To the Graduate Council:

I am submitting herewith a thesis written by Veronica S. Minsky entitled "A Multilayer Cooperative/Competitive Method For Grouping and Organizing Related Image Segments." I have examined the final copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Science, with a major in Computer Science.

Moonis Ali

Moonis Ali, Major Professor

We have read this thesis
and recommend its acceptance:

Bruce A. Whitehall

Accepted for the Council:

Carolyn Minkel

Vice Provost
and Dean of the Graduate School

STATEMENT OF PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a Master's degree at The University of Tennessee, Knoxville, I agree that the Library shall make it available to borrowers under the rules of the Library. Brief Quotations from this thesis are allowable without special permission, provided that accurate acknowledgement of the source is made.

Permission for extensive quotation from or reproduction of this thesis may be granted by my major professor, or in his absence, by the Head of Interlibrary Services when, in the opinion of either, the proposed use of the material is for scholarly purposes. Any copying or use of the material in this thesis for financial gain shall not be allowed without my written permission.

Signature Veronica S. Minsky

Date July 23, 1990

A MULTILAYER COOPERATIVE/COMPETITIVE METHOD FOR GROUPING AND ORGANIZING RELATED IMAGE SEGMENTS

A Thesis

Presented for the

Master of Science

Degree

The University of Tennessee, Knoxville

Veronica S. Minsky

August 1990

DEDICATION

This thesis is dedicated to my husband Mark. His encouragement and support made this possible.

ACKNOWLEDGMENTS

I wish to express my sincere appreciation to Dr. Moonis Ali, the chairman of my committee, for his assistance and suggestions. I also wish to thank Dr. Bruce Whitehead and Dr. Seung-Chul Lee for reviewing this thesis.

Special thanks are also due to Dr. Alianna Maren, who supervised and guided this research.

I am also grateful to the supervisor of my graduate research assistantship, Mr. Jim Nichols, of Sverdrup Inc., for his interest, enthusiasm and knowledge.

ABSTRACT

Most high-level machine vision systems, perform poorly when confronted with real world images. Occlusions, specularities and actual physical differences between segments cause objects or related information to be partitioned into multiple segments.

This thesis presents a methodology for joining together related regions and creating an organized representation of the image. This methodology uses perceptual features and a cooperative/competitive process to group regions and produce perceptually-based hierarchical structures.

The hierarchical structures overcome many difficulties associated with segmentation. In this thesis hierarchical structures are created for three images. In all three examples, regions separated by occluding objects or faulty segmentation were grouped together.

The organized representation of an image created by this methodology will be a valuable tool in future vision systems and is one step towards more powerful vision capabilities.

TABLE OF CONTENTS

CHAPTER		PAGE
I.	**INTRODUCTION**	1
II.	**HIGH LEVEL MACHINE VISION SYSTEMS NEED ORGANIZED REPRESENTATIONS**	3
	The Need For Robust Grouping of Related Regions	3
	Perceptual Information Provides a Rich Source of Knowledge	4
III.	**PERCEPTUAL INFORMATION ORGANIZES IMAGES** . .	8
	Perceptual Features Influence the Grouping of Related Regions	8
	Similarities	9
	Smooth Continuation of Boundary Lines . .	10
	Common Fate	11
	Proximity	12
	Closure and Enclosure	12
	Symmetry	13
	Early Visual Processing Responds to Simple Perceptual Features	13
	The Importance of Feature Contrast in Perceptual Grouping	15
	Perceptual Features Cooperate and Compete To Influence the Grouping of Related Regions	16

CHAPTER PAGE

IV. AN EXAMPLE PROBLEM 19

V. A MULTI-LAYER COOPERATIVE/COMPETITIVE METHOD FOR CREATING HIERARCHICAL STRUCTURES 24

Introduction 24

The Grouping Process 24

Layer One - The Storage of Input Information 25

Layer Two - The Computation of Relationship Values 26

Layer Three - The Interaction Between the Relationships 26

Layer Four - The Summation of Relationship Values 28

Layer Five - The Propagation of Values Exceeding a Threshold 29

The Combination of Grouped Input Regions . 29

Comparison of This Method with Models of Visual Processing 31

VI. IMPLEMENTATION 33

Introduction 33

A Description of the Program Modules. . . . 33

Preprocessing the Image 35

Initializing the Data Structures 37

CHAPTER PAGE

Performing the Grouping 38
Computing relationship values 38
Computing relationship interactions . . 40
Computing thresholds and
propagating values. 41
Combining Grouped Regions 42
Terminating the Grouping Process. 43
VII. **APPLICATION TO SPECIFIC EXAMPLES** 46
Introduction 46
Example One: A Tank in the Woods 47
Summary of Processing and Results 47
Detailed Description of the Example . . . 51
Preprocessing 52
Goals 54
Summary of sixteen processing
iterations. 56
Detailed Analysis of Results 60
Analysis of relationship values 60
The structures formed 60
The tank structure 61
The gun barrel and extreme tank
top structures. 71
The background and tree structures . 73
Regions not grouped into structures . 77

CHAPTER PAGE

The goals evaluated 77

Object Recognition Using Hierarchical Structures 78

Template matching 82

Conclusions 90

Example Two: Wrenches and Pliers 91

Detailed Description of the Processing. . 92

Preprocessing 92

Goals 97

Summary of six processing iterations. . 98

Detailed Analysis of Results 99

Analysis of relationship values 99

The structures formed 99

The wrench structure. 99

The pliers structure. 102

The bolt structures and background structures 102

Evaluation of Goals and Conclusions . . . 107

Example Three: Wrenches and Pliers Modified 109

Detailed Description of the Processing. . 110

Preprocessing 110

Summary of three processing iterations. 110

CHAPTER PAGE

Detailed Analysis of Results 113
The structures formed 113
The wrench structures 113
The plier structures 113
The bolt structure, shadow structures and background structures 120
Conclusions 124
VIII. CONCLUSIONS AND RECOMMENDATIONS 126
IX. FUTURE RESEARCH 131
BIBLIOGRAPHY 134
VITA 140

LIST OF TABLES

TABLE		PAGE
7.1.	Tank Template	79
7.2.	Example One Hierarchical Structures with Feature and Relationship Values	83
7.3.	Example One Features and Relationship Values of the Regions Forming Structure 120	85

LIST OF FIGURES

FIGURES		PAGE
3.1.	Different Leveled Parallel Bars Demonstrate Boundary Line Continuation	10
4.1.	Synthetic Image Containing Parallel Bars . . .	19
4.2.	The Hierarchical Structure of the Parallel Bars in Figure 4.1.	21
5.1.	The Five-Layered Architecture Which Produces Hierarchical Structures	25
5.2.	Detailed View of the Connections Between Layer One and Layer Two	27
5.3.	Detailed View of Layer Five.	30
5.4.	A Linked List of Input Regions (a) Before the Grouping Process and (b) After the Grouping Process	30
6.1.	An Overview of the Processing Sequence . . .	34
6.2.	Diagram of Eight Orientations	37
7.1.	Image of a Russian Tank (Soviet Military Power, 1988)	47
7.2.	Image of a Russian Tank with Labeled Components	48
7.3.	Tank Components Formed By the Grouping Program	48
7.4.	Structures Forming the Tank	50
7.5.	Background and Tree Structures	50
7.6.	Image of a Russian Tank After Preprocessing. .	53
7.7.	Seventy-eight Segmented Regions of the Tank Image	53
7.8.	Large Region Manually Separated Into Seven Smaller Regions, 20, 70, 71, 72, 73, 74 and 75	55

FIGURES PAGE

7.9. Two Large Regions Separated Into Three Smaller Regions Each, 1, 2, and 76, and 7, 77, 78 55

7.10. Hierarchical Structure of the Tank 62

7.11. Labeled Regions Which Comprise the Tank Structure 62

7.12. Hierarchical Structure of the Tank with Region Descriptions 63

7.13. Hierarchical Structure of the Tread Cover . . 65

7.14. Hierarchical Structure of the Tread Cover with Region Descriptions 66

7.15. Labeled Regions Which Comprise the Tread Cover 67

7.16. Hierarchical Structure of the Upper Tank Area 68

7.17. Hierarchical Structure of the Upper Tank Area with Region Descriptions 69

7.18. Labeled Regions Which Comprise the Upper Tank Area. 70

7.19. Hierarchical Structure of the Tank Tread . . . 70

7.20. Hierarchical Structure of the Dark Middle Area 70

7.21. Labeled Regions Which Comprise the Tank Tread and Dark Middle Components 71

7.22. Hierarchical Structure of the Gun Barrel . . . 72

7.23. Hierarchical Structure of the Extreme Tank Top 72

7.24. Labeled Regions Which Comprise the Gun Barrel 72

7.25. Labeled Regions Which Comprise the Extreme Tank Top Component 74

7.26. Hierarchical Structure of Background 75

FIGURES PAGE

7.27. Labeled Regions Which Comprise the Background 75

7.28. Hierarchical Structure of the Tree Area. . . . 76

7.29. Labeled Regions Which Comprise the Tree Area 76

7.30. Image of Wrenches and Pliers Source: Fu, K.S., Gonzalez, R.C., Lee, C.S., Robotics Control, Sensing, Vision, and Intelligence, (New York: McGraw-Hill;1987) . . 91

7.31. Image of Wrenches and Pliers After Preprocessing 93

7.32. Fifty-Three Segmented Regions of the Wrenches and Pliers Image 93

7.33. Image of Wrenches and Pliers with Labeled Tools 94

7.34. Whole Tools Segmented Into Single Regions. . . 94

7.35. Manually Divided Regions 95

7.36. Background Regions 96

7.37. Hierarchical Structure of Wrenches 100

7.38. Hierarchical Structure of Wrenches with Description 100

7.39. Labeled Regions of Wrench Structure 101

7.40. Hierarchical Structure of the Right Lower Pliers 103

7.41. Labeled Regions Comprising the Right Lower Pliers Structure, 57, and the Bolt Structure, 63. 103

7.42. Hierarchical Structure of the Bolts 104

7.43. Hierarchical Structure of Inside Background Area 104

7.44. Labeled Regions of Inside Background Area . . 105

FIGURES PAGE

7.45. Hierarchical Structure of Shadow Area Surrounding the Drill Key 105

7.46. Labeled Regions of Shadow Structure Surrounding the Drill Key 106

7.47. Hierarchical Structure of Background Near the Horizontal Wrench 106

7.48. Labeled Regions of Background Structure Near the Horizontal Wrench 107

7.49. Image of Wrenches and Pliers with Four Added Overlapping Wrenches 109

7.50. Regions Separated By Overlapping Wrenches - I. 111

7.51. Regions Separated By Overlapping Wrenches - II. 111

7.52. Regions Separated By Overlapping Wrenches - III 112

7.53. Hierarchical Structure of the Horizontal Wrench 114

7.54. Labeled Regions of the Horizontal Wrench Structure 114

7.55. Hierarchical Structure of the Lower Diagonal Wrench 115

7.56. Hierarchical Structure of the Right Diagonal Wrench 115

7.57. Hierarchical Structure of the Left Diagonal Wrench 115

7.58. Labeled Regions of Three Wrenches 116

7.59. Hierarchical Structure of Left Pliers Upper Handle and Nose 117

7.60. Labeled Regions of Left Pliers, Upper Handle and Nose Structure - 72, Bolts Structure- 87, and Right Lower Pliers Structure - 80 117

7.61. Hierarchical Structure of Lower Handle of Left Pliers 118

FIGURES PAGE

7.62. Labeled Regions of Lower Handle of Left Pliers Structure - 74, Shadow Along Left Diagonal Wrench Structure - 77, and Shadow Along Right Diagonal Wrench Structure - 84 118

7.63. Hierarchical Structure of Right Lower Pliers . 119

7.64. Hierarchical Structure of the Bolts 120

7.65. Hierarchical Structure of the Shadow Along the Right Diagonal Wrench 121

7.66. Hierarchical Structure of the Shadow Along the Left Diagonal Wrench 121

7.67. Hierarchical Structure of the Background Near the Horizontal Wrench 122

7.68. Hierarchical Structure of the Corner Background 122

7.69. Hierarchical Structure of Inside Background . 122

7.70. Labeled Regions of Background Structures . . . 123

7.71. Labeled Regions of Inside Background Structure 124

CHAPTER I

INTRODUCTION

Organizing related image segments improves the ability of machine vision systems to recognize objects and interpret scenes. Most high-level machine vision systems, perform poorly when confronted with real world images, rather than controlled or synthetic images. One reason for this is poor segmentation resulting from occlusions, specularities and actual physical differences between segments. This causes objects or related information to be partitioned into multiple disjoint segments. High level processes often lack any proven method for grouping these related segments.

There is a need to create an organized representation of the segmented scene which groups these related segments before high level processing begins. Several researchers in machine vision have pointed out that such an organization needs to be "perceptual" or "scene organizing" in nature [Barrow & Tenenbaum, 1981; Palmer, 1983; Witkin & Tenenbaum, 1983; Tsotsos, 1984]. Perceptual information, used by humans to organize images, is a powerful tool for grouping and organizing visual data. As of yet, few existing machine vision systems explicitly incorporate a perceptually-based scene organizing representation, [See, e.g. Lowe, 1985 & 1987; Lawton & McClellan, 1987; Maren & Ali, 1988].

This thesis explores the problem of scene organization and presents a methodology which creates an organized, hierarchical, perceptually-based representation of a segmented scene. This method is used to create hierarchical structures, useful for object recognition, for three real world images. Specifically this thesis will:

- o review the need for an organized representation of segmented images in machine vision systems,
- o discuss briefly research and theories of perceptual organization,
- o describe in detail the methodology for creating the scene organization,
- o present three examples of the hierarchical structures created for segmented images, and
- o discuss how these structures assist object recognition and scene analysis tasks.

CHAPTER II

HIGH LEVEL MACHINE VISION SYSTEMS NEED ORGANIZED REPRESENTATIONS

The Need For Robust Grouping of Related Regions

Effective object recognition and scene analysis tasks are often hindered by low level representations which partition objects or related information into disjoint regions. These regions must be joined into coherent meaningful groups before high level processing occurs. This thesis will specifically consider the need to organize segmented output. Other types of low level representations such as Marr's 2/3 sketch also require organization, but are not as extensively used in current vision systems and therefore will not be considered in this thesis.

Segmentation difficulties prevent perceptual grouping during the segmentation process. No improvement in segmentation algorithms alone will overcome two ever-present problems:

- o differences in illumination, shading, surface reflectivity, and object coloration may cause an object to be segmented into disjoint regions, and
- o occluding objects may partially obscure an object, and cause it to be segmented into disjoint regions.

It is not possible to address these issues strictly by refining segmentation algorithms. Instead, the regions which result from segmentation must be combined into perceptually meaningful groups.

Perceptual Information Provides a Rich Source of Knowledge

A realization of the need for perceptual organization in computer vision has grown out of work in model-based object identification. Early approaches to model-based vision (e.g. Brooks et al, 1979; Brooks, 1981, 1983), led to the development of systems with powerful matching capabilities of projected appearances of objects with their actual appearance. These systems were limited in their operations as objects had to be fully specified in advance. Further, such systems were particularly sensitive to difficulties caused by occlusions, shadows, and faulty segmentation.

The deficiencies of such model-based vision systems may be linked to the difficulty of their research problem, which searches among many arguments for object matches. Since they do not incorporate any form of perceptual organization, search time is lengthy and the "intelligence" which can be used to guide the search is limited.

Recently, various investigators have begun to explore methods for perceptual grouping to aid model-based vision. Lowe [1985, 1987] uses an intermediate processing level

which identifies certain perceptual relationships among the edges extracted during segmentation. Lowe points out that:

> "...an important function of this early stage of visual grouping is to distinguish as accurately as possible between....accidental and significant structures.

This extra "perceptual" information allows Lowe's SCERPO system to locate complex objects even when partially obscured, or when presented at a variety of angles and orientations. However, SCERPO only extracts a limited number of line-based perceptual features, and performs only a limited amount of edge-grouping based on these features or relationships. As a result, it is not readily extensible to complex (e.g., natural) environments.

Lawton and McConnel [1987] similarly explore the use of perceptual information to aid object recognition. Their LISP-based system is slow, and not readily extensible to real-time scenarios. Also, their approach does not create a new perceptually-based representation of the scene. This limits the ability of their approach to detect novel objects or process images with significant clutter.

An alternative to model-based vision systems are biological-based vision systems which use perceptual information. The perceptual information is derived by identifying primitives in the image. This approach, originally proposed by Marr [1982], is described as data-driven or bottom-up and

is a sharp contrast to systems using models of specific objects which are knowledge-driven or top-down.

Data driven systems are generally more flexible and able to process previously unknown objects and information. They are also more resistant to noise than are systems which use specific object models. Marr [1982], Pentland [1986], Rearick & Frawley [1987], and Havens and Mackworth [1983] have incorporated data-driven perceptual information into their systems. These researchers, however, use limited perceptual information and process all perceptual information equally.

Grossberg [1987 (a),(b)&(c)] uses perceptual information in his Boundary Contour System by organizing images based on real and illusory contours and filling in features across the boundary contours. But no further grouping of these boundaries is done, nor is any ranking of the most significant areas of an image performed.

Rosenfeld [1988] presents a low level grouping technique based on biological vision. His system segments an image using either similarities, good continuation, proximity, or enclosure. Each of these perceptual features may segment an image differently, and his approach does not attempt to combine the features and produce a single segmented image.

After assessing the use of perceptual organization to improve machine vision capabilities, it appears that no system has as yet attempted to incorporate robust, generic perceptual grouping leading to a new representation of the perceptual organization of the image.

A long-range goal for machine vision systems is the ability to "perceive" the environment in a manner analogous to human perception. If this goal could be accomplished, then humans could interact with an intelligent controller for the machine vision system to discuss mutually perceived images in a natural manner. This need increases the importance of perceptual grouping capabilities in machine vision systems.

CHAPTER III

PERCEPTUAL INFORMATION ORGANIZES IMAGES

Perceptual Features Influence the Grouping of Related Regions

The first researchers in the area of perceptual psychophysics identified several different inter-region relationships, referred to as perceptual features, which contribute to perceptual organization [Koffka, 1935; Kohler, 1947; Werthiemer,1958]. The early researchers did not distinguish among the relative importance of these features, nor did they succinctly determine the effects of their interactions. Current research focuses on identifying the relative significance and interactions of these features.

The features or relationships considered generally important are [Werthiemer, 1958; Tsotsos, 1984]:

- similarity,
- smooth continuation of boundary lines,
- common fate,
- proximity,
- closure, and
- symmetry.

Each of these features is discussed in terms of the recent research clarifying their nature and their significance in forming perceptual groups.

Similarities

Similarities in orientation, intensity, color, size, and shape are recognized as important features in the perceptual process. Olson and Attneave [1970], and Beck, Prazdny & Rosenfeld [1983] found similarity of orientation to be a significant factor in determining perceptual grouping in a number of experimental studies. Treisman and Patterson [1984] found similarities in shape, color and size to be significant. Julesz [1975] studied textures and found similarity in illuminance and density to be significant features when grouping textures (the combined effect of texture, illuminance, and density corresponds roughly to the intensity of a segmented region.)

Of all the different similarity relationships, researchers have consistently identified similarities in orientation and intensity as the most important. These two features may be particularly important in perceptual grouping because they reflect early processes in the human visual system. Specifically, they describe the functions of orientationally selective receptive fields which respond to a given orientation and location for patterns of contrasting brightness, such as bars and lines [Hubel, 1980]. Grossberg [1987 (a),(b)&(c)] found similarity of intensity and orientation so significant that his Boundary Contour System starts with values indicating orientation and direction of

intensity contrast. Much biological as well as psychophysical evidence points to the importance of orientation and intensity in perceptual grouping.

Smooth Continuation of Boundary Lines

Boundary line continuation strongly influences perceptual grouping. The regions in Figure 3.1 have similar intensities, orientations, and are proximal to each other. The only distinguishing factor between these regions is that if continued, the horizontal boundary lines of regions A and B would coincide, and the horizontal boundary lines of regions C and D would coincide. This continuation of boundary lines causes the perceptual grouping of A with B, and C with D.

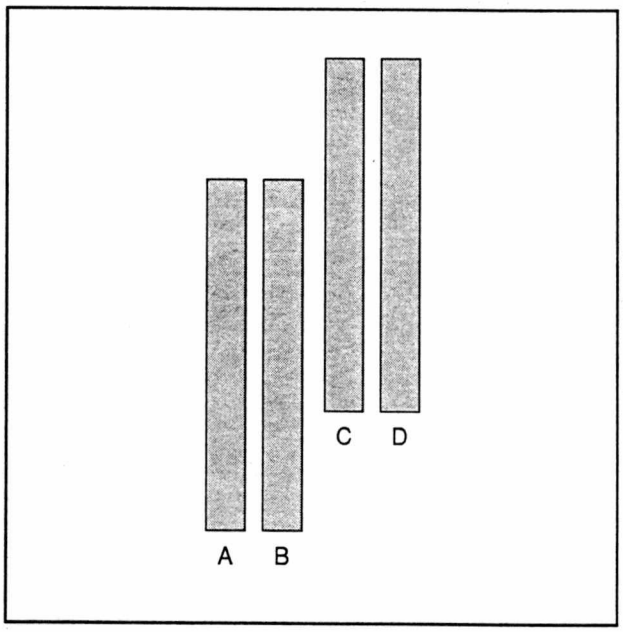


Figure 3.1. Different Leveled Parallel Bars Demonstrate Boundary Line Continuation.

Collinearity and curvilinearity are often key elements in the perception of boundary line continuation. The importance of collinearity is demonstrated in studies by Beck, Prazdny, and Rosenfeld [1983], where collinear lines produce stronger groupings than other configurations of similarly sloped lines. Grossberg [1987 (a),(b) &(c)] uses boundary contours, determined by similarities of brightness and orientation, to create a Boundary Contour System for low-level image processing. In his system, boundary line continuation, both observed and illusory, is the major factor in the segmentation process.

Common Fate

The vertex or cotermination of two lines, referred to as common fate, is the most perceptually important part of a line. Biederman [1985] showed that in line drawings with portions of the lines deleted, subjects were able to "fill-in" the missing areas in the middle of a line. Subjects were not as able to "fill-in" the missing areas in the vertex of a line. The vertex provides unique perceptual information. Studies by Walters [1985] also demonstrate the importance of vertices in determining perceptual significance and figure interpretation. The presence of line vertices needs to be considered as a factor in forming perceptual groupings.

Proximity

The proximity of regions influences perceptual grouping. Rarely does a grouping occur based upon proximity alone, usually it occurs based upon proximity as well as other features. The most persuasive evidence of this is presented in studies by Beck [1972] and Beck, Prazdny and Rosenfeld [1983] of texture grouping. In their studies observers are increasingly unable to segment texture regions as the texture elements are spaced further apart. Pomerantz [1986] found similar results by varying the proximity of sets of parentheses. These studies demonstrated the supportive role of proximity in perceptual grouping. Work by von Grunau et al. [1984] confirms that most perceptual groups are proximal.

Closure and Enclosure

Closure occurs when a line drawing or object has a continuous boundary without endpoints. The object or line drawing with closure may be perceived as a boundary between an enclosed area and the non-enclosed area. The boundary encourages the perceptual grouping of the boundary and the enclosed area.

The occurrence of a boundary does not actually need to be present to cause the perception of closure. Such examples of illusory closures have been studied by Treisman [1986]. Treisman found that the existence a closed shape such as a circle in an image, would cause subjects to report viewing

triangles when only disjoint line segments were present. Likewise, the lack of a closed shape such as a circle in the image, would cause subjects to less frequently report viewing triangles. Treisman suggests that in preattentive processing subjects sense the closure feature and incorrectly place it in the image. These studies provide evidence that closure is an important perceptual feature.

Symmetry

Humans not only place a value judgment of "appreciation" on highly symmetric figures, they also will reorganize a figure so as to represent it as a whole with maximal possible symmetry, [Garner, 1975; Vernon, 1952]. The symmetry between components of many objects yields relationships which are often invariant, or nearly invariant, when viewed over many different angles. This "transformational invariance," described by Palmer [1983] may play important roles not only in perceptual grouping, but in how humans store visual information. Hierarchical structures enable the consideration of high-level features such as symmetry.

Early Visual Processing Responds to Simple Perceptual Features

Many psychologists, physiologists and computer scientists accept a visual model consisting of two or more stages. The first stage extracts perceptual features, such

as those discussed in the previous section. The latter stages perform higher level functions leading up to object recognition.

Treisman [1986] defines two stages in visual processing. The first stage, preattentive processing, extracts features quickly, in a seemingly parallel manner. The second stage, focused attention processing, extracts features slowly, in a seemingly serial manner.

Tresiman studied the features processed at each of these stages by measuring the reaction times of subjects identifying objects. Specifically, Treisman noticed that during focused attention processing, the subject's reaction time increased as more distracter objects were added to an image. The increase in reaction time seems to indicate serial processing.

In contrast, Treisman noticed that preattentive processing is characterized by a constant reaction time regardless of the number of distracter objects present. The constant reaction time seems to indicate parallel processing. Treisman believes the features processed preattentively occur during the earliest stages of vision. Some of these features are, color, size, tilt, curvature and line ends.

Treisman found these features to be processed preattentively as long as they uniquely identify objects. When an object could only be identified by a combination of

these features, then focused attention processing is required. These studies indicate that simple, contrasting perceptual features are processed in earliest stages of vision.

The Importance of Feature Contrast in Perceptual Grouping

The main objective in perceptual grouping is to distinguish a group of regions from all other regions. The grouping process requires maximal relatedness between group members accompanied by maximal distinction between group members and all other regions. If, for example, all the regions in an image have low intensity values, then the "similarity of intensity" relationship feature will not distinguish any group of regions from the others, and will not be useful in perceptual grouping. This is illustrated in Figure 3.1 where the feature "similarity of orientation" is not distinctive and therefore does not influence the grouping process.

Studies by Beck, Prazdny and Rosenfeld (1983) support the hypothesis that perceptual grouping is based on the response of the visual system to stimulus differences as well as similarities. Beck et al. studied the ability of subjects to identify perceptual groups in an image of characters in which the characters were all identical except for their orientation. Beck et al. found the orientation differences which created the most distinctions between characters

were the easiest for subjects to identify. Pomerantz [1986] recognized contrast as a driving factor in perceptual grouping. He concluded that a feature must provide a necessary level of distinction in order to be useful in perceptual grouping.

Perceptual Features Cooperate and Compete To Influence the Grouping of Related Regions

While many important aspects of perceptual grouping have been identified, the relative importance of these features and how they interact still remain as major questions. Recently, researchers have come to recognize the diversity of processes in perceptual grouping and have found that not all perceptual features have the same importance in every image [Zucker, 1987; Pomerantz, 1986]. Prinzmetal's [1981] research indicates the perceptual features in an image will combine differently in different images. Studies in mental imagery by Kosslyn [1988] show that we recall images by their perceptually salient parts. Perceptually salient parts are the most distinctive parts of the image. Rosenfeld [1988] uses this concept in his approach to recognizing unexpected objects. His system broadcasts only the most "conspicuous" image parts to reduce the amount of data broadcast.

An additional factor influencing perceptual grouping is the distribution of regions which are to be considered. Zucker [1987] has discussed this problem, setting it in the broader context of the complexity of perceptual organization. He suggests that different grouping methods need to be invoked when the regions to be considered are either sparsely or densely distributed.

Either multiple grouping methods may be invoked, as suggested by Zucker, or a single grouping process itself is able to adapt to the vast diversity found in visual images. This thesis presents an adaptive grouping method which can flexibly group regions based on the most appropriate perceptual features.

