

# **Seeing, Sensing, and Selection: Modeling Visual Perception in Complex Environments**

Roxanne Canosa

A dissertation submitted in partial fulfillment of the  
requirements for the degree of Doctor of Philosophy  
in the Center for Imaging Science in the College of Science  
Rochester Institute of Technology

September 29, 2003

Signature of Author \_\_\_\_\_  
Roxanne Canosa

Accepted by \_\_\_\_\_  
Coordinator, Ph.D. Degree Program Date



THESIS RELEASE PERMISSION  
ROCHESTER INSTITUTE OF TECHNOLOGY  
COLLEGE OF SCIENCE

---

**Seeing, Sensing, and Selection:  
Modeling Visual Perception in Complex Environments**

---

I, Roxanne Canosa, hereby grant permission to the Wallace Memorial Library of RIT to reproduce my dissertation in whole or part. Any reproduction will not be for commercial use or profit.

Signature of Author \_\_\_\_\_  
Roxanne Canosa

Date \_\_\_\_\_



CHESTER F. CARLSON  
CENTER FOR IMAGING SCIENCE  
COLLEGE OF SCIENCE  
ROCHESTER INSTITUTE OF TECHNOLOGY  
ROCHESTER, NEW YORK

CERTIFICATE OF APPROVAL

---

Ph.D. DEGREE THESIS

---

The Ph.D. degree of Roxanne Canosa  
has been examined and approved by the dissertation committee  
as satisfactory for the dissertation requirement for the  
Doctor of Philosophy degree in Imaging Science

---

Dr. Jeff B. Pelz, Dissertation Advisor

---

Dr. Marc Marschark, Committee Chair

---

Dr. Julie A. Adams

---

Dr. Dana H. Ballard

---

Dr. Roger S. Gaborski

---

Date



## Acknowledgments

I would like to thank and acknowledge all those who helped to make this dissertation possible. Thanks to Jeff Pelz for his guidance, support, and encouragement over the many years of work that were required to bring this research to fruition. None of this would have been possible without his vast knowledge of the subject matter, technical expertise, and cheerful willingness to discuss and explore new ideas. Thanks also to Jason Babcock for the countless hours spent developing and fine-tuning the eye-tracker, and for sharing his ideas on experimental design and data analysis. This work was built upon earlier work conducted by Jeff Pelz, Mary Hayhoe, and Dana Ballard at the University of Rochester, and also Roger Gaborski at the Rochester Institute of Technology. Thanks for their insights and for providing the “shoulders upon which I stand.” I am also greatly indebted to Julie Adams for her thorough review of this work and her many useful and enlightening comments. The skills and ideas of Marianne Lipps, Constantin Rothkopf, and Vishal Vaingankar have proved to be invaluable, and I am grateful for our many discussions that helped to form and focus this dissertation. Also, thanks to Marc Marschark for his support. This work was supported in part by the Naval Research Laboratories, the New York State Office of Science, Technology, and Academic Research, and the Xerox Corporation. Finally, I owe my greatest appreciation to John, Elyse, Sandra, and Dennis Canosa, for their continued love and support through the many years of my academic pursuits.



## Curriculum Vitae

The author was born in Rochester, New York and has spent virtually her entire life there. She attended college as a first-generation college student at the State University of New York College at Potsdam, earning a Bachelor of Arts degree in Art History in 1980. She then returned to Rochester and earned an Associate of Applied Science degree in Optical Engineering Technology from Monroe Community College in 1983, and studied Electrical Engineering part-time at the Rochester Institute of Technology. While raising three children and working at JML Optical Industries, Inc., and Eastman Kodak Company, she returned to school to earn yet another degree. In May of 1998 she earned the Bachelor of Science degree in Computer Science at the State University of New York College at Brockport, and the following fall began graduate studies in Imaging Science at RIT. In October of 2000 she earned the Master of Science degree in Imaging Science, and began doctoral studies under the direction of Professor Jeff Pelz. She is currently an Assistant Professor of Computer Science at the Rochester Institute of Technology.



