

Senior Research

# Investigation of Binocular Eye Movements in the Real World

Final Report

Steven R Broskey  
Chester F. Carlson Center for Imaging Science  
Rochester Institute of Technology  
May, 2005

Copyright © 2005  
Steven Broskey  
Center for Imaging Science  
Rochester Institute of Technology  
Rochester, NY 14623-5604

This work is copyrighted and may not be reproduced in whole or part without permission of Steven Broskey or the Center for Imaging Science at the Rochester Institute of Technology.

This report is accepted in partial fulfillment of the requirements of the course 1051-503 Senior Research.

Title: Investigation of Binocular Eye Movements in the Real World  
Author: Steven Broskey  
Project Advisor: Jeff Pelz  
1051503 Instructor: Joseph P. Hornak

# Investigation of Binocular Eye Movements in the Real World

Steven Broskey

Center for Imaging Science  
Rochester Institute of Technology  
Rochester, NY 14623-5604

May, 2005

## **Abstract**

Eye movements are tied to specific tasks or strategies, so monitoring those movements can provide a valuable insight into our methods of perception. Taking advantage of this window, scientists have done eye tracking experiments in an attempt to characterize our visual perception of the world around us. Many eye trackers are laboratory based and immobile. The Visual Perception Lab at RIT utilizes portable monocular eye trackers developed within the lab. [Babcock & Pelz 2004] While tracking one eye provides good data to examine our vision and scene perception, humans are equipped with two eyes, which provide clues we use in our world in addition to those gathered by only one eye. By observing both eyes using the portable eye tracking system we are attempting to look at these additional clues. These new observations often occur in a laboratory setting within defined parameters; it has been shown (Smeets, Hayhoe, and Ballard, 1996), however, that well-constrained experiments may tell more about the constraints than about the properties being observed. The binocular eye tracker was analyzed and a technique was devised to calibrate the binocular eye tracker. The noise and resolution fall-off of the system were characterized and explained.

## Table of Contents

	Page
Copyright	2
Abstract	3
Table of Contents	4
Table of Figures	5
Background	6
Introduction	10
Experimental	12
Results	15
Conclusion	19
Biographical Sketch	20
References	22

## Table of Figures

	Figure
Distribution of photoreceptors in the human eye, rods and cones	1
Illustration of Binocular Disparity using two targets, including labels for equation 1	2
Illustration of Binocular Disparity using two real world example targets	3
The wearable eye tracker	4
Close view of the monocular eye tracking system with parts labeled	5
Illustrative sketch of the calibration target used in this project	6
Diagram of calibration target used in project including distance and angular measurements for targets	7
Example of graph used by experimenters to verify data collection	8
Example of graph used to interpret calibrated data	9
Results graphed as angular comparisons, expected vs. recorded	10
Results graphed as distance comparisons, expected vs. recorded	11

## Background

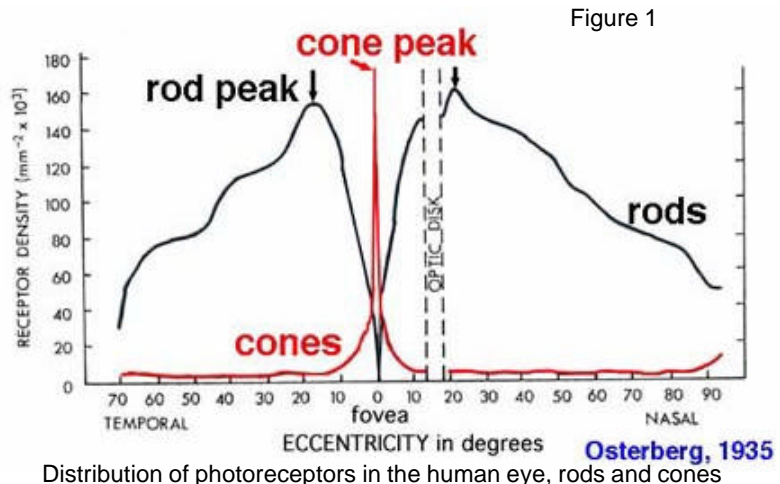
The human visual system is a very complex system. Even the complexity of the physiology is minor compared to the processes of high-level visual perception. On a

basic level, the eye can be considered analogous to a camera; it has an aperture, a lens, and a detector array. The eye's detector array is the retina. Within the retina are the individual

photoreceptors of the eye, the rods and cones. The

uneven distribution of cones in the human retina requires that people move their eyes to perform even simple tasks. Figure 1 shows the distribution of photoreceptors in the human eye. The solid line shaded in blue is the distribution of rods, the low-light photoreceptors, which are disabled in bright viewing environments. The dotted line shaded in red is the distribution of cones in the eye. The peak in cone density designates the center of the retina and area of highest resolution, known as the fovea. The fovea covers less than two degrees of visual angle. To construct a scene for the viewer, the eyes shift the viewer's gaze around the scene, moving the projection of the world across the fovea. In doing this, the fovea is used to sense key regions of the image. The time when the eye stops moving are called fixations. The movements between these pauses are called saccades.