This method discovers those regions with the strongest perceptual features relative to others in that image. Because the strength of perceptual features is relative, features which are strong in one image may be weak in another. Perceptually strong features may be defined as those features which create the greatest distinction and discrimination between a group of regions and the rest of the regions in the image. Measuring and determining the strength of perceptual features involves the following components:

- feature relatedness, this refers to existence of relationships such as similarity, boundary line continuation, enclosure, and proximity,

- o feature contrast, this refers to the uniqueness in a particular image of the relationships mentioned above, and
- o feature interactions, this refers to the influence one relationship (such as enclosure) may exert over another relationship (such as boundary line continuation).

In summary, regions whose features are maximally related to each other and maximally contrasting with other regions in the image will emerge as strong perceptual groups. In joining these regions, the grouping process is identifying the dominant features in an image and the most compelling aspects of the image.

CHAPTER IV

AN EXAMPLE PROBLEM

This Chapter illustrates how perceptual relationships may be used to group regions in an image and how these groups would assist machine vision systems. For demonstration purposes, this Chapter presents a simple synthetic image, shown in Figure 4.1, more complex real world images are presented in Chapter VII. This image contains 16 regions which could represent a portion of a control panel or the insertion slots for boards or tools. A robotic vision system may guide a robotic arm to specific slots by recognizing the pattern of controls or slots.

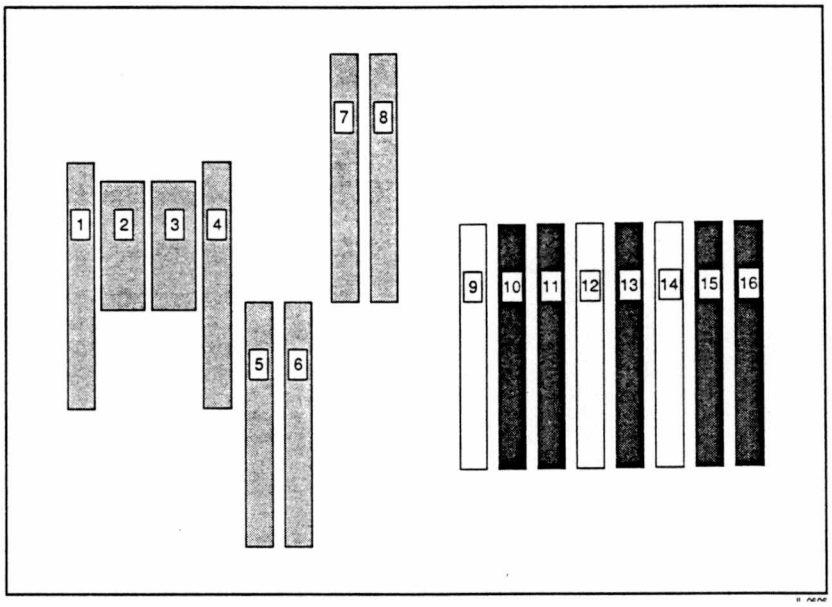


Figure 4.1. Synthetic Image Containing Parallel Bars.

In order to recognize the pattern of controls, a vision system needs an organized representation of the controls. A perceptual based organization offer advantages over appearance based organization. Appearance based models often perform poorly due to discrepancies between the actual appearances and the projected appearances. These problems would be aggravated by a mobile robotic system viewing the control panel from different positions.

A perceptual organization uses perceptual features to group regions. These groups often overcome segmentation problems by joining related regions which are separated by occluding objects or which have differences in shading or illumination. A perceptual organization is also less sensitive to changes in viewing positions.

A vision system could identify the sixteen controls or slots in this diagram by matching the perceptual groups and features in this image with those in a model. A grouping process created an organized representation of this image using the perceptual relationships similarity of intensity, similarity of orientation, continuation of boundary lines, enclosure and proximity. The hierarchical structure for this image is shown in Figure 4.2.

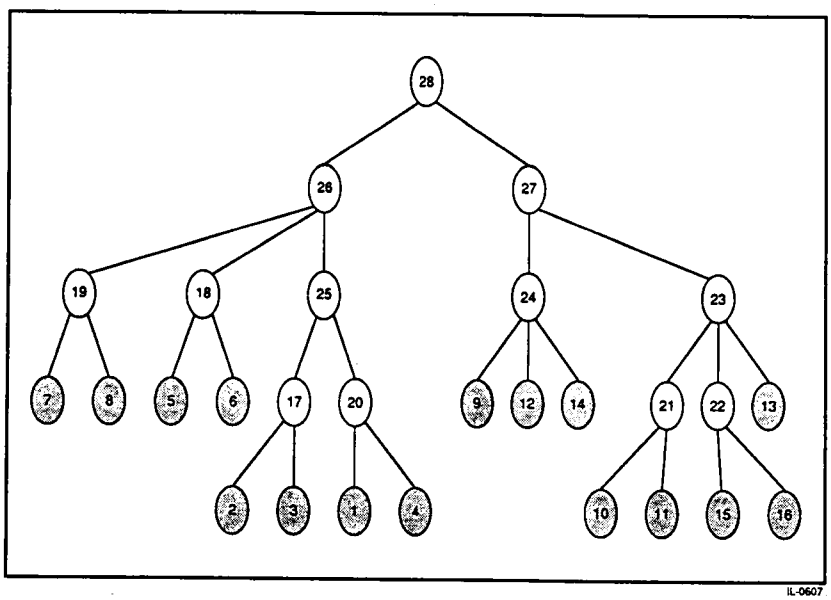


Figure 4.2. The Hierarchal Structure of the Parallel Bars in Figure 4.1.

The hierarchical structure shows two major perceptual groups consisting of regions 1-8 and regions 9-16. The features "similarity of intensities" and "continuation of boundary lines" influenced these groups.

Within each of these major groups, it is possible to identify further subdivisions. Within regions 1-8, regions 7 and 8, regions 5 and 6 and regions 1-4 are grouped due to continuation of boundary lines and proximity. In the case of regions 1-4, enclosure also influences the group. Within regions 1-4 there is a further division of groups as regions 1 and 4, and regions 2 and 3 group separately due to continuation of boundary lines.

Regions 9-16 may also be broken down into smaller groups based on the similarity of intensities. Regions 9,12 and 14 are similar in intensity as are regions 10, 11, 13, 15 and 16. A further division of groups occurs due to proximity as regions 10 and 11 and regions 15 and 16 group separately.

A vision system would search the hierarchical structure shown in Figure 4.2 and match the groups and their perceptual features with those in a model. The following example describes a possible search for control or slot number 5.

A perceptual model for these controls may indicate that control 5 is the left control in a group of 2 controls, both with light intensities. This group of two controls is contained in a larger group. This larger group contains three groups. The group containing control 5 is the middle group and is positioned lower than the other groups. This larger group joins one other group to form the entire control panel shown in this image. The other group contains eight controls of dark and light intensities. The perceptual features intensity, proximity and relative location identify these groups.

The matching process would identify groups 26 and 27, shown in Figure 4.2 as the two large groups described in the model. Group 26 contains the three groups described in the model and group 18 is the middle group, it has a light

intensity and is positioned lower than the other groups. The desired control is located in group 18.

Using a perceptual organization of the image, a vision system is able to match features and components and locate individual controls. If there were noise, occluding objects, differences in illumination, or variations in the viewing position, the perceptual grouping program would still group related regions and enable the vision system to locate the individual controls. This aspect of the perceptual grouping program is illustrated in the examples described in Chapter VII. A perceptual organization joins regions based upon perceptual relatedness to overcome segmentation problems and create an organized representation of the image. The following Chapter describes a methodology which creates perceptual organizations for real world images.

CHAPTER V

A MULTI-LAYER COOPERATIVE/COMPETITIVE METHOD FOR CREATING HIERARCHICAL STRUCTURES

Introduction

This chapter presents a five layer architecture which uses perceptual relationships to create hierarchical structures. The architecture uses a cooperative/competitive grouping process and considers multiple perceptual relationships in logical concurrence. The grouping process computes relationship values for pairs of regions, identifies the most strongly related regions, and joins the strongly related regions into a single region.

The Grouping Process

The grouping process occurs in five layers. These layers, shown in Figure 5.1, perform the following functions:

- o Layer One - The storage of input information,
- o Layer Two - The computation of relationship values,
- o Layer Three - The interaction of the relationship values,
- o Layer Four - The summation of the relationship values,
- o Layer Five - The propagation of values exceeding a threshold.

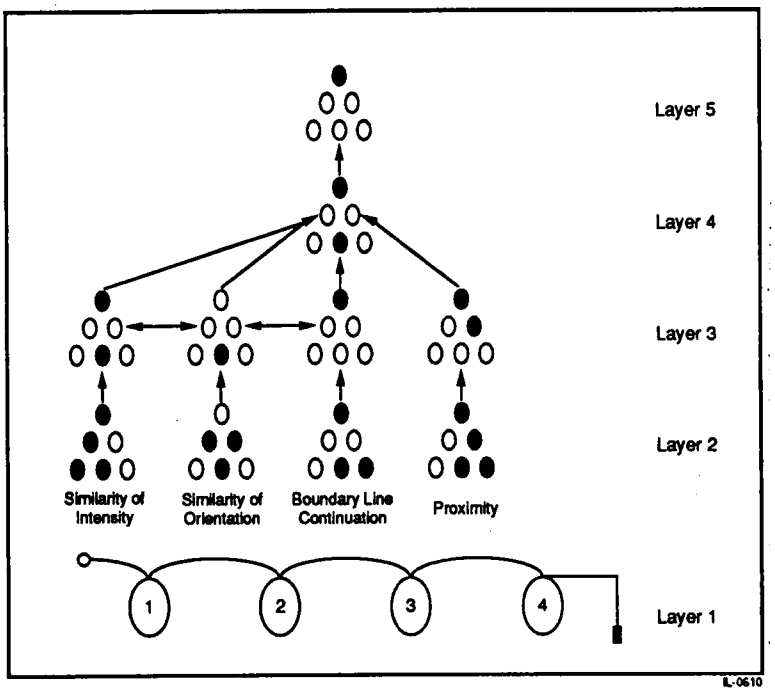


Figure 5.1. The Five-Layered Architecture Which Produces Hierarchical Structures.

Layer One - The Storage of Input Information

A preprocessing program segments images into regions and extracts feature information about these regions. The first layer lists these segmented input regions and the extracted feature information. Feature information such as intensity value, size, location and orientation will be stored in layer one and used to compute the perceptual relationships. During the grouping process, the regions will be grouped together and the feature information will be recomputed to represent the averaged values of the grouped regions.

Layer Two - The Computation of Relationship Values

Relationship values are computed for each of the following relationships:

- o Similarity of intensity,
- o Similarity of orientation,
- o Continuation of boundary lines, and
- o Proximity.

The grouping process first computes relationship values for each pair of input regions. Chapter VI describes the algorithms used to compute the relationship values. A large relationship value indicates similar intensities, similar orientations, continuation of boundary lines, or proximity. A small relationship value indicates dissimilar intensities, dissimilar orientations, non-continuation of boundary lines, or lack of proximity.

The architecture computes relationship values for each possible pair of input regions and stores them in layer two. As shown in Figure 5.2, in an image with four regions, relationship values would be computed for the pairs, (1-2), (1-3), (1-4), (2-3), (2-4), (3-4).

Layer Three - The Interactions Between the Relationships

In layer three all the relationships (except the proximity relationship) interact. Each relationship influences other relationships to group the same regions that it

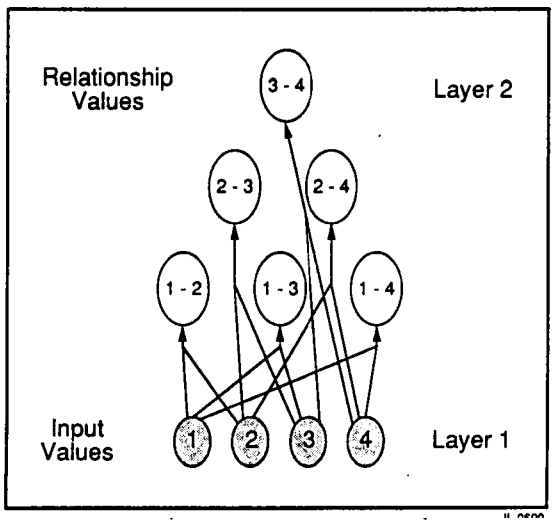


Figure 5.2. Detailed View of the Connections Between Layer One and Layer Two.

has grouped. Relationship values exert influence by increasing and decreasing other relationship values. The amount of influence a relationship exerts is proportionate to the size of the relationship value, the larger the relationship value the more influence it will exert.

Each relationship sends inhibitory and excitatory signals to increase or decrease other relationship values. The signal value is computed as the sending relationship value multiplied by a excitation/inhibition factor. Chapter VI discusses selecting an excitation/inhibition factor. An inhibitory signal sends a negative value which decreases the target relationship value. An excitatory signal sends a positive value which increases the target

relationship value. Positive relationship values send excitatory signals to other relationships of the same pair of input regions. Negative relationship values send inhibitory signals to other relationships of the same pair of input regions. Zero relationship values do not send signals and do not influence grouping.

Layer Four - The Summation of Relationship Values

Layer four combines the separate relationship values into a single value. The three relationships - similarity of orientation, similarity or intensity, and continuation of boundary lines are summed into a single value. Because of its supportive role, the proximity relationship value does not interact with the other relationships in layer three and is not included in the summation. The proximity relationship influences the other relationship after they are summed.

The proximity relationship will send an excitatory signal to all positive summed values in layer four. The size of the signal depends on the proximity of the pair of input regions. If two input regions are close together in the image, the signal for that pair will be large and if two regions are far apart, the signal for that pair will be small or zero. Only positive summed values receive signals from the proximity relationship which reflects the supportive role of proximity.

Layer Five - The Propagation of Values Exceeding a Threshold

Layer five computes thresholds and propagates the values exceeding these thresholds. The final values represent the pairs of regions to be grouped. A threshold is computed and each pair of input regions whose value exceeds the threshold is propagated. This process is repeated three times as is shown in Figure 5.3. The pairs of input regions whose values survive the third threshold in layer five will be grouped together and combined.

The Combination of Grouped Input Regions

Combining the grouped regions involves creating a new input region, computing new feature values and removing the grouped regions from the input list. The new input region is added to the linked list of input regions. Feature information is computed including a new average intensity value, a new orientation, new edges and new location information. The grouped regions are then removed from the list of input regions and added as children of the new input region.

The grouped regions may be original input regions or parent regions representing regions grouped in previous iterations. When a grouped input region is removed from the input list and added as a child of the new input region, all the branches below it are also moved. Figure 5.4 (a) and (b) illustrate the result of grouping three regions under a

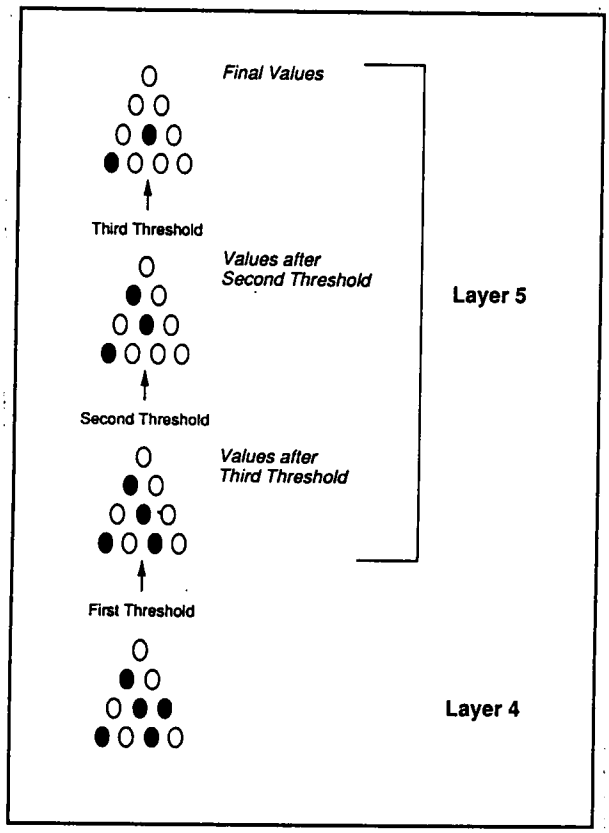


Figure 5.3. Detailed View of Layer Five.

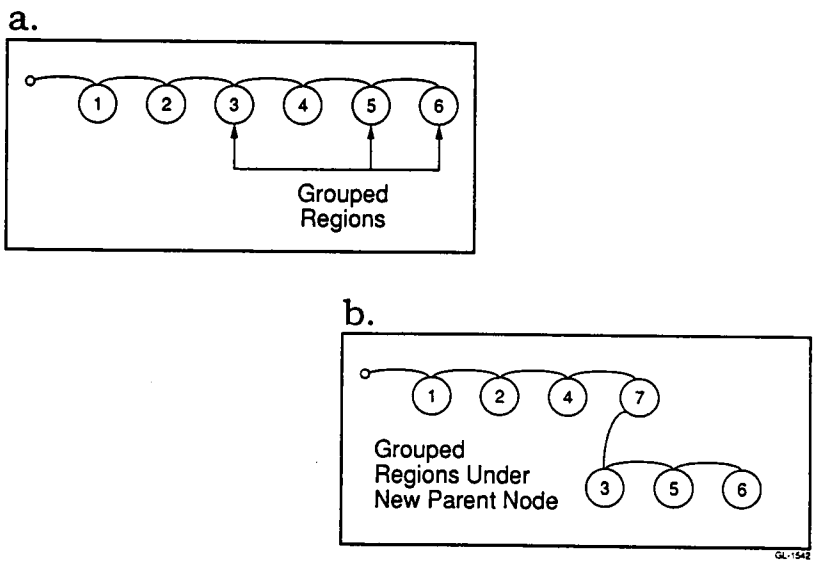


Figure 5.4. A Linked List of Input Regions (a) Before the Grouping Process and (b) After the Grouping Process.

new parent region. Through this process, the structural information of the grouping process is retained.

Using the new list of input regions, the iterative grouping process repeats itself. The process of grouping and combining continues until the threshold values computed in layer five fall below a computed cut-off value. Chapter VI discusses the cut-off value and terminating the grouping process. The hierarchical structures are complete when the grouping process terminates.

Comparison of This Method With Models of Visual Processing

The multi-layer cooperative/competitive method described in this chapter shares a number of characteristics with the model of visual processing presented by Treisman [1986]. Treisman's model consists of several levels of processing. The first level, representing early vision, extracts simple image features such as color, orientation, size and stereo distance. The next level of processing, referred to as the focused attention level, looks for discontinuity of the features. Higher levels of processing compare features corresponding to a selected location with features in object descriptions.

Like the Treisman model, this method extracts simple features, such as orientation, intensity, boundary lines and proximity. The cooperative/competitive method identifies

discontinuity or contrast in the image. Regions are grouped which are proximal, have similar features and are distinct from non-grouped regions.

Because of the proximity relationship typically only regions which are physically proximal are grouped. This characteristic corresponds to the level in Treisman's model which associates features with specific locations in the image. The features of the grouped regions may be compared with features in object templates corresponding to Treisman's object identification level.

CHAPTER VI

IMPLEMENTATION

Introduction

This chapter describes the processing sequence, data structures, programming modules and algorithms used to implement the hierarchical grouping architecture described in Chapter V. The programming modules are presented in four groups performing one of the following functions: preprocessing the original image, initializing the data structures, grouping the most related regions, and combining the grouped regions. A final section in this chapter discusses selecting an algorithm to terminate the grouping process.

A Description of the Program Modules

The hierarchical grouping programs are all written in C and designed to be highly portable. The data files may be processed as either linked lists or disk files depending on the size of the image and the capacity of the hardware. The hierarchical grouping system is comprised of 56 modular programs performing the four major functions shown in Figure 6.1 and described below.

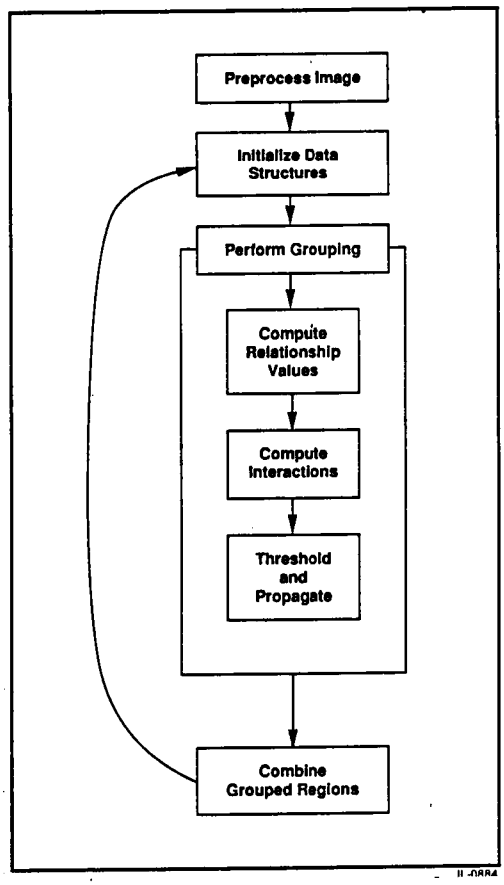


Figure 6.1. An Overview of the Processing Sequence.

Preprocessing the Image

Preprocessing involves extracting information from the original image and preparing it for use in the grouping program. Preprocessing routines written specifically for this implementation work in conjunction with routines from the VISILOG package by NOESIS to digitize, filter, reduce and segment the original image.

The VISILOG software running on a SUN microsystem digitized the images using a grab routine and a video camera. A 256x256 frame with 256 gray scale value was frozen. Three passes of a median filter routine were used to reduce the noise in the image. The median filter routine uses a 3x3 kernel and computes the gray level distribution within an elementary neighborhood, the median value of the distribution is used to filter the noise.

After the median filter, a VISILOG sampling routine is used to reduce the image. Every x pixel in the image is sampled to reduce the image to an 86x86 image. The image was reduced to expedite processing in the grouping program. Previous examples processed by the grouping program were small synthetic images of 20X20, (Minsky & Maren, 1989), processing actual digitized images at 86x86 is a significant challenge over the small synthetic images.

A preprocessing program written for this implementation uses the output from the VISILOG software to segment the image into distinct regions. The program performs

segmentation using the region growing algorithm described by Gonzalez and Wintz [1977]. After segmenting the image, the preprocessing program extracts feature information for the regions. The average intensity, the size and the region edges are computed.

The final preprocessing tasks involve creating an edge file and a continuation of boundary line file. These files will be used to reduce computation time during the grouping process and eliminate the redundant storage of edge information. The edge file contains a list of the edges bounding the segmented regions. A program computes the length, slope and orientation of each edge. The edges are categorized as having one of eight orientations shown in Figure 6.2. Eight was selected because studies of the human visual system indicate that we detect approximately eight orientations, [Hubel, 1980].

The edge file is used to create a continuation of boundary line file. This file identifies pairs of regions which are continuous. Storing this information in a file saves processing time when computing the continuation of boundary line relationship. A pair of edges is continuous if the edges are less than 10 pixels apart, have the same orientation, and the line connecting the closest endpoints has the same orientation as the edges.

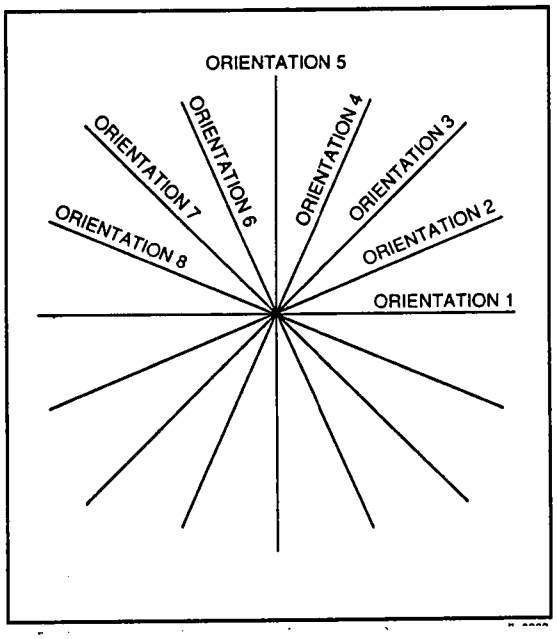


Figure 6.2. Diagram of Eight Orientations.

Initializing the Data Structures

After preprocessing, the input regions are loaded into a linked list and a network node is created for each pair of input regions. The segmented regions created during preprocessing and their extracted feature information are entered into the input region linked list. This list represents layer one in the architecture. As the regions are loaded the program computes the orientation of the region as the orientation of the longest edge.

After the input regions are entered in the input list, a network structure is created which contains one node for

each pair of input regions. The network structure nodes contain all the computed relationship values and the computed values for layers 2-5 of the architecture.

Performing the Grouping

The grouping process involves computing the relationship values, computing the relationship interactions, and computing threshold values. The program HNNMAIN directs the grouping process and calls the programs RELATIONSHIP_PROCESS, THRESHOLD and LEVEL_EXCITE to perform these functions.

Computing relationship values. The program RELATIONSHIP_PROCESS calls routines to compute the four relationship values. The routines are: COMPUTE_INTENSITY, COMPUTE_ORIENTATION, COMPUTE_SMOOTH, and COMPUTE_NEARNESS. The relationship values are stored in the network data structure and represent the second layer of the architecture. Each of the routines is discussed below.

COMPUTE_INTENSITY compares the intensity values of two input regions and returns a similarity of intensity value. If the absolute difference between the intensity values of the two regions is less then 10, then the maximum value, 1.0, is returned as the similarity of intensity value. For absolute differences greater than 10, the inverse of the difference divided by 10 is returned as the similarity of intensity value. Therefore a difference of 20 will return a

similarity of intensity value of 0.5. The relationship value decreases as the difference between the intensity values increases.

The COMPUTE_ORIENTATION routine compares the orientation values of two input regions and returns a similarity of orientation value. The orientation values of the input regions may range from one to eight, with each number representing 22.5 degrees as shown in Figure 6.2. The orientation value of a region matches the orientation value of the longest edge in the region. If the orientations of the two input regions are the same or up to 22.5 degrees apart 1.0 is returned as the similarity of orientation value. If the orientations are up to 45 degrees apart, 0.5 is returned and if they are up to 67.5 degrees apart, a 0.33 similarity of orientation values is returned. For all other differences a zero relationship value is returned. The relationship values represent the inverse of the difference between the orientation numbers, 1-8.

COMPUTE_SMOOTH uses the continuation of boundary line file to compare the edges of the two input regions and compute a continuation of boundary lines value. The program looks up a value in the continuation of boundary lines file for each pair of edges combining one edge from each input region. When a pair of edges is found which are continuous, and the orientation of the continuous edges matches the orientation of the input regions, a 1.0 continuation of

boundary lines value is returned. When the orientation of the continuous edges is not the same as the orientation of the region, the importance of the continuous edge decreases. Therefore, if the orientation of the continuous edges matches only one of the input regions, 0.65 is returned and if the orientation of the continuous edges matches neither input region, 0.3 is returned. If no continuous edges are found a 0 relationship value is returned. These values were selected to reflect equal increments in value.

COMPUTE_NEARNESS determines the proximity value as the inverse of the smallest distance between two regions. The proximity value decreases as the distance between the regions increases. The grid positions in a 20x20 grid are used to compute the distance between two region.

