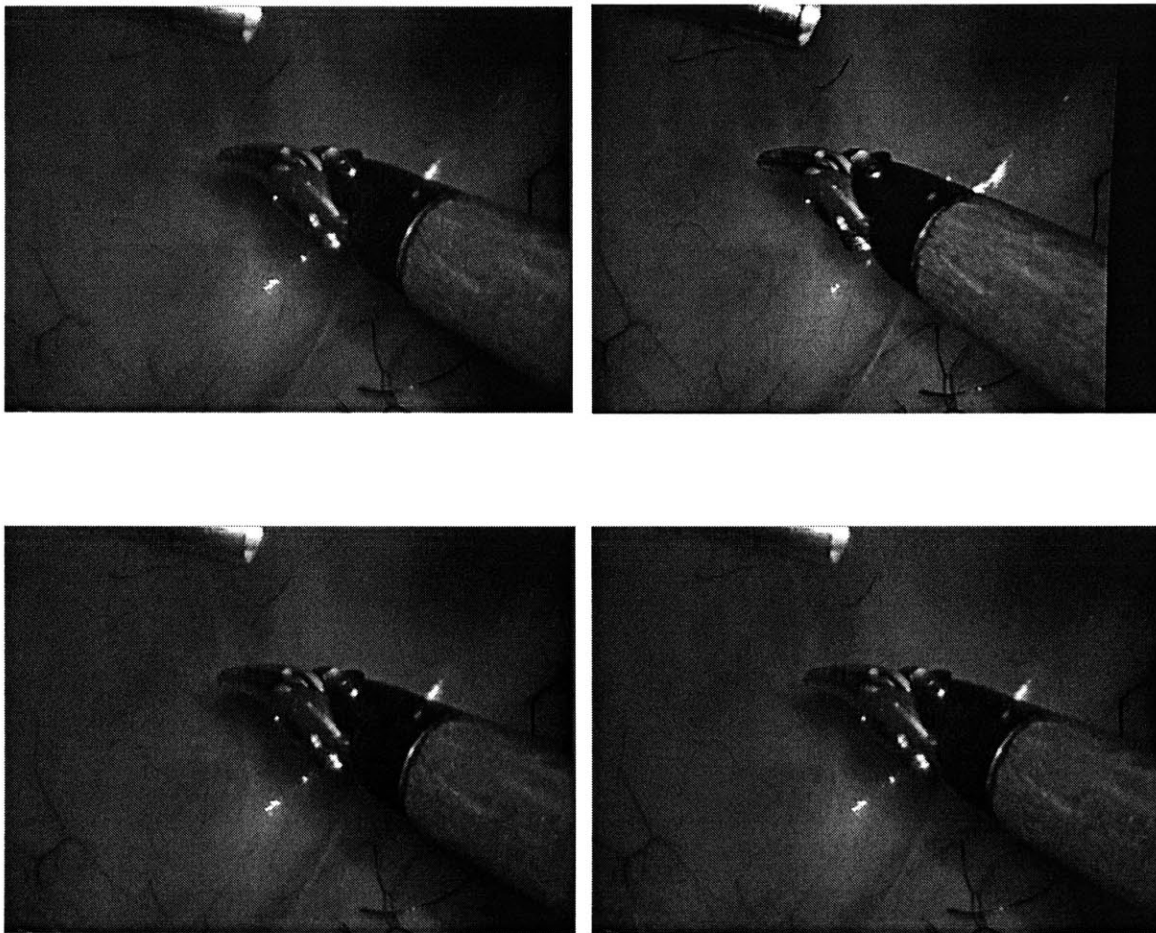


*Figure 4.23: Top figure shows a raw stereo pair with intensity variations and vignetting that are extremely distracting when fusing the stereo. Bottom figure shows a much improved processed image pair corrected out for intensity variations.*



## CHAPTER 5: QUANTIFIED 3D IMAGING AND THE PERCEPTION OF DEPTH BY THE HUMAN EYE

As discussed in chapter 2, a benchmark commercial stereo-endoscope system such as the Schoelly endoscope is composed of two lenses and cameras separated by a baseline on the order of 6mm while an AWS system made with an identical lens and camera might only have a sampling diameter of 2mm. All other things being equal, therefore, an AWS-scope system using only two sampling positions would only have 1/3<sup>th</sup> the depth sensitivity of the stereo system. Figure 5.1 below illustrates this scenario.



*Figure 5.1: The stereo pair on top is captured using the Schoelly while the stereo pair on the bottom is representative of a stereo pair captured using the AWS-scope. Notice the significant difference in disparity sensitivity perceived between the two pairs.*



Qualitatively, equations 5.5 and 5.6 can be interpreted as follows:

- 1) For a stereo pair (with positive convergent parallax) placed at distances close to the eye such that the second term in equation 5.5 is dominant, we have both good sensitivity to disparity in depth as well as disparity in convergence. Therefore, theoretically if the human eye could exhibit convergence up to 90 degrees, it would be ideal to view things as close up as possible.
- 2) As the screen is moved away, convergence decreases for the human eye and so does sensitivity of disparity to changes in depth. However, disparity sensitivity to convergence first decreases to a minimum and then rises with a slope of unity. Intuitively, at large distances, small changes in convergence will cause large changes in rate of change of disparity because of a longer 'lever arm' (longer focal length by virtue of accommodation) multiplying the effects of the change.

The aforementioned interpretations seem to agree with the mechanism of sight in human beings. For objects placed very close to the eye, as long as the eyes can accommodate and generate the necessary vergence to fuse the left and right images, a very good sense of depth is experienced assisted by large angles of vergence and good disparity sensitivity to depth. This is the same reason why it is possible to thread a needle at close quarters. However, as we keep moving away, the ability to perceive small changes in depth (which is directly related to the vergence angle) drops off rapidly due to decreased vergence. In fact we lose the so called parallax effect when the distance of the object goes to infinity as the parallax (and sensitivity of parallax to depth) then goes to zero. This happens for

In reality, the difference in disparity sensitivity of the Schoelly and the AWS-scope is higher than the expected 200% due to inherent construction limitations of off-the-shelf monoscopes and custom-made stereoscopes. This difference in performance would at first seem to rule out the use of AWS-scope in favor of the Schoelly, but the comparison requires further examination. It must be remembered that the larger baseline present in a stereo-endoscope comes at the cost of greater occlusion (which is when a target feature is visible in one camera but not in the other), higher hardware cost and greater difficulty with calibration. Though it is certainly true that the Schoelly may be on the order of three times more sensitive to depth than our laboratory built AWS-scope that uses two sampling positions, this performance differential can be significantly reduced by using a combination of an increased number of sampling positions and robust algorithms that can track target features from one image to another. In other words, the performance of the AWS-scope can be digitally enhanced to generate synthetic stereo which exhibits disparity sensitivity characteristics similar to those of the Schoelly endoscope. Before delving into the details of generation of synthetic stereo however, it is instructive to understand the key drivers of depth perception in human vision especially with respect to disparity sensitivity. These drivers would enable the creation of better informed synthetic stereo image pairs.

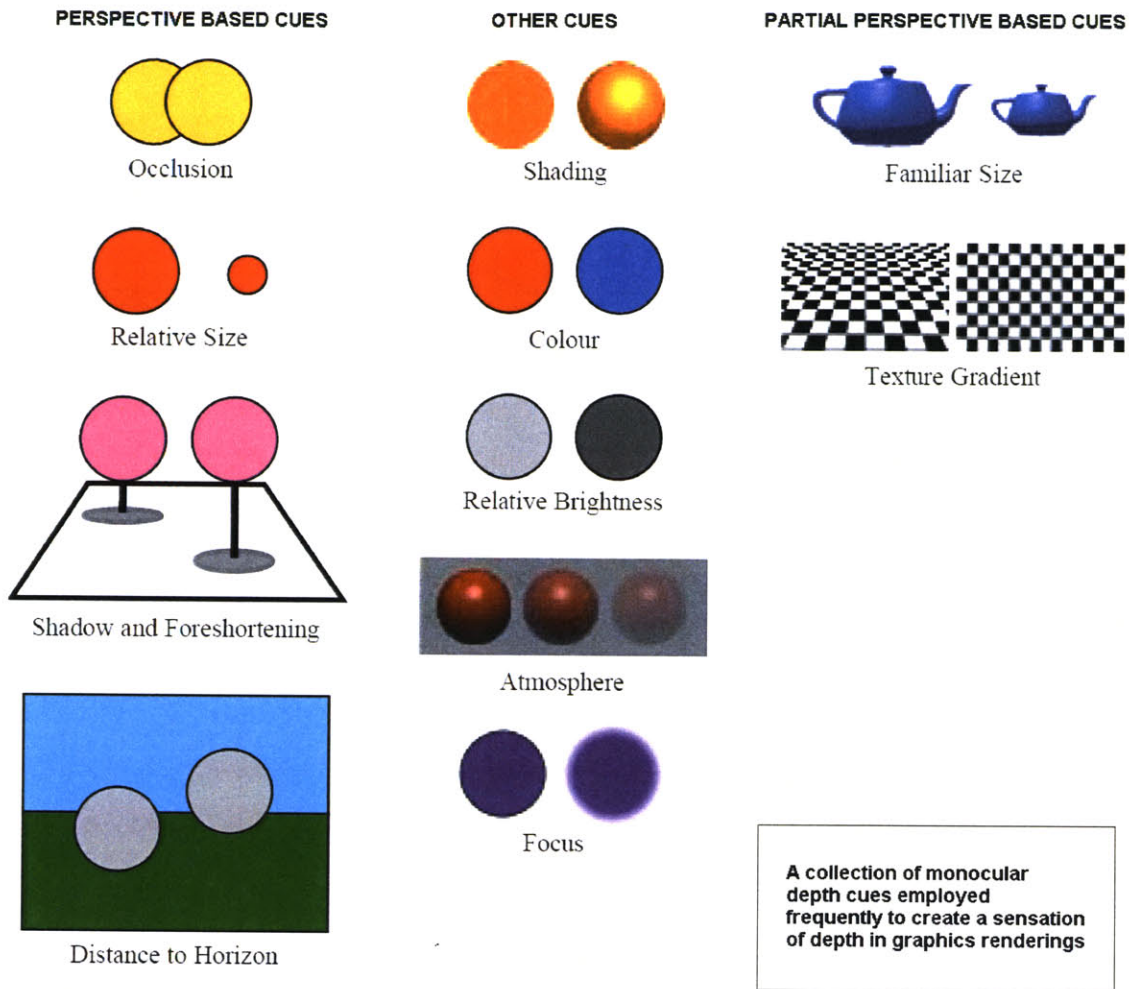
## **6.1 Depth Perception**

The world we live in is three dimensional. We see this world through images projected in our retinas, and although the world is three dimensional, the images themselves are flat. The way we create depth perception from flat images is through depth cues. Depth cues are different sources of information that combine to give a viewer the 3D layout of a

scene. There are two primary categories of depth cues – Pictorial, Oculomotor and Binocular depth cues [37].

### **6.1.1 Pictorial Depth Cues**

Pictorial Depth cues or monocular depth cues are 2D sources of information that the viewer interprets as three-dimensional. These are best experienced if we close one eye and see the world with the other. The world does not appear to be flat like a picture, and we still have a perception of depth, although it is more difficult to judge distances. This means that there are other sources of depth information that exist on a single image. These sources of information are called *monocular visual cues*. Notwithstanding the fact that these are powerful cues, by nature they could give rise to significant ambiguities. Figure 5.2 illustrates some of the more common pictorial depth cues encountered.



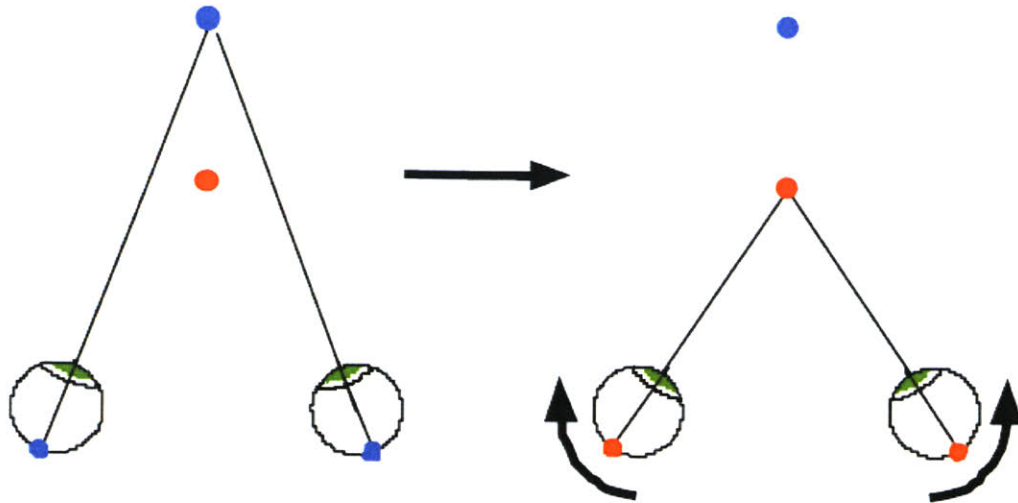
*Figure 5.2: The various types of pictorial or monocular depth cues. [37]*

### 6.1.2 Oculomotor and Binocular Depth Cues

Oculomotor depth cues include Convergence and Accommodation.

#### Convergence

The human eyes are connected to muscles that allow them to rotate in their sockets. When we look at an object, the muscles contract and force the eyes to converge and look directly at the object. The closer the object is to the eyes, the more these muscles have to contract to converge the eyes. Figure 5.3 illustrates Convergence.



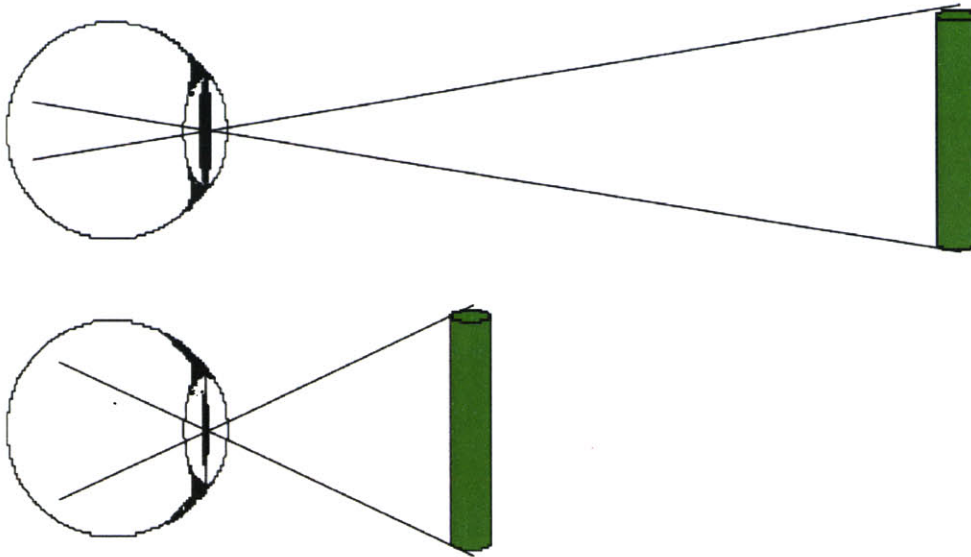
*Figure 5.3: Convergence results in the rotation of eyeballs aligning optical axes toward target object of focus. [37]*

### **Accommodation**

Besides the muscles that move the eyes, there is a set of muscles attached to the lens that make it change shape to focus on objects at different distances. When the object is far away the muscles relax, and the lens become more spherical. When the object is nearby, the muscles pull on the edges of the lens making it flatten out. This capability of the human eye is called Accommodation. Figure 5.4 illustrates Accommodation.