Saccades and fixations have been studied since the work of Yarbus in the 1960's. [Dodge & Cline, 1901; Buswell, 1935] They have been analyzed as an indication of perception, and studied to determine if there is a link between the task assigned to a person viewing a scene, or if scene viewing patterns are task independent. [Pelz 1995] As an observer views a scene, the observer makes version and vergence eye movements; version movements are when both eyes shift across the scene together, and vergence eye



movements are when one eye moves differently than the other, such as when fixating on a close target versus a far target.

On an anatomical level, humans are equipped with two eyes and a certain degree of mental processing to extract the difference in viewing angle between them. The difference in the angles of the two eyes, or vergence, is controlled by three sets of muscles surrounding the eye. These muscles rotate the eye in its socket up, down, left, right, and can also rotate the eye around the axis of gaze (the axis passing straight through the pupil, lens and retina). Vergence eye movements are special among eye movements because the eyes move different amounts depending on how close the object is to the viewer. At close ranges, the axis of gaze crosses close to the viewer. As the distance between the object and the viewer increase, the eyes move to make the two gaze axes closer to parallel to each other. The nerves controlling these muscles pass information to the brain about the position of the eye. Through lower level processing of difference in vergence angle between the two eyes positions (viewing near relative to viewing far), estimates can be determined for distances of objects from the viewer.

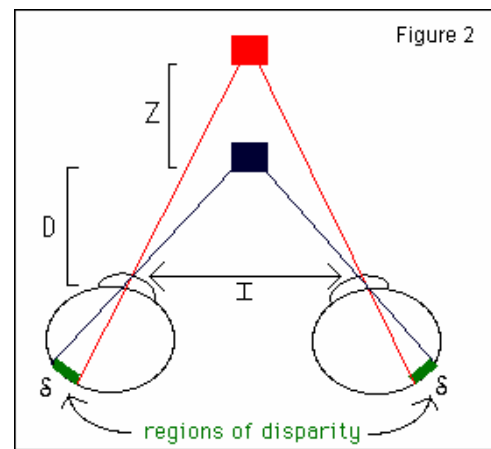


Figure 2  
Illustration of Binocular Disparity using two targets, including labels for equation 1

Difference in vergence angle should not be confused with differences in disparity. An example of disparity is given in Figure 2, which corresponds to equation 1. Disparity

$$\delta \approx \frac{IZ}{D^2}$$

Equation 1

is the difference in retinal position of the images of two different objects. In figure 3, the tree is further from the viewer than the police officer. Notice how the officer and the tree appear at different positions on the retina; this is binocular disparity.

This ability to estimate distance using vergence information is an additional cue humans have to interpret the world around them. Other clues may be more important in determining distance (such as occlusion), disparity is still used as a way to determine lateral distances, or make educated assumptions. [Knill, 2005] These assumptions are

not without other in-scene clues, including viewing distance estimation. The viewing distance estimation is required because depth is related to disparity as in equation 1: In Equation 1,  $I$  is interocular distance,  $Z$  is the distance between two points in depth,  $D$  is the viewing distance, and  $(d)$  is the disparity. It has been established in Foley (1980) that vergence state could contribute as a useable extra-retinal cue for estimating distance, which makes the vergence state of eyes during an eye tracking session an interesting variable to examine.

Where does the required initial estimate of viewing distance come from? The vergence position of an observer's eyes is potentially useful, thanks to the direct relationship between vergence angle and object distance. According to Foley (1980) object distance can be estimated from 10cm to 6m using vergence angle. Erkelens & Collewijn (1985, 1990) have published material noting the possibility of lack of usefulness of vergence information as a clue to distance. This may be a topic of research later, but the usefulness of this eye tracker must first be analyzed. Brenner and associates (1996, 2000) have rebutted these suggestions, and brought to light similarities between extra-retinal clues in visual direction, and extra-retinal clues in distance perception, both of which appear equally flawed; therefore because extra-retinal clues for visual direction are reliably accurate, extra-retinal clues for distance perception have been upheld as well. In examining these conclusions,

There are three methods observers could use to take advantage of vergence eye movement and information to extract depth distances from a scene. First, Enright (1996) showed that observers judging relative distances of objects look at them in turn. This method allows for the observer to extract data by locking the vergence state of the eyes from fixation to fixation so that when the eyes landed on a target, the object would have a