Computing relationship interactions. Once the relationship values have been computed, RELATIONSHIP_PROCESS weighs the relationship values, and then computes the interactions between the relationships. The interaction for each of the three relationships, similarity of orientation, similarity of intensity, and continuation of boundary lines may be computed using an excitation/inhibition factor. Each relationship value equals itself plus the excitation/inhibition factor multiplied by the two other relationships values computed for that pair of input regions.

A large excitation/inhibition factor enables a strong relationship value to have more influence in the grouping

process. A small excitation/inhibition factor allows small relationship values to more readily prevent the grouping of a pair of regions. After experimentation, the moderate value of 0.16 was selected as the excitation/inhibition value. In this program, each relationship value was set equal to itself plus 0.16 multiplied by the other two relationship values and is stored in the network data structure representing the third layer in the architecture.

Once the relationship interactions are computed, the new relationship values are summed and the total represents layer four in the architecture. This total combines the similarity of intensity, similarity of orientation and the continuation of boundary lines relationship values for each pair of input regions.

The proximity relationship influences grouping in the fourth layer. The proximity value is multiplied by the excitation/inhibition factor and added to the total value in the fourth layer.

Computing thresholds and propagating values. The grouping program thresholds and propagates values. The program THRESHOLD computes the threshold as one standard deviation above the mean value. The program LEVEL_EXCITE propagates all values exceeding this threshold. The propagated values are stored in the network data structure as the fifth layer. A threshold is computed and values are

propagated three times in the fifth layer; Figure 5.3 on page 30 shows a detailed illustration of the threshold sequence. Pairs of regions whose values survive the third threshold will be grouped together. If no values exceed the threshold then the threshold is reduced by 0.5 to avoid premature termination of the grouping process.

Combining Grouped Regions

The values remaining in the fifth layer represent the pairs of input regions to be grouped. A program HNNCOMB identifies which regions will be grouped together and then combines these regions. Combining these regions involves creating a new input region and computing the new feature values.

HNNCOMB searches the fifth layer for surviving pairs of input regions. When one input region exists in more than one pair of regions, that region plus all the regions it was paired with will be grouped together. For example, if the pairs A and C, and B and C survived in the fifth layer, then regions A, B, and C will form a single group.

HNNCOMB combines the grouped regions by creating a new input region and computing new feature values. The program computes an average intensity value weighted by the size of each region. It sums the region sizes and joins the list of grid positions from all the combined regions into a single linked list. HNNCOMB also joins the edges from the combined regions into a single linked list of edges in the new input

region. The longest edge in the list determines the orientation of the new input region. After computing the features in the new input region, HNNCOMB adds this new region to the linked list of input regions and removes the combined regions from the list. The combined regions form a linked list of children to the new input region.

Using this new input list, the program repeats the grouping process, starting with creating a network structure. A new structure must be created because the number of input regions has changed. The iterative process continues until a threshold in the fifth layer falls below the computed cut-off value.

Terminating the Grouping Process

The hierarchical grouping architecture is iterative, and as the number of iterations increases the grouping values decrease. As the grouping values decrease, the perceptual relatedness of the regions also decreases. When the grouping process continues until only two input regions remain grouping values become very low and the program joins regions whose features are very different. In addition, as the grouping process builds large regions comprised of smaller regions, the feature values of these conglomerate regions will tend towards middle values. It is therefore ideal to stop the grouping process before unrelated regions group and extremely large conglomerate regions form.

The termination point, or cut-off value, should reflect the average values in the particular image. For example, if an image has highly related regions, then the cut-off will be a high value. For another image with less related regions, the cut-off will be a lower value.

The first iteration of the grouping program produces the largest grouping values of any iteration. The fifth layer of the first iteration contains many of the pairs of regions which will group in later iterations. The fifth layer computes three threshold values shown in Figure 5.3 on page 30. The values which survive the second but not the third threshold reflect the next regions to be grouped, unless a newly formed region creates a stronger relationship with one of the pair. Because the first iteration contains the largest values in the entire network, the cut-off value is derived from this group of values. The third threshold is one standard deviation above the mean values that survive the second threshold. The cut-off value is one standard deviation below the mean of values surviving the second threshold in the first iteration. Therefore, the cut-off value for each image is computed during the first iteration.

During subsequent iterations of the grouping process the thresholds in layer five are checked. If the third threshold in any subsequent iteration falls below the cut-off value the grouping process will terminate.

The next chapter presents three images processed by the grouping program. The termination of the grouping process was studied in these examples. The cut-off value described above was determined to provide a reasonable termination point and is used in all these examples.

CHAPTER VII

APPLICATION TO SPECIFIC EXAMPLES

Introduction

Three images were successfully processed by the hierarchical grouping implementation described in Chapter VI. The images are diverse and require the grouping program to form large complex structures and small simple structures.

The first image, a tank in the woods, contains one large object comprised of many different featured regions. In order to create a hierarchical structure for this image, the grouping program joins regions forming the tank components. The components are then grouped together to form the tank structure.

The second and third images are of tools lying on a plain background. These images contain numerous small objects comprised of generally similar regions. In order to create hierarchical structures for these images, the grouping program joins regions to form several individual object structures.

This Chapter describes the grouping process and hierarchical structures for these images. A conceptual object recognition system is presented and applied to the hierarchical structures created for the first image of a tank in the woods.

Example One: A Tank in the Woods

The first image processed by the grouping program is the photograph of a tank in the woods shown in Figure 7.1. The following discussion will first present an overview of the processing and results followed by a more detailed description and analysis.

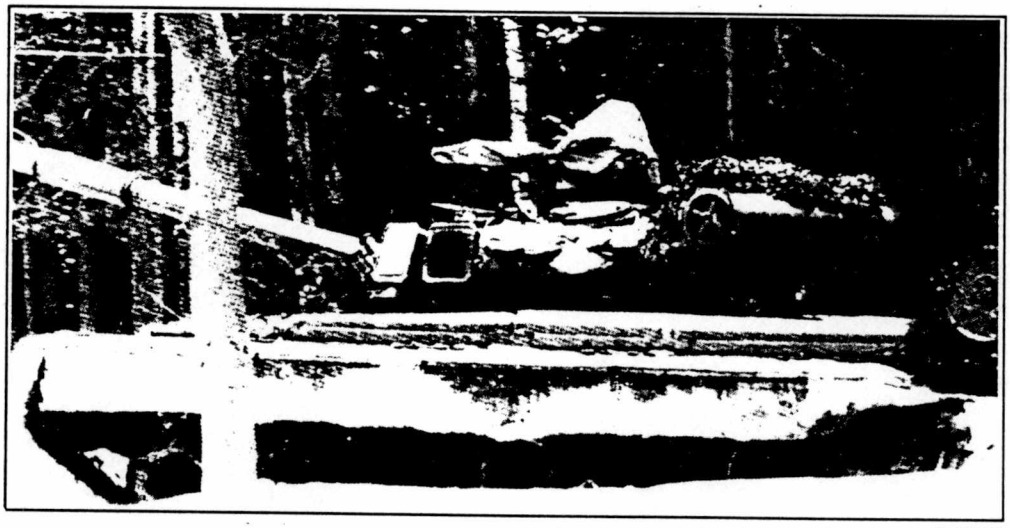


Figure 7.1. Image of a Russian Tank (Soviet Military Power, 1988.)

Summary of Processing and Results

The grouping program joined regions to form the tank components and then grouped many of these components into a single tank structure. Figure 7.2 identifies the tank components in the original photograph and Figure 7.3 shows

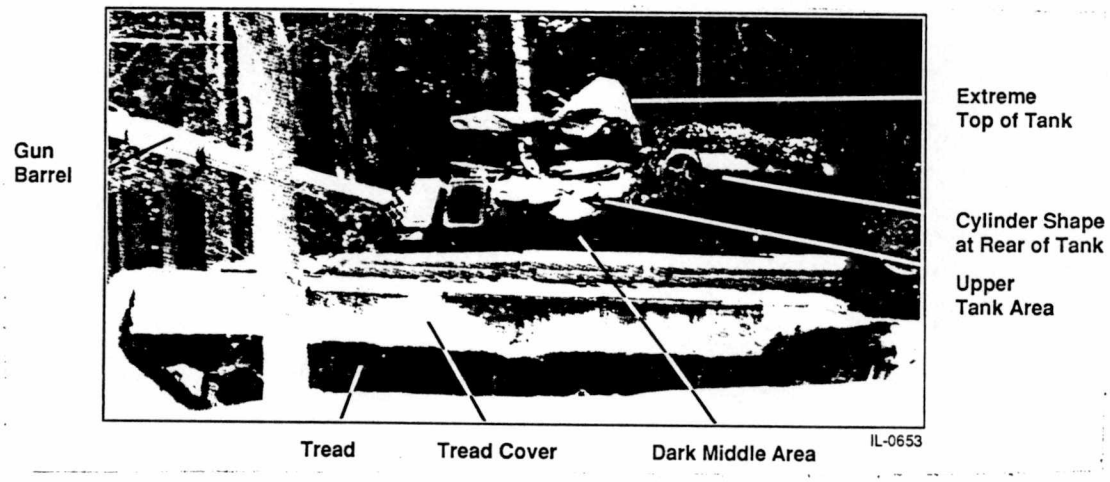


Figure 7.2. Image of a Russian Tank With Labeled Components.

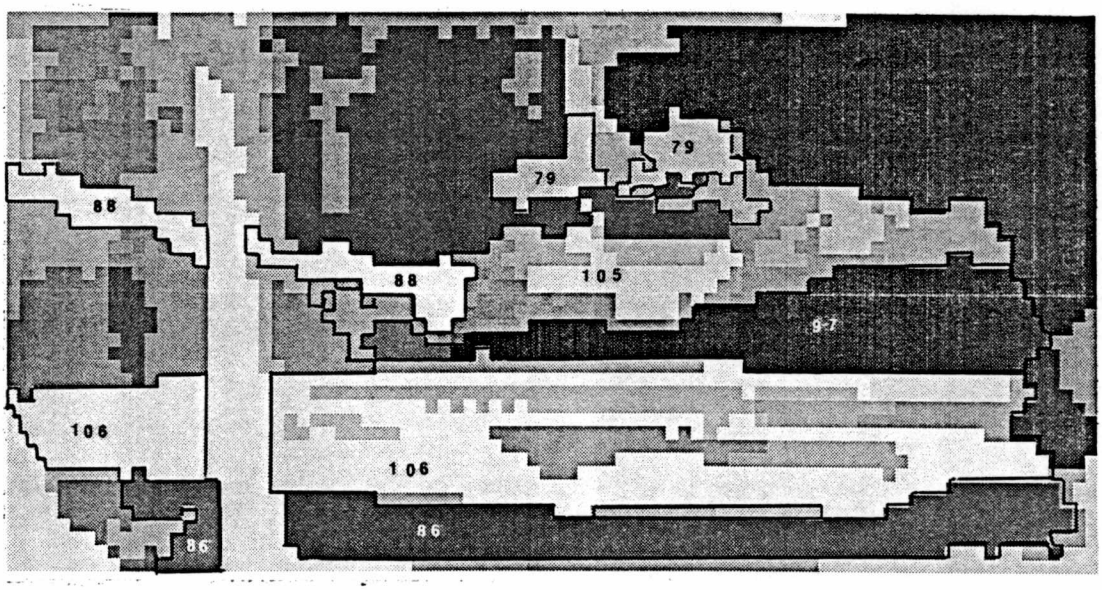


Figure 7.3. Tank Components Formed By the Grouping Program
79 - Extreme Top of Tank,
88 - Gun Barrel,
105 - Upper Tank Area,
97 - Dark Middle Area,
106 - Tread Cover, and
86 - Tread.

the tank components formed by the grouping process. The components are:

- o 79 - The extreme top of the tank,
- o 88 - The gun barrel,
- o 105 - The upper tank area,
- o 97 - The dark middle area,
- o 106 - The tread cover, and
- o 86 - The tank tread.

The grouping program joined four of these components into a single hierarchical structure. The upper tank area, the dark middle area, the tread cover and the tread join together to form structure 120. This structure and the two remaining tank structures are shown in Figure 7.4.

Because most of the tank components grouped together, an object recognition system could match structure 120 with a tank template structure. (See the discussion on object recognition in this chapter for more detail on structure matching.) The grouping program also joined some of the trees and background regions. These structures are shown in Figure 7.5, structure 114 represents the small trees in the background surrounding the large tree in the foreground. Structure 103 contains background regions above the tank. Almost all of the regions not grouped into structures were portions of the background.

The first iteration of the grouping program joins the most strongly related regions. In this example all the

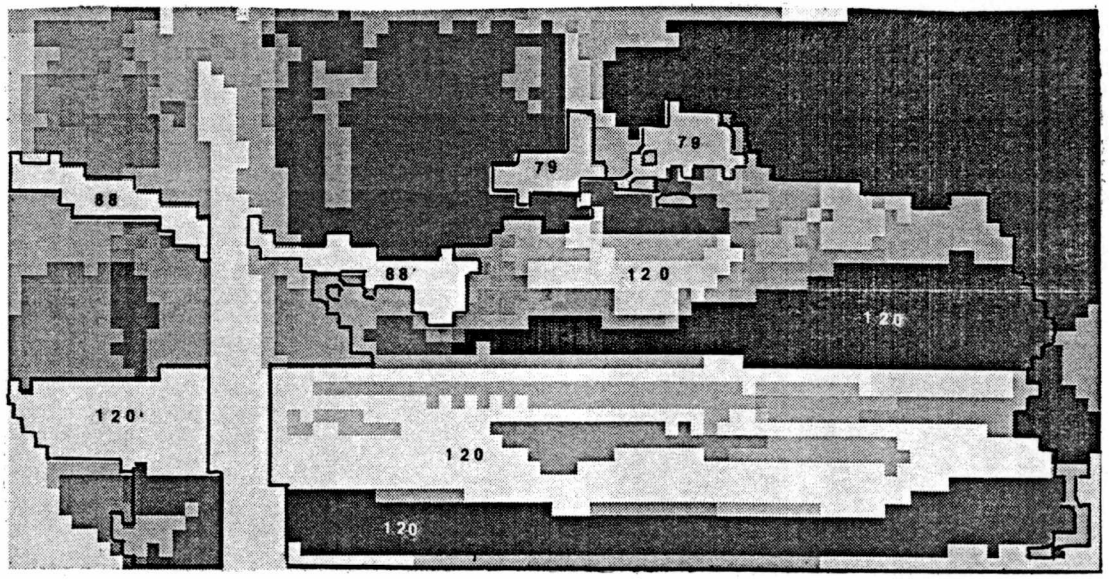


Figure 7.4. Structures Forming the Tank,
120 - The Tank Structure,
79 - The Extreme Top of Tank, and
88 - The Gun Barrel.

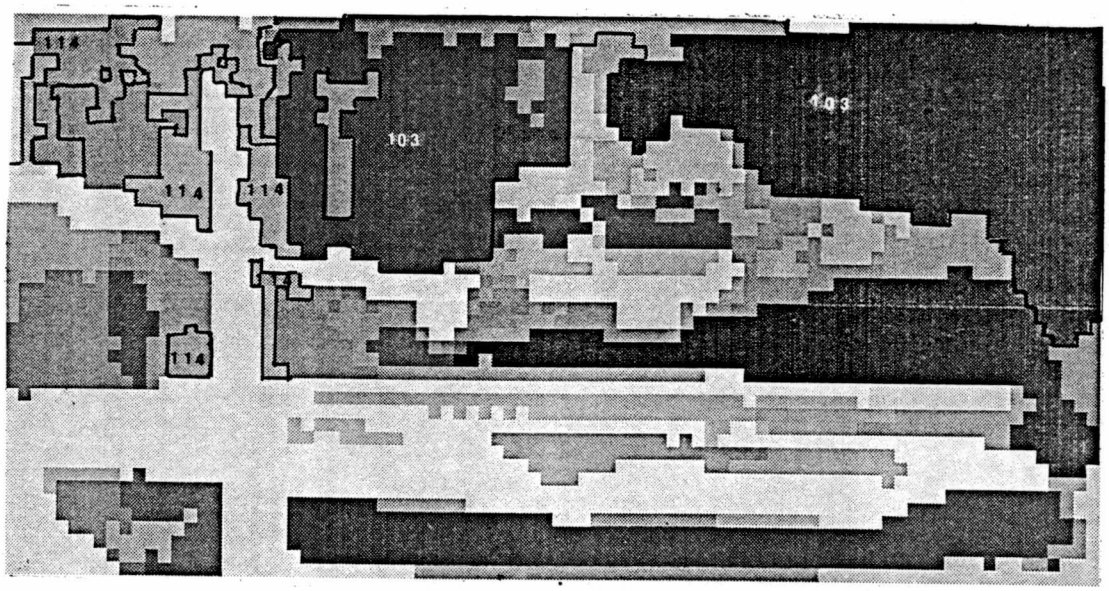


Figure 7.5. Background and Tree Structures,
103 - Background Above Tank, and
114 - Background Trees.

regions separated by the occluding tree were grouped in the first iteration, specifically, the gun barrel, the tank tread, and the tread cover.

Even though all relationship weights were set to 1.0, the relationships varied in their influence upon the grouping process. The similarity of orientation relationship strongly influenced grouping and almost all grouped regions had similar orientations. Neither the continuation of boundary line nor the similarity of intensities relationships influenced the grouping as strongly, because pairs of regions exhibited a greater range of values for these relationships. The proximity relationship proved important for setting a maximum distance of 8 pixels between grouped region pairs. Proximity appeared to have little influence on grouping when the regions were between 1-8 pixels apart.

Detailed Description of the Example

This detailed section describes the preprocessing, the goals for grouping, the processing and the resulting structures of the tank image. The goals for grouping are based on the regions segmented during preprocessing, and are therefore presented after a discussion of the preprocessing. This is followed by a summary of the 16 iterations of the grouping program and a description of the hierarchical structures produced. The goals are then evaluated and compared to the actual structures created.

Preprocessing. The tank image was digitized, filtered and segmented before the grouping process began. Figure 7.6 displays the digitized, filtered, reduced image and Figure 7.7 displays the segmented image. Chapter VI discusses the implementation and describes the general preprocessing procedures. Unique to this image are the parameters used in segmentation, the removal of a white border before segmentation and the manual division of three segments.

A large white border above and below the photograph of the tank was manually removed before segmentation. Removal of the border left a 42x86 image and allowed for faster processing.

The segmentation parameters determine the range of gray scale values allowed in a single region, and the minimum number of pixels required for a region. In this example, the parameters allowed a +/- 12 gray scale difference in regions and required at least 4 pixels to form a region.

The segmentation process created 69 regions. Three of the regions were unusually large and encompassed regions of varying orientations. The region growing algorithm used by the segmentation program may incorporate orientation to determine the segments (Gonzalez and Wintz, 1977.) Because orientation adds significant complexity to the segmentation

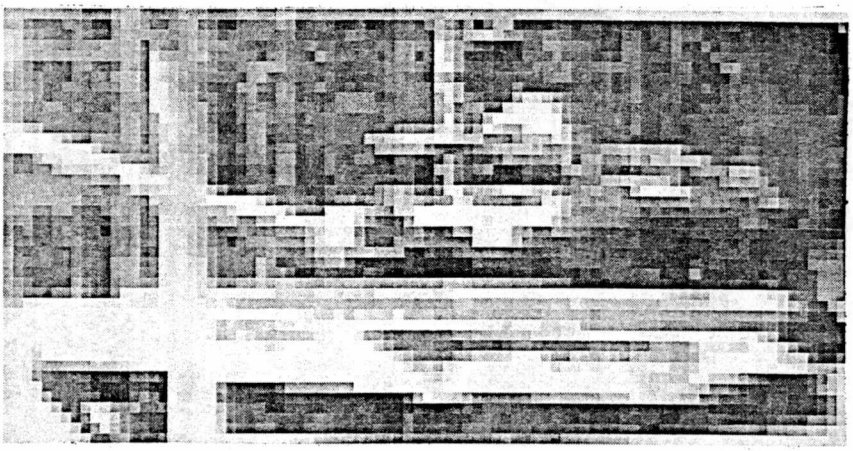


Figure 7.6. Image of a Russian Tank After Preprocessing.

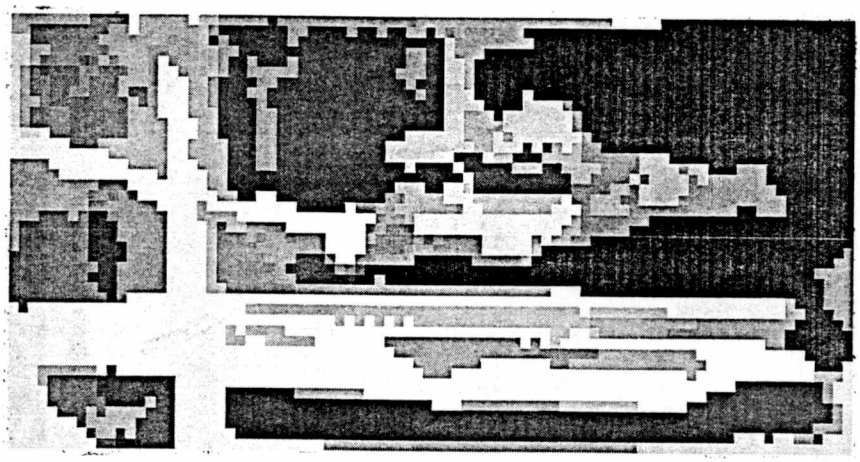


Figure 7.7. Seventy-eight Segmented Regions of the Tank Image.

process and as the focus of this thesis is on grouping rather than segmentation, orientation was not incorporated into the segmentation program. However, as orientation is often used with region growing segmentation, three regions were manually divided based upon differences in orientation. Figure 7.8 identifies the first region divided into the seven smaller regions, 20, 70, 71, 72, 73, 74, and 75. The dashed lines represent the manual division of the region. Figure 7.9 identifies the second region divided into the three smaller regions, 7, 77, and 78, and the third region divided into the three smaller regions, 1, 76, and 2.

Goals. Goals for the grouping process are based on the regions created by the segmentation process and how these regions may be grouped to best enhance object recognition. An object recognition system will match a template tank structure with structures created by the grouping program. The matching process will look for similar structure components. It will match the component features as well as the relationships which most strongly influence the grouping of the components. The discussion on object recognition at the end of this example provides a detailed description of the matching process.

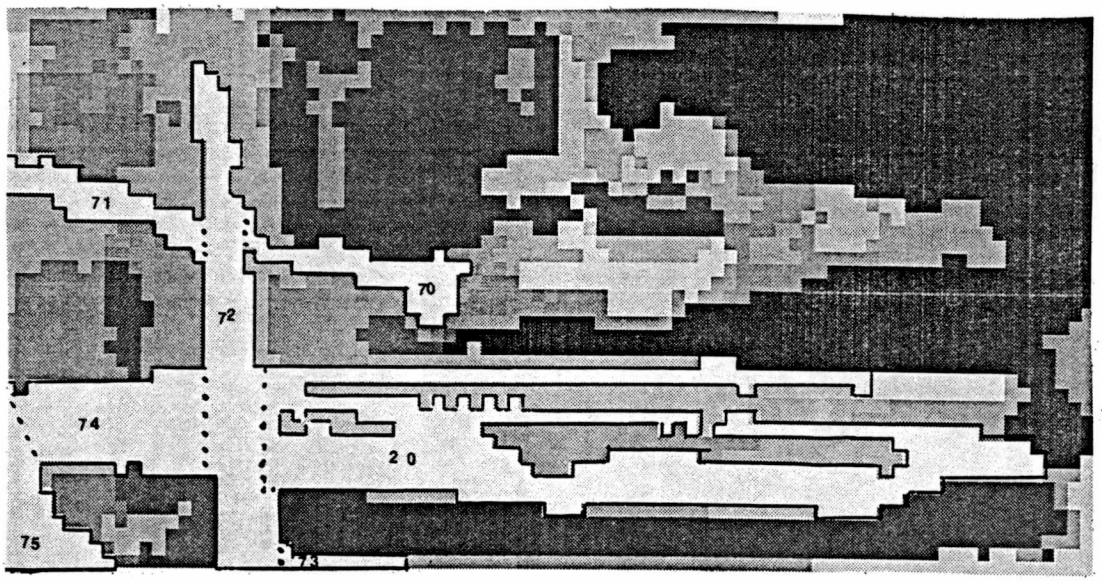


Figure 7.8. Large Region Manually Separated Into Seven Smaller Regions, 20, 70, 71, 72, 73, 74, and 75.

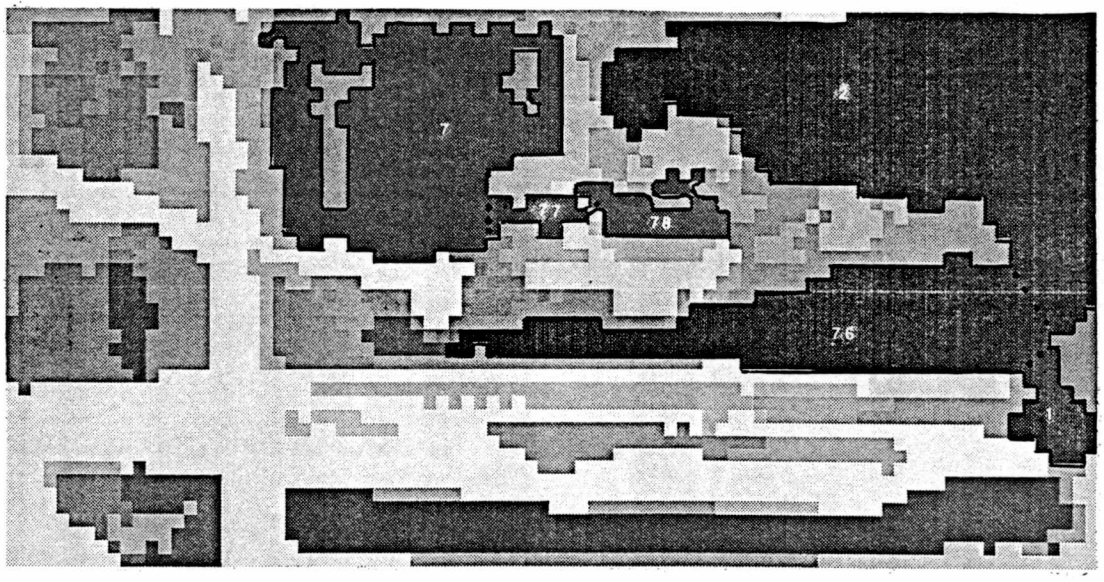


Figure 7.9. Two Large Regions Separated Into Three Smaller Regions Each, 1, 2, and 76, and 7, 77, and 78.

Because an object recognition system matches components, it is important for the grouping program to first join regions which form the components of the tank and then join these components into a larger structure. The program should also group other perceptually related groups, such as trees and background regions. Therefore, the following goals may be stated for the grouping program:

- o group regions to form tank components, specifically:
  (refer to Figure 7.2 for identification of these components)
  tread,
  tread cover,
  dark middle area,
  upper tank area,
  gun barrel,
  extreme top of tank,
- o group these tank components into a larger tank structure,
- o group the trees in the background, and
- o group the background regions.

Summary of sixteen processing iterations. At each iteration or pass, the grouping program generates relationship values and forms groups based on these values. These groups create a portion of the final hierarchical structures for the image. The tank structure was formed in 16

iterations, the following discussion summarizes the processing of these iterations.

The first iterations groups the most strongly related regions in the image. The grouping program joined regions to form three of the six tank components shown in Figure 7.3, the gun barrel (88), the extreme tank top (79), and the tread (86). All of these components, except the extreme tank top, consisted of regions separated by the occluding tree in the foreground.