### **Binocular Depth**

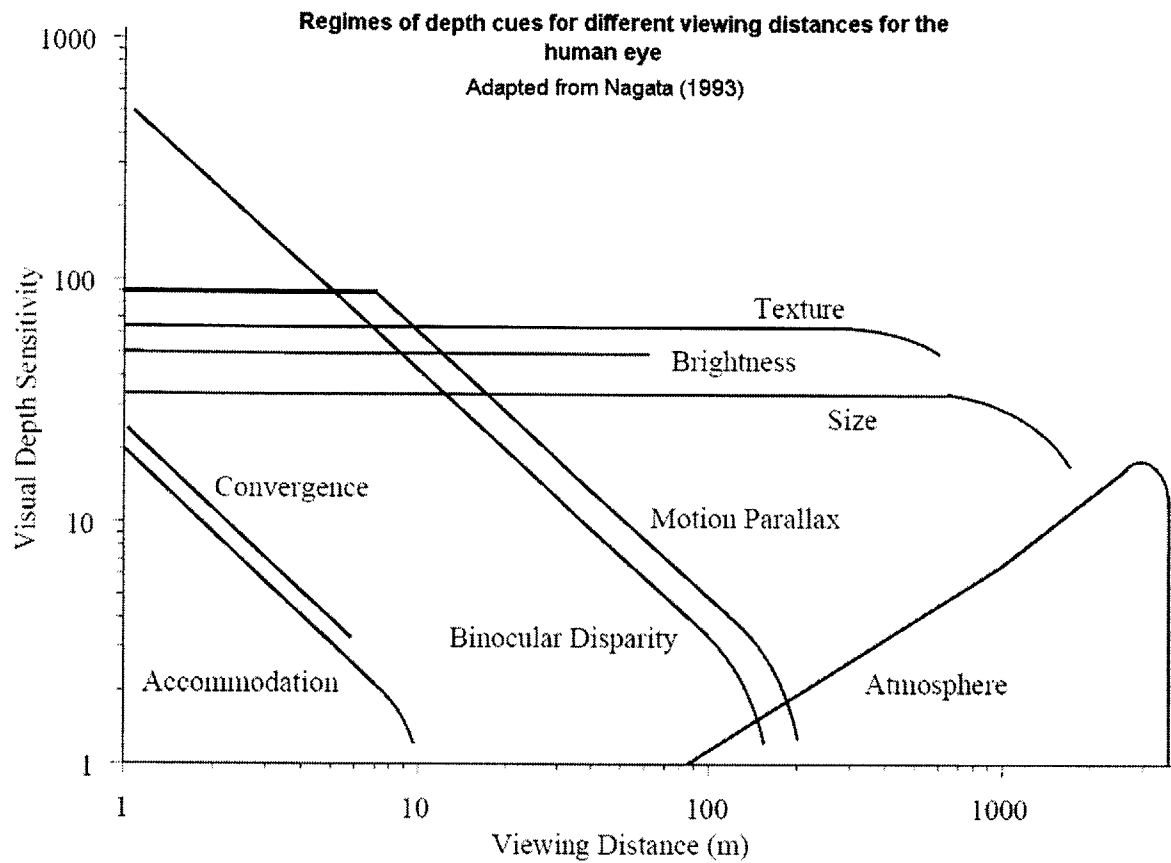
Binocular depth perception is generated by the 55 to 60 mm baseline separation between the two eyes which generates disparity between images observed by the left and right eyes. Binocular depth perception works typically in conjunction with Oculomotor cues to generate the sensation of depth.



*Figure 5.4: Accommodation is the ability of the eye lens to change shape and focus on objects at different distances. [37]*

### **6.1.3 Interplay between depth cues**

In addition to the depth cues presented above, other cues such as motion parallax etc. combine together to produce powerful three-dimensional effect in scenes. However, intuitively it is well known that the most effective (and disambiguating) depth cue is binocular stereo. In fact, much study has been performed to identify regions where different cues dominate. For example, a person threading a needle primarily uses binocular depth cues with strong vergence and accommodation to increase accuracy of perception. On the other hand, a truck driver on a highway will look at non-stereo cues for sensing distance to the vehicle ahead. In general, an important criterion for the dominance of one depth cue over another is the distance from the viewer to the object of interest. Figure 5.5 illustrates the dominance of different depth cues at different distances from the observer to the target of interest.



**Figure 5.5:** Depth perception cues that dominate at different distances of the target object from the eye. The Y axis is a normalized scale for comparing the relative importance of the various sources of depth perception. [50]

Figure 5.5 clearly indicates that in the case of endoscopy, wherein the viewing distances are of the order of a few centimeters, binocular disparity combined with convergence and accommodation plays the predominant role in the perception of depth. Subsequent sections of this chapter will attempt to develop an understanding of the variation of disparity sensitivity perceived by the human eye with binocular effects, convergence and accommodation [47]. While stereo imaging without convergence has previously been described in chapter 2, convergence, as shall be observed parametrically, significantly affects depth perception.

## 6.2 Convergence and Vision

The brain uses convergence to simultaneously move the two eyes in opposite directions to maintain binocular vision. When the human visual system observes an object which is nearby, the eyes rotate toward each other while they move out for a far-off object. In doing so, the eyes maintain positive convergent parallax (parallax is described later in this chapter). Exaggerated convergence is called negative parallax as the optical axes cross one other. When the eyes stare at infinity or into nothingness, the optical axes are parallel and the parallax is said to have zero parallax.

A geometrical model of the human visual system with accommodating convergence can be constructed as shown figure 5.6.

From figure 5.6 using similarity of triangles,  $p$  can be obtained as

$$p = \frac{Z_{xy}}{\left(\frac{e}{D} - 1\right)} \quad 5.1$$

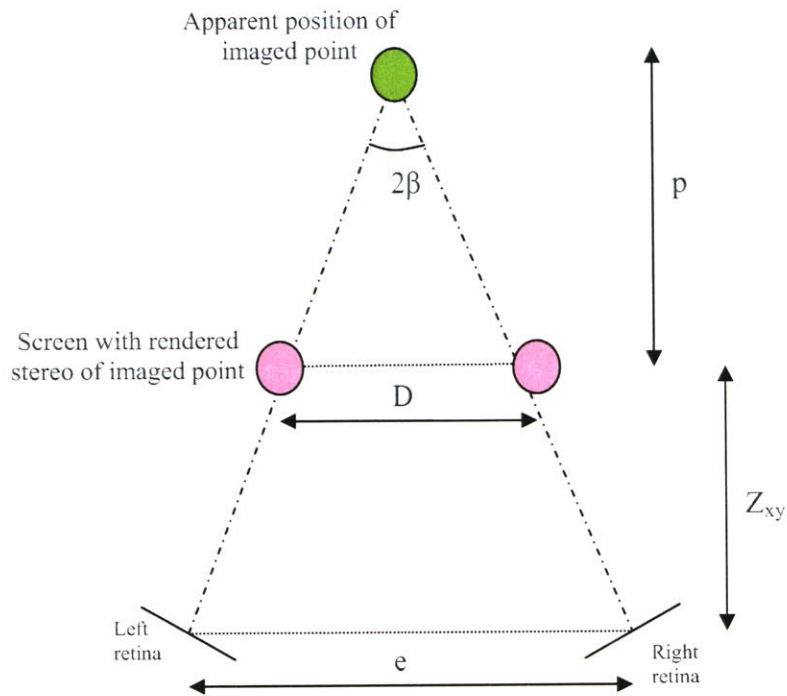
The angle of convergence can be written as,

$$\tan \beta = \frac{D}{2p} = \frac{e - D}{2Z_{xy}} \quad 5.2$$

Combining equations 5.1 and 5.2, the parallax (which is representative of disparity) can be expressed as,

$$D = e - 2Z_{xy} \tan(\beta) \quad 5.3$$





$\beta$  : The angle of vergence

$D$  : Parallax of the rendered stereo on screen (representative of disparity)

$e$  : Separation between the two eyes

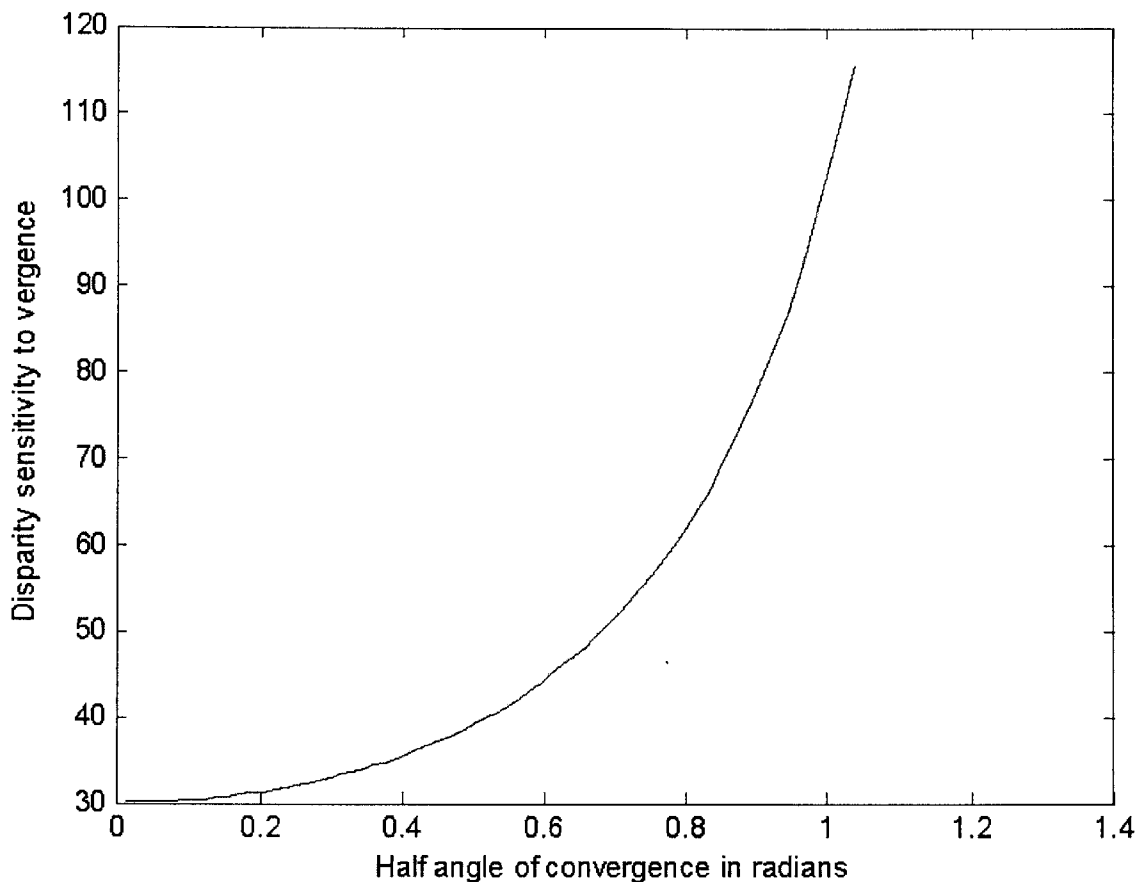
$Z_{xy}$ : Distance from eyes to screen on which stereo is displayed

**Figure 5.6:** Schematic of the geometric model used to illustrate the perception of disparity by the human visual system under convergence, accommodation and binocular effects.

The first derivative of equation 5.3 with the convergence angle yields the sensitivity of parallax to changes in the convergence angle  $\beta$ .

$$\left| \frac{dD}{d\beta} \right| = Z_{xy} \sec^2(\beta) \quad 5.4$$

Figure 5.7 illustrates the variation of disparity sensitivity with convergence angle for fixed distances of the eyes to the screen. As can be observed, the variation is a strong increasing function of the convergence angle.



**Figure 5.7:** *Convergence vs. Disparity sensitivity to convergence for a typical human visual system at a fixed distance to screen of 50mm*

Figure 5.7 indicates that large angles of vergence yield better sensitivity to depth perception for a fixed observer and screen position. While this seems intuitive enough, an interesting question that needs to be addressed is the issue of increasing or decreasing angle of vergence. In fact one of the levers for increasing vergence is the distance between the observer and the screen. If this distance is fixed, from equation 5.2 the angle of vergence can be increased or decreased by increasing the baseline ( $e$ ) between the two

eyes or by decreasing the parallax (D) between the rendered images. Since e is fixed (about ~60 mm for human eyes) , vergence angle can be improved by decreasing parallax (D) which could possibly be achieved to some extent by scaling down the rendered images.