Figure 3

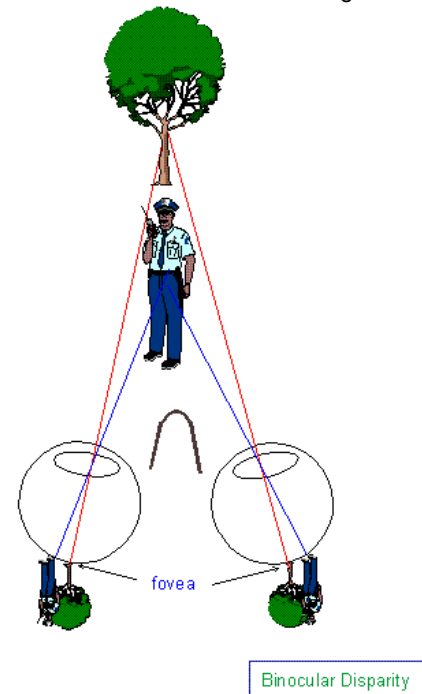


Illustration of Binocular Disparity using two real world example targets

disparity reading from the two retinas which would allow the observer to determine distance. Second, Foley (1980, 1985) put forth the suggestion that an observer's visual system would use a single estimate of distance, a notional reference point, to scale other measurements of disparity and extract depth estimates. With respect to this method, it is not known whether observers fixate on this reference point, or use some other means to determine the distance to this point. In the third method, observers combine changes of version with changes of vergence, as shown in Ono, Nakamizo & Steinbach (1978), and others.

## Introduction

Many apparatuses have been constructed in the history of researching eye movements. Yarbus attached mechanisms directly to his subjects' eyes in order to track their gaze, and intrusive methods like this are still used. While these methods allow for precise, high frequency sampling, comfort for the subject can be reached without sacrificing accuracy in some applications. Most eye trackers are stationary, mounted in a laboratory, and requiring a subject's head to be firmly held to maintain calibration and precision. The Visual Perception Laboratory at the Center for Imaging Science at Rochester Institute of Technology builds and uses portable, "wearable" eye tracking units that allow subjects to walk freely around the world.

An example of a monocular eye tracker built at the VPL is in figure 4. [Babcock & Pelz, 2004] The system is contained within a small backpack and processing of the tracking tape is done offline. Giving a subject maximum freedom of motion should minimize the outside influence that hardware might exert on a subject, including any effect that Smeets, Hayhoe, and Ballard (1996) refer to when commenting that experiments conducted under strict conditions can be more informative about the constraints placed on the observer, rather than about the eye movements themselves.

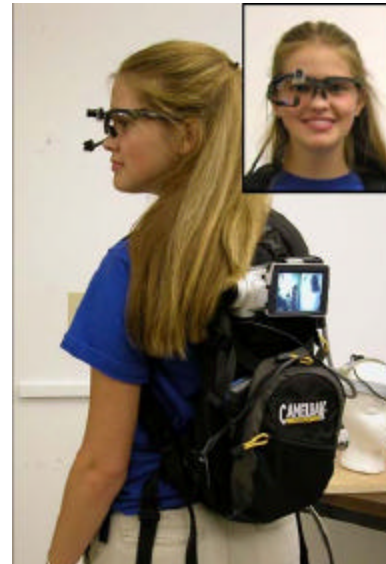
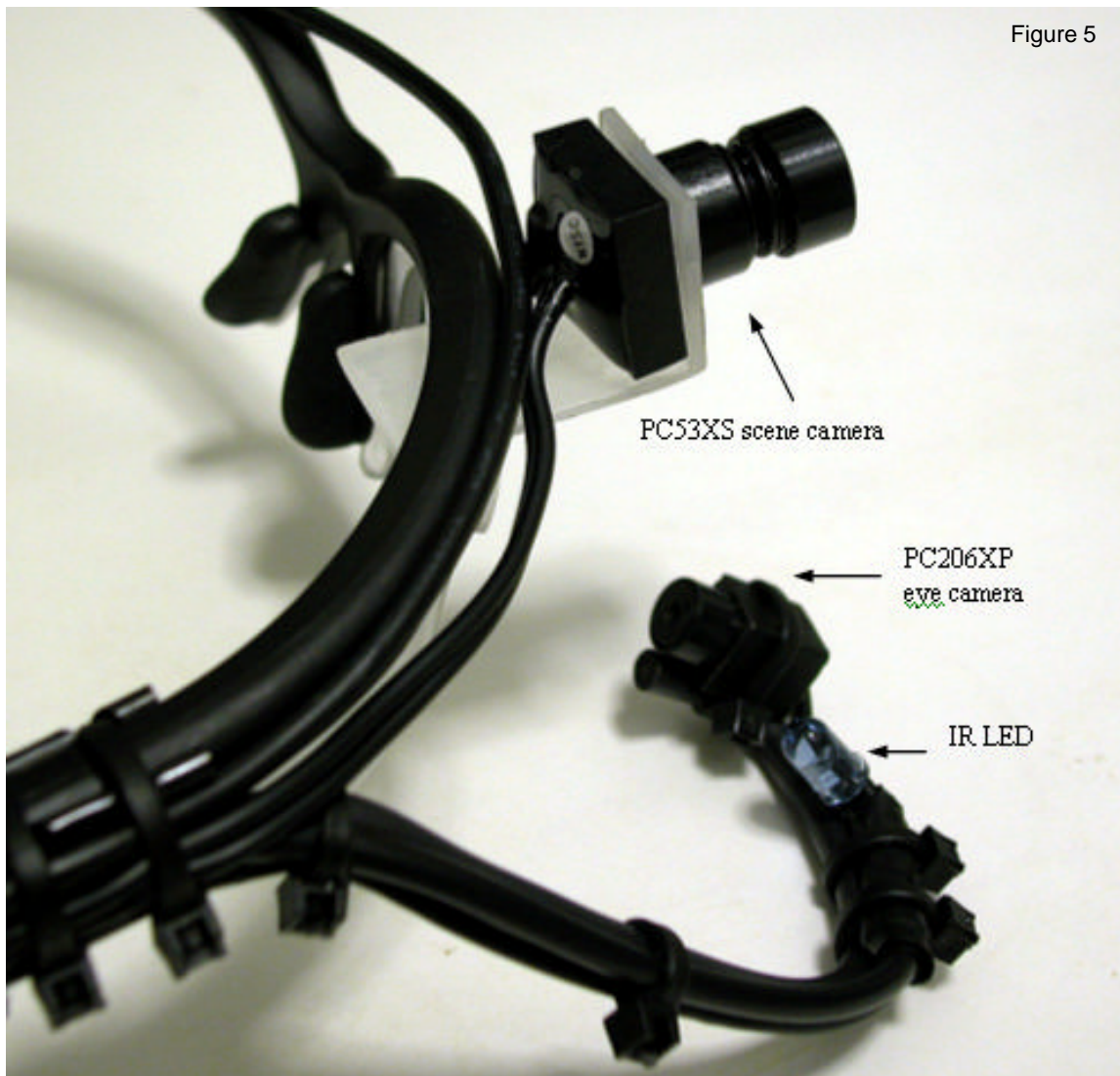