## Abstract

The purpose of this thesis is to investigate human visual perception at the level of eye movements by describing the interaction between vision and action during natural, everyday tasks in a real-world environment. The results of the investigation provide motivation for the development of a biologically-based model of selective visual perception that relies on the relative perceptual conspicuity of certain regions within the field of view. Several experiments were designed and conducted that form the basis for the model. The experiments provide evidence that the visual system is not passive, nor is it general-purpose, but rather it is active and specific, tightly coupled to the requirements of planned behavior and action. The implication for an active and task-specific visual system is that an explicit representation of the environment can be eschewed in favor of a compact representation with large potential savings in computational efficiency. The compact representation is in the form of a topographic map of relative perceptual conspicuity values. Other recent attempts at compact scene representations have focused mainly on low-level maps that code certain salient features of the scene including color, edges, and luminance. This study has found that the low-level maps do not correlate well with subjects' fixation locations, therefore, a map of perceptual conspicuity is presented that incorporates high-level information. The high-level information is in the form of figure/ground segmentation, potential object detection, and task-specific location bias. The resulting model correlates well with the fixation densities of human viewers of natural scenes, and can be used as a pre-processing module for image understanding or intelligent surveillance applications.



# Table of Contents

<b>List of Figures</b>	<b>xvi</b>
------------------------	------------

<b>List of Tables</b>	<b>xxiv</b>
-----------------------	-------------

<b>1. Introduction</b>	<b>1</b>
1.1 Overview.....	1
1.2 Problem statement.....	3
1.3 Outline of the presented work.....	5
<b>2. Background</b>	<b>9</b>
2.1 Historical perspective.....	9
2.2 The human visual system.....	15
2.2.1 Image formation.....	15
2.2.2 Center-surround organization of receptive fields.....	18
2.2.3 Contrast sensitivity function.....	20
2.2.4 Opponent processes.....	22
2.2.5 Eye movements.....	24
2.3 Visual attention and selectivity.....	26
2.3.1 The influence of attention on neural response.....	26
2.3.2 Orienting of attention.....	29
2.3.3 Behavioral data on selectivity and capacity limitations.....	32
2.4 Task-oriented vision.....	35
2.4.1 Task-dependency of visual scan paths.....	36
2.4.2 Limited memory representations.....	38
2.4.3 Natural tasks.....	42
2.5 Computational modeling of visual attention.....	45
2.5.1 Hierarchical models of attention.....	45
2.5.2 Connectionist models of attention.....	47
2.5.3 Graphical models of attention.....	49
2.5.4 Saliency models and guided search models.....	53

<b>3. Approach</b>	<b>57</b>
3.1 Overview.....	57
3.2 The benefits of eye-tracking.....	58
3.3 Eye-tracking – theory of operation.....	59
3.3.1 Bright-pupil detection.....	59
3.3.2 Calculation of eye position.....	61
3.4 The VPL portable eye-tracker.....	62
3.4.1 The optics module and mirror.....	64
3.4.2 The eye camera.....	65
3.4.3 The scene camera.....	66
3.4.4 The LASER.....	67
3.4.5 The control unit.....	67
3.4.6 Eye-tracker set up and calibration.....	68
3.4.7 Eye movement monitoring.....	71
3.4.8 Portable eye-tracker precision, accuracy, and noise.....	72
3.5 The ASL model 501 eye-tracker.....	76
3.5.1 Integrating head movements.....	77
3.5.2 ASL eye-tracker precision, accuracy, and noise.....	78
3.5.3 Estimation and correction of accuracy loss.....	80
3.5.4 Fixation finding.....	82
<b>4. Modular Visual Routines</b>	<b>87</b>
4.1 Introduction.....	87
4.2 Method.....	91
4.3 Results.....	95
4.3.1 Mean fixation durations of tasks – pooled data.....	96
4.3.2 Variance of fixation duration – pooled data.....	101
4.3.3 Statistical differences between subjects.....	106
4.3.4 Mean saccade amplitude of tasks – pooled data.....	111
4.3.5 Variance of saccade amplitude – pooled data.....	117
4.3.6 Statistical differences between subjects.....	119
4.4 Discussion.....	121
<b>5. Task-dependencies of Fixation Locations</b>	<b>127</b>
5.1 Introduction.....	127
5.2 Fixation locations in a simple environment.....	133
5.2.1 Method.....	134
5.2.2 Results.....	138
5.2.3 Discussion.....	145
5.3 Fixation locations in an extended environment.....	145
5.3.1 Method.....	146
5.3.2 Results.....	150
5.4 General discussion and conclusions.....	163