A more general approach to understanding disparity sensitivity to vergence would be to accommodate the fact that (e-D) remains a constant and  $Z_{xy}$  varies and thereby controls vergence as well as sensitivity to vergence. When  $Z_{xy}$  varies, the sensitivity to vergence angle can be rewritten as,

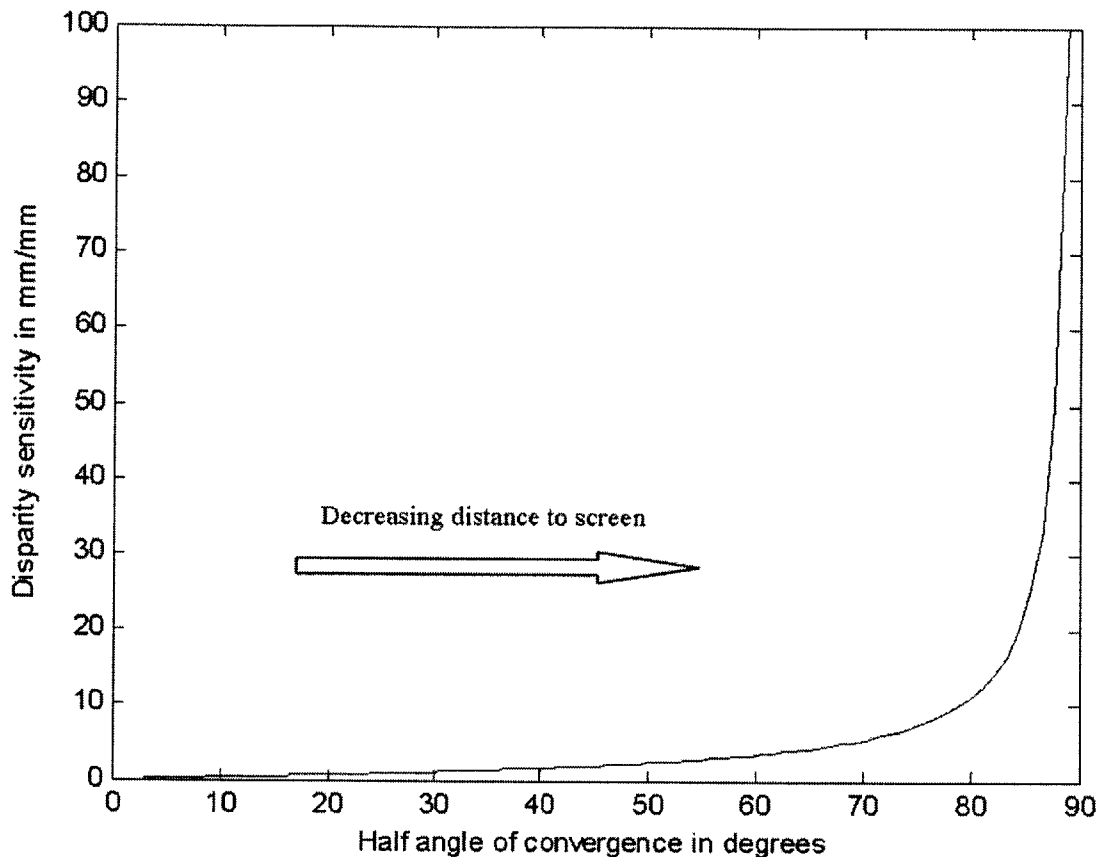
$$\left| \frac{dD}{d\beta} \right| = Z_{xy} + \frac{\left( \frac{(e-D)}{2} \right)^2}{Z_{xy}} \quad 5.5$$

The variation of disparity sensitivity with depth from the screen can be expressed as,

$$\left| \frac{dD}{dZ_{xy}} \right| = 2 \tan(\beta) = \frac{(e-D)}{Z_{xy}} \quad 5.6$$

In order to fulfill the objective maximizing disparity sensitivity to depth as well as angle of convergence, the two equations above must ‘supplement’ each other and not ‘trade-off’ with one another. In addition, from the nature of the two equations, it is evident that if the hyperbolic term in the sensitivity to convergence equation is dominant, the two sensitivities ‘move in the same direction’.

the human eyes at distances of 30 feet. Figure 5.8 illustrates disparity sensitivity to depth with changing angle of convergence.



**Figure 5.8:** Convergence vs. Disparity Sensitivity to Depth for the human visual system for changing target object distance.

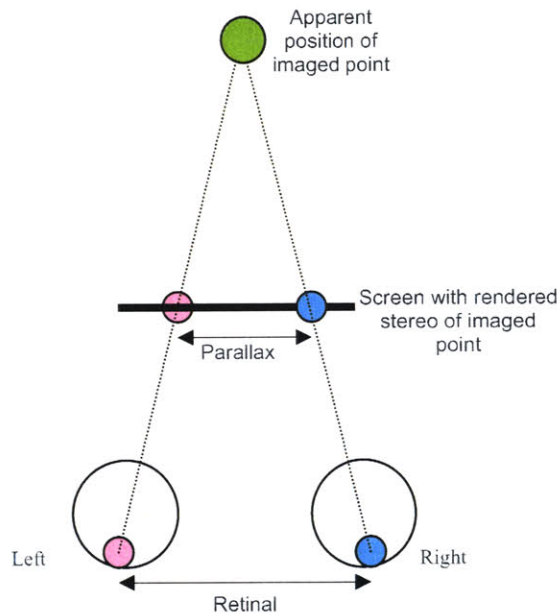
As can be observed, sensitivity drops-off rapidly in a non-linear fashion with increasing distance from screen.

## 6.3 Display of stereo images for effective binocular vision

The preceding analyses tried to address the factors affecting quality of depth perceived by the human eye. The following section attempts to throw clarity on the notion of parallax and its role in depth perception [37]

### 6.3.1 Parallax

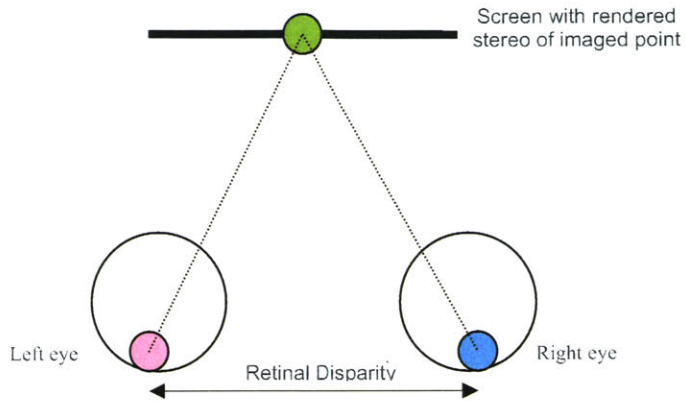
Consider figure 5.9 below.



**Figure 5.9:** Schematic illustrating positive convergent parallax. By virtue of its definition, parallax is representative of the disparity perceived by the human visual system

As can be seen, parallax (produced on the screen) produces disparity (measured on the retina) which in turn produces stereopsis (which creates the sensation of depth). Parallax can be classified into four categories.

### 6.3.2 Zero Parallax

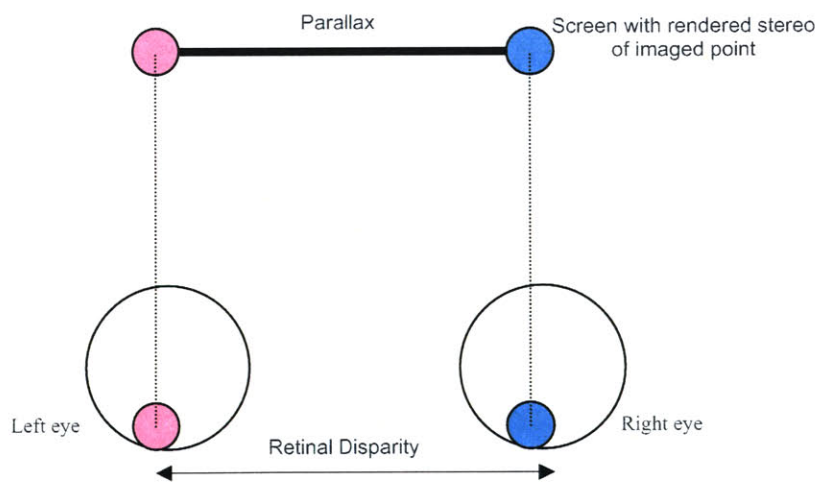


*Figure 5.10: Schematic illustrating zero parallax.*

As is evident from figure 5.10, when the eyes observe an object with zero parallax, the optical axes of the eyes cross on the plane of the screen and the eyes are said to converge at the screen to a zero parallax setting.

### 6.3.3 Positive Parallax

Positive Parallel Parallax



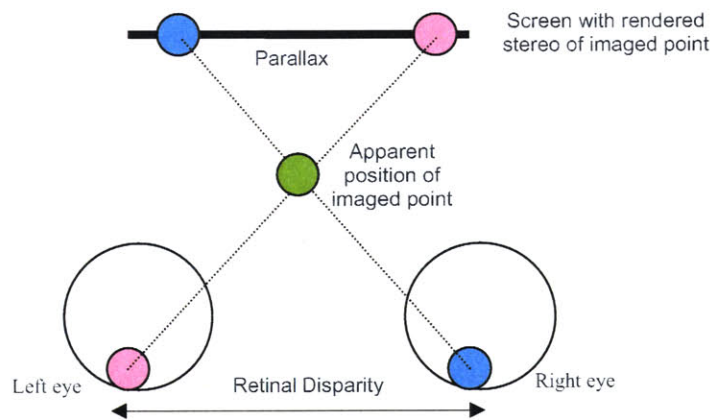
*Figure 5.11: Schematic illustrating positive parallel parallax.*

Positive parallel parallax (figure 5.11) is experienced when the eyes try to look at objects at a great distance thereby making their optical axes parallel. In the case of looking at a generated stereo scene, this form of parallax is experienced when the parallax on screen is equal to the separation between the eyes. This form of display produces a lot of discomfort especially on small screens.

Positive Converging Parallax

Figure 5.9 represents the case of positive converging parallax wherein the axes of the eyes cross beyond the screen with the value of parallax on screen being inferior to the separation between the eyes. This is the most effective and comfortable form of parallax

**6.3.4 Negative Parallax**

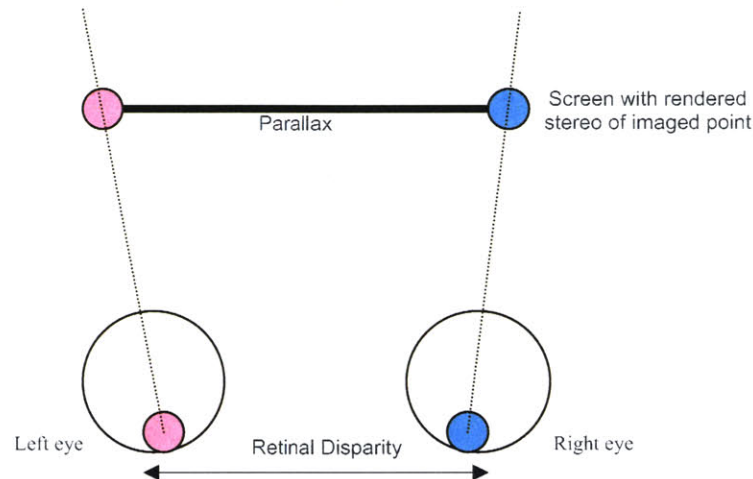


*Figure 5.12: Schematic illustrating negative parallax.*

Under the condition of negative parallax (figure 5.12), the optical axes of the eyes are crossed in front of the screen to produce an apparent point which appears to be ‘floating’ in the viewer space.

### 6.3.5 Divergent Parallax

Divergent parallax is experienced when the separation between the points is larger than the separation between the eyes resulting in optical axes which are 'diverging' (figure 5.13). The divergence does not appear when viewing real world objects. The unusual muscular effort necessary to fuse such image pairs causes immense discomfort and thereby destroys quality of stereo.



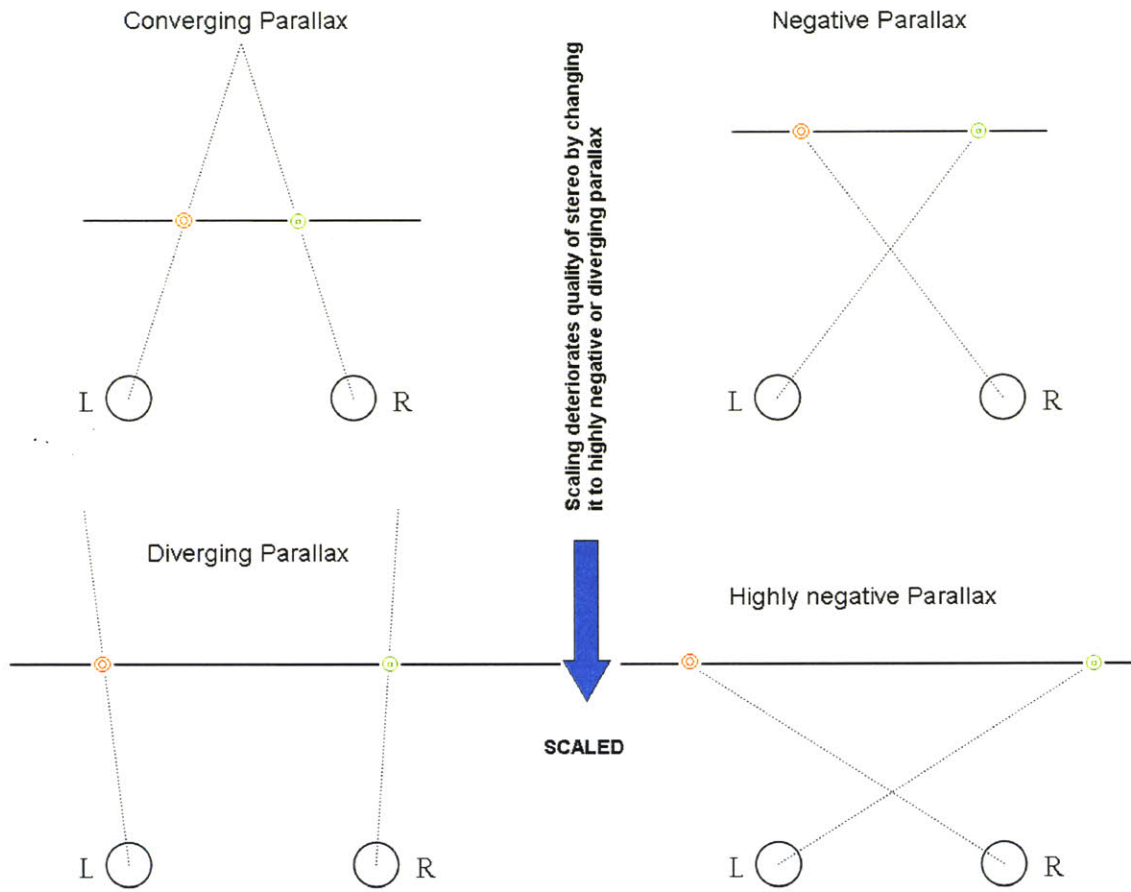
*Figure 5.13: Schematic illustrating divergent parallax.*

Parallax and scaling of stereo images have direct correlation in deciding the quality of stereo. The schematic in figures 5.14 and 5.15 illustrate this correlation. As explained in the previous sections, it is preferable to have positive converging or moderately negative parallax. However, if the parallax changes to diverging or highly negative parallax, the eyes find it almost impossible to fuse the stereo.

As shown in figure 5.14, as the scaling is gradually increased, there reaches a point where the retinal disparity falls below parallax (leading to diverging parallax) or extensive crossing of optical axes before screen (leading to strongly negative parallax).