Figure 4

The wearable eye tracker

The parts used in an eye tracker are labeled in figure 5. The binocular eye tracker was similar to a monocular eye tracker (Figures 4 and 5), but the binocular tracker has two scene cameras, two eye cameras, and two IR LEDs. The scene cameras were detached for this project, because it was only focusing on the eye cameras. The scene cameras should be integrated in the steps following the completion of this project. All hardware other than the headgear is located in the camelback® backpack the subject wears (Figure 4). The wearable tracker provides the advantage of going out into the world, which (as pointed out by Smeets et. al) is a handy thing, allowing the subject to be in the environment naturally.

Figure 5

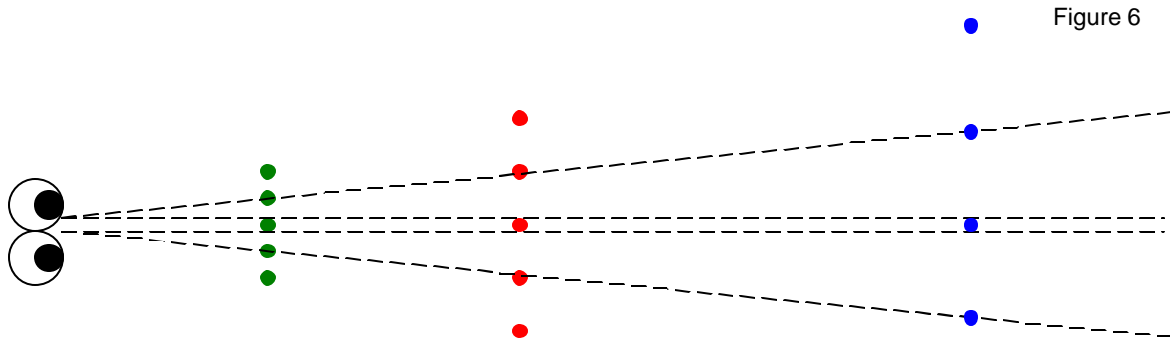


Close view of the monocular eye tracking system with parts labeled

Offline tracking was performed on the video signal recorded during the trial. On the first and second generation wearable eye trackers, the data was processed in real time. However, for the third generation eye trackers (the ones in use now), the video multiplexer combines the signals from the scene camera and the eye camera to output an image with both signals in it. During the analysis, the signal is de-multiplexed, and two images are returned. Because the signals are combined, and this combined output written to a digital videocassette, part of the resolution of the signal is lost. This reduction in signal is dealt with by simply ignoring it, there is enough resolution remaining to do the eye tracking analysis. Using a multiplexer allows the two video streams to be synchronized and remain synchronized the for the whole trial.