The upper tank area and the tread cover are not formed in this pass because they are the most complex tank components. They contain the widest range in intensity values and boundary lines. The grouping program must first join the most related regions of these components and then join these groups to form the whole component. This will require more than a single iteration of the program.

The second pass of the grouping program joins regions which form portions of the most complex tank components, the upper tank area and the tread cover. All the medium intensity regions in the tread cover group and many of the medium intensity regions in the upper tank area group. The third pass groups regions with a greater range in intensity values, such as the dark middle component of the tank, bright portions of the upper tank area and the background above the tank. Pass 4 allows a greater variation in both

intensity values and boundary line continuation values. This results in the grouping of all the bright and medium intensity portions of the upper tank area and the tree areas to the right of the large tree and the medium intensity edges under the tread.

In the fifth pass the grouping program follows a pattern of grouping the largest portions of a component or structure and then grouping the small regions on the borders. These small regions are transition regions between one component and another. Pass 5 groups the larger regions representing the bright and medium bright portions of the tread cover and a smaller border region representing a thin strip between the tread cover and the dark middle components.

The grouping program completes the most complex tank components, the upper tank area and the tread cover in the sixth pass, and pass 7 begins grouping these components into a single tank structure. The upper tank area and the dark middle group together in the seventh pass. These two components are contiguous, share continuous boundaries and are closer in intensity than other possible pairs such as the tread cover and the dark middle or the tread cover and the tread. Other groups generated in the seventh pass are the tree regions in the upper left corner, a thin strip between

the tread and tread cover, and bright regions in the front and under the tread. Once again lower similarity of intensity values enabled these groups to form.

The eighth, ninth and tenth passes concentrate on building the tank structure and the tree structure. The pattern discussed earlier of first grouping large portions of a structure or component and then grouping the smaller transitional regions is seen as small regions between the dark middle and upper tank components group with the tank structure.

The tread cover and tread join with the all the tank components except the extreme top and the gun barrel to form the tank structure in passes 11 and 12, and then transition regions complete the tank structure in passes 13, 14, 15 and 16.

For most of these transition regions, the intensity values contrast with the surrounding tank components. As the components are grouped, the average intensity value moves towards a medium value, which is closer to that of the contrasting transition region. For most groups, except in pass 16, the similarity of intensity was the weakest relationship, and all other relationships were maximal, 1.0. Pass 16 presents the only group where the similarity of orientation value is not maximal.

Detailed Analysis of Results

This analysis considers the range and influence of the relationship values, analyzes the hierarchical structures and evaluates the program goals.

Analysis of relationship values. Similar orientations were a minimum requirement for grouping. All but one group had a maximal value for this relationship. The continuation of boundary lines relationships were maximal for 41 of 52 pairs and the similarity of intensity values were maximal for only 19 of 52 pairs. Lower values for continuation of boundary lines and similarity of intensities relationships allowed regions with greater differences in these features to group. Proximity supported grouping pairs by setting a maximum pair distance. All pairs were between 1 and 8 pixels apart, and 33 of 52 pairs were contiguous.

The regions that grouped might then be characterized as having the same orientation, as being 1.34 pixels apart, as having continuous lines with the same orientation as the grouped regions, and as having a gray scale difference of 15. The intensity difference increases with successive iterations. The greatest intensity difference is in pass 11 of 59 gray scale values.

The structures formed. The grouping program formed 5 structures ranging in complexity from a single pair formed in a single pass to 34 pairs formed in 16 passes. The tank

structure is the most complex. It contains 4 major tank components, and 7 transition regions. Within the tank structure some components are more complex than others. The following discuses each structure and its components and how they were formed.

The tank structure. The tank components and transition regions which form the tank structure are listed in Figure 7.10. The darkened circles indicate that a region is one of the original seventy-eight regions. The regions at the top of the structure were the last groups to form. The regions at the bottom, level 10, were the first groups to form. Each of the darkened regions as well as the tank components is identified in Figure 7.11.

The diagram of the tank structure appears in Figure 7.12 with the addition of region descriptions and the collapsing of levels which join only transition regions. These two figures illustrate the pattern of grouping tank components and then transition regions to form the tank structure. In levels 2-5, shown in Figure 7.12, the last groups to form consist of transition areas, regions between components previously grouped. Transition regions below the tread, between the tread and the tread cover, and between the upper area and the extreme tank top group in these levels. These regions have intensities that contrast with nearby components. It is only after the component has

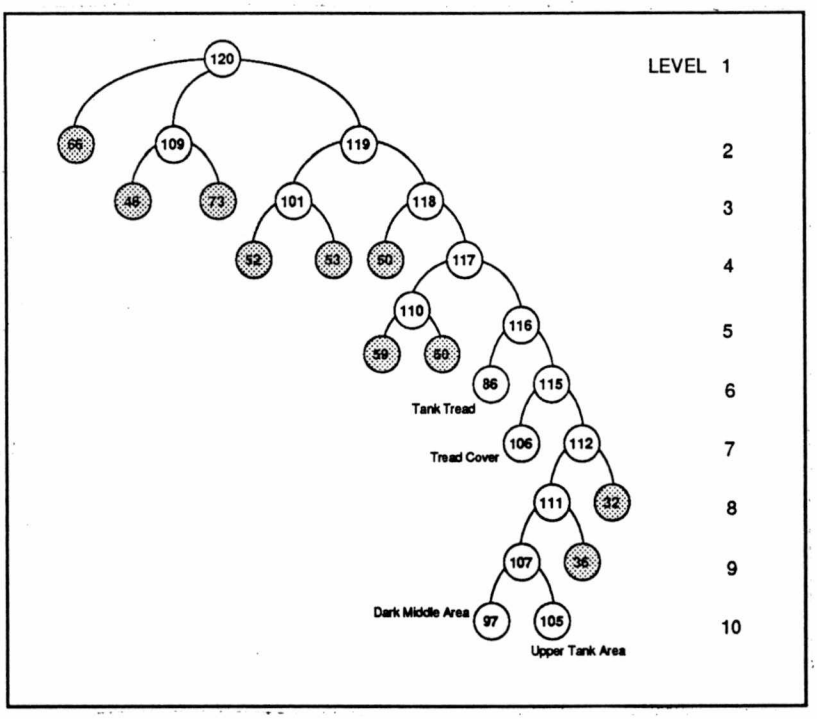


Figure 7.10. Hierarchical Structure of the Tank.

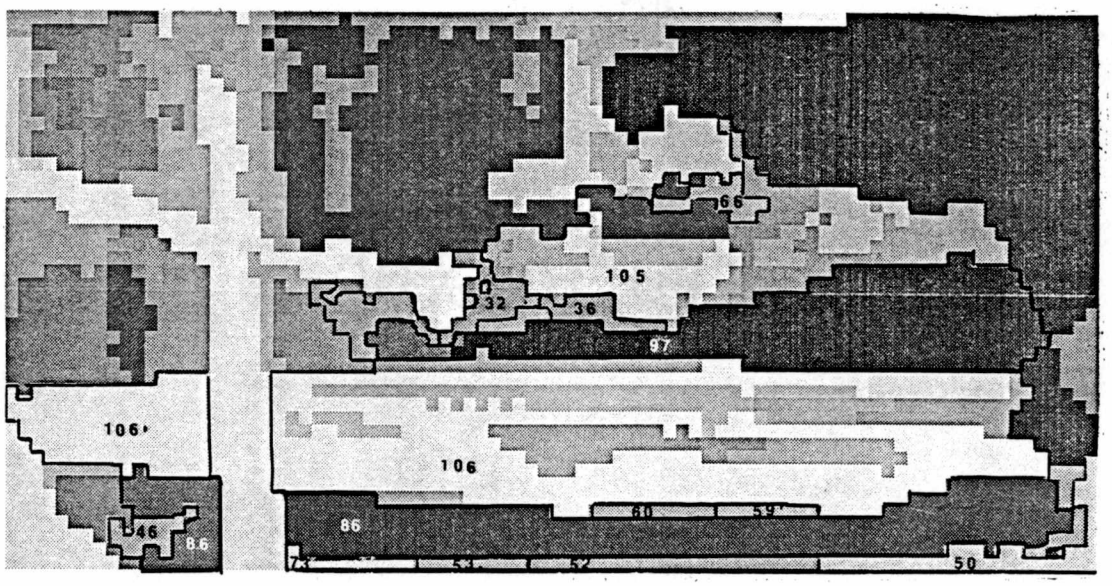


Figure 7.11. Labeled Regions Which Comprise the Tank Structure.

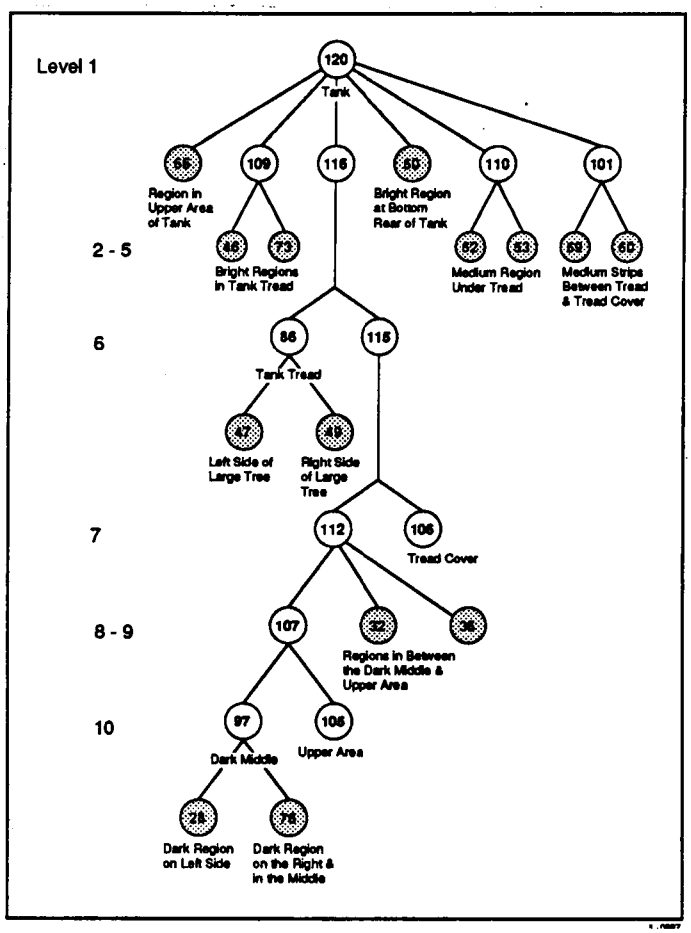


Figure 7.12. Hierarchical Structure of the Tank with Region Descriptions.

grouped and changed the average intensity of the tank structure towards a middle value, that these transition areas are able to group.

Levels 6 and 7 in Figure 7.12 show the grouping of two major components, the tread cover and the tread with the tank structure. Levels 8 and 9 group transition regions between the upper tank area and the dark middle tank area and finally, level 10 groups the upper tank area and dark middle components.

This same pattern of grouping components followed by transition areas may be seen in the complex tank components. Complex components consist of three or more regions and have large variations in feature values. The tread cover, a complex component, consists of 14 regions which vary in intensity values. Figures 7.13 and 7.14 illustrate the grouping structure of the tread cover and Figure 7.15 identifies the regions in the tank image.

The first level in the tread cover structure groups a transition region between the tread and tread cover. The second level breaks the tread cover into the bright and medium-bright intensity areas. This illustrates how complex components are formed by first joining the most similar portions of the component and then grouping these regions as well as transition regions to form the component.

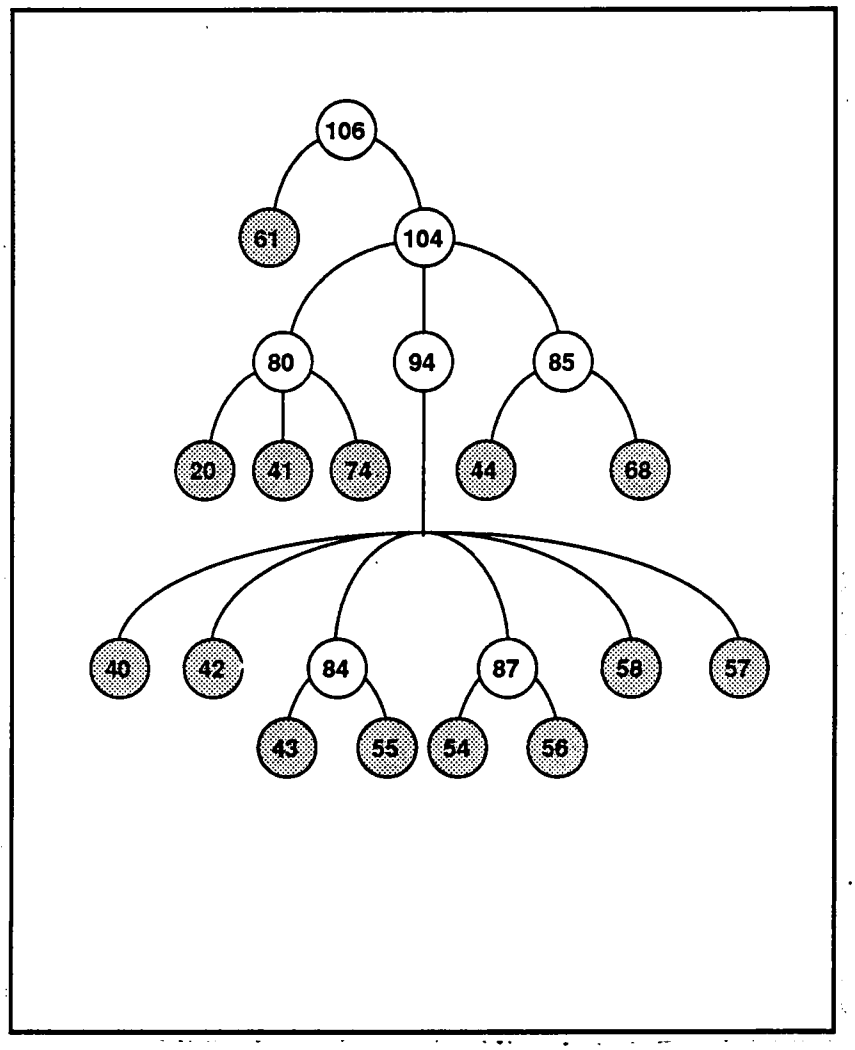


Figure 7.13. Hierarchical Structure of the Tread Cover.

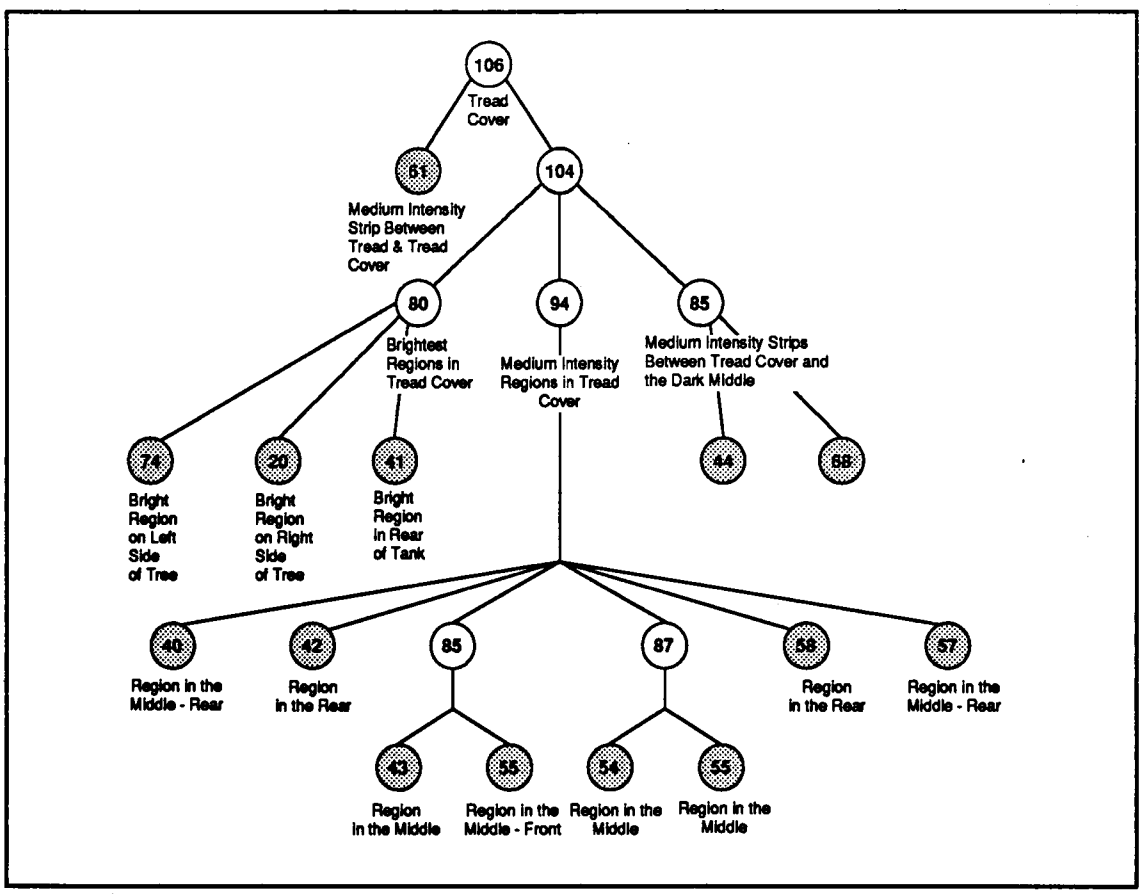


Figure 7.14. Hierarchical Structure of the Tread Cover with Region Descriptions.

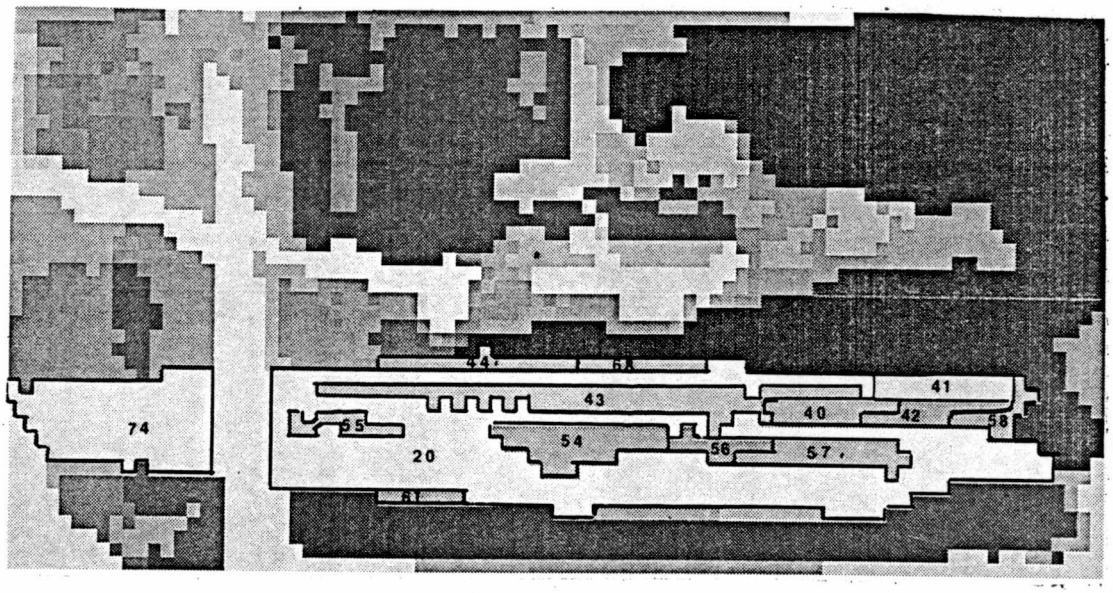


Figure 7.15. Labeled Regions Which Comprise the Tread Cover.

The upper component displays a wider range of intensity values than the tread cover component. Figures 7.16 and 7.17 illustrate the structure of the upper tank area and Figure 7.18 identifies the regions in the tank image. The top of the structure groups a transition region between the upper area and the extreme tank top. The next level divides the component into medium dark and medium to bright regions. This illustrates how complex components are formed by first joining the most similar portions of the component and then grouping these regions as well as transition regions to form the component.

The remaining tank components, the tank tread and the dark middle area are simple components as shown in Figures 7.19 and 7.20. They consist of a single pair of similarly

featured regions. Both groups were formed early in the process, the tank tread in the first pass and the dark middle area in the third pass. Figure 7.21 identifies the regions forming these components in the tank image. The simpler structures form earlier and faster in the grouping process than the more complex structures.

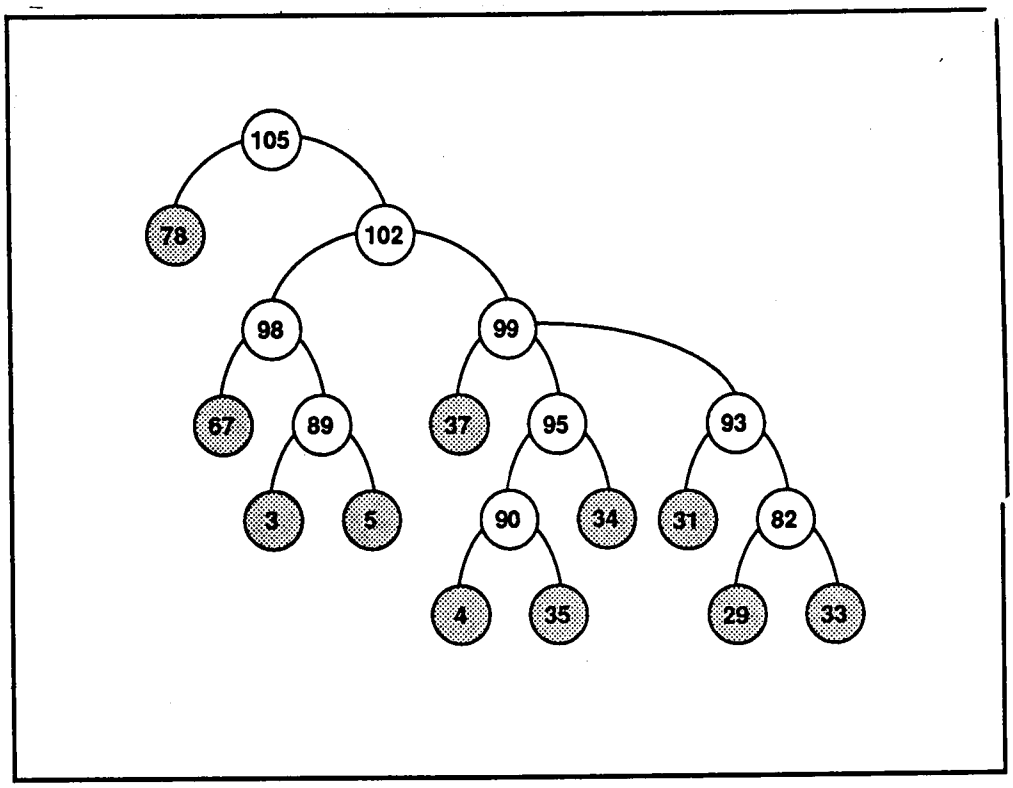


Figure 7.16. Hierarchical Structure of the Upper Tank Area.

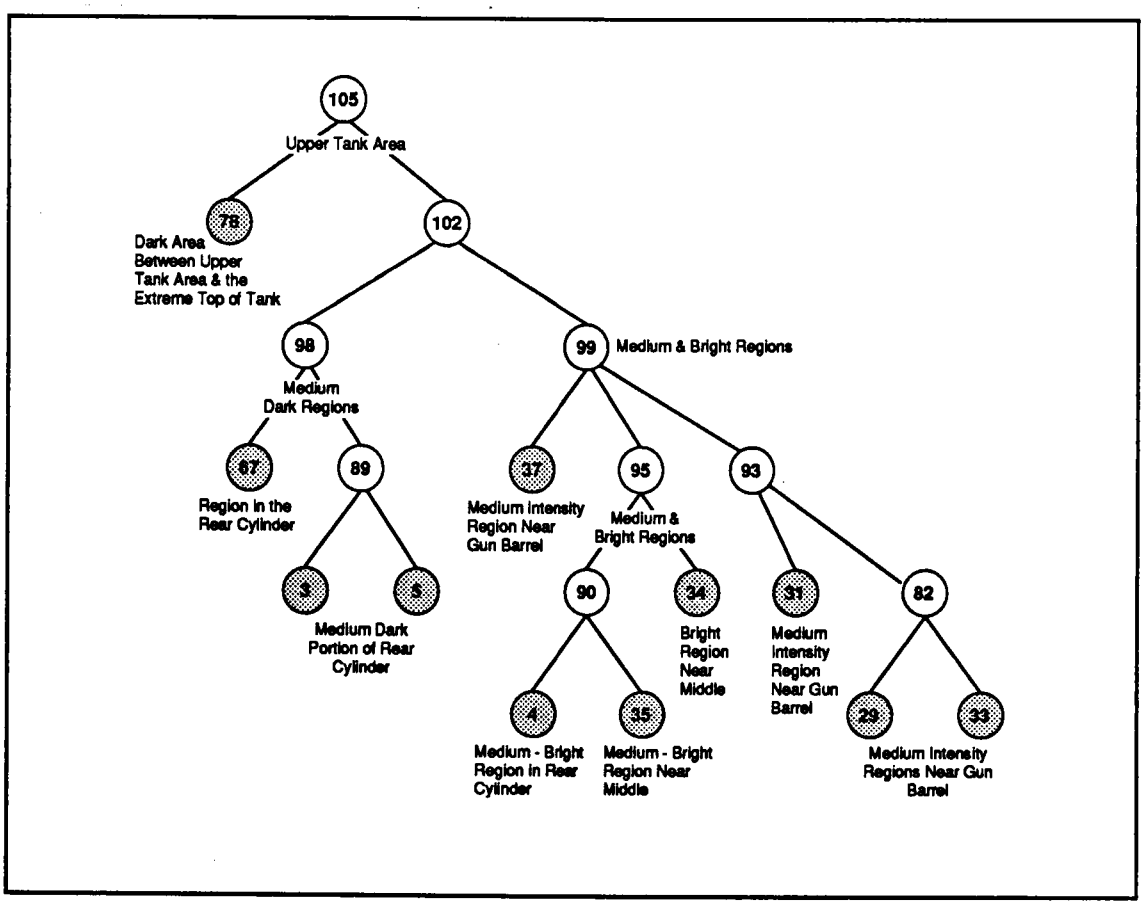


Figure 7.17. Hierarchical Structure of the Upper Tank Area with Region Descriptions.