<b>6. The Conspicuity Map</b>	<b>167</b>
6.1 Overview.....	167
6.2 Model description.....	169
6.2.1 Input image processing.....	169
6.2.2 The low-level saliency map.....	175
6.2.3 High-level proto-object map.....	187
6.2.4 Expected location mask.....	189
6.3 Verification of model using eye-tracking methods.....	191
6.3.1 Data collection.....	191
6.3.2 Comparison of fixation densities to model predictions.....	192
6.3.3 Determination of map weights.....	196
6.4 Natural-task images.....	206
6.4.1 Comparison to extended environment.....	206
6.4.2 Free-view and multi-view.....	213
6.4.3 Estimation of location bias.....	215
6.4.4 Expected locations.....	221
6.5 General discussion and conclusion.....	224
<b>7. Conclusion</b>	<b>227</b>
<b>References</b>	<b>239</b>
<b>Appendix</b>	<b>249</b>

## List of Figures

- Figure 2-1 Cross-section of the human eye, depicting image formation components...16
- Figure 2-2 Spectral sensitivities of the three types of cones. Measurements include light loss due to absorption from the cornea, lens, and other pigments in the eye.....17
- Figure 2-3 Receptive fields of two types of retinal neurons: on-center/off-surround and off-center/on-surround. Yellow areas indicate locations of light stimulus.....19
- Figure 2-4 Contrast sensitivity function with example spatial frequencies and on-center/off-surround neuron tuned to the peak response.....21
- Figure 2-5 Study showing that scan paths are task dependent. Original painting of I. E. Repin's *Unexpected Return* is at upper left, with five example scanpaths for a single subject who viewed the painting while being asked to formulate answers to the various questions.....38
- Figure 3-1 Image of the pupil (white) and corneal reflection (black) as detected by the eye camera. Centers are indicated by crosshairs. A slight offset between the actual centers of the images and the displayed centers is due to a timing offset during data capture, and does not affect calculation of eye movement amplitude and direction...60
- Figure 3-2 Calculation of the line-of-gaze.....61
- Figure 3-3 Portable eye-tracking headgear and backpack.....63
- Figure 3-4 Top view of headgear.....64
- Figure 3-5 Optics module.....64
- Figure 3-6 Diffraction pattern used for calibration.....69
- Figure 3-7 Eye movement trace after calibration. The subject was instructed to look at each of the nine target points, from upper left to bottom right, for approximately two

seconds each.....	71
Figure 3-8 Vertical eye position.....	73
Figure 3-9 Horizontal eye position.....	73
Figure 3-10 Expanded view of Figure 3-8.....	73
Figure 3-11 Expanded view of Figure 3-9.....	73
Figure 3-12 Eye-tracker noise, no averaging.....	74
Figure 3-13 Eye-tracker noise, two field ave.....	74
Figure 3-14 Eye-tracker noise, four field ave.....	74
Figure 3-15 Eye-tracker noise, eight field ave.....	74
Figure 3-16 Average angular deviation for each of the nine calibration points at the start of the experiment, across eight subjects.....	75
Figure 3-17 Average angular deviation for each of eight subjects at the start of the start of the experiment, across nine calibration points.....	75
Figure 3-18 Average angular deviation for each of the nine calibration points at mid-experiment, across six subjects.....	76
Figure 3-19 Average angular deviation for each of six subjects at mid-experiment, across nine calibration points.....	76
Figure 3-20 ASL model 501 eye-tracker.....	77
Figure 3-21 Deviations from calibration target points at the start of the experiment, before and after correction across eleven subjects.....	80
Figure 3-22 Deviations from calibration target points at the end of the experiment, before and after correction, across eleven subjects.....	80
Figure 3-23 Deviations from calibration target points at the start of the experiment, before and after correction, across nine points.....	80
Figure 3-24 Deviations from calibration target points at the end of the experiment, before and after correction, across nine points.....	80
Figure 3-25 Raw ASL eye-head data.....	85
Figure 3-26 Fixation locations after error correction.....	85

Figure 4-1	Relative frequency of fixation durations for subjects JB and JP for Reading, Search, and Manipulations during rocket-building.....	89
Figure 4-2	Fixation sequences for three sub-tasks in the rocket-building task – Reading, Search, and Manipulation. Bars indicate periods of fixation, spaces indicate gaze changes between fixation points.....	90
Figure 4-3	Walking along a hallway.....	92
Figure 4-4	Having a face-to-face conversation.....	92
Figure 4-5	Telephone conversation.....	92
Figure 4-6	Sorting cards.....	92
Figure 4-7	Sorting blocks.....	92
Figure 4-8	Reading poster.....	92
Figure 4-9	Reading form.....	93
Figure 4-10	Counting change.....	93
Figure 4-11	Counting red blocks.....	93
Figure 4-12	Mean fixation duration for each of the nine tasks, pooled across all eight subjects.....	97
Figure 4-13	95% confidence interval of the mean fixation durations for each of the tasks. A statistically significant difference between two tasks exists if there is no overlap of the corresponding confidence intervals. Center dots represent the mean values.....	99
Figure 4-14	The gamma density function with $a = 2$ and $\beta = 1$ .....	103
Figure 4-15	Walk hall pooled data.....	103
Figure 4-16	Talk conversation pooled data.....	103
Figure 4-17	Talk telephone pooled data.....	104
Figure 4-18	Sort cards pooled data.....	104
Figure 4-19	Sort blocks pooled data.....	104
Figure 4-20	Read poster pooled data.....	104