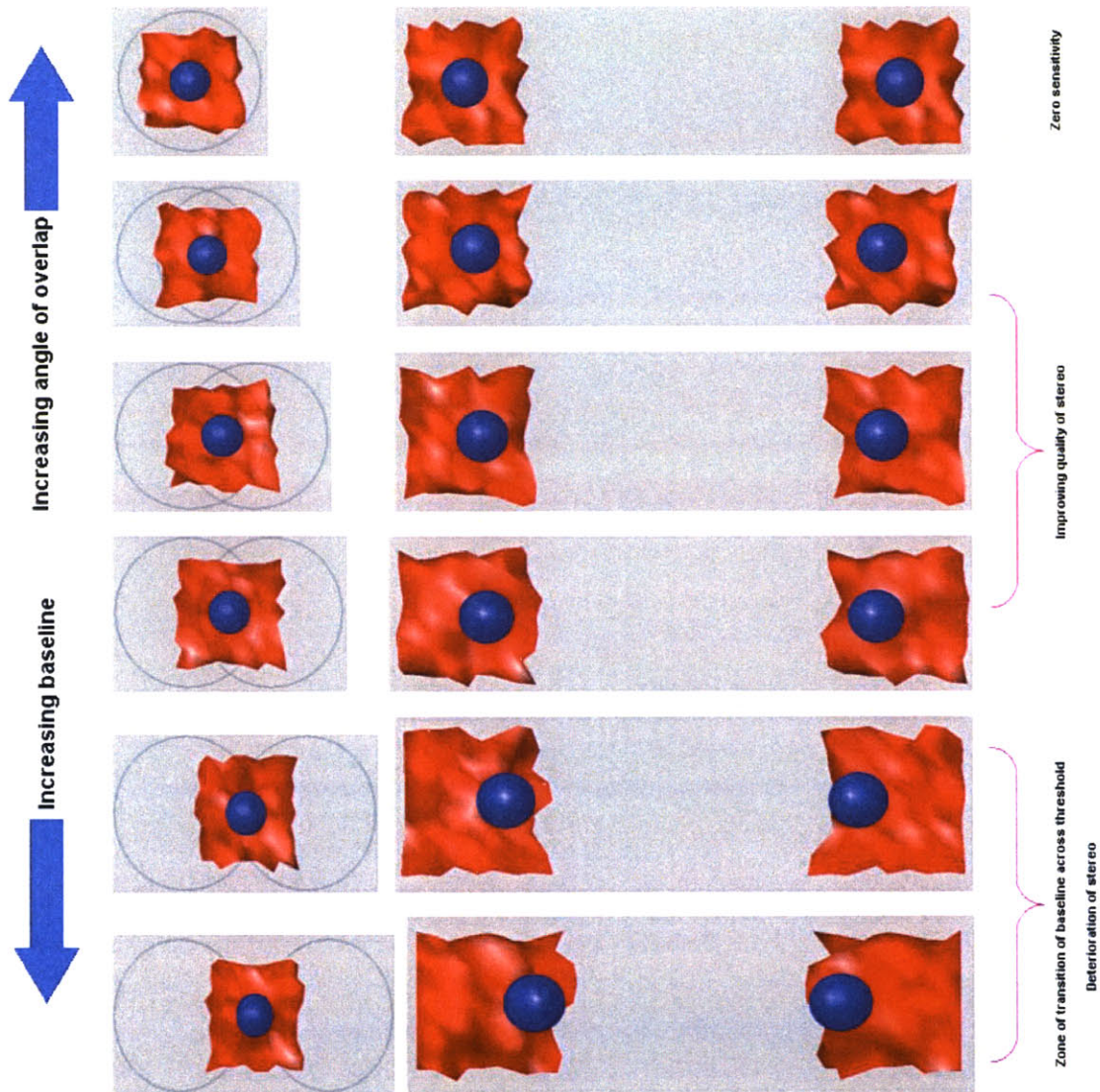


Similarly, if scaling is gradually decreased, diverging parallax can be transformed to positive converging parallax which highly preferable.



**Figure 5.14:** Schematic illustrating the effect of scaling of stereo image pairs on parallax.

Figure 5.15 illustrates the effect of increased baseline on perceived stereo. As indicated on the figure, the stereo effect improves strongly with increased baseline before dramatically deteriorating due to a change in the nature of parallax from positive to negative to ultimately divergent parallax.



**Figure 5.15:** A synthetically generated schematic of a blue sphere on an irregular red surface illustrating the effect of baseline on the overlap of the fields of view of the two eyes thereby resulting in the changing nature of parallax.

## 6.4 Design Guidelines for Synthetic Stereo Generation

Based on the discussions on the parametric variation of disparity sensitivity perceived by the human eye and the notion of parallax a set of simple design rules can be listed when generating synthetic stereo. The rules have been listed in the decreasing order of priority.

1. In generating synthetic stereo from a surface model, the key driver of sensitivity is the baseline used for the purposes of projection. While a larger baseline is preferable, beyond a limit the quality of depth perception degrades dramatically due to a change in the nature of parallax to divergent parallax.
2. Increased base line at closer target distances translates to greater rotation of the eye-balls. As a result, it is important to adjust the convergence of the projection line so as to ensure greater convenience for the observer perceiving depth through stereo-fusion.
3. Finally, the nature of convergence as well as the baseline should be adjusted to yield positive or zero parallax. The acceptable minimum parallax should be set at slightly negative parallax.

In Chapter 6, these aforementioned guidelines are applied to synthetically generate high quality (high disparity sensitivity) stereo from raw stereo which does not in and of itself demonstrate high quality depth perception visually.

## CHAPTER 6: GENERATING SYNTHETIC STEREO WITH THE AWS-SCOPE

Chapter 5 outlined reasons for the need to generate synthetic stereo. In addition, it addressed the question of what constitutes good quality stereo under imaging conditions prevalent in endoscopy. This chapter attempts to implement ideas discussed in chapter 5 through the generation of visually enhanced synthetic stereo.

Synthetic generation of stereo involves three major steps:

- 1) Raw images are captured using the AWS-scope system at several angular positions of the sampling aperture. Using error-correlation coupled with optical flow techniques, these raw images are correlated to obtain a dense depth map of the scenario being imaged. The surface model of the surgical scenario is then recreated using the triangulation technique.
- 2) Keeping in mind the general design principles discussed in the concluding section of chapter 5, perspective left and right eye views of the surface model are generated.
- 3) Finally, color is mapped onto the synthetic stereo pair to recreate the original surgical scenario.

The main focus of this chapter is on steps 2 and 3 as step 1 has been successfully demonstrated previously by my colleague at MIT [6]. For the purposes of this demonstration, a coral model with a highly textured surface and a variety of coloring was



*Figure 6.1: Coral scene used for demonstrating generation of enhanced synthetic stereo.*

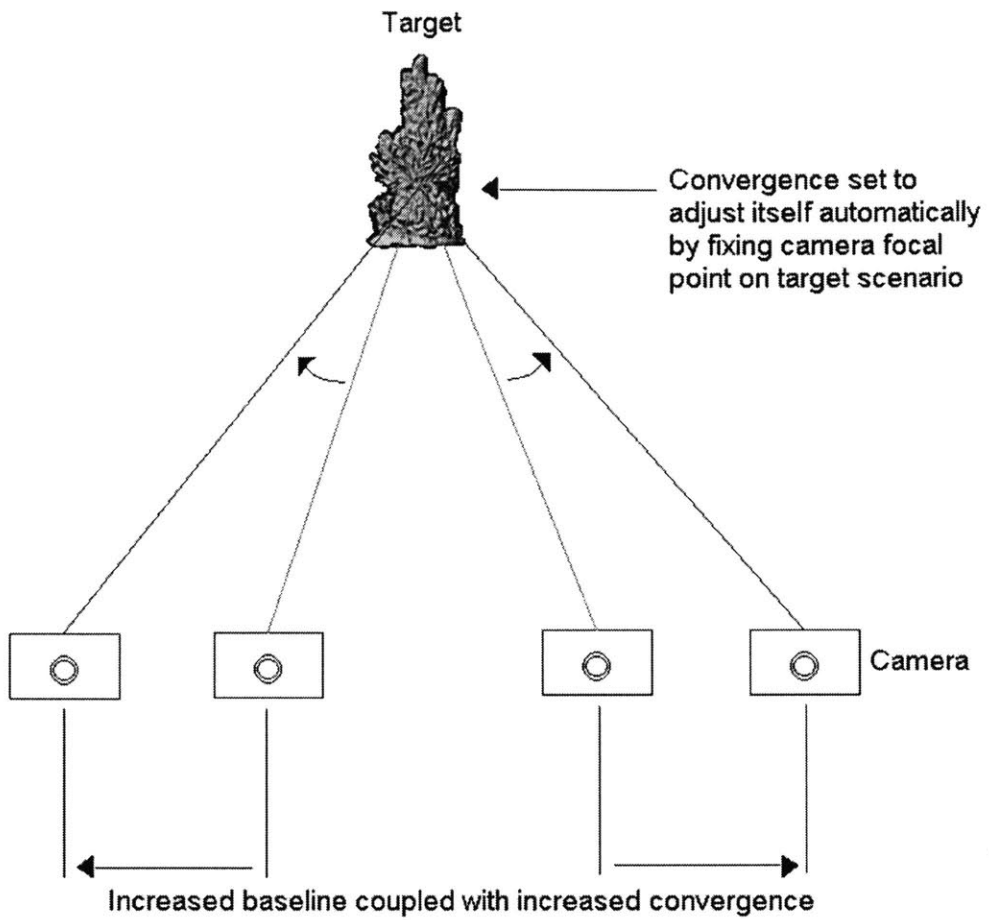
## 6.1 Generating Stereo Views

Matlab<sup>TM</sup> was used to generate synthetic stereo from the surface model of the imaged scenario. Appendix B lists a sample code used to generate synthetic stereo using Matlab.

- As a first step, the surface model and its associated texture map were stored as Matlab objects. The Camera Toolbar function of Matlab was then used to position the surface model with respect to the virtual camera thereby facilitating flexible capture of different perspectives of the surface. Figure 6.2 summarizes the methodology adopted in generating synthetic stereo image pairs of the surface.
- Several stereo image pairs of increasing disparity sensitivities were generated as can be seen in figures 6.3a-f. The stereo pairs were generated under imaging conditions identical to those in real-world stereo-endoscopy i.e. the field of view of the virtual camera was set at 85 degrees and the distance to the reference focal plane was set at 50mm. In generating these disparity sensitivities, care was taken to ensure that the same point on the surface was maintained at zero parallax in generating every synthetic stereo pair. This way, the convergence would

plane was set at 50mm. In generating these disparity sensitivities, care was taken to ensure that the same point on the surface was maintained at zero parallax in generating every synthetic stereo pair. This way, the convergence would automatically adjust itself with changing baseline and distance to target object (figure 6.2).

- As discussed in chapter 5, the quality of stereo (in figures 6.3a-e) improves with increasing baseline (with the result increasing convergence as a result). However there comes a situation while the increased baseline results in severely negative or divergent parallax at which point the perceived quality of stereo deteriorates dramatically (figure 6.3f).



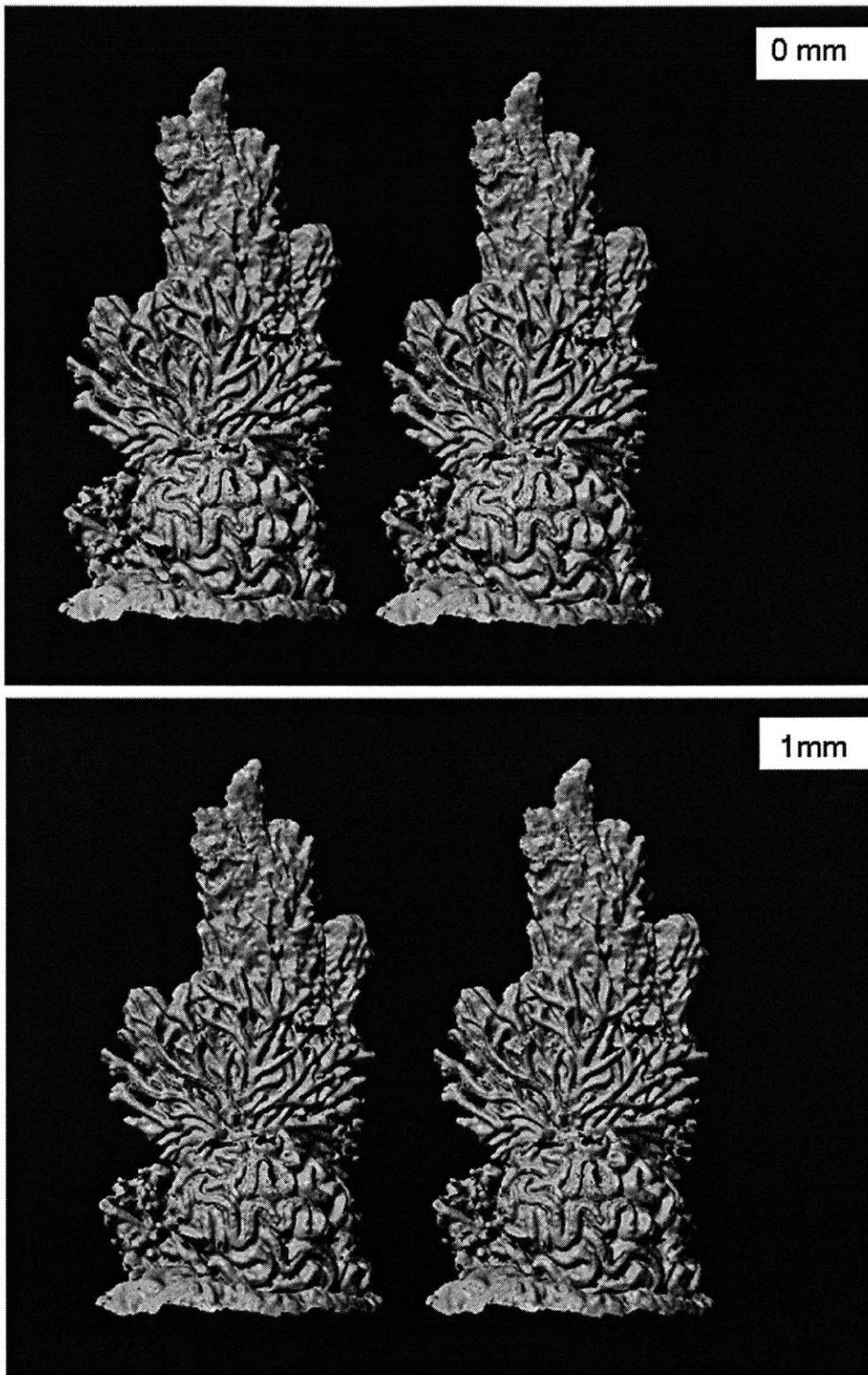
Imaging Conditions:

Field of View = 85 degrees

Distance to focal plane = 50mm

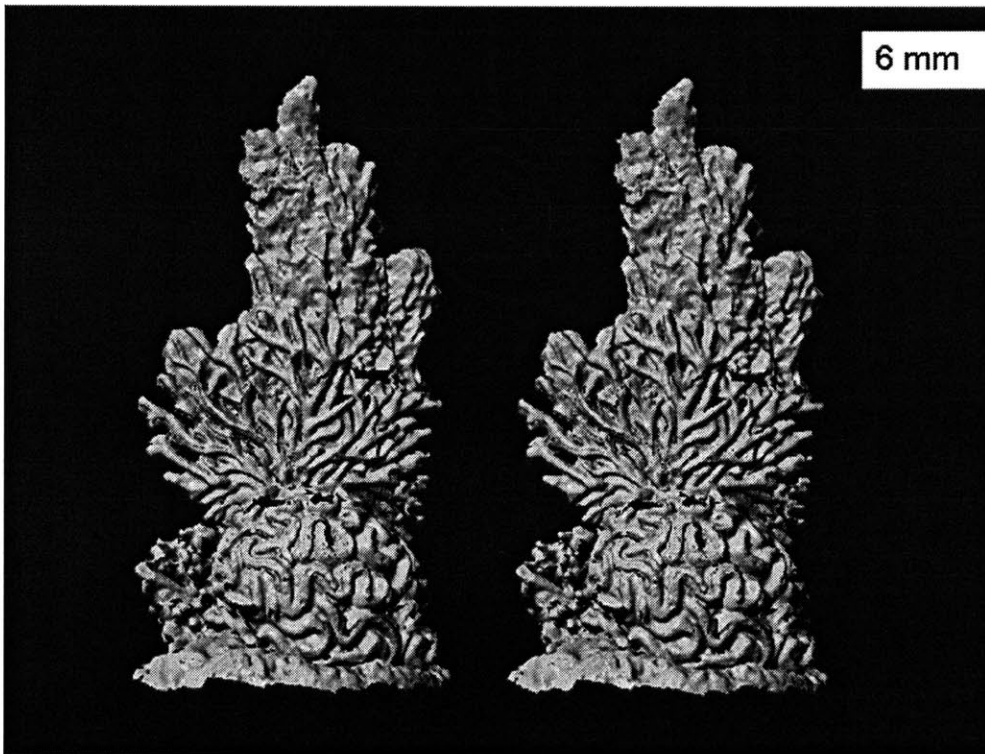
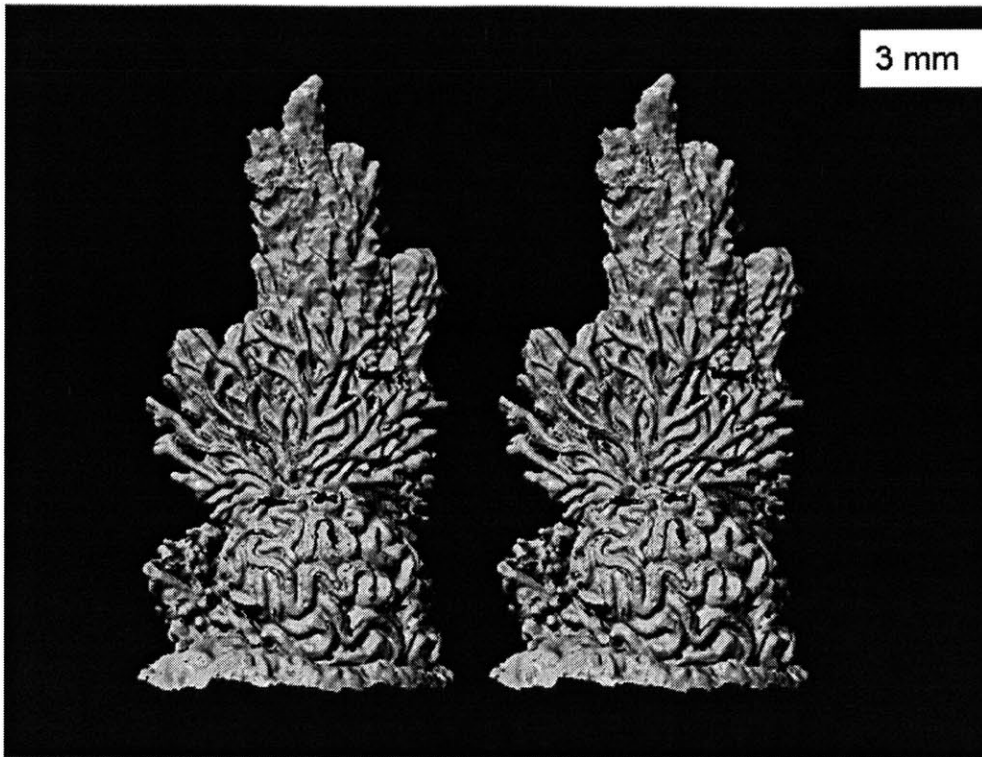
**Figure 6.2:** Adjusting disparity sensitivity by simultaneously varying baseline and angle of convergence.



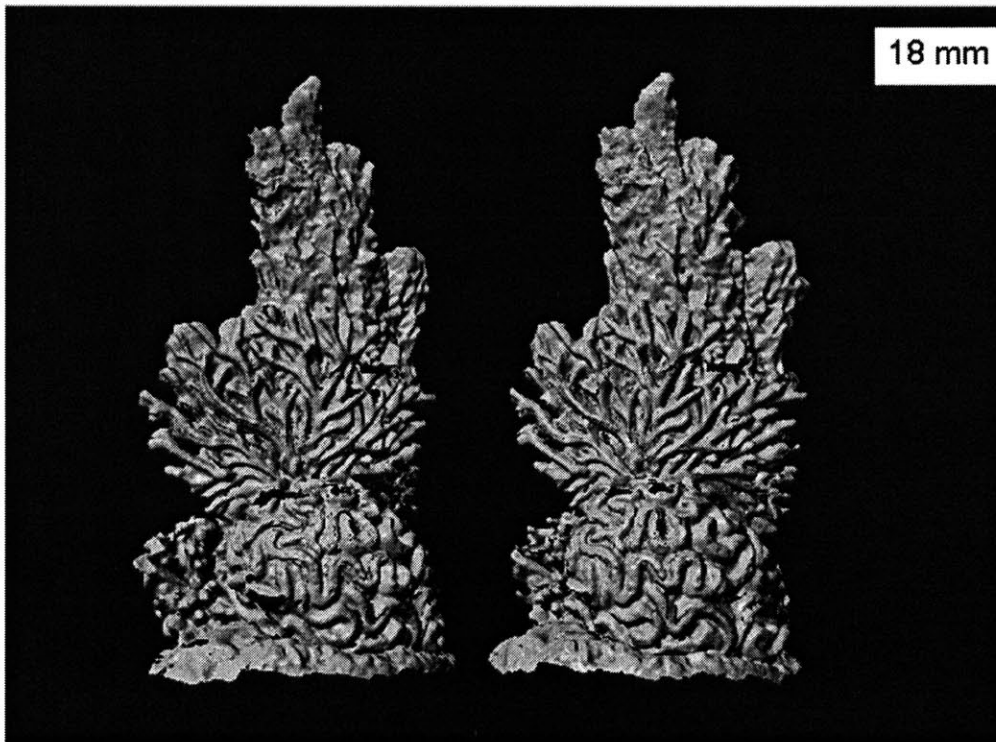
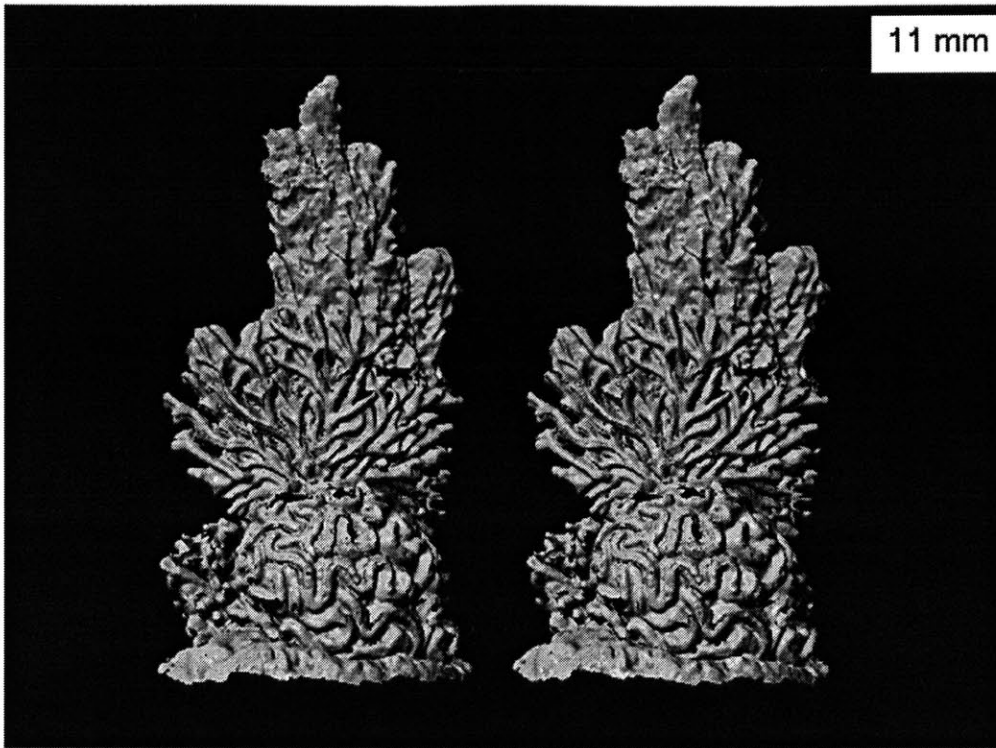


*Figure 6.3a & b: Enhanced synthetic stereo pairs imaged with baselines of 0 and 1mm respectively*





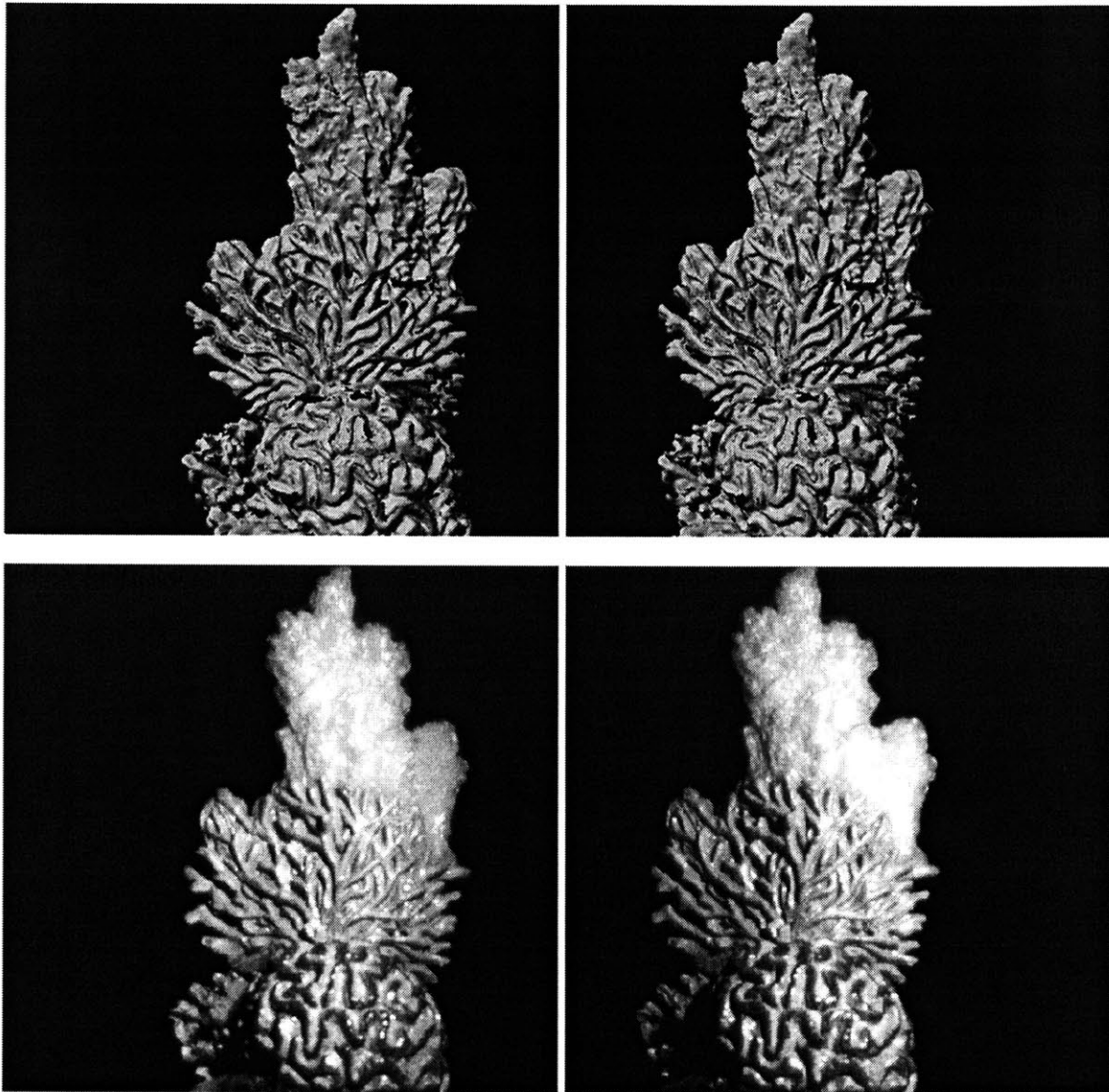
*Figure 6.3c & d: Enhanced synthetic stereo pairs imaged with baselines of 3 and 6mm respectively*



*Figure 6.3e &f: Enhanced synthetic stereo pairs imaged with baselines of 11 and 18mm respectively*

## 6.2 Comparison of Synthetic Stereo and Raw Stereo

It is instructive to compare synthetic stereo and real stereo generated under identical imaging conditions. Figure 6.4 illustrates this comparison. The distance to the reference focal plane, the field of view of the imaging system and the baseline are held constant at 50 mm, 85 degrees and 11 mm respectively in capturing the two image pairs.