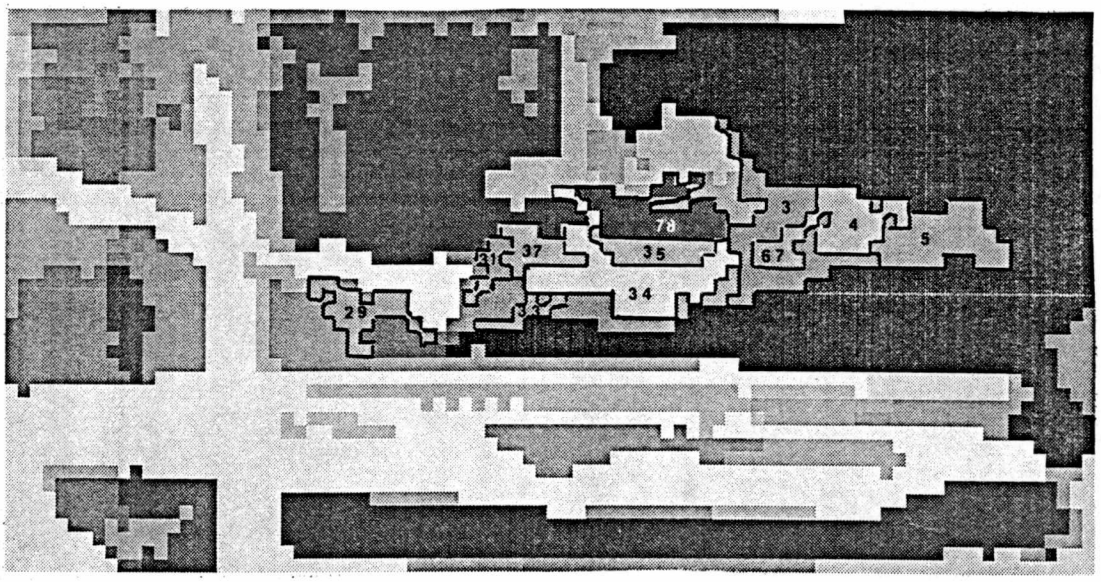


Figure 7.18. Labeled Regions Which Comprise the Upper Tank Area.

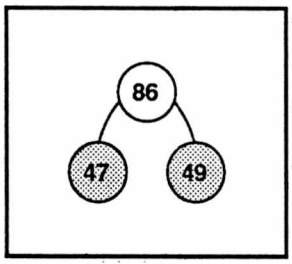


Figure 7.19. Hierarchical Structure of the Tank Tread.

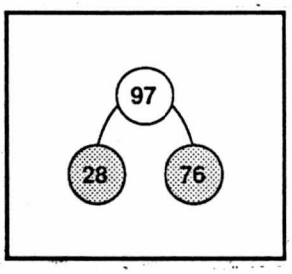


Figure 7.20. Hierarchical Structure of the Dark Middle Area.

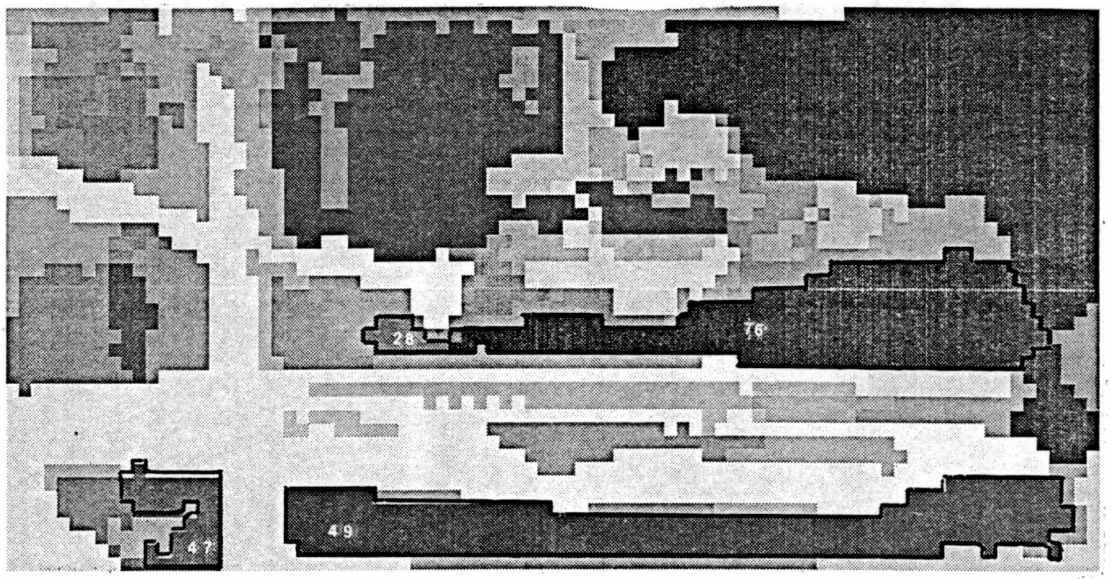


Figure 7.21. Labeled Regions Which Comprise the Tank Tread and Dark Middle Components.

The gun barrel and extreme tank top structures. These components are distinctive and for this reason the grouping program recognized them quickly, but left them standing alone rather than grouping them with the tank structure. Figures 7.22 and 7.23 illustrate the simple structure of these components.

The orientation and boundary lines of the gun barrel distinguish this component. These features enabled the grouping program to join regions separated by an occluding object. These features also prevent the gun barrel from joining any other portion of the tank. Figure 7.24 identifies the regions which form the gun barrel.

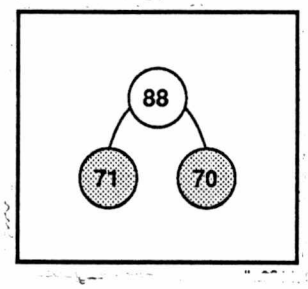


Figure 7.22. Hierarchical Structure of the Gun Barrel.

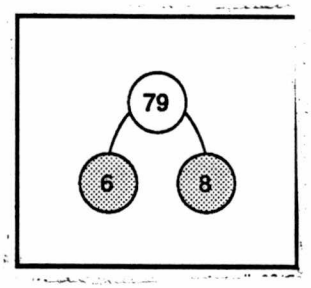


Figure 7.23. Hierarchical Structure of the Extreme Tank Top.

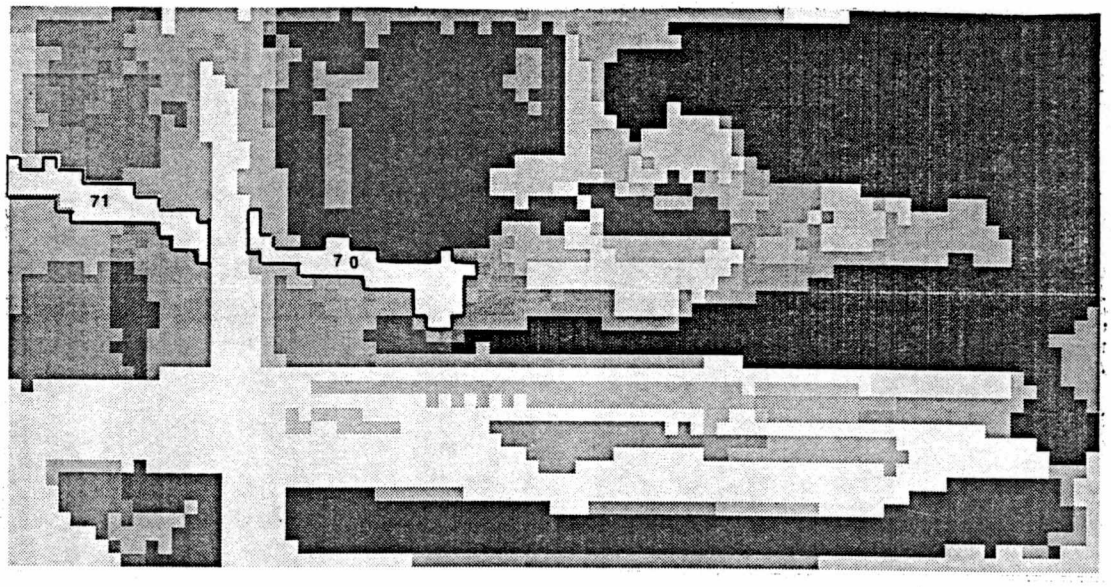


Figure 7.24. Labeled Regions Which Comprise the Gun Barrel.

Similarity of orientation is strong in all grouped pairs. The gun barrel's unique orientation produces a weak similarity of orientation value with other regions. The continuous boundary lines of grouped pairs is usually along the lines of orientation of those regions. The gun barrel's edges are continuous with few regions, and never along the lines of orientation. For those regions it does share continuous boundary lines, the brightness of the gun barrel creates a low similarity of intensity relationship. As the upper tank region is formed, the intensity moves towards a middle value further away from the bright value of the gun barrel. These factors prevent a strong grouping value between the gun barrel and other regions.

The extreme tank top's bright intensity surrounded by dark regions isolate this component as shown in Figure 7.25. Even though the extreme tank top shares an orientation and continuous boundary lines with surrounding regions these regions are dark in intensity, having gray scale differences from 60 to 84.

The background and tree structures. The background consists of a large dark area above and to the rear of the tank and the tree area consists of an area dense with trees near the front of the tank. These two areas yield the background and tree structures.

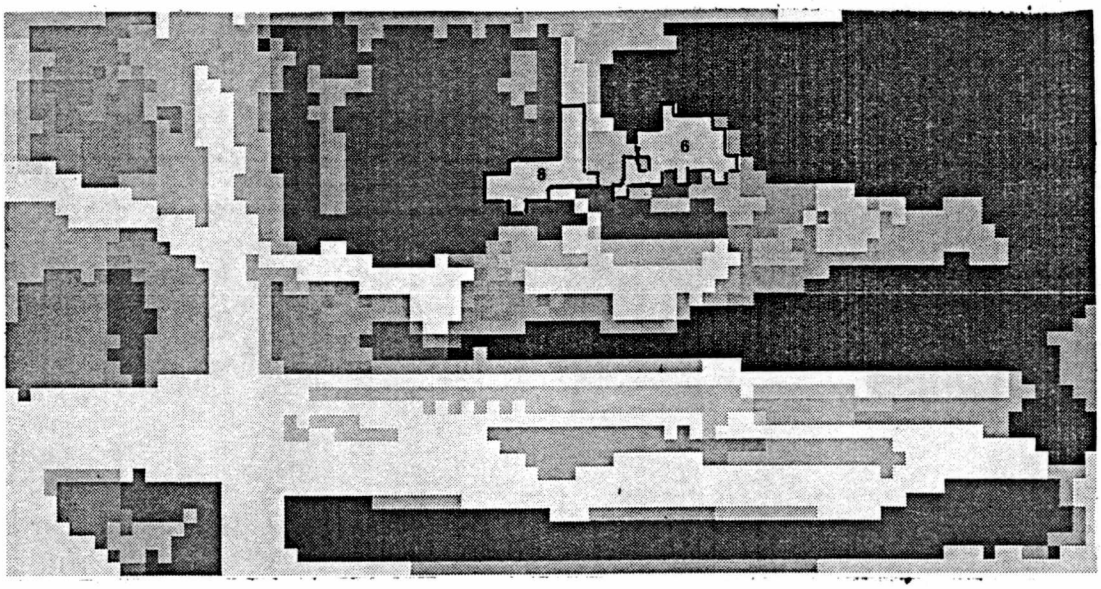


Figure 7.25. Labeled Regions Which Comprise the Extreme Tank Top Component.

The background structure is illustrated in Figure 7.26. This structure groups the bright and dark background regions shown in Figure 7.27. The bright background regions are thin strips along the top edge. These may reflect the white border removed from the image. The dark regions are above the tank and to the rear. An object recognition system could identify this structure as background because of the dark intensity and location of the regions.

The tree structure joins regions with medium to medium bright intensities in the upper left portion of the image. Figure 7.28 diagrams the structure of the trees. Groups form on the right and left of the large tree as shown in Figure 7.29. A group also forms in the upper left corner of the image. This structure may be characterized as vertical

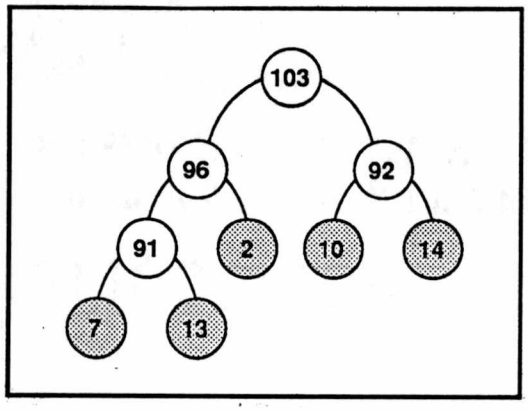


Figure 7.26. Hierarchical Structure of Background.

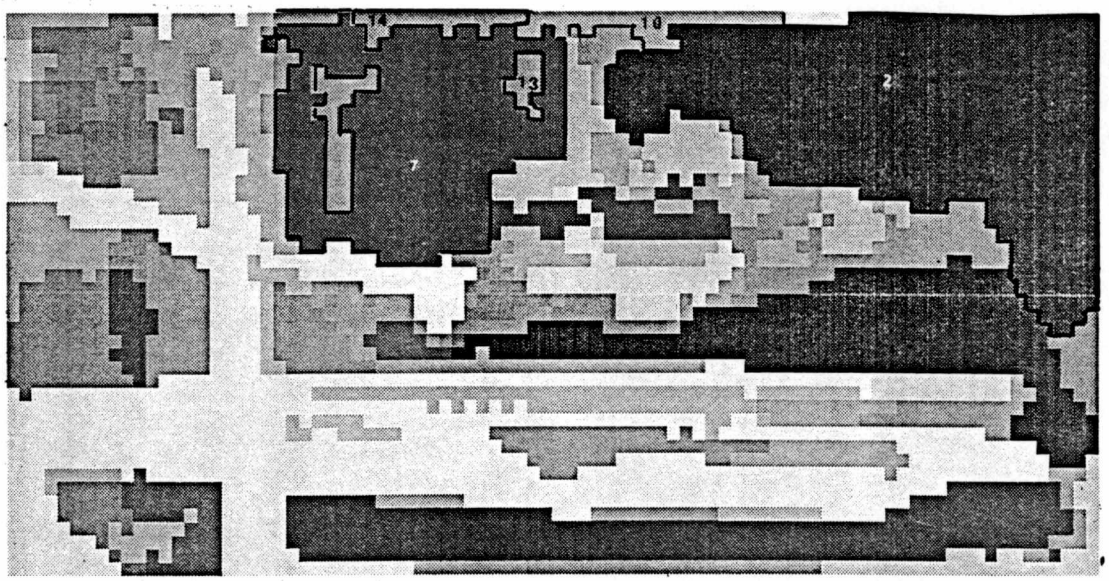


Figure 7.27. Labeled Regions Which Comprise the Background.

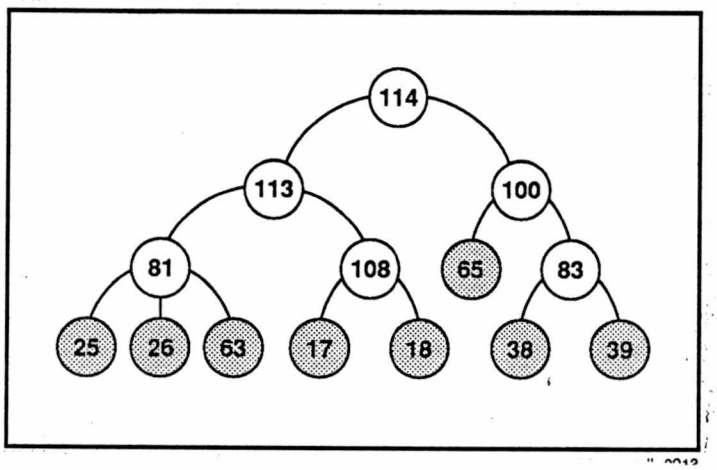


Figure 7.28. Hierarchical Structure of the Tree Area.

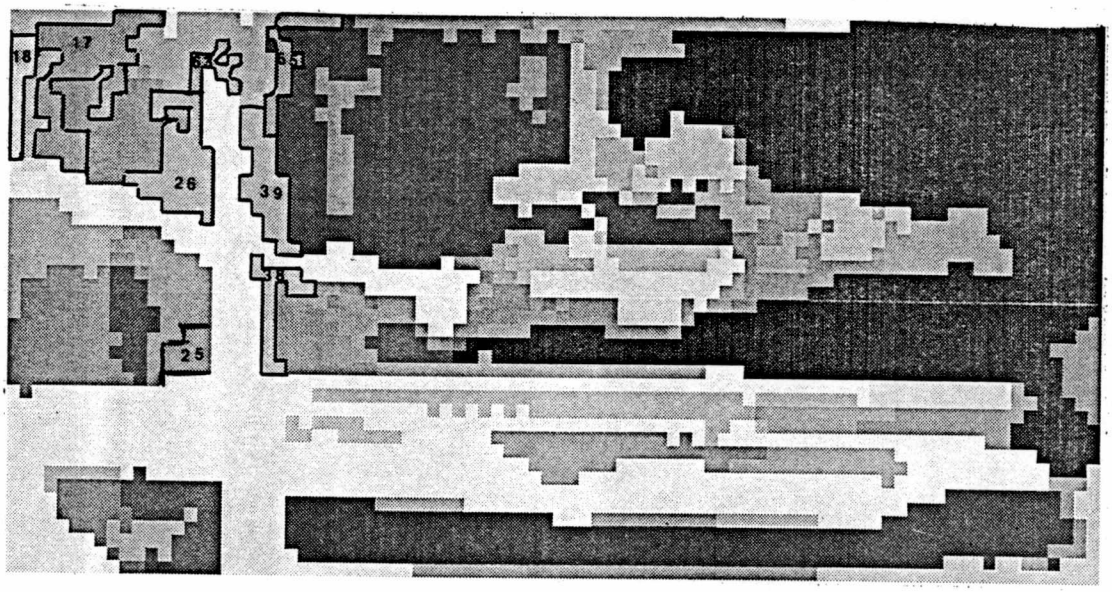


Figure 7.29. Labeled Regions Which Comprise the Tree Area.

with a medium bright intensity. Similarity of orientation and intensity were major influences in the formation of this structure. Because of the vertical orientation and medium bright intensity, an object recognition system could identify this structure as wooded background.

Regions not grouped into structures. The large tree, numerous background regions, and the extreme front edge of the tank tread did not join any group. The large tree is a significant object in this image, it therefore remains alone, not embedded in a structure. It does not group with other tree regions surrounding it because its intensity is much brighter than the other trees.

The background regions which stand alone generally have orientations which are neither horizontal or vertical. They may be diagonal and unlike other nearby regions. The region in the extreme front of the tread, has a unique orientation, preventing it from grouping with the rest of the tread.

The goals evaluated. The grouping program satisfied goals by forming components and then grouping these components into a tank structure. The tank structure contains the tread, tread cover, dark middle and upper tank area components. The grouping program successfully formed the gun barrel and extreme tank top components, however, these components were never grouped with the tank structure. Because these components were formed, and because most

components did group together, an object recognition system could use feedback to identify the gun barrel and extreme tank top structures as components of the tank.

Object Recognition Using Hierarchical Structures

The following discussion describes a conceptual object recognition system which uses perceptual features to identify objects. A hierarchical structure contains the feature information for an object and its components. An object recognition system matches these features with the features in an object template.

An example tank template is described below. The tank components, their perceptual features and the features which should influence grouping of regions are listed in Table 7.1. This template is based on viewing the tank image and all figures are general estimates. This discussion presents an example using only one tank image, multiple views of different tanks would be needed to test these concepts of object recognition.

The perceptual features described in the tank template include the intensity, orientation, location and size. Size of the components is given as a percentage of the total tank. Location of the components is given relative to other contiguous components, if the contiguous components have not been identified the system will look for surrounding regions which have features similar to the contiguous components

Table 7.1. Tank Template.

	Feature Values				Influential Relationships		
Id.	Int.	Or.	Location	Size %	Int.	Or.	C.B.L.
Whole Tank	80-110	0	Contiguous with ground	10-70%			
Tank Components							
Gun Barrel	110-225	0-45 135-180	Contiguous with middle of tank and extending outward	5-10%	no	yes	yes
Tread	50-0	0	At bottom of the tank	15-30%	yes	yes	no
Tread Cover	80-110	0	Above tread and below dark middle	30-50%	no	yes	no
Dark Middle	50-0	0	Above tread cover and below upper area	10-30%	yes	yes	no
Upper Area	80-110	0	Above dark middle and below extreme top	15-30%	no	yes	no
Extreme Top	110-225	0	At the top of the tank	5-10%	yes	no	no

Grouping features which are expected to strongly influence the formation of each component are also listed. The following describes the template in Table 7.1.

- o Whole Tank - Because there are more bright areas than dark, the tank may be characterized as medium bright in intensity. The tank is horizontal in orientation and will probably be neither an extremely small nor an extremely large portion of an image, therefore the estimated tank size will range from 10-70% of the image. The tank is always located contiguous to the ground. Another template may be necessary to identify the "ground".

Tank components and their perceptual features:

- o Gun Barrel - This component may be characterized by a bright intensity due to its metallic shine. The gun barrel may have a horizontal or diagonal orientation and will be located contiguous to the middle of the tank and extend outward in one of numerous possible directions. It is estimated the gun barrel will comprise approximately 5-10% of the tank. Regions which form the gun barrel will have similar orientations and continuation of boundary lines.
- o Tread - Characterized by a dark intensity and horizontal orientation, the tread will be located at the bottom of the tank, and comprise 15-30% of the tank.

Strong grouping relationships for this component are similarity of intensity and orientation.

- o Tread Cover - This component contains regions with intensities ranging from bright to dark. The larger number of bright regions will cause the average value to be medium-bright. Its orientation is horizontal and the tread cover is located above the tread and below the dark middle. It comprises 30-50% of the tank and similarity of orientation is the primary grouping relationship.
- o Dark middle - This component is dark, horizontal and located between the tread cover and the upper tank area components. It comprises 10-30% of the tank and its regions group because of strong similarity of orientation and intensity relationships.
- o Upper area - This component contains regions with intensities ranging from bright to dark. The larger number of bright regions will cause the average value to be medium-bright. Its orientation is horizontal and this component is located between the dark middle and the extreme tank top. It comprises 15-30% of the tank and orientation is the primary grouping relationship.
- o Extreme tank top - This bright, horizontal component is located at the top of the tank. It comprises

5-10% of the tank and its regions group because of a strong similarity of intensity relationship.

The hierarchical structures produced by the grouping program for the tank in the woods image are listed in Table 7.2. The Table lists each structure with a structure number, feature values and relationship values. Structure numbers 120, 79 and 88 are identified in Figure 7.4 on page 50, structures 114 and 103 are identified in Figure 7.5 on page 50. The features listed in Table 7.2 include intensity, orientation, size and location. The intensity value is given in gray scale values and the orientation is given in degrees. The location is described with text. Pixel locations would be used in an actual system. Size is given first as the number of pixels in the structure and second as the percentage of the total image. The table also includes the relationship values, similarity of intensity, similarity of orientation, and continuation of boundary lines, of the last group formed in each structure. An average is used if more than one pair formed the last group.

Template matching. The object recognition system starts by identifying the structure most likely to produce a match with the tank template. The system matches the features of each structure with the features for the whole tank listed in Table 7.1. The features that best match the whole tank features in the template will indicate the structure most likely to produce a match with the template.

Table 7.2. Example One Hierarchical Structures with Feature and Relationship Values.

Region	Feature Values					Relationship Values		
No.	Int.	Or.	Location	Size	%	Int.	Or.	C.B.L.
120	86	0	Middle to lower, front to back	1631	44%	0.65	0.75	0.82
88	136	135	Middle, front to center	86	2.3%	1.00	1.00	1.00
79	118	0	Top middle	64	1.7%	1.00	1.00	1.00
114	90	90	Top to center front	172	4.6%	1.00	1.00	0.30
103	42	0	Top to center middle to rear	855	23%	0.23	1.00	1.00

The tank template indicates the whole tank has medium-bright intensity, is horizontal, and consists of 10-70% of image. In the structure feature Table 7.2, the only structures with horizontal orientations are 120, 79 and 103. Structure 103 has a dark intensity value and structures 79 and 120 have a medium-bright intensities. The size of structure 79, 1.7% of the image, is under the 10-70% specified in the template for the whole tank. Structure 120 is therefore the best candidate to match the tank template.

The object recognition system will attempt to match the components in the tank template to the components in structure 120. The features of the components in region 120 are

listed in Table 7.3. The levels in Table 7.3 correspond to the levels shown in Figure 7.10 on page 62.

The object recognition system searches the hierarchical structure for matches with the template components. During this search the system recognizes the existence of small transition regions in the structure and attempts to discard them. The size of a region usually indicates that it is a transition region. In level 2 of Table 7.3, both regions 66 and 109 are small, consisting of less than 2% of the structure. The system concludes that these are transition regions. Region 119 in level 2 of Table 7.3 comprises 97% of the structure, because this size is larger than any of the template components the system searches the next level which contains the regions forming region 119.

Levels three, four and five contain transition regions and regions comprising over 90% of the structure. The transition regions are identified by their small size, 1.2, 2.4, and 1.1 percent of the structure respectively. The regions comprising over 90% of the structure are larger than any template components. In each case the system considers the next structure level which contains the groups forming the regions which comprise over 90% of the structure.

Level six contains a region that is a potential component. Region 86 has a dark intensity, a horizontal orientation and consists of 17% of the structure. Its strongest

Table 7.3. Example One Features and Relationship Values of the Regions Forming Structure 120.

Region			Feature Values			Relationship Values		
No.	Int.	Or.	Location	Size	%	Int.	Or.	C.B.L.
Level 1								
120	86	0	Middle to lower, front to back	1631	100%	0.65	0.75	0.82
Level 2								
66	83	0	Upper-middle middle-rear	13	0.7%	-	-	-
109	117	0	Lower front and middle	25	1.5%	0.28	1.00	1.00
119	85	0	Middle to lower, front to back	1594	97%	0.92	1.00	1.00
Level 3								
101	74	0	Lower center and rear	20	1.2%	0.29	1.00	1.00
118	85	0	Middle to lower front to back	1574	96%	0.27	1.00	1.00
Level 4								
50	122	135	Lower rear	40	2.4%	-	-	-
117	84	0	Middle to lower front to back	1534	94%	1.00	1.00	1.00
Level 5								
110	90	0	Lower center and rear	18	1.1%	0.28	1.00	1.00
116	84	0	Middle to lower front to back	1516	93%	0.19	1.00	1.00
Level 6								
86	43	0	Lower front to rear	285	17%	0.88	1.00	1.00
115	94	0	Middle to lower front to back	1231	75%	0.17	1.00	1.00

Table 7.3. (Continued)

Region			Feature Values			Relationship Values		
No.	Int.	Or.	Location	Size	%	Int.	Or.	C.B.L.
Level 7								
106	122	0	Middle-lower front to back	661	40%	0.34	1.00	1.00
112	62	0	Middle front to back	570	35%	1.00	1.00	0.30
Level 8								
111	62	0	Middle front to back	552	34%	1.00	1.00	0.65
32	67	0	Middle front	18	1.1%	-	-	-
Level 9								
107	62	0	Middle front to back	537	33%	0.21	1.00	1.00
36	62	135	Middle center front to back	15	0.9%	-	-	-
Level 10								
97	36	0	Middle front to back	247	15%	0.42	1.00	1.00
105	83	0	Middle-upper front to back	290	18%	0.34	1.00	1.00

grouping relationships are similarity of intensity and continuation of boundary lines. The only dark intensity components are the tread and dark middle components. Both components have a horizontal orientation and match the size feature of region 86. The tread is located at the bottom of the tank and region 86 is also located in the lower portion of the structure. The tread indicates that intensity should be a strong grouping relationship and the intensity relationship was not the strongest for region 86. However, because so many other features match, and because the intensity relationship value is high, region 86 is considered a match with the tread component.

The other region in level 6 is region 115 and it comprises 75% of the structure which is larger than any template component. The system therefore considers level 7.

Level 7 contains two regions within the size range of the tank components. Region 106 has a medium bright intensity and horizontal orientation which match the features of the extreme tank top, the tread cover and the upper area components in the template. A size of 40% of the structure, and a location above the tread and below the dark middle area indicate a match with the tread cover. Even though the dark middle area has not been identified, the regions above region 106 are dark and horizontal matching the features of the dark middle area. The similarity of orientation relationship values are strong in region 106, which

matches the template for the tread cover. Region 106 is therefore considered a match with the tread cover.