Figure 4-21 Read form pooled data.....	104
Figure 4-22 Count change pooled data.....	104
Figure 4-22 Count blocks pooled data.....	104
Figure 4-24 Relationship between mean and standard deviation for all of the tasks. From left, tasks are: RF, CC, RP, SC, SB, TT, WH, CB, TC.....	104
Figure 4-25 Mean fixation duration for each subject, all tasks.....	107
Figure 4-26 Calculation of visual angle from field of view.....	113
Figure 4-27 Mean saccade amplitude for each of the nine tasks, pooled across all eight subjects.....	113
Figure 4-28 Mean saccade amplitude for each subject, all tasks, with standard error bars.....	114
Figure 4-29 95% confidence intervals of the mean saccade amplitudes for each of the nine tasks. An overlap between two or more intervals indicates that there is no statistically significant difference between the corresponding mean values.....	115
Figure 4-30 Walk hall pooled data.....	117
Figure 4-31 Talk conversation pooled data.....	117
Figure 4-32 Talk telephone pooled data.....	117
Figure 4-33 Sort cards pooled data.....	117
Figure 4-34 Sort blocks pooled data.....	117
Figure 4-35 Read poster pooled data.....	117
Figure 4-36 Read form pooled data.....	117
Figure 4-37 Count change pooled data.....	117
Figure 4-38 Count blocks pooled data.....	117
Figure 4-39 Relationship between the mean and the standard deviation of saccade amplitude. From the left, tasks are RP, CC, CB (lower), RF, SB (lower), TC, TT, and WH (same), and SC.....	119

Figure 4-40 Comparison of tasks in terms of mean fixation duration and mean saccade amplitude.....	123
Figure 5-1 Block copying task. This is the display that was shown on the computer screen. The display subtended an area of 17° x 13° visual angle. A trace of the eye movement and of the hand movement is shown as arrows connecting the different regions.....	129
Figure 5-2 Eye movement strategies used for block copying task. Relative frequencies of each strategy from a sample containing approximately fifty block moves for each of seven subjects.....	130
Figure 5-3 Fixation duration as a function of task difficulty for a driving task.....	131
Figure 5-4 Sorting Cards.....	135
Figure 5-5 Sorting Blocks.....	135
Figure 5-6 Copy-model-same-room.....	135
Figure 5-7 Model from copy-model-different-room.....	136
Figure 5-8 Resource and Workspace from copy-model-different-room.....	136
Figure 5-9 Amount of time taken by each subject to complete each of the four tasks. The tasks along the x-axis are ordered according to the order of performance by Group 1 (subjects B, D, F, and H). The first four bars for each task correspond to the Group 1 subjects, and the second four bars correspond to the Group 2 subjects (A, C, E, and G) who performed the tasks in the reverse order.....	138
Figure 5-10 Division of time between the two different regions – sorting blocks and sorting cards.....	142
Figure 5-11 Division of time between three different regions – copy model same room and copy-model different room.....	143
Figure 5-12 Depiction of four extended environments used for the portable eye-tracking study. Clockwise from the top left, Washroom, Hallway, Office, and Vending.....	149
Figure 5-13 Relative amounts of time spent on different objects in the Washroom environment, pooled across all fixations and all subjects.....	151
Figure 5-14 Washroom environment. Time spent fixating objects as the tasks progress for Subject T. Tasks are, from the top, “Wash your hands,” “Fill a cup with water,” and “Comb your hair.”.....	153