## Experimental

A target to analyze both version and vergence eye movements was constructed. Figure 6 shows a top-down view of the target and illustrates how it was used to isolate vergence and version eye movements. By following a dotted line in the diagram, a



Illustrative sketch of the calibration target used in this project

subject maintains visual angle, or the subject does not look side-to-side. If a subject looks from the green point to the red point to the blue point on a given line, the subject makes only vergence movements, without version movements.

The goal is to find whether a few propositions are true: As vergence angle decreases, the measured angle will approach the magnitude of the noise, and the accuracy of the measure of vergence will drop. However, we estimate that within two meters, the system should detect the vergence angle fairly accurately, because it is likely humans will interact with something once it is within two arm's distance from a person.

To gather data the subjects were seated in a room with 78 suspended golf balls. The balls were hung so as to construct the target described in figure 6, with the addition of one more plane in distance. The golf balls in each distance plane were painted a uniform color to make them easier to differentiate. The closest plane to the subject was green, followed by red, then blue, and unpainted (white) golf balls made up the farthest

plane. The planes were situated 5, 10, 20, and 28 feet from the subject, and each contained 21 balls, with the exception of the closest, which contained 15 balls.

Figure 7 shows the entire target visible from the top, including distances from the subject and degrees of visual angle. Each line maintains a certain visual angle, in increments of 4° out from the center. Within each colored distance plane, the balls are

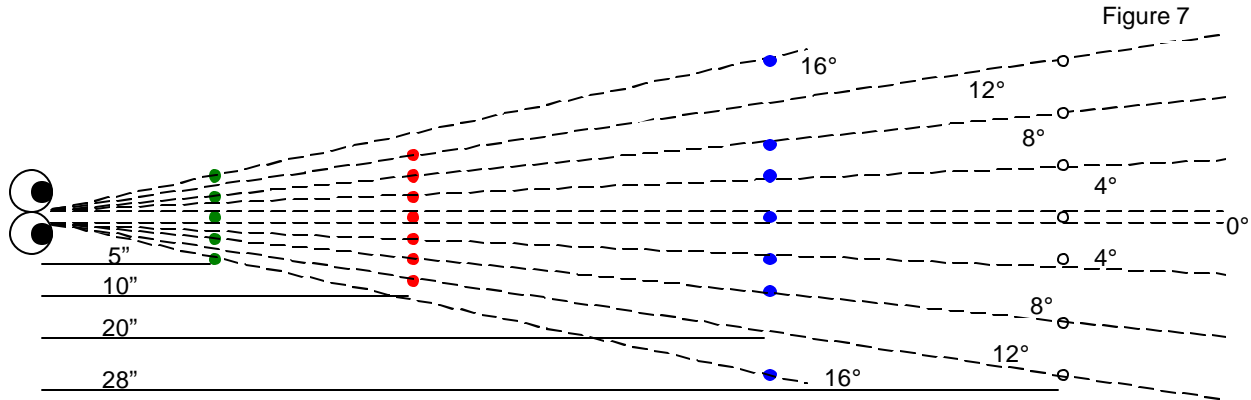


Diagram of calibration target used in project including distance and angular measurements for targets stacked 3 levels high, so that future efforts may also incorporate different target heights.

A chin rest was provided and adjusted to each subject, and used to give the subject a moderate amount of feedback about the position and movement of his/her head. It was essential that a subject's head be restrained; if the subject's head shifted, any subsequent measurements would be thrown off. To isolate the individual targets, a projector was mounted behind the subject and was used to illuminate individual targets. PowerPoint slides were created to allow a small point of light that is projected into the room, and a PowerPoint slideshow sequenced the slides so that any order could be constructed by the experimenter.

Recording the eye movements of the subject, an experimenter cycled through the slides which had been pre-arranged for the trial. Post-processing of the video recording utilized ISCAN eye tracking software, which gave the output of numerical data values.

This data was initially not calibrated, rather a calibration routine was developed and run at the beginning of each subject. This calibration routine began each trial with a centered gaze, allowing a center reference point. Next the subject looked left twelve degrees and right twelve degrees; this compensated for any difference in size between the two eye images recorded on the video track. Since the two eye cameras were not always the same distance from their respective eyes, the image of one eye may be different from the other eye. If the eye images are two different sizes, a twelve-degree movement may appear as different movements in the right eye versus the left eye.

To gather additional information, several trials required subjects to simply stare at a moving target as the target moved towards and away from them. Different distance increments were used. In one portion, a total distance of ten feet was used with marked intervals every foot, while the other trial used a total distance of eight meters with each meter having a marker to designate the distance interval.