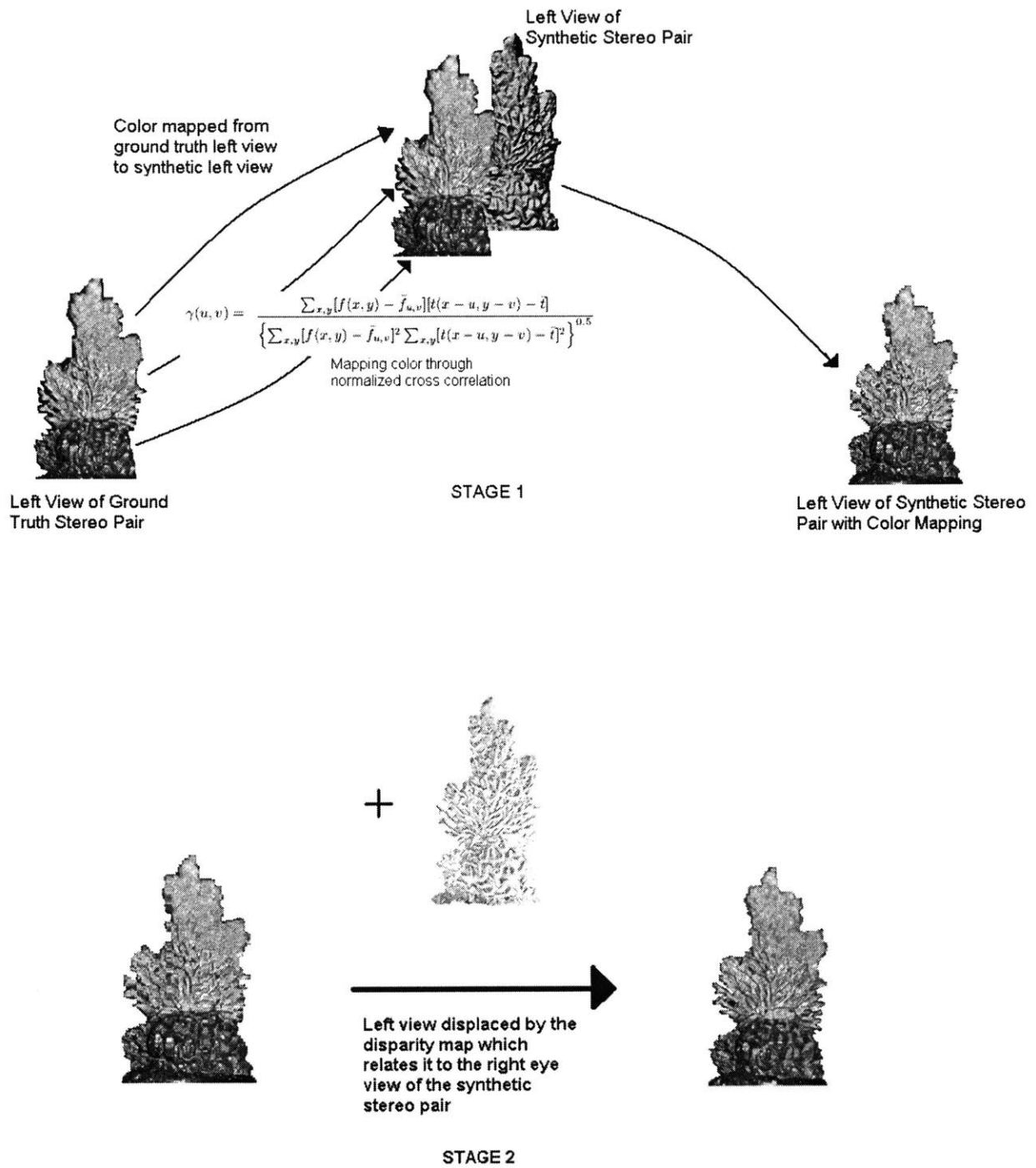
*Figure 6.4: Top figure is the synthetically enhanced texture mapped gray scale image of the coral scene and bottom figure is the gray scale version of the ground truth for comparison which possesses identical disparity sensitivity.*

Despite the fact that the disparity sensitivity to depth in both images equals 0.7 pixels/mm, to the naked eye, the quality of stereo perceived is better in the case of the synthetically generated stereo due to lesser illumination variation and other defects present in a real world image. However as shall be noticed in the subsequent section, by mapping color from the real world image onto the synthetically generated stereo, some of the 'perceived superior quality of depth' is lost.

### **6.3 Mapping Color onto Synthetic Stereo**

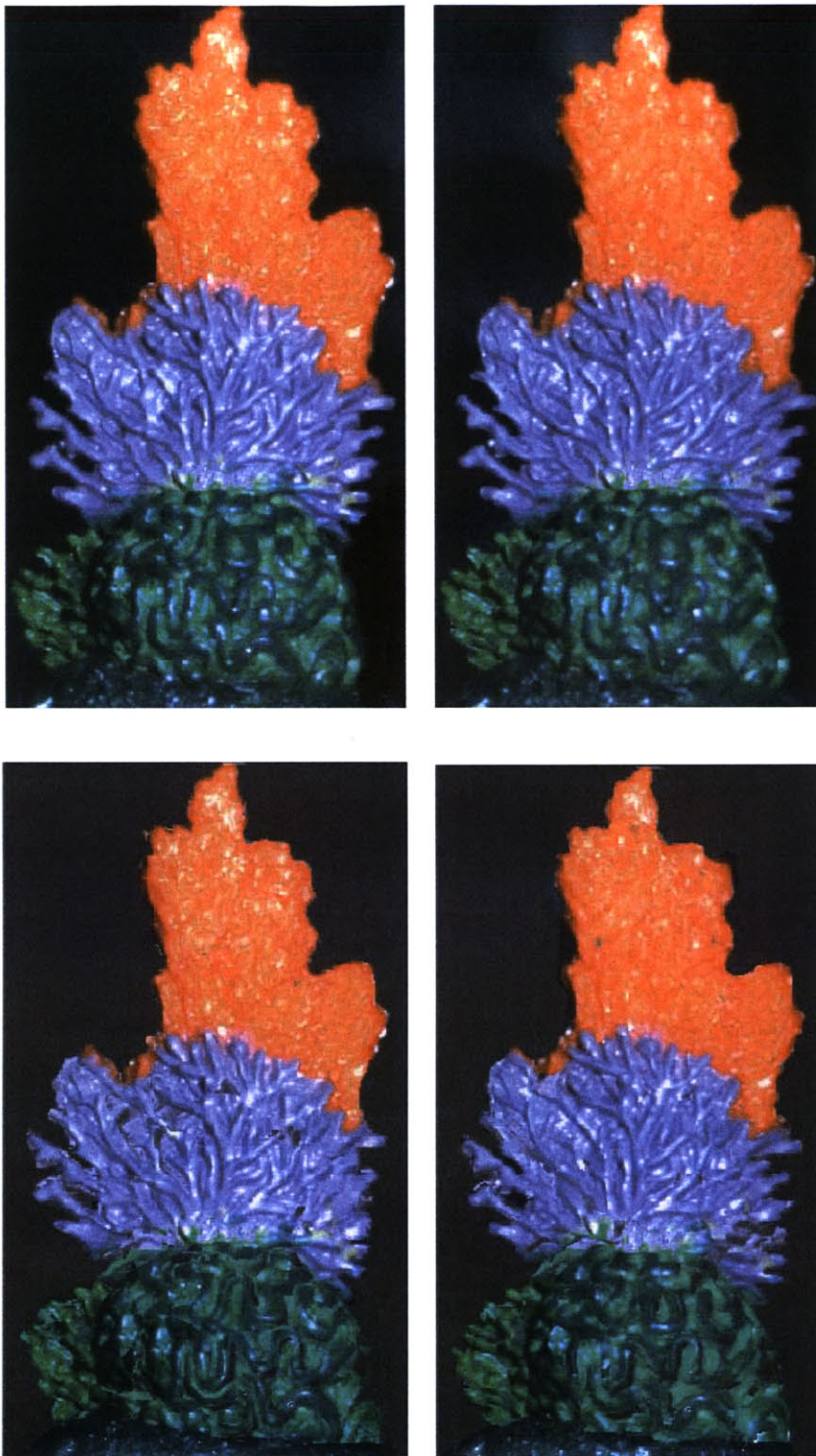
For the purpose of completeness in demonstrating the generation of enhanced synthetic stereo, it is useful to map color onto the synthetic stereo. The various stages employed in mapping color are summarized in figure 6.5. In the first stage, color is mapped from captured color image of the left view of the target onto the synthetically generated left view using the normalized cross-correlation technique (refer to appendix C for sample MATLAB code). In the second stage, the color mapped left eye view of the synthetic stereo pair is displaced by the disparity map which ties it to the right eye view so as to obtain a color version of the right eye view.

The results of the color mapping exercise are illustrated in figure 6.6. As can be observed, the quality of depth perceived in both the ground truth case and the synthetic case are very similar. The disparity sensitivity to depth was measured to be about 0.3 pixels/mm in both cases for the color stereo pairs.



**Figure 6.5:** The two stage process for adding color to the synthetically enhanced gray scale stereo pair.





**Figure 6.6:** Comparison of color mapped synthetic stereo and ground truth real stereo of identical disparity sensitivity. The real pair is displayed on top and the synthetic pair is displayed at the bottom.

The results listed in the preceding sections confirm the hypothesis in chapter 5 that limitations of the AWS-scope in terms of poor disparity sensitivity can be overcome digitally. By suitably adjusting the virtual baseline – and as a result the angle of convergence – disparity sensitivity can be improved from about 0.02 pixels/mm to about 0.3 pixels/mm. Having demonstrated a ‘proof of concept’ for the AWS-scope, the issue of developing a prototype that delivers enhanced synthetic stereo in real-time needs to be addressed. In chapter 7 which is the concluding chapter of this thesis, areas that need to be investigated for future work will be discussed in greater detail.

# CHAPTER 7: SUMMARY, CONCLUSION AND FUTURE

## WORK

This work is best concluded by revisiting the last six chapters in a succinct manner and extrapolating the thread of thought into work for the future.

- Chapter 1 introduced the concept of stereo-endoscopy besides discussing the limitations of the state-of-the-art and therefore laying out the need for a better quality stereo-endoscope – which is the proposed AWS-scope.
- Chapter 2 delved into the conceptualization of the AWS-scope examining among other things, the expected performance of its key optical attribute of disparity sensitivity.
- Chapter 3 realized the concept presented in chapter 2 by presenting a ‘proof-of-concept’ AWS-scope system installed at MIT. The concluding section of the chapter presented an experimental verification of the conceptual model.
- Chapter 4 presented a comparison of the optical attributes of the ‘proof-of-concept’ AWS-scope with the benchmark – state-of-the-art- Schoelly stereo-endoscope. The fundamental optical characteristics of system resolution, captured field of view, light transmission levels and depth of focus of the AWS-scope were experimentally shown to be identical or superior to those of the Schoelly. In addition, the issues of radiometric calibration as intensity correction of images captured by the AWS-scope were described in detail.
- Chapter 5 clarified the deficiency of disparity sensitivity to depth for stereo image pairs captured using the AWS-scope when compared to the Schoelly – a case was thereby made for overcoming this limitation digitally by generating synthetically



enhanced stereo pairs. In order to better understand synthetic enhancement of stereo, the notion of depth as perceived by the human visual system was examined in greater detail. In doing so, the key depth perception drivers of baseline, convergence and parallax were analyzed to ultimately yield generic design rules while constructing stereo for effective display.

- Chapter 6 finally demonstrated implementation of the design rules discussed in chapter 5 on the AWS-scope. Synthetically enhanced stereo was generated using images captured from the AWS-scope and juxtaposed in comparison to real ground truth stereo under identical imaging conditions.

The preceding chapters thus clearly demonstrate the technical feasibility of the ‘proof-of-concept’ AWS-scope as an effective alternative to the industry-standard Schoelly. However, it is worth re-iterating the fact that such an AWS-scope presents several other strategic and critical benefits over the Schoelly.

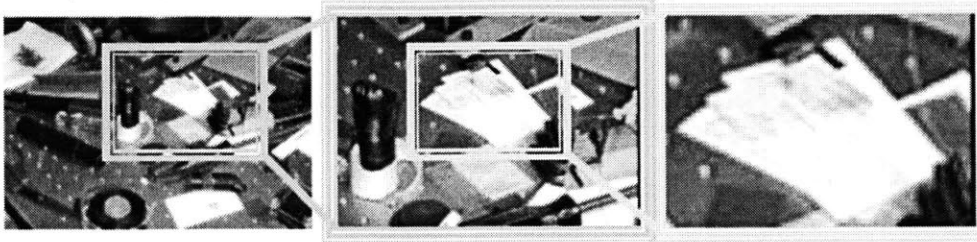
1. **Enhanced Performance and Additional Features due to Quantified 3D:** The AWS-scope system has the flexibility to provide both distortion corrected live stereo and quantitative real time three-dimensional images. This capability provides numerous attractive features never seen before in surgical endoscopy. AWS combined with a class of new processing algorithms opens the doors to generating millimeter accurate real time three-dimensional surface models of a surgical scenario without any pattern projection by sampling the optical wavefront at several angular positions; all this using a single camera and relatively much lesser computational resources.

Second, since AWS shifts the image a fraction of a pixel and provides over-sampling, it is possible to obtain super-resolution two-dimensional reconstructed images from the generated three-dimensional models. Super-resolution refers to the resolution of the reconstructed image being better than that of the recording device. This has tremendous benefit as it overcomes the inherent spatial resolution limits present in traditional imaging providing the surgeon with high spatial frequency detail needed to perform delicate tasks. Figure 7.2 clearly demonstrates the superiority of super-resolution.

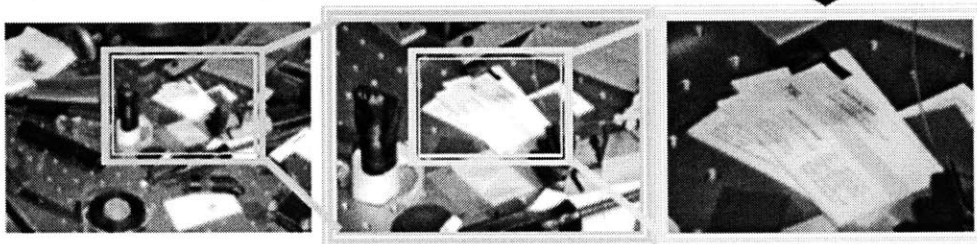
Third, it is possible to perform surface rendering from the generated three-dimensional surface model. Three-dimensional surface rendering allows surgeons to digitally alter camera position to obtain better perspectives of a scene. Experience has indicated that viewing angles of nearly 30 degrees from actual camera position can be satisfactorily achieved. This ability has already been demonstrated at video rates in other applications. Furthermore, viewer perspective can be altered in numerous other ways including altering left-right eye perspectives to obtain increased depth perception, changing contrast and brightness of illumination based on depth, and segmenting images to eliminate background detail thereby eliminating clutter from the surgeon's field of view. Ability to flexibly alter the viewer's perspective is thus an exceptionally powerful feature that AWS based three-dimensional surface modeling yields.

