COUNTERPROPAGATION NEURAL NETWORK DETECTION OF VISUAL PRIMITIVES

BY

CYNTHIA LYNN JOHNSON
B.S.E.E., University of Miami, 1986

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Computer Engineering
in the Graduate Studies Program
of the College of Engineering
University of Central Florida
Orlando, Florida

Summer Term
1990

ABSTRACT

Psychological testing has shown that there is an early preattentive stage in the human visual system. At this level, simple features and properties of objects known as visual primitives are detected spatially in parallel by groupings of cells in the visual cortex known as feature maps. In order to study this preattentive stage in a machine vision system, the biologically inspired, highly parallel architecture of the artificial neural network shows great promise. This paper describes how the unique architecture of the counterpropagation neural network was used to simulate the feature maps which detect visual primitives in the human visual system. The results of the research showed that artificial neural networks are able to reproduce the function of the feature maps with accuracy. The counterpropagation network was able to reproduce the feature maps as theorized, however, future research might investigate the abilities of other neural network algorithms in this area. Development of a method for combining the results of feature maps in a simulation of full scale early vision is also a topic for future research that would benefit from the results reported here.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1
PREATTENTIVE VISION

The concept of a machine that can identify and recognize objects as quickly and easily as humans can is an ambitious goal, but not an impossible one. The human visual system is one of the most complex neural assemblies known. Although it is not yet completely understood, psychologists and physiologists have obtained a great deal of insight into the structure and function of human vision. In all likelihood, the theories and principles of the biological vision system will provide the best starting point in the development of a machine visual system capable of equal performance. In this research, an attempt was made to understand a portion of the biological vision system and reproduce its function on a machine in an effort to develop a more thorough understanding of complex visual mechanisms.

From biological studies, it is clear that an image is registered by cells within the eye and transmitted to the brain by the optical nerve. In a computer based system, a camera can digitize a scene and convert it into a collection of signals which are transmitted to the computer. The next step is analyzing and accurately recognizing the component objects within the scene. The human recognition system seems to have a striking dichotomy at this point. Many discriminations appear to be made automatically without attention and spatially in parallel, while other discriminations require focused attention or scrutiny. These two types of processing were originally attributed to different levels by Neisser (Neisser 1967). He identified an early or preattentive stage where simple features were registered determining texture segmentation and figure ground groupings. This preattentive stage is separate from a second, attentive stage where focused attention recognizes specific objects within a complex grouping of objects. The understanding of and recreation of a portion of the early visual stage was the primary goal of the research presented here.

There are two types of psychological evidence that support the concept of preattentive processing. One is textural segmentation and the other is visual search. Textural segmentation is the division of an image into segments based on the texture of its component parts. Figure 1 shows an image which is easily partitioned into two segments. The zeros constitute one segment and the other is comprised of ones. Julesz (Julesz 1981) has proposed that texture segmentation is preattentively processed using simple features called textons. When effortless texture segregation occurs, it is because the two segments do not contain the same type of texton. He defines the texton classes as color, elongated blobs of specific widths, orientation and aspect ratios, and the terminators of these blobs.

1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

Figure 1. An Example of Texture Segmentation.

A second source of evidence that supports the early vision theory is visual search. When subjects are asked to identify a target object in displays containing a varying number of distracters, the target appears to "pop-out" of the image when the target is defined by a simple visual feature (Treisman 1985). Figure 2 illustrates this "pop-out" phenomenon. The circular zero among the distracters of ones appears to "pop-out" at the observer. The speed of target detection in these cases suggests spatially parallel processing at the stage prior to attentive vision. This conclusion supports the theory of early vision. Treisman and Gelade have developed an explanation known as feature integration theory (Treisman and Gelade 1980). This theory states that primitive elements are directly sensed by specialized populations of detectors called feature maps. Each feature map responds to a particular feature, and all maps operate in parallel. It is only when attention is