## Results

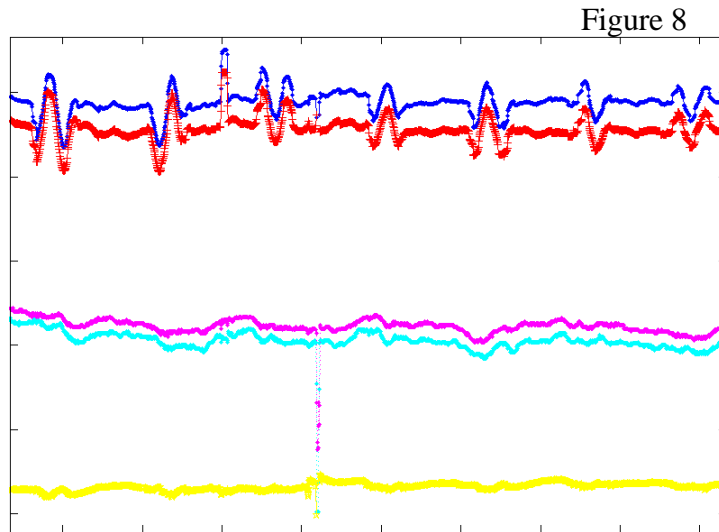
During post-processing, multiple streams of data were gathered including pupil position (horizontal and vertical), corneal reflection (horizontal and vertical), and pupil diameter. For this project only horizontal pupil position and pupil diameter were used. The pupil diameter was used to determine where a blink or track loss occurred, and the horizontal pupil position was

used as the data. Figure 8 shows a plot of raw data: red and blue are right and left (respectively) horizontal pupil position, magenta and cyan are right and left pupil diameter, and yellow is the uncalibrated vergence angle.

Plotting the data simply as an

arbitrary value on the y axis with time on the x axis allows the processor to synchronize the right and left tracks, and to ensure that there are no processing glitches.

Once the video had been tracked and raw data was gathered, custom MatLab code was run on the data to process it for final analysis. The MatLab code synchronized the right and left eye data, and eliminated any data corresponding to blinks or track losses. A blink or track loss was defined as any point where the pupil diameter dropped below 70% of the mean pupil diameter for the whole trial. The Vergence was roughly calculated at first by simply subtracting the left eye pupil data from the right eye pupil data.



Example of graph used by experimenters to verify data collection. The red and blue plots are right and left eye pupil positions. The magenta and blue are right and left eye pupil diameters. The yellow plot is uncalibrated vergence angle.

To obtain calibration, the track during the calibration routine was manually analyzed for the raw data values. The data values at the left and right twelve degree

marks were used to calculate a slope-intercept linear equation to transfer the raw data points to calibrated visual angles. Once the left and right horizontal pupil data had been calibrated, the left eye's visual angle could

be subtracted from the right eye's visual angle to obtain the vergence angle.

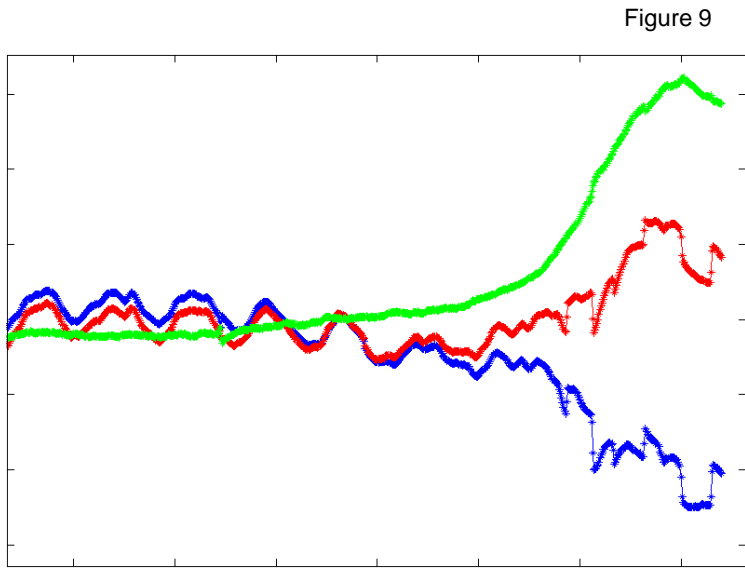


Figure 9

Example of graph used to interpret calibrated data. The green plot is vergence angle, the red and blue are right and left pupil measurements

Expressing the processing of the system was then shown by plotting the expected vergence angle on the x axis, and the measured vergence angle on the y axis – showing how well the system correctly predicted the vergence angle of the subject's eyes. The predictions expected high noise at small vergence angles and good system throughput at less than three meters. Figure

10 demonstrates the high noise at low vergence angles, however it is not obvious that the noise is relatively high, since the error bars are relatively constant. Instead