Finally, since AWS generates quantified three-dimensional models, it provides for intra-operative anatomy such as measuring the size of a tumor or a cyst by examining its three-dimensional model.

(a) Enlargement (conventional camera)



(b) Super resolution through Active Imaging

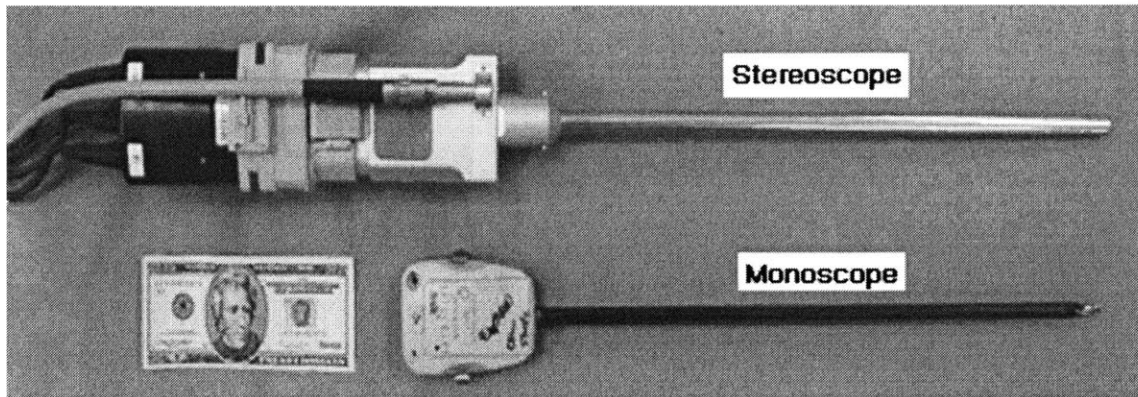


*Figure 7.1: Super resolution achieved using the AWSM module and its comparison to the resolution using a conventional camera [29,30].*

- 2. Cost Advantages:** Current commercial stereoscopes have non-standard, non-autoclavable dual channels which are more expensive and less readily available when compared to the widely available standard single channel autoclavable monoscope. Replacing the custom made stereoscope channel with an off-the-shelf monoscope (which is the case if an AWS-scope were to be realized) could result in substantial cost reduction. Second, a number of present day stereo-endoscope systems use multiple cameras for producing three-dimensional stereo. Since the AWS-scope generates three-dimensional stereo using just a single standard camera, it can lead to appreciable cut down in costs. Finally, the existing commercial stereo-endoscope systems demand dedicated and specially trained staff for operation and maintenance. Besides, they occupy a large amount of space and are non-ambulatory. These drawbacks translate into large costs in the long run. On the other hand, the AWS-scope by virtue of its design based on standard

components would be easier to handle and maintain without the need for separately trained personnel. And furthermore, it can easily be integrated into a pre-existing endoscopic facility with minimal changes in the electronics.

3. **Size Advantages:** Most existing commercial stereoscopes are extremely bulky and awkward to handle. Since these endoscopes are dual channel, their channel diameters are appreciably larger than those of monoscopes. Besides, a larger channel needs a larger incision to be made on the patient for insertion. And this is definitely contrary to the quintessential goal of endoscopy which is minimally invasive surgery. The use of a monoscope, whose channel is atleast about 20% smaller than the stereo-endoscope (Figure7.1), would certainly improve matters on this front. Second, bulky stereoscopes need bigger and stronger auxiliary systems such as robot arms etc. As a consequence, the entire endoscope plus imaging module occupies a large volume. Besides, its size precludes mobility which then rules out ambulatory surgery, a major incentive of endoscopy. Since the AWS-scope uses a standard monoscope, endoscope diameters would tend to get substantially smaller and this translates to gains on both the patients' side as well as the surgeons' side; the former benefiting from smaller incisions being made and latter benefiting from improved handling ease and quantified real time three-dimensional data.



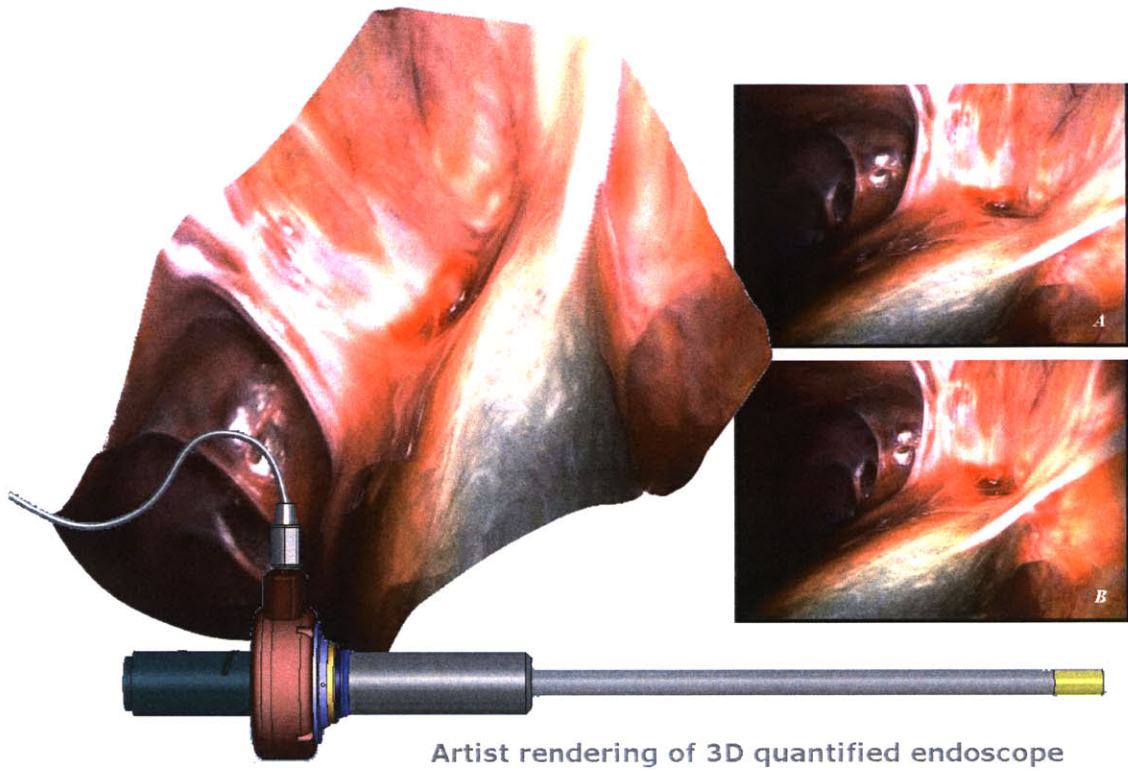
*Figure 7.2: Monoscope (used in the AWS-scope) juxtaposed with a stereo-endoscope (typical industry standard).*

## Future Work

Having realized a feasible ‘proof-of-concept’ AWS-scope, potential future deliverables in that order of priority include:

- Developing custom real-time software that integrates the AWS module with the imaging system, image processing unit and image rendering and display unit.
- Construction of a prototype of the AWS-scope in order to gain benefits of a better aligned and robust set-up that could be deployed in actual surgical clinical trials.
- Broadening the scope of applications of the AWS-scope by transforming robot assisted surgery into largely robot controlled surgery operating on a quantified surgical scenario.

Figure 7.3, an artist’s rendering, aptly sums up the AWS-scope.



Artist rendering of 3D quantified endoscope

*Figure 7.3: The AWS-scope prototype when in action*

## REFERENCES

1. **Breidenthal R S, Forkey R E, Smith J, Volk B E**, Stereoscopic endoscope with virtual reality viewing. US Patent 6,139,490 (2000)
2. **Brown M Z, Burschka D, Hager, G D**, Advances in computational stereo, IEEE Transactions on Pattern Analysis and Machine Intelligence (2003) (to appear)
3. **Chan M, Lin W, Zhou C, Qu J Y**, Miniaturized three-dimensional endoscopic imaging system based on active stereovision. Applied Optics 42 (10): 1888-1898 (2003)
4. **Costales B**, Single-lens 3D method, microscope, and video adapter, US Patent 6,275,335 (2001)
5. **Fitch J**, Sparse aperture endoscope. US Patent 5,919,128 (1999)
6. **Frigerio F**, PhD thesis on TV Rate 3D imaging, MIT (2006)
7. **Guthart G S, Salisbury J K**, The Intuitive™ telesurgery system: overview and Application Proc. of the IEEE Int. Conf. on Robotics and Automation (ICRA2000), San Francisco, CA, April 2000, Page(s): 618-621 vol.1 (2000)
8. **Hart D P**, Sparse Array Image Correlation, selected for publication in the reviewed archival proceedings of the 8th International Symposium on Applications of Laser Techniques to Fluid Mechanics, Lisbon, Portugal," July 1996, pp. 21.1.1-21.1.8. (1996)
9. **Hart D P**, High-Speed PIV Analysis Using Compressed Image Correlation, *Journal of Fluids Engineering*, Vol. 120, pp. 463-470 (1998)
10. **Hart D P**, Successive Approximation PIV Analysis to Achieve High Accuracy, Resolution, and Speed, 13th U.S. National Congress of Applied Mechanics, June 21-24, (1998)

11. **Hart D P**, Super-Resolution PIV through Recursive Local-Correlation, VSJ-SPIE International Conference on Optical Technology and Image Processing in Fluid, Thermal, and Combustion Flow, Dec. 6-10, 1998, Yokohama, Japan (1998)
12. **Hart D P**, The Elimination of Correlation Errors in PIV Processing, 9th International Symposium on Applications of Laser Techniques to Fluid Mechanics, July 13-16, 1998, Lisbon, Portugal (1998)
13. **Hart D P** Second-Order Correlation, YAYOI Proceedings Kogakuin Univ., December 11, 1998, pp. 30-36 (1998)
14. **Hart D P**, Super-Resolution PIV Processing by Recursive Correlation, *Journal of Visualization*, Vol.2, No.3/4, 1999 (1999)
15. **Hart D P**, PIV Error Correction, *Experiments in Fluids*, 1999 (1999)
16. **Hart D P**, PIV Processing Using Multidimensional Correlation,” ISFV-10 10<sup>th</sup> International Symposium of Flow Visualization, Kyoto, Japan, July, 2002. (2002)
17. **Hasegawa K, Sato Y**, Endoscope system for high-speed 3D measurement, *Systems and Computers in Japan* 32 (8): 30-39 (2001)
18. **Hayashibe M, Nakamura Y**, Laser-pointing endoscope system for intra-operative 3D geometric registration, IEEE International Conference on Robotics and Automation, 2001, Proceedings 2001 ICRA, Page(s): 1543 -1548 vol.2 (2001)
19. **Igarashi T**, Rigid endoscope optical system. US Patent 6,471,642 (2000)
20. **Kim J, Hwang D, Jeong H, Song C, Lee K, Lee M** Development of depth extraction algorithm for the stereo endoscopic image. Engineering in Medicine and Biology Society, Proceedings of the 20th Annual International Conference of the IEEE, Page(s): 884-887 vol.2 (1998)



21. **Lia RA**, Endoscope or borescope stereo viewing system, US Patent 5,222,477 (1991)
22. **Loranger F, Laurendeau D , and Houde R**, A fast and accurate 3-D rangefinder using the Biris technology: the TRID sensor, Proceedings of 3-D Digital Imaging and Modeling, in International Conference on Recent Advances, pp.51 –58,1997
23. **Mihalca G, Kazakevich Y E**, Stereoscopic imaging by alternately blocking light. US Patent 5,964,696 (1999)
24. **MIT**-owned US Patent 6,108,458
25. **Mourgues F, Devemay F, Coste-Maniere E**, 3D reconstruction of the operating field for image overlay in 3D-endoscopic surgery, IEEE and ACM International Symposium on Augmented Reality Proceedings, Page(s): 191-192 (2001)
26. **Nakamura M, Kubota T**, Stereoscopic endoscope having image transmitting optical system and pupil dividing unit that are axially movable with respect to each other. US Patent 5,743,847 (1995)
27. **Okatani T, Deguchi K**, Shape reconstruction from an endoscope image by shape from shading technique for a point light source at the projection center, Computer Vision and Image Understanding 66 (2): 119-131 (1997)
28. **Quartucci Forster C H, Tozzi C L**, Towards 3D reconstruction of endoscope images using shape from shading. Proceedings XIII Brazilian Symposium on Computer Graphics and Image Processing, Page(s): 90-96 (2000)
29. **Rohaly J, Hart D P**, High-resolution Ultrafast 3D Imaging,” Proceedings of Photonics West 2000; Three-Dimensional Image Capture and Applications III, SPIE vol. 3958 March 2000, San Jose, CA. (2000)

30. **Rohaly J, Hart D P**, Monocular 3-D Active micro-PTV, 4th International Symposium on Particle Image Velocimetry, Gottingen, Germany, September 17-18, 2001. (2001)
31. **Rohaly J, Frigerio F, Hart D P**, Reverse Hierarchical PIV Processing, *Measurement Science & Technology*, 13 (2002)
32. **Roth G, Hart D, Katz J**, Feasibility of using the L64720 video motion estimation processor (MEP) to increase efficiency of velocity map generation for Particle Image Velocimetry (PIV), Proceedings of the ASME/JSME Fluids Engineering and Laser Anemometry Conference, Hilton Head, SC, pp.387-393, 1995
33. **Sekiya T, Eguchi M**, Stereoscopic endoscope, US Patent 6,338,711 (2002)
34. **Shipp J I**, Single lens stereoscopic video camera, US Patent 5,471,237 (1993)
35. **Sjödahl M**, Accuracy in electronic speckle photography, *Applied Optics* 36, pp.2875-2885, 1997
36. **Takahashi S et al.**, Stereoscopic endoscope. US Patent 5,720,706 (1995)
37. **Teixeira dos Santos R J**, Scaling of rendered stereoscopic scenes, University of Lisbon, Faculty of Sciences (2005)
38. **Verri A and Torre V**, Absolute estimate of depth in stereopsis, *Opt. Soc. Of Am. A* Vol 3 No.3, pp. 297-299, 1986
39. **Viguer A, Clement G, Trotter Y**, Distance perception within near visual space, *Perception* volume 30, pp. 115-124, 2001
40. **Watts J R**, Stereoscopic imaging arrangement and viewing arrangement, US Patent 5,914,810 (1999)
41. **Willert C E and Gharib M**, Three-dimensional particle imaging with a single camera, *Experiments in Fluids* 12, pp.353-358, 1992