Figure 5-15	Relative amounts of time spent on different objects in the Hallway environment, pooled across all fixations and all subjects.....	155
Figure 5-16	Relative amounts of time spent on different objects in the Office environment, pooled across all fixations and all subjects.....	157
Figure 5-17	Relative amounts of time spent on different objects in the Vending environment, pooled across all fixations, and all subjects.....	158
Figure 5-18	Hallway environment. Time spent fixating objects as the tasks progress for Subject T. Tasks are, from the top, “Throw something in the garbage,” “The fire alarm just went off,” and “Find a bathroom.”.....	160
Figure 5-19	Office environment. Time spent fixating objects as the tasks progress for Subject U. Tasks are, from the top, “Get supplies from the closet,” “Work at the computer,” and “Make a photocopy.”.....	161
Figure 5-20	Vending machine environment. Time spent fixating objects as the tasks progress for Subject U. Tasks are, from the top, “Check for Skittles,” “Buy a Snickers bar,” and “Check for change.”.....	162
Figure 5-21	Comparison of fixation types.....	164
Figure 6-1	Construction of the Conspicuity Map.....	170
Figure 6-2	Creation of the color map from photoreceptor responses. Upper left is input image, upper right is C1 (red/green) signal, lower left is C2 (blue/yellow) signal, and lower right is the resulting color map. Dark blue areas in the signal maps correspond to low signal values, yellow corresponds to medium signal values, and red corresponds to high signal values.....	176
Figure 6-3	Intensity map for example input image of Figure 6-2.....	177
Figure 6-4	The seven levels of the multi-resolution Gaussian pyramid for the example input image.....	178
Figure 6-5	The seven Gaussian convolution kernels of the Laplacian pyramid. They are used to create the bandpass filters that detect a specific range of frequencies in the input image. The spatial domain representation is shown in the top row, and the corresponding frequency domain representation is shown in the bottom row.....	180
Figure 6-6	Six bandpass filters used to detect frequencies of a particular range in the input image. F1-F2 is elliptical in shape because f1 is odd-sized (5x5) and f2 is even-sized (10x10).....	181

- Figure 6-7 One-dimensional frequency response characteristics of the six bandpass filters shown in Figure 6-6. Note that only the right half of the response curves are shown, *i.e.*, they are symmetrical about the origin.....181
- Figure 6-8 Six levels of the Laplacian edge cube (difference-of-Gaussians) derived from the seven levels of the Gaussian pyramid (labeled L0 – L6) after weighting each Laplacian level by the response from the contrast sensitivity function. Top row from left L0-L1, L1-L2, and L2-L3. Bottom row from left L3-L4, L4-L5, and L5-L6....183
- Figure 6-9 Basis functions of the Gabor filters used to model the tuning of receptive fields in area V1 of striate cortex. From left, 0°, 45°, 90°, and 135°.....184
- Figure 6-10 Four oriented edge signals and resulting oriented edge map. Top row from left, 0°, 45°. Middle row, from left, 90°, 135°. Bottom row, false-colored oriented edge map.....185
- Figure 6-11 Low-level feature maps and resulting low-level saliency map. Top row, from left, color map and intensity map. Bottom row from left, oriented edge map and saliency map.....186
- Figure 6-12 Creation of the binary proto-object map. Top row from left, input image, estimation of background, and foreground segmentation. Bottom row from left, after thresholding the foreground image and Canny edge detection, after dilation, and after hole filling and erosion.....188
- Figure 6-13 Comparison of F/M ratios for 76 images in set A, free-view condition. Three maps were generated for each image, as given in Equations 6-15 through 6-17. Images numbered 44-47, 64-66, 69-71, and 80-83 are duplicate images, and not shown here.....195
- Figure 6-14 Comparison F/M ratios for 76 images in set B, free-view condition. Three maps were generated for each image, as given in Equations 6-15 through 6-17. Images numbered 1-4, 44-47, 67, 68, and 80-84 are duplicate images, and are not shown here.....196
- Figure 6-15 Mean F/M ratios for the three different maps, averaged across all 152 images.....200
- Figure 6-16 Example images and overlaid fixation plot for which the optimal weights were found using the random weight generation method. The corresponding weighted conspicuity map (C-Map) is shown beneath each image. From left, A1, A28, B17.....200
- Figure 6-17 F/M ratio for set A images, free-view condition, using the C-Map. The F/M ratios for the other three maps are the same as shown in Figure 6-13, and are included for comparison to the C-Map. Note that the C-Map F/M ratio for Image