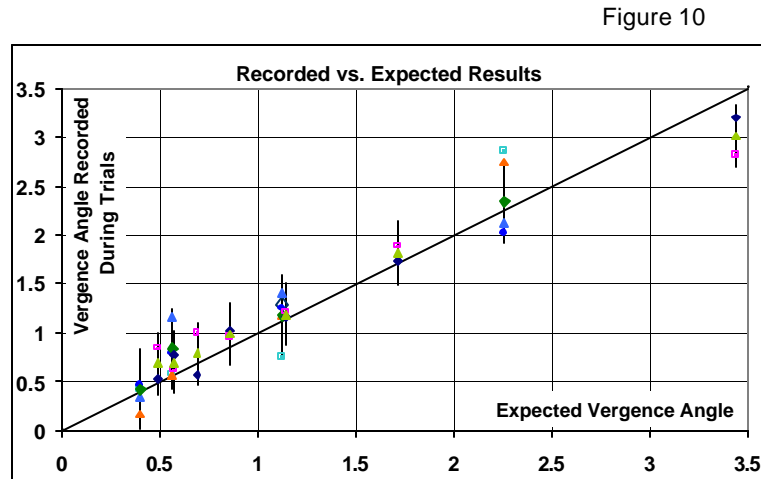


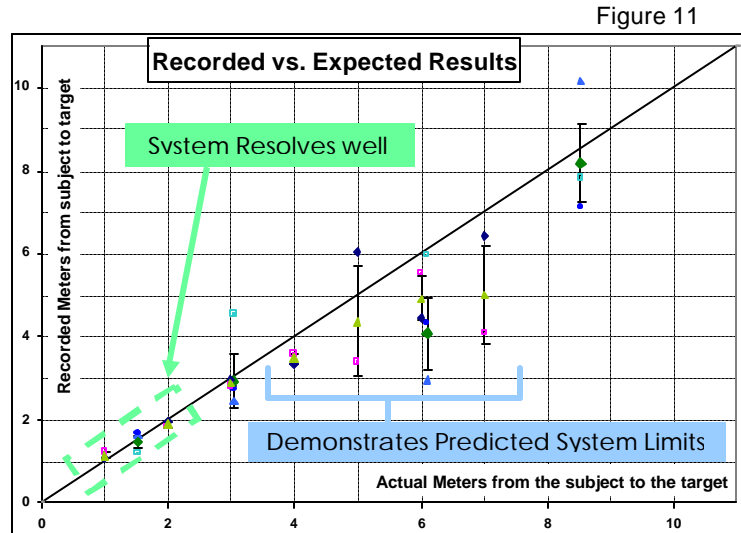
Figure 10

Results graphed as angular comparisons, expected vs. recorded. Each type of data point in Figure 10 corresponds to a different subject.

the noise and resolving power of the system is better shown when graphing distances.

Plotting the data in distance space shows the noise fairly well at distances of three meters or greater. Figure 11 shows the plot of the data in distances instead of angles.

Between one meter and two meters the system's measured distance is almost the same as the actual distance between subject and target. At three meters the errors begin to manifest themselves, and the error bars continue to be large,



Results graphed as distance comparisons, expected vs. recorded. Each type of data point in Figure 11 corresponds to a different subject.

with the exception of the four meter point. However, even at the four meter point, which has good precision, the clustering represents an inaccurate prediction of the distance between subject and target. Within the four to eight meter range, the resolved distance continues to drop away from the actual distance. This can likely be explained as a possible limit in system resolution.

As the distance between target and subject grows, the vergence angle decreases. Therefore, at large distances, a small vergence angle – which is the quantity the system is measuring – is expected and is seen. As this measurable quantity grows smaller, it approaches the noise threshold that is present in any electronic system. The noise therefore contributes larger and larger amounts of variance to the measurement as the system resolves smaller and smaller data values, or small angles of vergence. Thus when

variation increases as vergence angle decreases, the variation will increase as distance between subject and target increases.

## **Conclusion**

The system performed up to the expectations of this experimenter, accurately detecting when subjects were looking small distances (between one and two meters away), and began to fail where expected (at distances above three meters from subject to target). This means that the system should be used to gather data in ranges closer than three meters from the subject. If data is needed closer to the subject than three meters, further analysis is necessary.

It is possible that analysis of the noise's interference with the signal, and the limit of distance detection could be combined in the future to extract more information regarding the noise. The system is now ready to be integrated as originally built, with scene cameras. Now that a calibration scheme has been worked out, a portable method of calibration should also be devised. One possible way to do this includes a diffraction grating interfaced with a laser diode to create a grid to be projected on a surface in the real world.