42. **Yeung S Y, Tsui H T, Yim A**, Global shape from shading for an endoscope image Medical Image Computing And Computer-Assisted Intervention, Miccai'99, Proceedings Lecture Notes in Computer Science 1679: 318-327 (1999)
43. **[www.vision.caltech.edu/bouguetj/calib\\_doc/htmls/links.html](http://www.vision.caltech.edu/bouguetj/calib_doc/htmls/links.html)**
44. **<http://imaginis.com/endoscopy/>**
45. **[www.olympus-europa.com/endoscopy/425\\_2813.htm](http://www.olympus-europa.com/endoscopy/425_2813.htm)**
46. **<http://ubimon.doc.ic.ac.uk>**
47. **Buser P, Imbert M**, Vision, MIT Press
48. **Davies E R**, Machine Vision, Academic Press
49. **Hecht E**, Optics 4<sup>th</sup> Ed., Pearson Education
50. **Nagata**, 1993

# APPENDIX A: CAMERA CALIBRATION PARAMETERS

```
% Intrinsic and Extrinsic Camera Parameters
%
% This script file can be directly executed under Matlab to recover the camera intrinsic and extrinsic parameters.
% IMPORTANT: This file contains neither the structure of the calibration objects nor the image coordinates of the calibration points.
% All those complementary variables are saved in the complete matlab data file Calib_Results.mat.
% For more information regarding the calibration model visit http://www.vision.caltech.edu/bouguetj/calib\_doc/

%-- Focal length:
fc = [ 1051.853096801346100 ; 1055.426559922483200 ];

%-- Principal point:
cc = [ 627.945891829681050 ; 479.893783704729170 ];

%-- Skew coefficient:
alpha_c = 0.000000000000000;

%-- Distortion coefficients:
kc = [ -0.116631911154849 ; 0.047094316518577 ; 0.000273470912964 ; -0.002446841187132 ; 0.000000000000000 ];

%-- Focal length uncertainty:
fc_error = [ 3.819023791856214 ; 4.081458725528470 ];

%-- Principal point uncertainty:
cc_error = [ 6.910184361236114 ; 5.294343115227970 ];

%-- Skew coefficient uncertainty:
alpha_c_error = 0.000000000000000;

%-- Distortion coefficients uncertainty:
kc_error = [ 0.013522063944659 ; 0.028165345455071 ; 0.001257349138992 ; 0.001542044035415 ; 0.000000000000000 ];

%-- Image size:
nx = 1280;
ny = 1024;

%-- Various other variables (may be ignored if you do not use the Matlab Calibration Toolbox):
%-- Those variables are used to control which intrinsic parameters should be optimized

n_ima = 10; % Number of calibration images
est_fc = [ 1 ; 1 ]; % Estimation indicator of the two focal variables
est_aspect_ratio = 1; % Estimation indicator of the aspect ratio fc(2)/fc(1)
center_optim = 1; % Estimation indicator of the principal point
est_alpha = 0; % Estimation indicator of the skew coefficient
est_dist = [ 1 ; 1 ; 1 ; 1 ; 0 ]; % Estimation indicator of the distortion coefficients

%-- Extrinsic parameters:
%-- The rotation (omc_kk) and the translation (Tc_kk) vectors for every calibration image and their uncertainties

%-- Image #1:
omc_1 = [ -2.240746e+000 ; -2.155675e+000 ; 1.599752e-002 ];
Tc_1 = [ -5.528240e+001 ; -2.681693e+001 ; 1.028584e+002 ];
omc_error_1 = [ 4.560032e-003 ; 4.831531e-003 ; 9.879281e-003 ];
```

```

Tc_error_1 = [ 6.957208e-001 ; 5.430834e-001 ; 5.760162e-001 ];

%-- Image #2:
omc_2 = [ -2.101659e+000 ; -2.031495e+000 ; 2.663056e-001 ];
Tc_2 = [ -5.863445e+001 ; -2.712413e+001 ; 1.217955e+002 ];
omc_error_2 = [ 4.902299e-003 ; 4.987952e-003 ; 9.289128e-003 ];
Tc_error_2 = [ 8.050426e-001 ; 6.254946e-001 ; 6.079747e-001 ];

%-- Image #3:
omc_3 = [ 1.994943e+000 ; 1.907319e+000 ; 4.629383e-001 ];
Tc_3 = [ -3.648017e+001 ; -3.463975e+001 ; 7.522258e+001 ];
omc_error_3 = [ 5.183249e-003 ; 4.303491e-003 ; 8.361708e-003 ];
Tc_error_3 = [ 5.361277e-001 ; 3.988997e-001 ; 4.755492e-001 ];

%-- Image #4:
omc_4 = [ -1.968201e+000 ; -1.831393e+000 ; -3.276074e-001 ];
Tc_4 = [ -4.516327e+001 ; -2.095120e+001 ; 9.483379e+001 ];
omc_error_4 = [ 4.017260e-003 ; 5.616823e-003 ; 8.779322e-003 ];
Tc_error_4 = [ 6.370294e-001 ; 5.011236e-001 ; 5.134377e-001 ];

%-- Image #5:
omc_5 = [ 1.756668e+000 ; 1.597693e+000 ; -6.545812e-001 ];
Tc_5 = [ -5.572071e+001 ; -8.670079e+000 ; 1.462632e+002 ];
omc_error_5 = [ 4.357997e-003 ; 5.514058e-003 ; 8.046289e-003 ];
Tc_error_5 = [ 9.598539e-001 ; 7.518268e-001 ; 6.121312e-001 ];

%-- Image #6:
omc_6 = [ 1.879971e+000 ; 8.702180e-001 ; -4.321700e-001 ];
Tc_6 = [ -5.941231e+001 ; -1.552138e+000 ; 9.880677e+001 ];
omc_error_6 = [ 5.009843e-003 ; 4.893142e-003 ; 6.871990e-003 ];
Tc_error_6 = [ 6.568303e-001 ; 5.174598e-001 ; 5.290008e-001 ];

%-- Image #7:
omc_7 = [ -1.820955e+000 ; -1.809406e+000 ; -6.200421e-001 ];
Tc_7 = [ -5.177743e+001 ; -2.072549e+001 ; 9.569861e+001 ];
omc_error_7 = [ 3.679425e-003 ; 5.735458e-003 ; 8.430748e-003 ];
Tc_error_7 = [ 6.440819e-001 ; 5.140159e-001 ; 5.700119e-001 ];

%-- Image #8:
omc_8 = [ 1.893331e+000 ; 1.212356e+000 ; -5.891028e-001 ];
Tc_8 = [ -7.480444e+001 ; -4.452326e+001 ; 1.619550e+002 ];
omc_error_8 = [ 4.864352e-003 ; 5.831290e-003 ; 7.721023e-003 ];
Tc_error_8 = [ 1.089433e+000 ; 8.638092e-001 ; 7.806940e-001 ];

%-- Image #9:
omc_9 = [ 1.736143e+000 ; 1.635786e+000 ; -6.179776e-001 ];
Tc_9 = [ -4.688419e+001 ; -6.271341e+001 ; 1.592679e+002 ];
omc_error_9 = [ 4.264641e-003 ; 6.160751e-003 ; 8.032508e-003 ];
Tc_error_9 = [ 1.086194e+000 ; 8.330432e-001 ; 6.718559e-001 ];

%-- Image #10:
omc_10 = [ -1.562766e+000 ; -2.028131e+000 ; 1.314505e+000 ];
Tc_10 = [ -4.043786e+001 ; -3.611143e+001 ; 1.681064e+002 ];
omc_error_10 = [ 6.924947e-003 ; 4.891941e-003 ; 7.526541e-003 ];
Tc_error_10 = [ 1.117795e+000 ; 8.654492e-001 ; 5.310534e-001 ];

```

# APPENDIX B: GENERATING SYNTHETIC STEREO FROM SURFACE MODEL

```
%This code first converts point data captured using the AWS into a regular
%grid and generates a surface model from this regular grid mesh. Finally,
%the Camera toolbox functions are used to generate the left and right
%perspectives from the surface model.
```

```
% Reading data depth data captured using AWS-scope
```

```
clc;clear all;
fid=fopen('coral4.txt');
A = fscanf(fid,'%g %g %g',[3 inf]);
A=A';
```

```
*****
%Generating Regular Grid from irregular grid
```

```
x1=min(A(:,1));
x2=max(A(:,1));
y1=min(A(:,2));
y2=max(A(:,2));
```

```
A_mon=sortrows(A,[1,2]);
A=A_mon;clear A_mon;
```

```
%Setting up blocks for interpolation
```

```
Xsteps=(x2-x1)/0.25;
iloop=floor(Xsteps);
iextra=Xsteps-iloop;
```

```
Ysteps=(y2-y1)/0.25;
jloop=floor(Ysteps);
jextra=Ysteps-jloop;
xmin=x1;ymin=y1;
```

```
for i=1:iloop
    A_piecx=xmin:0.0001:(xmin+0.25-0.0001);
```

```
    for j=1:jloop
```

```
        %define grid for allocating re-allocating the scattered data
        A_piecy=ymin:0.0001:(ymin+0.25-0.0001);
        [X,Y]=meshgrid(A_piecx,A_piecy);
        Z=zeros(size(X));
```

```
        %determine search length in the main matrix
        indexx_min=find(A(:,1)>xmin,1);
        indexx_max=find(A(:,1)>(xmin+0.2499),1);
        indexx_min=indexx_min-1;
        indexx_max=indexx_max-1;
```

```
        %match and place from scattered array to regular grid
        for ia=indexx_min:indexx_max
```

```
            ind_x=find(floor(X(1,:)*10000)==floor(A(ia,1)*10000))
```

```
            if (A(ia,2)>=ymin && A(ia,2)<=(ymin+0.2499))
                ind_y=find(floor(Y(:,1)*10000)==floor(A(ia,2)*10000))
                Z(ind_x,ind_y)=A(ia,3);
            end
```

```
        end
```

```

end

%Interpolate Z to an undersampled surface
yi=xmin:0.01:xmin+0.25-0.01;
xi=ymin:0.01:ymin+0.25-0.01;
[XI,YI]=meshgrid(xi,yi);
ZI=zeros(size(XI));
ZI=interp2(X,Y,Z,XI,YI,'nearest');

Z_consolidated((i-1)*25+1:(i-1)*25+25,(j-1)*25+1:(j-1)*25+25)=ZI;

clear ZI; clear yi;clear xi;clear XI;clear YI;
%increment jloop minimum
ymin=ymin+0.25;
end

clear A_piecx;clear A_piecy;clear X;clear Y;clear Z;
xmin=xmin+0.25;
end

%-----
% Generating stereo views

% defining surface
[X,Y]=meshgrid(size(Z_consolidated));
z=Z_consolidated;

%defining rotation parameters
zdir = [0 0 3];
center = [10 10 0];

%creating left eye
% subplot(1,2,1);
hleft=surfc(X,Y,z,'Edgecolor','None','facecolor','Yellow');
rotate(hleft,zdir,0,center);
set(gca,'CameraTarget',[9.1,10,0]);
%camtarget([10,10,0]);
set(gca,'CameraPosition',[9.1,10,10]);
%campos([5,10,10]);
set(gca,'Projection','Perspective');
%camproj('Perspective');
camlight left;lighting phong;
set(gca,'DataAspectRatioMode','Manual');
%set(gca,'XTickLabel','', 'YTickLabel','', 'ZTickLabel','');
%set(gca,'XTick',[], 'YTick',[], 'ZTick',[]);
axis off;
grid off;
F_left=getframe;
I_left=frame2im(F_left);
I_left=rgb2gray(I_left);
imwrite(I_left,'left_eye.bmp');

```

```

% %creating the right eye
% subplot(1,2,2);
hright=surfc(X,Y,z,'Edgecolor','None','Facecolor','Yellow');
rotate(hright,zdir,0,center);
set(gca,'CameraTarget',[14,10,0]);
%camtarget([10,10,0]);
set(gca,'CameraPosition',[14,10,10]);
%campos([5,10,10]);
set(gca,'Projection','Perspective');
%camproj('Perspective');
camlight left;lighting phong;
set(gca,'DataAspectRatioMode','Manual');
%set(gca,'XTickLabel','', 'YTickLabel','', 'ZTickLabel','');
%set(gca,'XTick',[], 'YTick',[], 'ZTick',[]);
grid off;
axis off;
F_right=getframe;
I_right=frame2im(F_right);
I_right=rgb2gray(I_right);
imwrite(I_right,'right_eye.bmp');

```



# APPENDIX C: CORRELATING LEFT AND RIGHT IMAGES FOR MAPPING COLOR

```

clc;clear;

%Sample code for two dimensional normalized cross-correlation correlating the real and
%synthetic left views of the coral scenario

%-----
%Acquiring images to be matched
f=imread('synthetic_left.bmp');
f=im2gray(f);
f=im2double(f);
htot=imread('real_left.gif');
htot=im2double(htot);

%Initializing template and obtaining image and template sizes
h=zeros(16,16);
[K L]=size(htot);
[P Q]=size(h);

%Function call to record disparities using normalized cross-correlation
%with progressively finer search windows
for iouter=1:(K/P)
    for jouter=1:(L/Q)
        %Generating template to be matched
        h=htot(((iouter-1)*P+1):((iouter-1)*P+P),((jouter-1)*Q+1):((jouter-1)*Q+Q));

        %Segmentation of reference image for matching
        Qindex=(jouter-1)*Q+3*Q;
        if (Qindex > L)
            Qindex=L;
        end
        f3=f2((iouter-1)*P+1:(iouter-1)*P+P,(jouter-1)*Q+1:Qindex);

        %Performing the normalized cross correlation
        cc = normxcorr2(h,f3);
        [max_cc, imax] = max(abs(cc(:)));
        [ypeak, xpeak] = ind2sub(size(cc),imax(1));

        %Recording disparities
        corr_offset(iouter,jouter,1:2) = [ (ypeak-size(h,1)) (xpeak-size(h,2)) ];
    end
end

%-----
%Generating the mapped color image of the synthetic left view
disparity=corr_offset(:, :, 2);
htot1=htot;
for iouter=1:(K/P)
    for jouter=(1:L/Q)
        f_mapped(((iouter-1)*P+1+250):((iouter-1)*P+P+250),((jouter-1)*Q+1+250-...
            disparity(iouter,jouter)):((jouter-1)*Q+Q+250-disparity(iouter,jouter)))...
            =htot1((iouter-1)*P+1:(iouter-1)*P+P,(jouter-1)*Q+1:(jouter-1)*Q+Q);
    end
end
end

```