In level 7, region 112 has a medium dark intensity and comprises 35% of the structure. These features do not match any of the components and therefore the regions forming region 112 will be searched.

Levels 8 and 9 both contain small transition regions consisting of less than 2% of the structure. The remaining region is medium dark in intensity and does not match any template components.

Level 10 presents two possible component matches. Region 97 has a dark intensity, horizontal orientation, comprises 15% of the structure and is located in the middle of the structure. These features match the dark middle component. Region 105 has a medium bright intensity, a horizontal orientation, comprises 18% of the structure, and is located above the dark middle component. These features match the upper area component.

Four of the six tank components have now been matched. Structure 120 has been completely searched, and the system must therefore search elsewhere for the missing components. A feedback program directs the object recognition system to search the remaining structures for a match with the gun barrel and the extreme tank top.

The tank template in Table 7.1 indicates the gun barrel is bright with a horizontal or diagonal orientation. It will be located contiguous to the tank and extend outward. Table 7.2 lists the structures and their features. Regions 88 and 79 both match the intensity and orientation of the gun barrel, however, only region 88 matches the location. The object recognition system considers region 88 to be a match with the gun barrel component.

Region 79 provides the only possible match with the extreme tank top. It has a bright intensity, a horizontal orientation and a location at the top of the image above the upper area component. Its size also matches the extreme tank top component. The object recognition system concludes that region 79 matches the extreme tank top component.

The object recognition system has matched all six of the tank components and concludes that it has a tank in the image. Since all the components were identified the probability for this identification is high. The object templates may include probability factors for the identification of an object without specific components and for the identification of a component without specific features. Using these factors, an object recognition system provides the user with an object match and a factor indicating how good that match really is.

Conclusions

The grouping program uses perceptual features to form tank components and create useful hierarchical structures. The first iteration of the program joined regions separated by an occluding tree and demonstrated the program's ability to group regions across boundaries. The tank in this image is comprised of many heterogeneous regions. The grouping process demonstrated its ability to form a structure representing this complex object. The program divided the object into significant perceptual components which themselves created complex structures. It was shown that an object recognition system could identify the object components using their perceptual features and conclude fairly confidently that there is a tank in the image.

Example Two: Wrenches and Pliers

In the second example, the grouping program processed the image containing various tools shown in Figure 7.30. This example contrasts with the tank in the woods example because this image contains numerous small objects, the tools, rather than a single large object, the tank. The grouping program processes this example in fewer iterations, six, compared to 16 for the tank image. Six structures were produced, all simpler than the complex structures created for the tank image. These six structures and the processing of this image are described below.

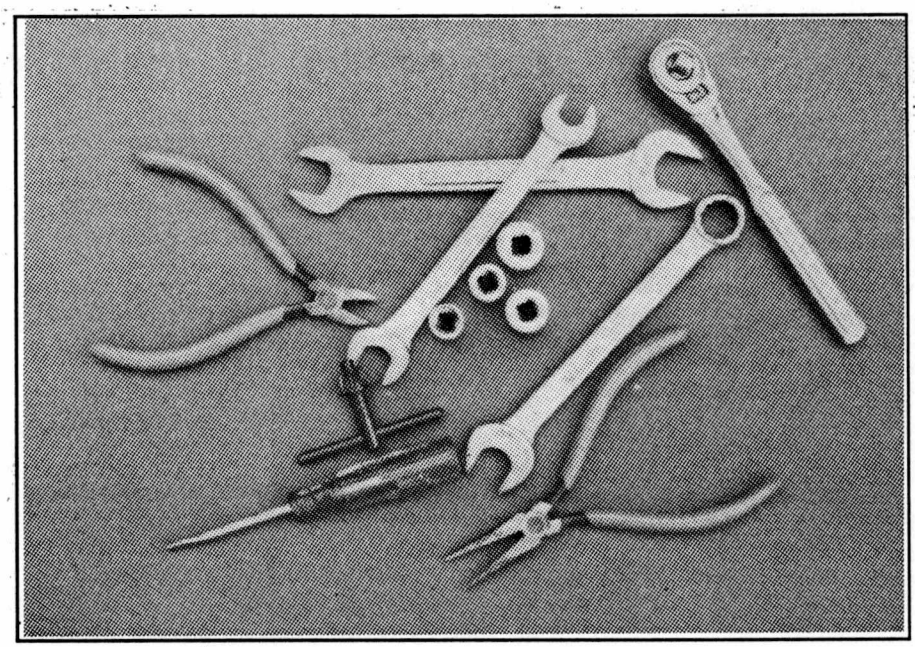


Figure 7.30. Image of Wrenches and Pliers
Source: Fu, K.S., Gonzalez, R.C., Lee, C.S., Robotics Control, Sensing, Vision, and Intelligence, (New York: McGraw-Hill;1987)

Detailed Description of the Processing

Preprocessing. Preprocessing routines reduced, filtered, and segmented this image. Figure 7.31 displays the digitized, filtered and reduced image. From this image the segmentation process identified the fifty-three regions shown in Figure 7.32. Many of the tools segmented into single regions. Figure 7.33 identifies the tools in the image and the single tools are outlined in Figure 7.34. The grouping program joined regions to form the horizontal wrench separated by the overlapping diagonal wrench and portions of the right lower pliers.

The unique preprocessing of this image includes using specific segmentation parameters, removing a background region, and manually dividing one segment. The segmentation parameters determine the range of gray scale values allowed in a single region and the minimum number of pixels required for a region. In this example, the parameters allowed a +/- 12 gray scale difference in regions and required at least 4 pixels to form a region.

The segmentation process created 50 regions. One of the regions was unusually large and encompassed areas of varying orientations. The region growing algorithm used by the segmentation program uses significant features to determine the segments (Gonzalez and Wintz, 1977.) The version of

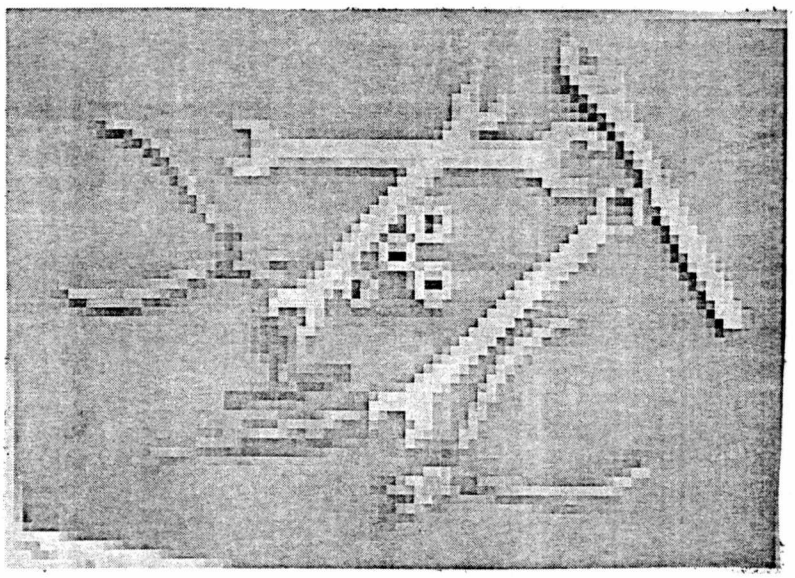


Figure 7.31. Image of Wrenches and Pliers After Preprocessing.

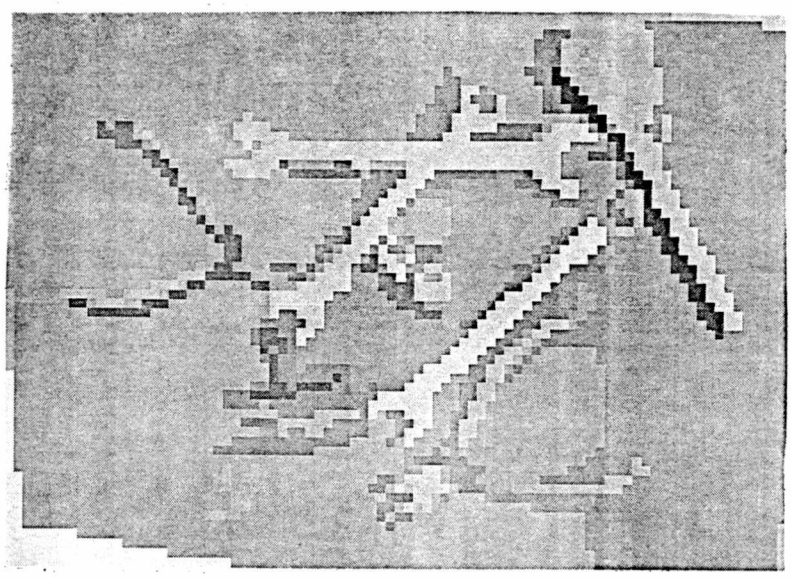


Figure 7.32. Fifty-Three Segmented Regions of the Wrenches and Pliers Image.

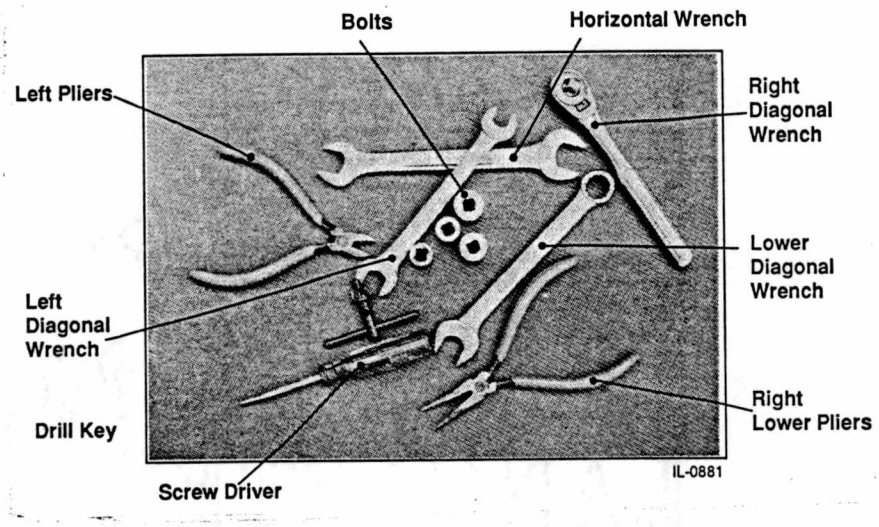


Figure 7.33. Image of Wrenches and Pliers with Labeled Tools.

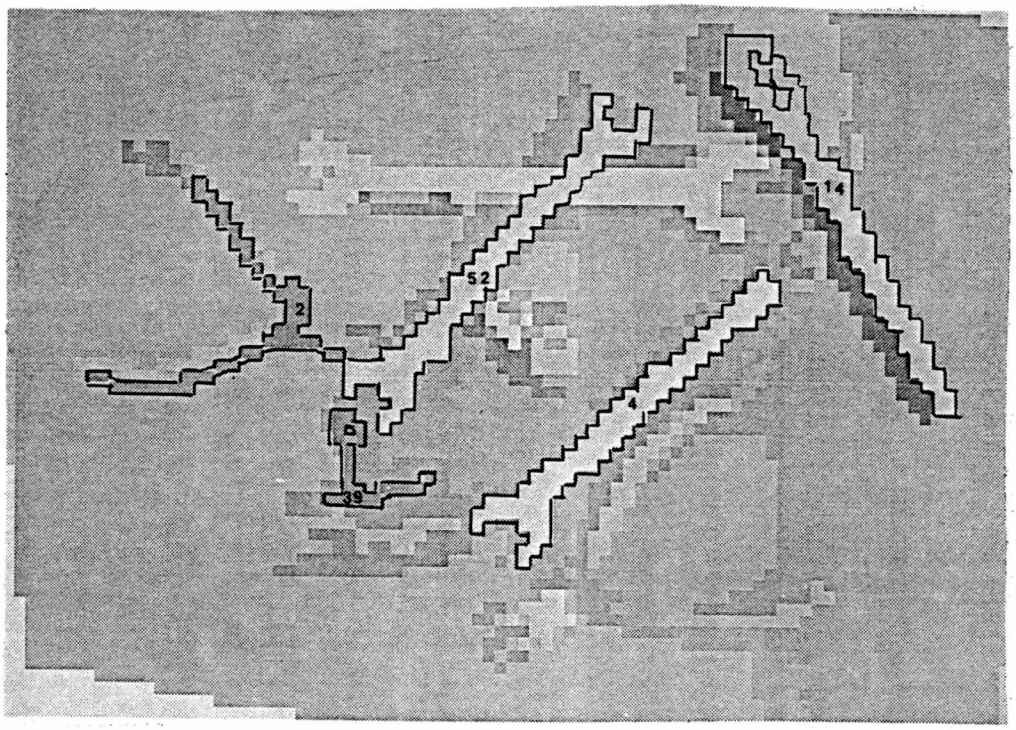


Figure 7.34. Whole Tools Segmented into Single Regions:
Region 2 = Left Pliers,
Region 52 = Left Diagonal Wrench,
Region 4 = Lower Diagonal Wrench,
Region 14 = Right Diagonal Wrench,
Region 39 = Drill Key.

this algorithm implemented in this program uses the intensity feature only. However, as orientation is a significant feature, one region was manually divided based upon differences in orientation. The single segmented region is shown in Figure 7.35, the dashed lines represent the manual division of regions.

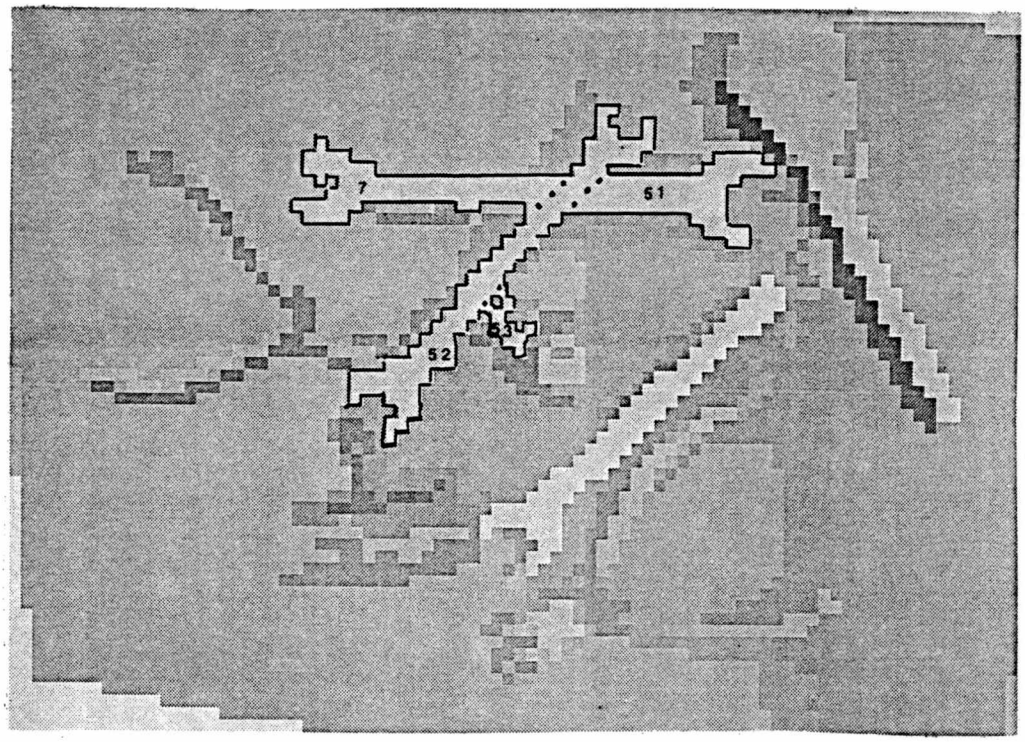


Figure 7.35. Manually Divided Regions.

A single background region encompasses nearly all the background area except the area surrounding the bolts and the area in the far left. Region 1, shown in Figure 7.36, extends from the top to the bottom edge and from the right to within 13 pixels of the left edge. Using the a priori

knowledge that this image contains tools which are smaller than the length or width of the image, it is possible to deduce that region 1 contains background. If regions can be identified as background regions, they may be eliminated from the grouping process. Because of the number of edges and proximity positions in a large background region such as region 1, their elimination reduces processing time for the

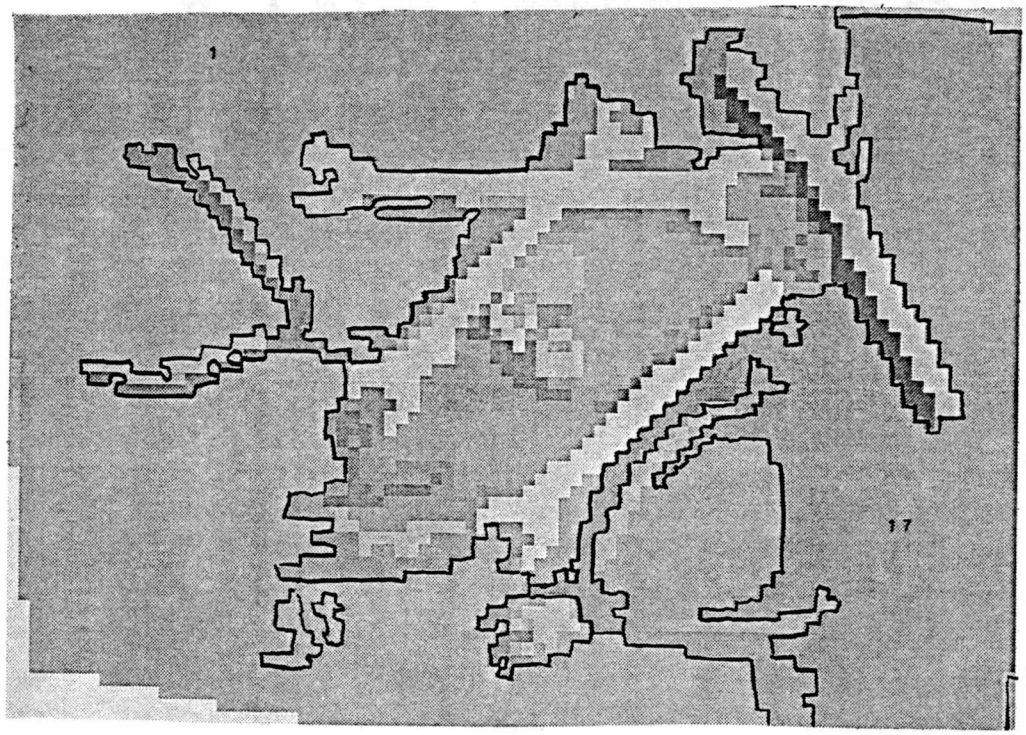


Figure 7.36. Background Regions.

grouping program. Background regions generally hold no value in a hierarchical structure to an object recognition system. For images with large background areas, the saved processing time justifies an inspection of large regions to identify

and eliminate background regions. In the wrenches and pliers image, region 1 was identified as background and not considered in the grouping process.

Goals. Goals for the grouping process are based on the regions created by the segmentation process and how these regions may be grouped to best enhance object recognition. The region segmentation program successfully formed regions containing whole tools. Specifically it created the following regions representing whole tools shown in Figure 7.34:

- o left pliers - region 2
- o overlapping diagonal wrench - region 52
- o lower diagonal wrench - region 4
- o right diagonal wrench - region 14
- o drill key - region 39

The remaining tools are the horizontal wrench, the screwdriver, and the right lower pliers. The screwdriver blade is missing due to preprocessing filters and not available for grouping. Therefore, the remaining tasks for the grouping program are to join the regions forming the horizontal wrench and join the regions forming the right lower pliers. Other goals include joining shadow regions with the appropriate tool region and grouping background regions with other background regions.

Summary of six processing iterations. At each iteration or pass, the grouping program generates relationship values and forms groups based on these values. The structures for this image were formed in 6 passes, the following discussion summarizes the groups formed in each pass.

The first pass groups the most strongly related regions in the image, regions 7 and 51, representing the two disjoint regions of the horizontal wrench. Two regions in the right lower pliers also grouped together. Other groups formed in this pass are background regions surrounding the drill key and regions in the upper right corner of the inside background area near the bolts.

Pass 2 joins a bright region with the horizontal wrench and pass 3 groups medium dark shadow regions with the horizontal and overlapping diagonal wrenches. In both of these passes the similarity of intensity relationship values were medium and low enabling regions with a greater variety in intensity values to group.

In the fourth pass, weak continuation of boundary lines relationships enable groups without strong boundary line continuations to group, such as bright regions representing the bolts, shadow regions in the upper right corner, and the background regions in the center. Pass 5 joins background regions in the center of the image. Pass 6 groups the horizontal wrench with the overlapping diagonal wrench and groups shadow regions surrounding the drill key.

Detailed Analysis of Results

This analysis considers the range and influence of the relationship values, analyzes the hierarchical structures and evaluates the program goals.

Analysis of relationship values. The similarity of orientation relationship proved to be the dominant relationship, followed by similarity of intensities, and finally continuation of boundary lines as the weakest relationship. The weights for all relationships were 1.0. The maximum distance between grouped regions was 8 pixels. The proximity relationship proved important for setting this maximum distance, but was not decisive in grouping regions which were between 1-8 pixels apart.

The structures formed. The grouping program created six simple structures in 6 iterations. The following discuses each structure.

The wrench structure. The grouping program joined the horizontal wrench divided by the overlapping diagonal wrench in the first iteration. Figures 7.37 and 7.38 illustrates the wrench structure and Figure 7.39 identifies the regions in the image.

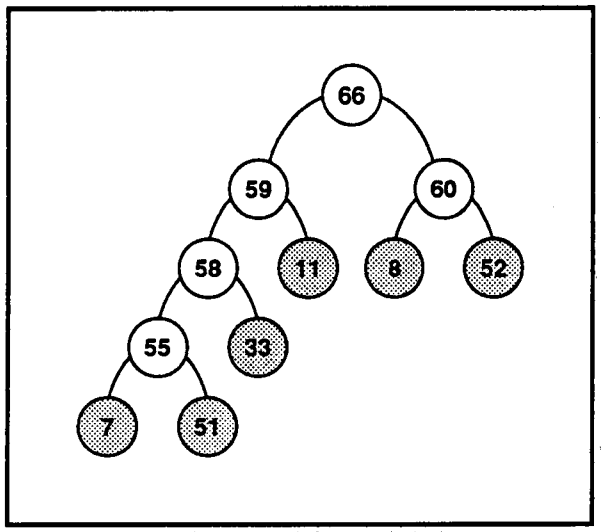


Figure 7.37. Hierarchical Structure of Wrenches.

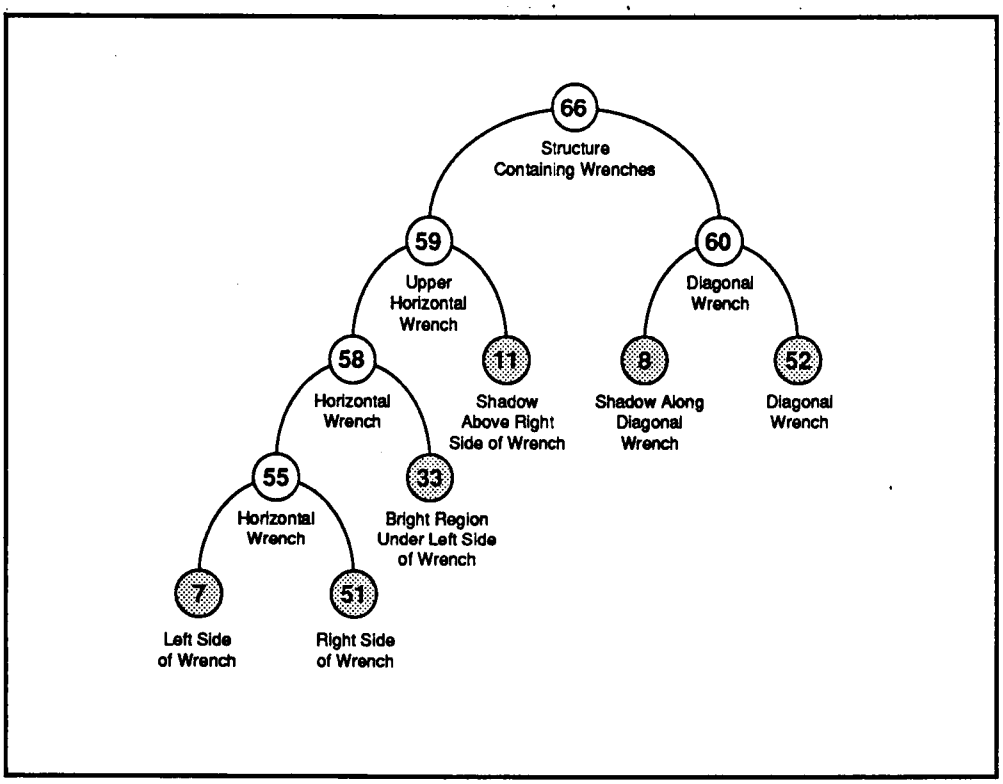


Figure 7.38. Hierarchical Structure of Wrenches with Description.

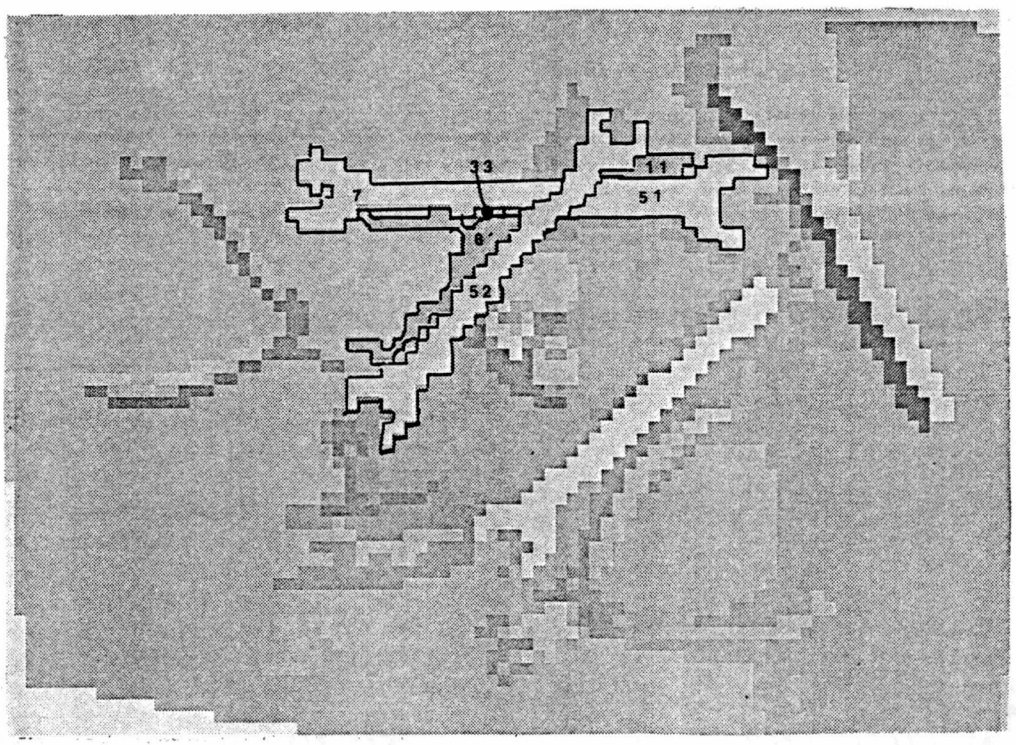


Figure 7.39. Labeled Regions of Wrench Structure.