## Resume and Biographical Sketch

### PERMANENT ADDRESS

1871 Dolphin Dr  
Allison Park, PA 15101  
412-366-7166

### LOCAL ADDRESS

149D Perkins Rd  
Rochester, NY 14623  
585-424-8300

---

### **OBJECTIVE**

Obtain a full-time job designing, developing, testing, or manufacturing imaging systems

### **EDUCATION**

Rochester Institute of Technology;  
Bachelor of Science in Imaging Science  
Current GPA: 3.0

College of Science  
Expected: 5/2005

Curriculum sample:

Digital Image Processing I and II

Radiometry with Lab

Optics for Imaging with Lab

Imaging Systems Laboratory I and II

Interaction between Light and Matter

Electronic Measurements with Lab

Vision and Psychophysics

Programming for Imaging Scientists I and II

Addition Course Information Available on Request

### **SKILLS**

Digital Image Processing, Sensor Characterization, Optical System Design, Circuit Construction  
Computer programming: IDL (Interactive Data Language), Java, Pascal, HTML, Visual Basic  
Language Skills: German (Eight years, verbal and literal), American Sign (Three years)

### **EXPERIENCE**

Binocular Eye-tracking Primary Researcher

(2004 - present) Rochester, NY

Image Processing Technician

(Summer 2004) Rochester, NY

Ultrasound Research Technician

(Spring 2004) Rochester, NY

Optics Stockroom Attendant

(2002 - 2004) Rochester, NY

Various Service Industry Jobs

(1995 - 2002) Pittsburgh, PA

### **VOLUNTEER WORK**

Habitat for Humanity™

March 2003, 2004

Project H.O.P.E.

June 2002

### **ADDITIONAL INFORMATION**

Active Member of RIT student chapter of Imaging Science and Technology

Student Research Scientist in Ultrasonic System Characterization Group

RIT Swing Dance Club Publicist 2002-2003

Tripled Attendance during tenure

### **REFERENCES**

Available on Request, also available online

*This resume is available online at [www.cis.rit.edu/~srb0036/Resume.htm](http://www.cis.rit.edu/~srb0036/Resume.htm)*

The author of this paper is completing his Bachelor of Science in Imaging Science from the Rochester Institute of Technology in May, 2005. He plans to enter the professional field in Imaging Systems, and later study towards a Masters of Science in Systems Engineering or Mechanical Engineering.

Jeff Pelz served as the advisor for this project, and is supervisor of the Visual Perception Laboratory at the Chester F. Carlson Center for Imaging Science at RIT.

## References

- Brenner, E. & van Damme, W. J. M. 1998, Judging distance from ocular convergence, *Vision Research*, 38(4), 493-498
- Buswell, G. T. (1935). *How people look at pictures: a study of the psychology and perception in art*. Oxford, England: Univ. Of Chicago Press
- Collewijn, H. & Erkelens, C. J. 1990, Binocular eye movements and the perception of depth, in E. Kowler ed., *Eye movements and their role in visual and cognitive processes*. Amsterdam
- Dodge, R., & Cline, T. S. (1901). The angle velocity of eye movements. *Psychological Review*, 8(2), 145-157
- Enright, J. T. Enright, 1996, Sequential stereopsis: A simple demonstration, *Vision Research*, 36(2), 307-312
- Enright, J. T. 1991, Exploring the 3rd-dimension with eye-movements - better than stereopsis, *Vision Research*, 31(9), 1549-1562
- Erkelens, C. J. & Collewijn, H. 1985, Motion perception during dichoptic viewing of moving random-dot stereograms. *Vision Research*, 25(4), 583-588
- Foley, J. M. 1985, Binocular distance perception: egocentric distance tasks, *Journal of Experimental Psychology: Human Perception and Performance*, 2, 133-149
- Foley, J. M. Foley, 1980, Binocular distance perception, *Psychological Review*, 87(5), 411-434
- Harris, J. M. & Welchman, A. E. 2003, Task demands and binocular eye movements, *Journal of Vision*, 3:817-830
- Pelz, J. B. 1995, *Visual Representations in a Natural Visuo-motor Task*, Thesis, Carlson Center for Imaging Science, Rochester Institute of Technology
- Knill, D. C. 2005, Reaching for visual cues to depth: The brain combines depth cues differently for motor control and perception, *Journal of Vision* (2005) 5, 103-115
- Smeets, J. B. J., Hayhoe, M. M., & Ballard, D. H. 1996, Goal-directed arm movements change eye-head coordination, *Experimental Brain Research*, 109(3), 434-440
- Wright, W. D. 1951, The role of convergence in stereoscopic vision, *The Proceedings of the Physical Society*, 64(376B), 289-297

Yarbus, A. L. 1967, Eye movements during perception of complex objects, in L. A. Riggs, ed., 'Eye Movements and Vision', Plenum Press, New York

### **Exterior Source Image Credits**

Figure 1: [http://www.phys.ufl.edu/~avery/course/3400/gallery/gallery\\_vision.html](http://www.phys.ufl.edu/~avery/course/3400/gallery/gallery_vision.html)

Figures 4, 5: Babcock & Pelz, Building a Lightweight Eyetracer