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

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 Technolgy. 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

List of Figures

List of Tables