74 is off the chart and has a value of 6.89.....	204
Figure 6-18 F/M ratio for set B images, free-view condition, using the C-Map. The F/M ratios for the other three maps are the same as shown in Figure 6-14, and are included for comparison to the C-Map.....	205
Figure 6-19 Mean F/M ratios for all 152 images using four different maps.....	206
Figure 6-20 Four natural-task images with overlaid fixation plots from one subject, free-view condition, and corresponding maps. From left, Washroom (A1), Hallway (A2), Office (A3), and Vending (A4). Maps are, from top to bottom, the CIE map, the P map, the CIEP map, and the C-Map.....	208
Figure 6-21 Fixation density plots for free-view and three multi-view conditions for four images. Images are from top, Washroom, Hallway, Office, and Vending Machine.....	210
Figure 6-22 F/M ratios for free-view and multi-view conditions for the four natural-task images. A comparison is shown between the low-level CIE map and the high-level perceptual conspicuity C-Map for each image.....	214
Figure 6-23 F/M ratios for 1000 randomly generated fixation locations.....	215
Figure 6-24 F/M ratios computed for mixed image and fixation data. Each chart is for one of the four images for which two maps were computed, CIE map (saliency) and C-Map (conspicuity). The free-view fixation data is indicated along the x-axis.....	217
Figure 6-25 Histograms of fixation distances from the center of each image.....	219
Figure 6-26 F/M ratios for random fixations restricted to $\frac{1}{4}$ image size distance from center, and $\frac{1}{16}$ image size distance from center.....	220
Figure 6-27 Nine grid locations used to compute the expected location map.....	221
Figure 6-28 F/M ratios for different locations in the C-Map, found by turning on a single grid element and turning off all other elements.....	223

## List of Tables

Table 3-1	Newell's temporal hierarchy of brain organization.....	58
Table 4-1	Time in seconds and number of fixations (in parenthesis) per task for the eight subjects who performed the experiment.....	94
Table 4-2	Order of tasks for Group 1 and Group 2.....	94
Table 4-3	Task abbreviations.....	96
Table 4-4	Pairwise comparisons for significant differences in fixation durations between tasks. An X indicates that a statistically significant difference exists between the corresponding tasks in the row and column.....	98
Table 4-5	Hallway Walking (WH).....	107
Table 4-6	Conversation (TC).....	107
Table 4-7	Telephone Talking (TT).....	107
Table 4-8	Sorting Cards (SC).....	107
Table 4-9	Sorting Blocks (SB).....	107
Table 4-10	Reading Poster (RP).....	107
Table 4-11	Reading Form (RF).....	107
Table 4-12	Counting Change (CC).....	107
Table 4-13	Counting Blocks (CB).....	107

Table 4-14	Subject A task differences.....	109
Table 4-15	Subject F task differences.....	109
Table 4-16	Subject C task differences.....	109
Table 4-17	Subject H task differences.....	109
Table 4-18	Subject E task differences.....	109
Table 4-19	Subject G task differences.....	109
Table 4-20	Subject B task differences.....	109
Table 4-21	Subject D task differences.....	109
Table 4-22	Pairwise comparisons for significant differences in saccade amplitude between tasks. An X indicates that a statistically significant difference exists between the corresponding tasks in the row and column.....	114
Table 4-23	Hallway walking (WH).....	119
Table 4-24	Conversation (TC).....	119
Table 4-25	Telephone talking (TT).....	120
Table 4-26	Sorting cards (SC).....	120
Table 4-27	Sorting blocks (SB).....	120
Table 4-28	Reading poster (RP).....	120
Table 4-29	Reading form (RF).....	120
Table 4-30	Counting change (CC).....	120
Table 4-31	Counting blocks (CB).....	120
Table 4-32	Summary of results from study of natural tasks.....	122
Table 5-1	Statistical comparison of completion times for the subjects of Group 1 and Group 2. In each case the null hypothesis is rejected ( $h=0$ ), indicating that there is no statistically significant difference between the ordering of the tasks in terms of completion times.....	139

Table 5-2	Order of instructions for Group A and Group B during extended environment study.....	147
Table 6-1	Maximum F/M ratios and associated weights for three example images using the random weight generation method. 10,000 trials.....	200
Table 6-2	Maximum F/M ratios and associated weights for three example images using the genetic algorithm method. 2,400 trials. # Gens refers to the actual number of trials required before a solution converged. Images A30, A32, A76, B30, and B88 are not included in the range data because the weights were greater than $\pm 50$ , due to many mutations.....	203
Table 6-3	Instructions for multi-view part of the experiment.....	207
Table 6-4	Three most frequently fixated objects and percentage of time spent looking at those objects for each of the tasks in the extended environment study from Section 5.3, over all subjects.....	212
Table 7-1	Classification of tasks into feature vector corresponding to both the level of visual engagement with the environment and amount of strategic planning required.....	230