Regions 7 and 51 forming the horizontal wrench have the strongest relationship values of any group in all iterations which demonstrates the program's ability to join regions separated by occluding objects. A shadow region, 11, above to the right of the horizontal wrench grouped with the wrench as well as a bright region, 33, in the lower center of the wrench. The overlapping diagonal wrench, 52, grouped with a shadow region, 8, along its side.

These two wrenches, the horizontal wrench and the overlapping diagonal wrench, joined together into a single structure. Because they are overlapping and share many features, the grouping program considered them as one

structure with two components. Figures 7.37 and 7.38 illustrate the single structure comprised of two distinct components. To an object recognition system, a single structure comprised of two tools indicates a close relationship between the tools such as they are overlapping.

The pliers structure. The lower handle and the darker portions of the right lower plier nose grouped together in a simple structure shown in Figure 7.40. The regions of this structure are identified in Figure 7.41. The upper handle and the brightest portion of the plier nose did not group with the rest of the plier structure, they remained as single regions. This is because the regions forming the pliers display different intensity values, orientation values and orientations of boundary lines. These differences produce weak relationship values and prevent these regions from grouping. The perceptual features of common endpoints and symmetry may be used by humans to group these tools. Incorporating these features into the grouping program would enhance the ability of the program to join these regions and form a single structure for each tool. In this example, an object recognition system could use a feedback system to locate the missing regions of the pliers.

The bolt structures and background structures. The grouping program formed a structure consisting of the bolts in the center of the image shown in Figure 7.42, and identified in Figure 7.41. The program also formed background

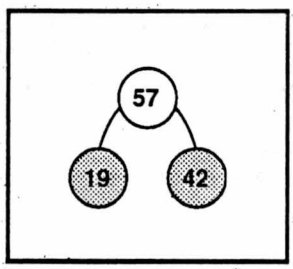


Figure 7.40. Hierarchical Structure of the Right Lower Pliers.

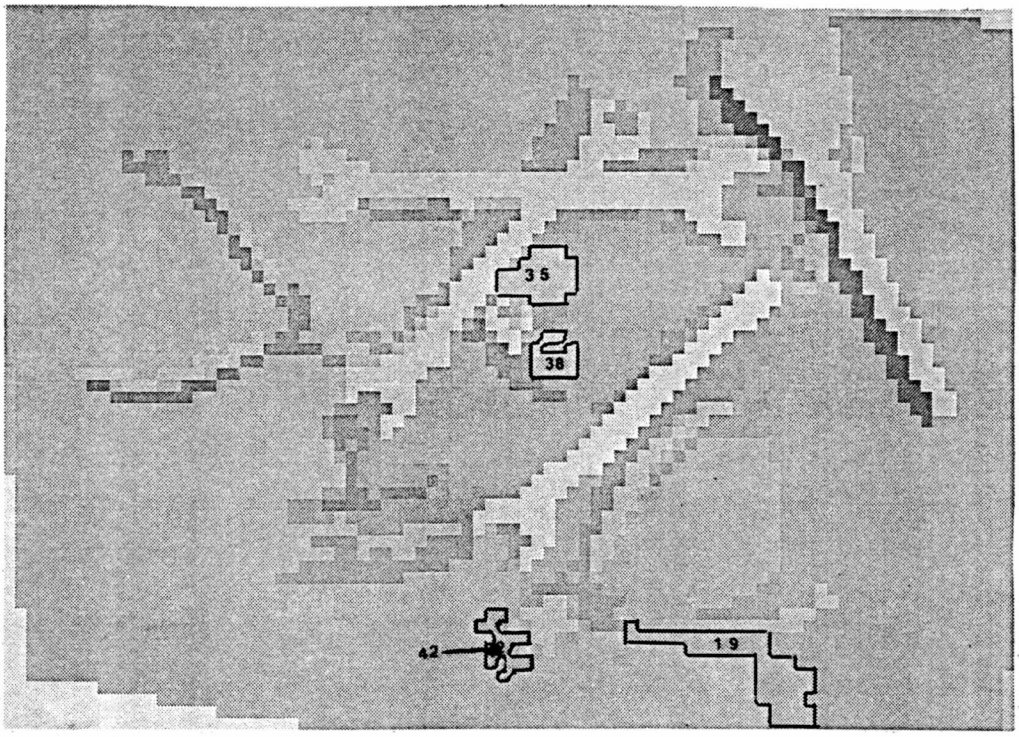


Figure 7.41. Labeled Regions Comprising the Right Lower Pliers Structure, 57, and the Bolt Structure, 63.

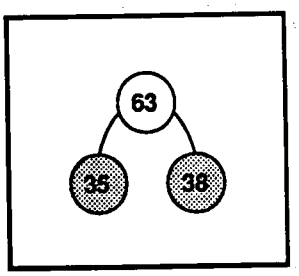


Figure 7.42. Hierarchical Structure of the Bolts.

structures. Figures 7.43 and 7.44 represent the background on the inside area surrounding the bolts. Figures 7.45 and 7.46 represent the area surrounding the drill key and Figures 7.47 and 7.48 represent a background area near the horizontal wrench.

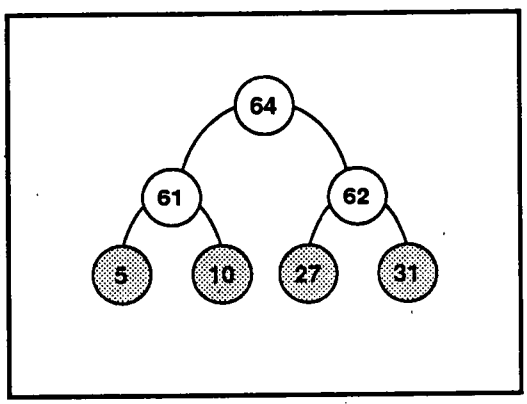


Figure 7.43. Hierarchical Structure of Inside Background Area.

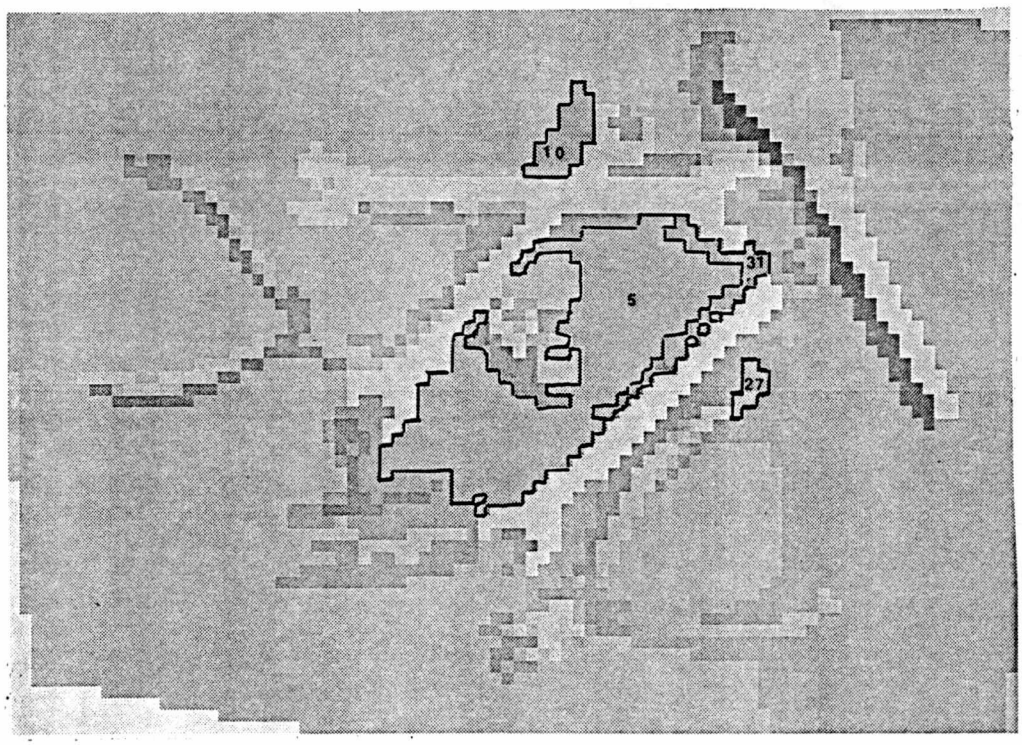


Figure 7.44. Labeled Regions of Inside Background Area.

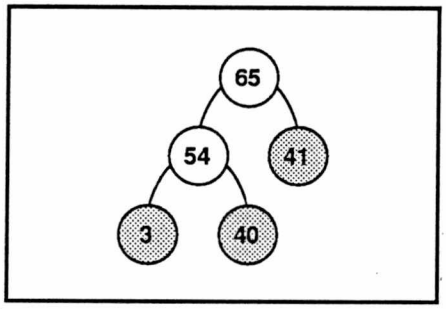


Figure 7.45. Hierarchical Structure of Shadow Area Surrounding the Drill Key.

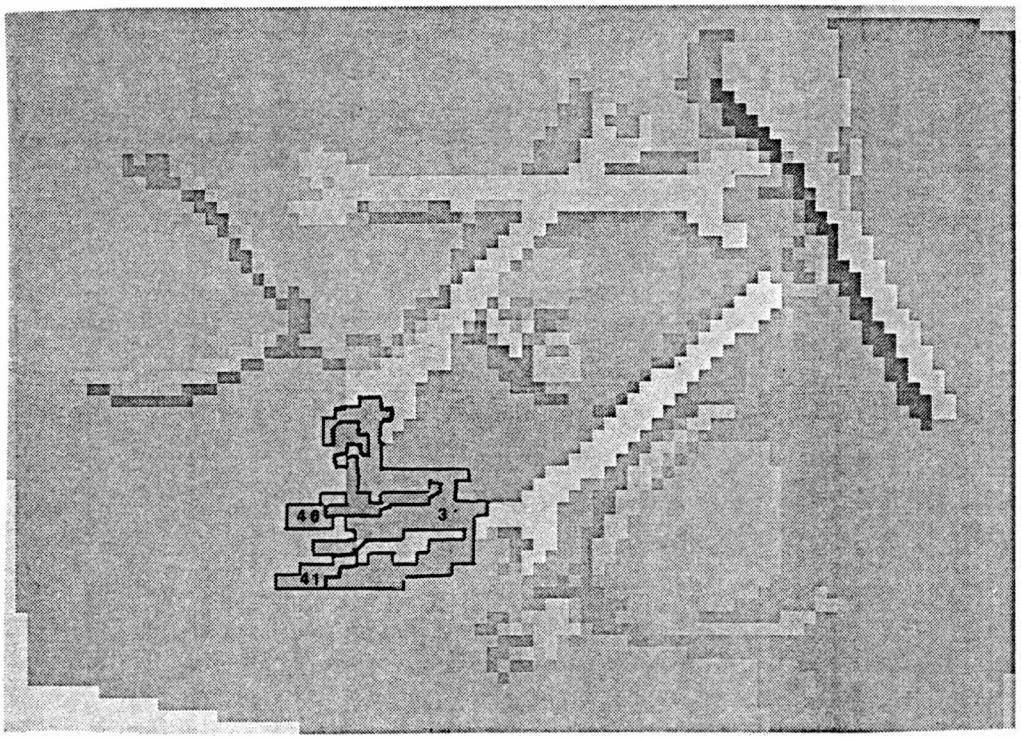


Figure 7.46. Labeled Regions of Shadow Structure Surrounding the Drill Key.

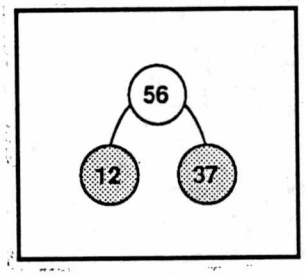


Figure 7.47. Hierarchical Structure of Background Near the Horizontal Wrench.

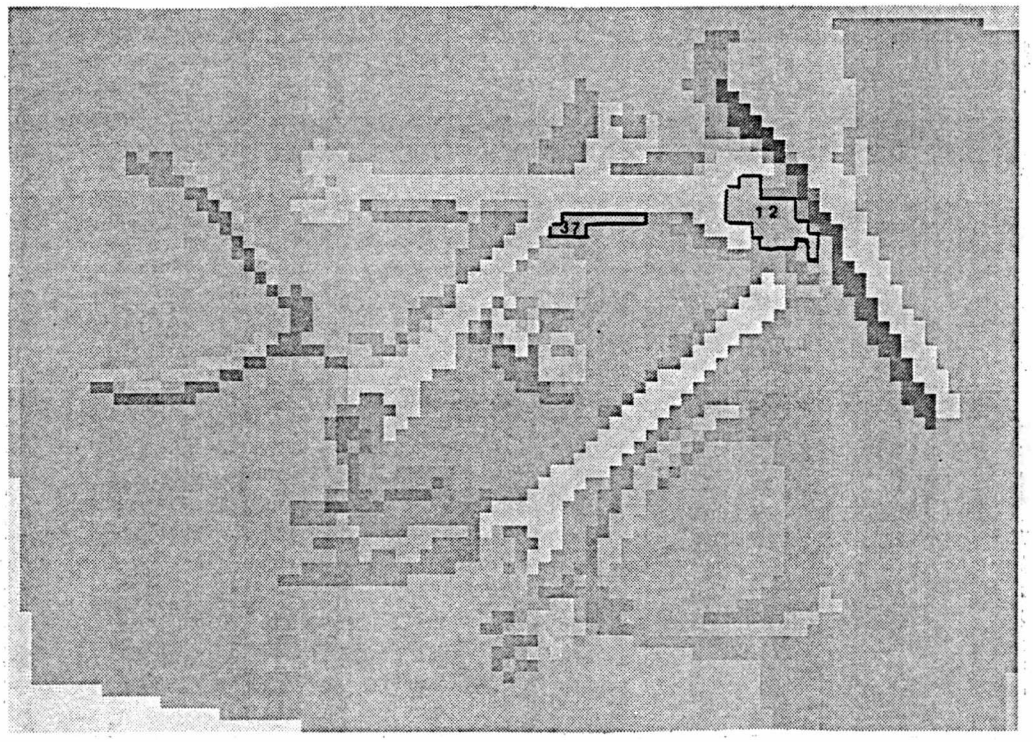


Figure 7.48. Labeled Regions of Background Structure Near the Horizontal Wrench.

Evaluation of Goals and Conclusions

The grouping program successfully joined the disjoint regions of the horizontal wrench and demonstrated its ability to join regions separated by occluding objects. It joined portions of the lower right pliers, but for reason discussed previously, all the regions of the tool were not grouped together.

In this example the region segmentation program did most of the work when it successfully formed regions representing whole tools. The only tools consisting of disjoint regions are the horizontal wrench and the right lower pliers. The program grouped what it could and left most of the tools alone as single regions. This is appropriate since these regions represent whole tools. In this sense, the grouping program was successful at what it did do and at what it did not do.

Example Three: Wrenches and Pliers Modified

In order to further test the grouping program's ability to rejoin disjoint tool regions, four overlapping wrenches were added to the image in the second example. The overlapping wrenches, shown in Figure 7.49, cause the tool regions to break up into disjoint segments. This tests the program's ability to produce several small structures representing whole tools.

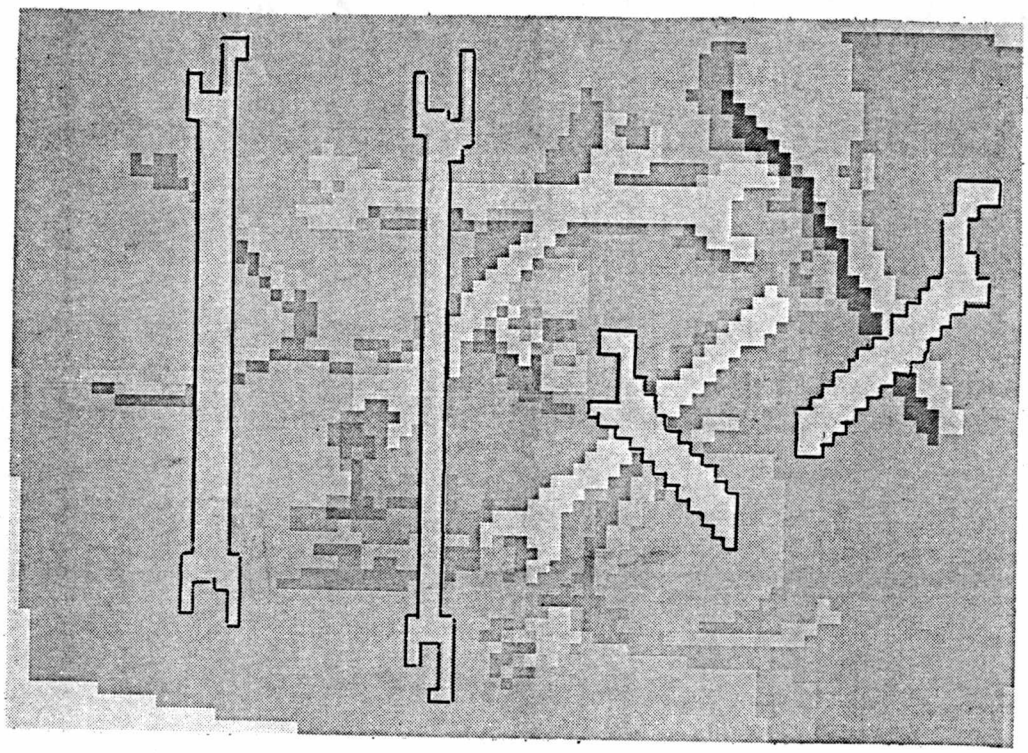


Figure 7.49. Image of Wrenches and Pliers with Four Added Overlapping Wrenches.

The grouping program generated 13 simple structures representing the separate tools in three iterations. This example demonstrates the versatility of the grouping program. It successfully forms numerous simple structures for images with many small objects, such as the tool structures in this example, and forms complex structures for images with few large objects, such as the tank structure in the first example. The following will discuss the processing of this image and the 13 structures created.

Detailed Description of the Processing

Preprocessing. The preprocessing of example three involves manually adding four new regions to the preprocessed image from example two. The four new regions represent the overlapping wrenches shown in Figure 7.49. Twenty-two new regions resulted from previously whole regions being split into two or more regions by the overlapping wrenches. The new regions were manually created wherever portions of a region were no longer contiguous, reflecting the requirement of the region segmentation program that regions be contiguous. The previously whole regions which split into multiple new regions are shown in Figures 7.50, 7.51 and 7.52.

Summary of three processing iterations. At each iteration or pass, the grouping program computes relationship values and forms groups based on these values. These groups create a portion of the final hierarchical structure.

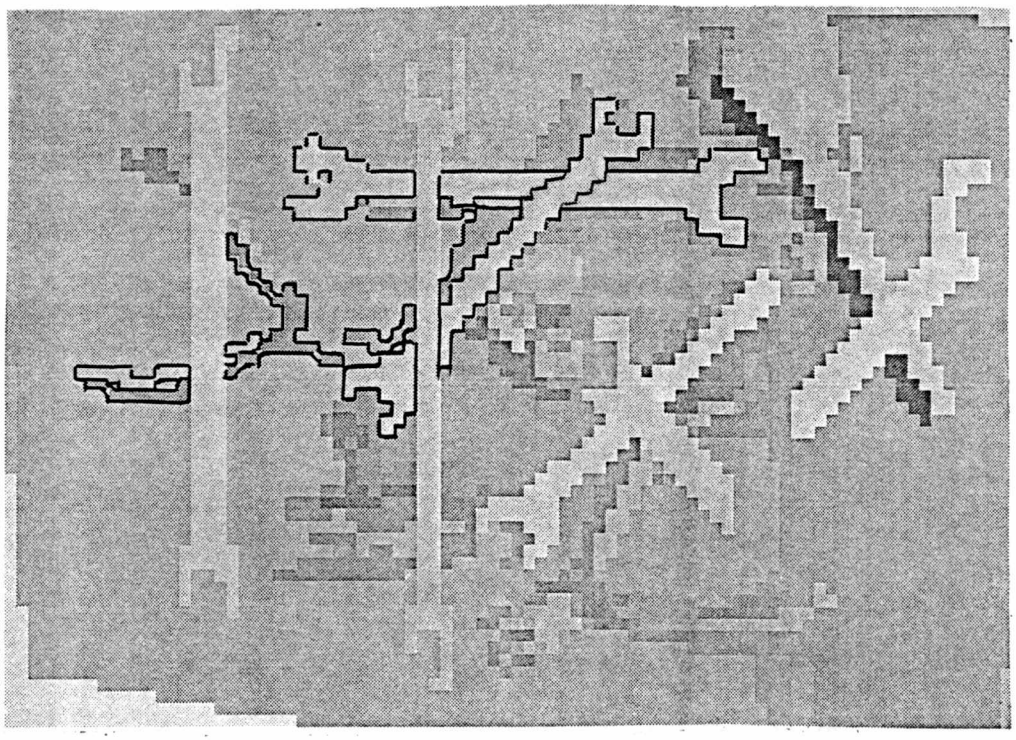


Figure 7.50. Regions Separated By Overlapping Wrenches - I.

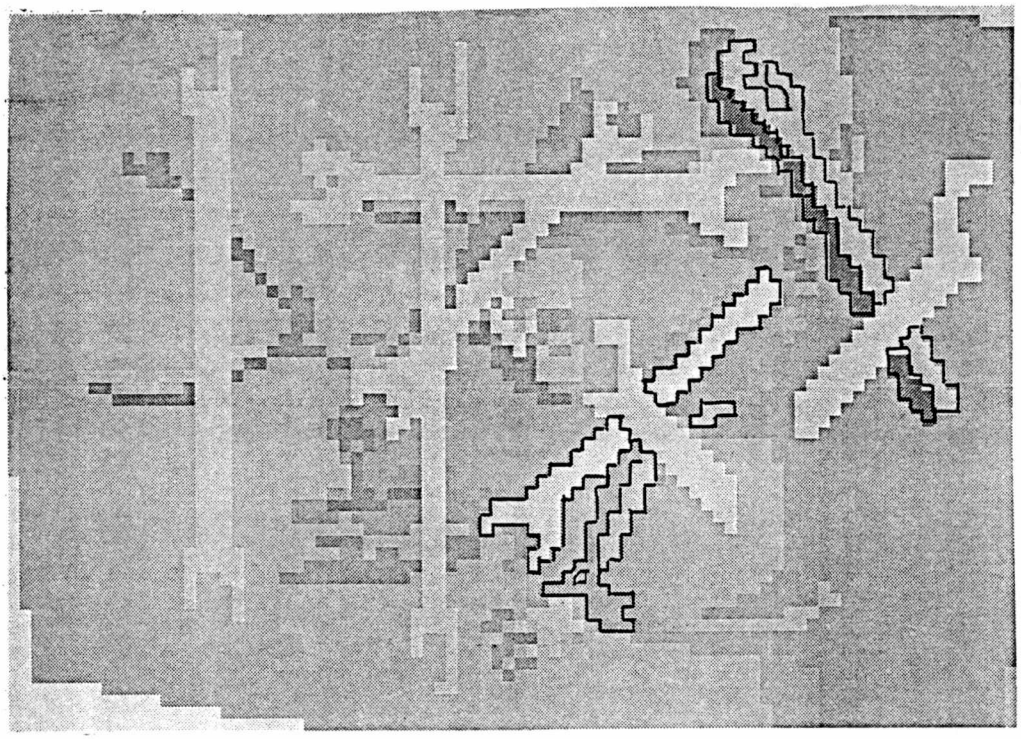


Figure 7.51 Regions Separated By Overlapping Wrenches - II.

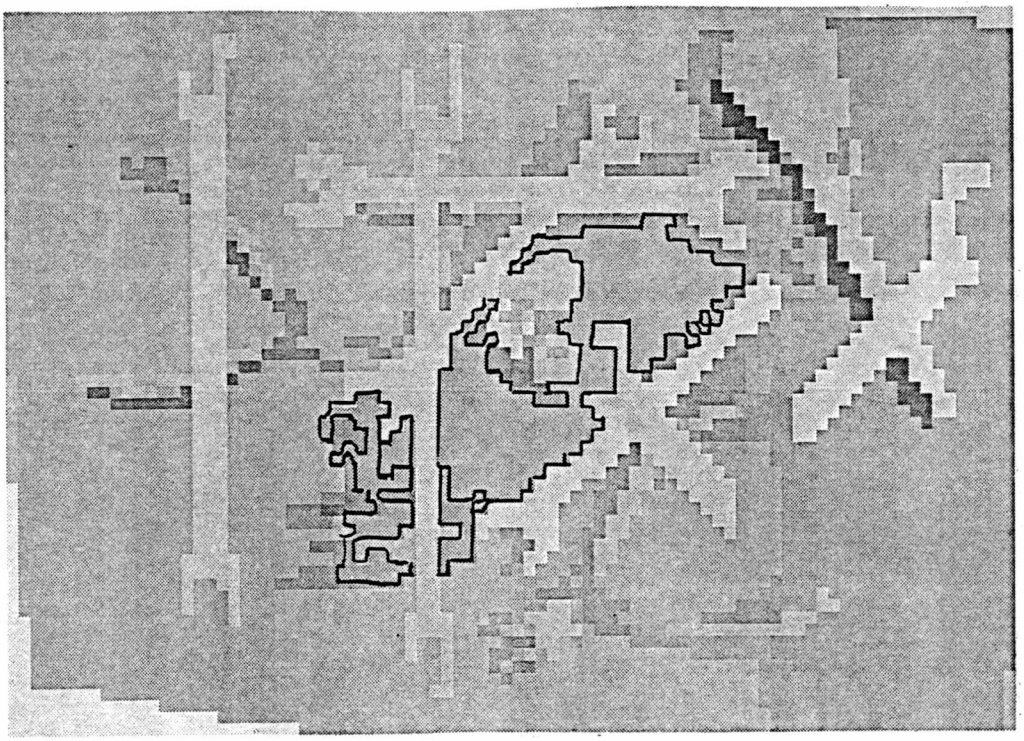


Figure 7.52. Regions Separated By Overlapping Wrenches - III.

The structures for this image formed in three iterations. The following discussion summarizes the processing of these iterations.

The first pass forms single tool structures for most of the disjoint tool regions. Regions with the strongest relationships group in the first pass and in this example all regions separated by overlapping tools grouped. The horizontal wrench and all three diagonal wrenches formed individual tool structures. Portions of the left pliers

grouped as did portions of the right pliers. Passes two and three group background regions, small shadow regions and two bolt regions.

Detailed Analysis of Results

The structures formed. The grouping program created thirteen structures representing the tools divided by the overlapping wrenches. The following describes these structures.

The wrench structures. The grouping program grouped the disjoint regions of all the wrenches in the first iteration because of their strong relationship values. These structures could be identified as wrenches because of their continuous boundary lines and bright intensity values. Figures 7.53 displays the structure of the horizontal wrench and Figure 7.54 identifies the regions forming this structure in the image. Figures 7.55-7.57 illustrate the structures of the remaining three wrenches, and Figure 7.58 identifies the regions forming the wrenches in the image.

The plier structures. Figure 7.59 diagrams the structure of the upper handle and nose of the left pliers. Figure 7.60 labels these regions in the image. Two regions forming the lower handle of the left pliers grouped together as shown in Figure 7.61 and labeled in Figure 7.62.

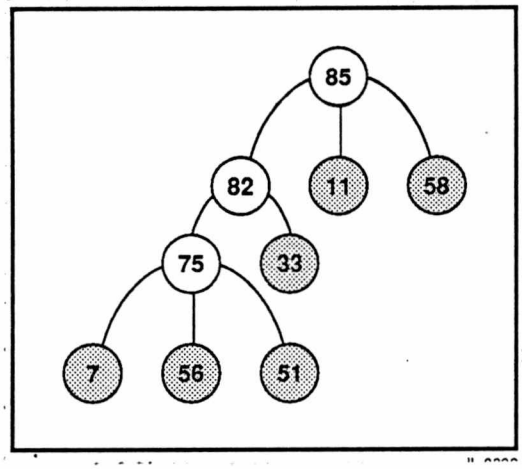


Figure 7.53. Hierarchical Structure of the Horizontal Wrench.

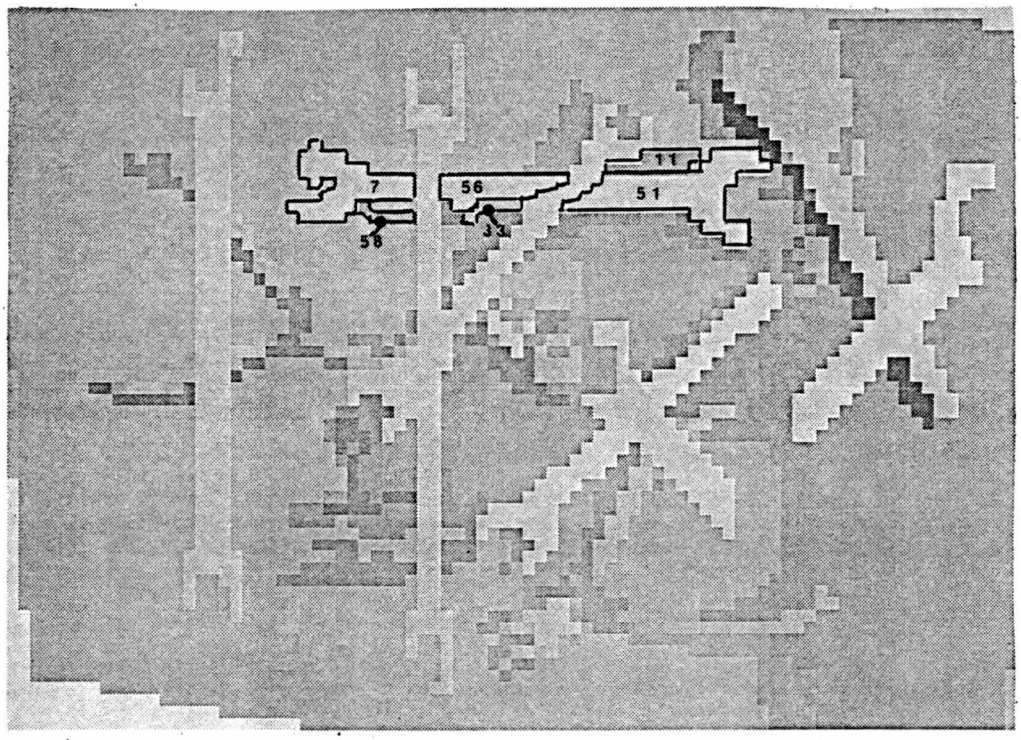


Figure 7.54. Labeled Regions of the Horizontal Wrench Structure.

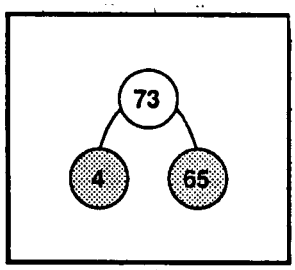


Figure 7.55. Hierarchical Structure of the Lower Diagonal Wrench.

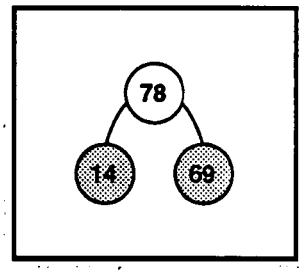


Figure 7.56. Hierarchical Structure of the Right Diagonal Wrench.

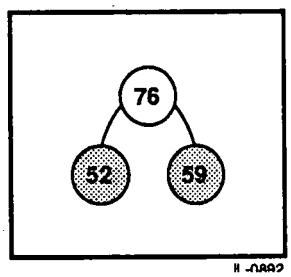


Figure 7.57. Hierarchical Structure of the Left Diagonal Wrench.

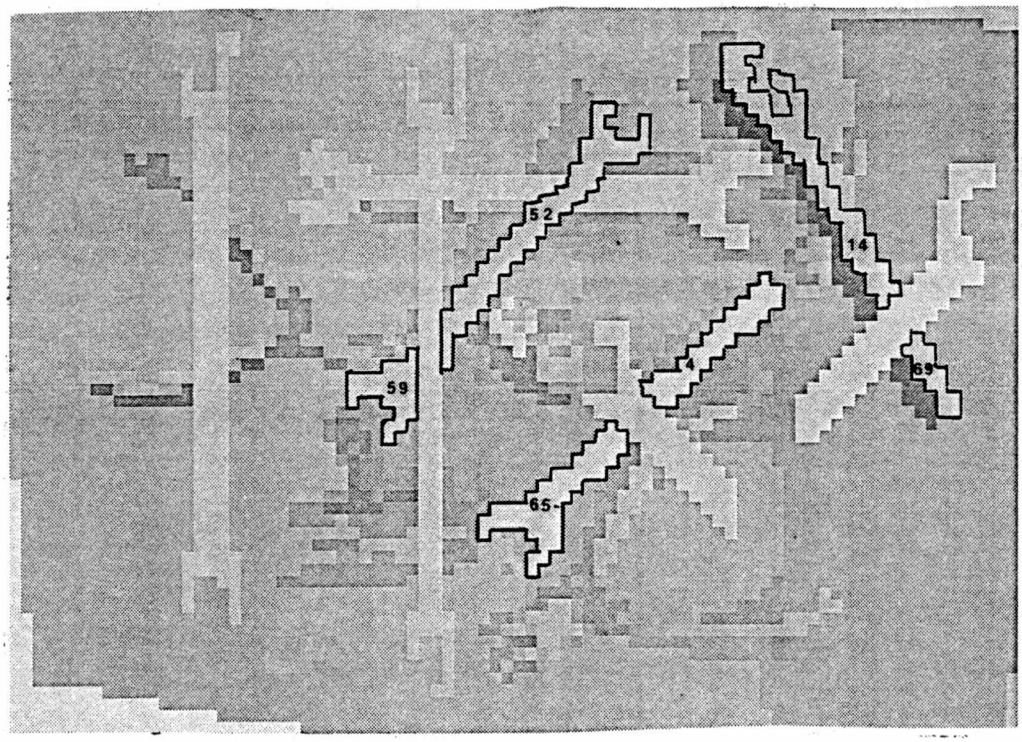


Figure 7.58. Labeled Regions of Three Wrenches.

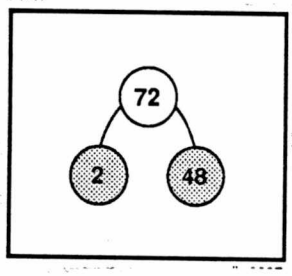


Figure 7.59. Hierarchical Structure of Left Pliers Upper Handle and Nose.

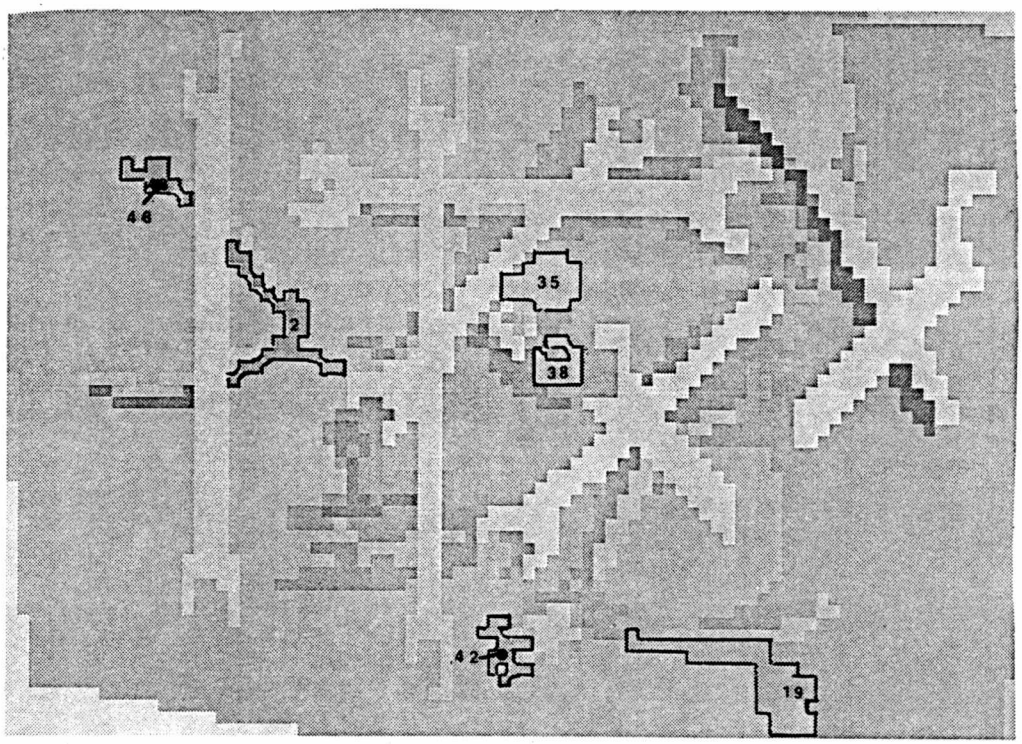


Figure 7.60. Labeled Regions of Left Pliers Upper Handle and Nose Structure- 72, Bolts Structure-87, and Right Lower Pliers Structure -80

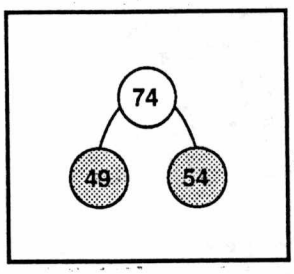


Figure 7.61. Hierarchical Structure of Lower Handle of Left Pliers.

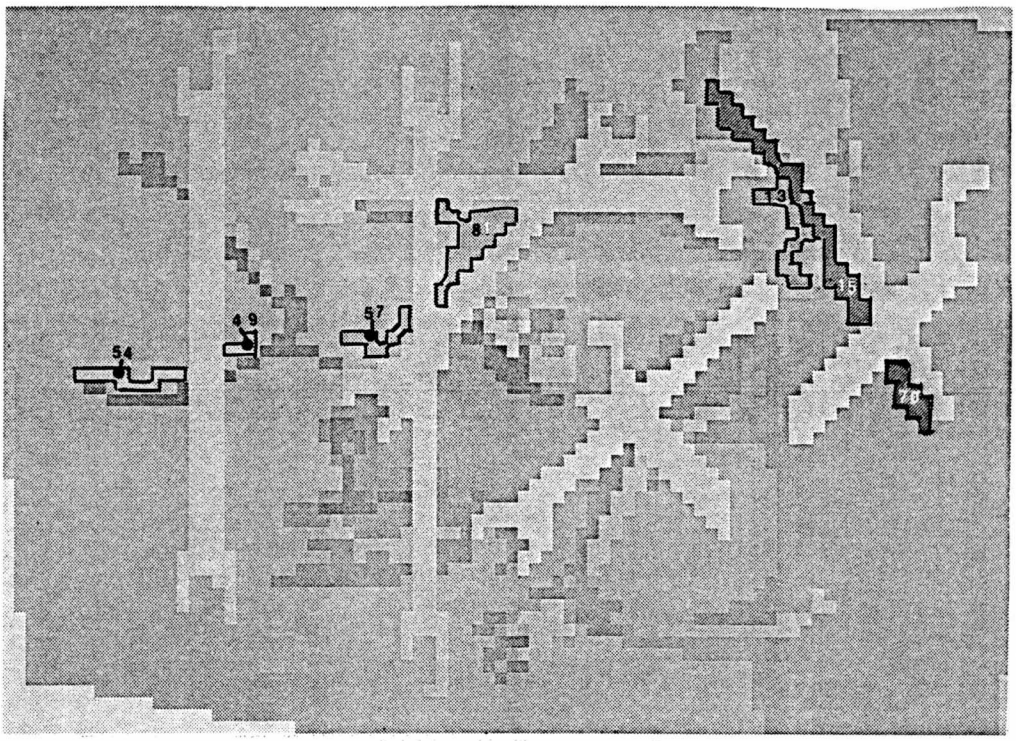


Figure 7.62. Labeled Regions of Lower Handle of Left Pliers Structure - 74, Shadow Along Left Diagonal Wrench Structure - 77, and Shadow Along Right Diagonal Wrench Structure - 84.

The upper handle and nose structure 72, and the lower handle structure 74, did not group together because the regions forming the pliers have weak relationship values. Similarly, the right lower pliers shown in Figure 7.63 and labeled in Figure 7.60, contains the lower handle and portions of the plier nose, but did not group with the upper handle and the bright portion of the plier nose.

The regions forming the pliers display different intensity values, orientation values and orientations of boundary lines. These differences produce weak relationship values and prevent these regions from grouping. The perceptual features of common endpoints and symmetry may be used by humans to group these tools. Incorporating these features into the grouping program would enhance the ability of the grouping program to join these regions and form a single structure for each tool. In this example, an object recognition system could use a feedback system to locate the missing regions of the pliers.

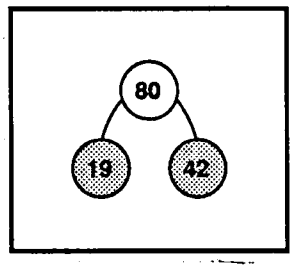


Figure 7.63. Hierarchical Structure of Right Lower Pliers.

The bolt structure, shadow structures and background structures. Figure 7.64 illustrates the structure containing the grouped bolt regions. Figure 7.60 labels these regions in the image. The brightness, small size and boundary lines would identify these regions as bolts rather than wrenches, pliers, or background.

Figures 7.65 and 7.66 show structures representing shadows of tools with regions separated by an overlapping wrench. Figure 7.62 identifies these regions in the image. An object recognition system could identify these regions as shadows by their dark intensities.

The structures shown in Figures 7.67, 7.68, and 7.69 represent background areas. Figures 7.70 and 7.71 label the regions forming these structures. The orientation, intensity and boundary lines of these regions indicate they are background.

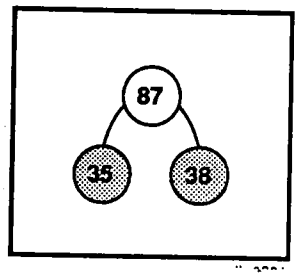


Figure 7.64. Hierarchical Structure of the Bolts.

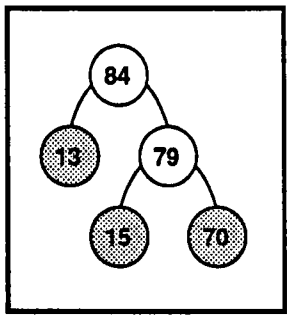


Figure 7.65. Hierarchical Structure of the Shadow Along the Right Diagonal Wrench.

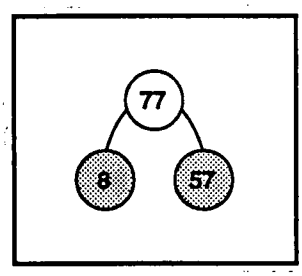


Figure 7.66. Hierarchical Structure of the Shadow Along the Left Diagonal Wrench.

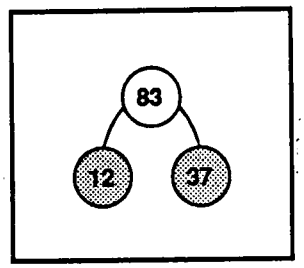


Figure 7.67. Hierarchical Structure of the Background Near the Horizontal Wrench.

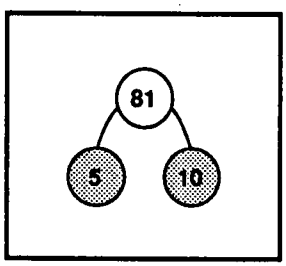


Figure 7.68. Hierarchical Structure of Corner Background.

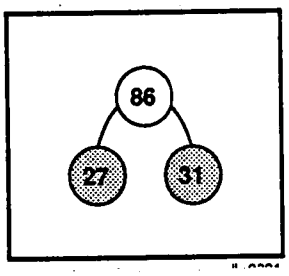


Figure 7.69. Hierarchical Structure of Inside Background.

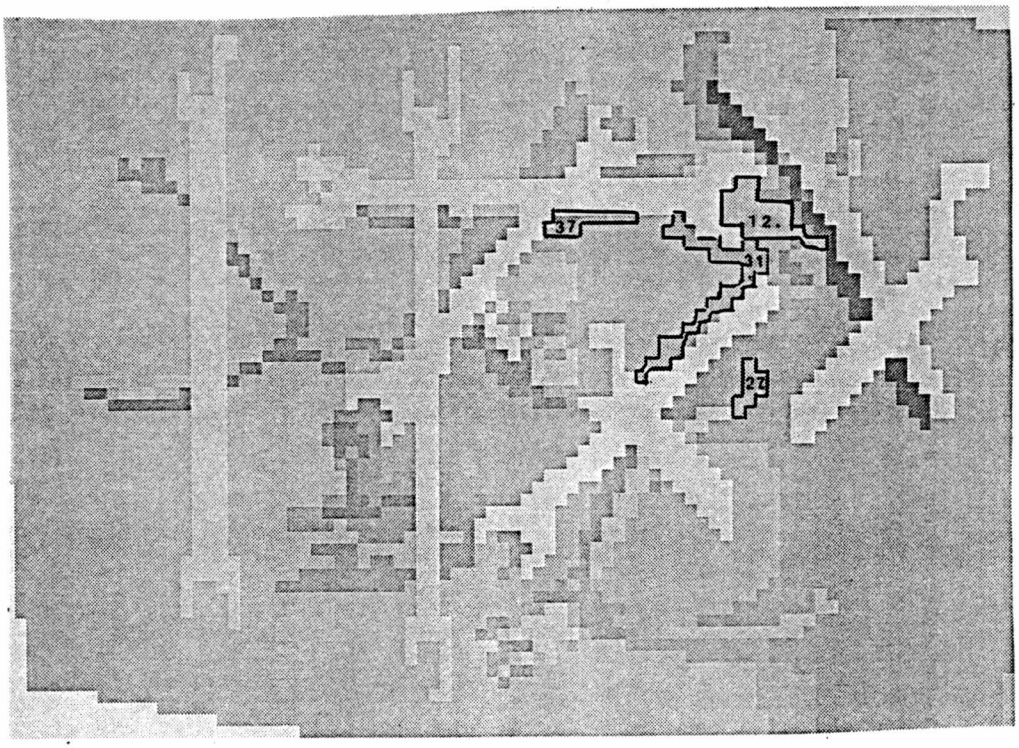


Figure 7.70. Labeled Regions of Background Structures.



Figure 7.71. Labeled Regions of Inside Background Structure.

Conclusions

The grouping program successfully joined the disjoint regions of all wrenches in the image and some but not all the regions corresponding to the pliers. This points to the strength of the grouping program when joining regions separated by an occluding object. However, more perceptual

features such as common endpoints and symmetry may be needed to group the diverse regions and form a single structure for the pliers.

The grouping program created numerous simple structures representing the tools in the image. These structures would enable an objects recognition system to identify many of the tools in the image. The perceptual features of intensity, orientation and continuous boundary lines could be used to identify the tools. Additionally, the perceptual features used to group regions could also be used, for example, in the case of the wrenches, grouped regions shared continuous boundary lines, similar intensity values and similar orientations.

This example shows the versatility of the grouping program and its ability to form simple structures for images with small objects. This compliments the ability of the grouping program to form complex structure for images with large complex objects as the tank in example one.

CHAPTER VIII

CONCLUSIONS AND RECOMMENDATIONS

An organized representation of the segmented scene improves the ability of machine vision systems to recognize objects and interpret scenes. This thesis presents a methodology for representing the organization of a scene. This methodology uses the perceptual features of similarity of intensity, similarity of orientation, continuation of boundary lines and proximity, to group and organize images. A cooperative/competitive process computes these four features in logical concurrence and produces perceptually-based hierarchical structures of the image.

Perceptual organization improves machine vision systems by overcoming many difficulties associated with segmentation such as occlusions, shadows and faulty segmentation. The examples presented in Chapter VII clearly demonstrate the ability of this methodology to join regions separated by occluding objects. In all three examples regions separated by occluding objects were joined in the first iteration of the grouping process and were computed to have the largest relationship values of any groups. In example one the tank regions separated by the large tree were joined together. In examples two and three the tool regions separated by

overlapping tools were joined together. Example one also demonstrated the grouping of related regions separated by faulty segmentation.

Perceptual organization enables machine vision systems to overcome problems associated with model-based object recognition. By matching the perceptual features and components of objects with those in a template, the system will be less sensitive to viewpoint differences and noisy images. The examples presented in Chapter VII demonstrate the ability of the grouping program to create structures using the perceptual features of objects. The program generated both complex and simple structures depending on the complexity of the objects.

In example one the tank is a large and complex object which contains many diverse regions. The program formed a tank structure by first grouping those regions which form the tank components. Many of the components were then grouped into a single tank structure which could be used by an object recognition system to identify the tank.

Examples two and three contrast with example one and contain numerous smaller objects with homogeneous regions. The objects are tools and the grouping program formed small simple structures. In most cases the tools remain as individual structures which would enable an object recognition system to identify them using their perceptual features. An exception occurred in example two when two overlapping

wrenches were joined into a single structure. This structure was comprised of two separate components, one for each wrench. In this image, these were the only overlapping tools and joining them into a single structure indicates the unique and strong relationship between these tools. Example three also contained overlapping tools, however, none of these tools joined together. This may be because the relationship was not unique.

Perceptual structures reduce the search time usually associated with model-based matching. If the features at the top level of a hierarchical structure do not match those of a template there is no need to search the detail of the structure. Perceptual structures also allow the identification of novel objects. Structures may be formed for objects or scenes without templates.

The perceptual organization of images is closer to human methods of perception than appearance based organizations. Studies by Kosslyn [1988] have demonstrated that humans use perceptual features and components to recall and visualize objects. Perceptual organizations may provide a common basis for human and machine interpretation of scenes and allow a more natural interface concerning mutually perceived images.

In order to produce a human-like scene structure the methodology must incorporate and accurately compute the significant perceptual features in an image. This points to

the need to include more perceptual features in the system. The second and third examples presented in Chapter VII, contained an object for which the grouping program was unable to form a single structure, a set of pliers. The regions forming the pliers have low relationship values for all the relationships studied in this grouping program. The inclusion of perceptual features such as common endpoints and symmetry would enhance the ability of the grouping program to form a single structure for objects such as the pliers. This inclusion of additional perceptual relationships is an area for future study.

The current results indicate this method will be useful for mediating the transition from low-level to high-level image representation. Such a transition would improve machine vision capabilities and benefit many applications. Some of these applications are:

- o autonomous vehicle operations, hierarchical structures would help match images to expected objects in structured conditions and in unstructured or field conditions the hierarchical structure would group and help identify related regions such as vegetation and water,

- automatic target recognition, the hierarchical structure will be valuable in distinguishing the regular structure of man-made objects from the irregular structure in nature, and
- automated robotic assembly and inspection, and
- medical imaging and remote sensing.

In summary, numerous application would benefit from more robust vision capabilities. The methodology presented in this thesis uses perceptual information to create an organized representation of an image. Such a representation will be a valuable tool in future vision systems and is one step towards more powerful vision capabilities.

CHAPTER IX

FUTURE RESEARCH

Future research on the cooperative/competitive grouping methodology should concentrate on three areas:

1. improving the grouping criteria,
2. improving the utility of the hierarchical structures produced, and
3. improving the implementation of the methodology.

The implementation of the methodology would be improved by using an enhanced region segmenter. The region segmentation algorithm used in this implementation considers only similarities of intensity when creating regions, orientations should also be considered. This would eliminate the need for manual separation of regions, and produce regions from pixels of similar intensity and orientation. As mentioned in Chapter II, these two similarities, intensity and orientation, appear to influence grouping in early vision processing.

The implementation of this methodology could also be improved by reducing the number of regions which are compared in levels two, three and four of the architecture. Currently all regions are compared to all other regions, resulting in $n(n-1)/2$ comparisons. During experimentation it was found that the proximity relationship allowed only regions within eight pixels to group. This knowledge could

be used to reduce the processing time by only comparing those region which are within eight pixels of each other in levels two, three and four of the architecture.

The threshold determination between the levels could also be improved. Currently a simple standard deviation measurement is used. Other threshold determining methods should be explored. One possible method involves the employment of inhibitory connections to eliminate all but the most distinct and related groups.

Another area for improvement includes the grouping criteria. In examples two and three discussed in Chapter VII, the regions comprising the plier tools were not grouped together. The ability of the methodology to group these regions would be enhanced if the perceptual features of common fate and symmetry were incorporated into the implementation. These features may be used by humans in grouping these regions. These features would also enhance the ability of the program to group the gun barrel with the tank in the first example discussed in Chapter VII.

As more perceptual features are added to the system, the need to weigh the features differently arises. Determining the weight for a particular feature may be an area for future research. One possible method involves determining which features provide the greatest discrimination.

The third area for improvement involves enhancing the hierarchical structures produced by the program. When

regions are compared, the orientation is measured on an absolute scale. When object descriptions are used, the orientation of the object components is generally in relative terms. Therefore, after the hierarchical structures are created, the relative orientation of each component could be computed. These orientations will be used to compare with the object descriptions.

Research in these areas would improve the grouping methodology. The implementation of the methodology and the grouping criteria would more closely reflect human vision processing. Then the hierarchical structures created could more easily be compared with object descriptions.

BIBLIOGRAPHY

BIBLIOGRAPHY

Arbib, M.A., & Hanson, A.R. "Overview." In M.A. Arbib & A.R. Hanson (Eds.), Vision, Brain, and Cooperative Computation Cambridge, M.A.; MIT Press, 1987), 1 - 83.

Barrow, H.G., & Tenenbaum, J.M. "Computational vision," Proc. of IEEE}, 69 (1981), 572 - 595.

Beck, J. "Similarity grouping and perceptual discriminability under uncertainty," Am. J. Psychology, 85 (1972), 1-19.

Beck, J.; Prazdny, K.; and Rosenfeld, A. "A theory of textural segmentation." In J. Beck, B. Hope, and A. Rosenfeld (Eds.), Human and Machine Vision (New York: Academic, 1983), 1-38.

Biederman, I., "Human image understanding; Recent Research and a Theory," Computer Vision and Graphics, 32 (Oct 1985), 29-73

Brooks, R.A. "Symbolic reasoning among 3-D models and 2-D images", Artificial Intelligence, 17 (1981), 285-348.

Brooks, R.A. "Model-based three dimensional interpretations of two-dimensional images," IEEE Trans. PAMI, 5} (1983), 140-150.

Brooks, R.A., Greiner, R., & Binford, T.O. "The ACRONYM model-based vision system," Proc. IJCAI-79, 105-113.

Cowan, C.K., Chelberg, D.M., & Lim, H.S. "ACRONYM model-based vision in the Intelligent Task Automation project," Proc. IEEE First Intl. Conf. on AI Applications (Denver, CO, Dec. 5-7, 1984), 176-183.

Garner, W.R. "Good patterns have few alternatives," American Scientist, 58 (1970), 34-42.

Grossberg, S., "Neural dyamics of perceptual grouping: Textures, boundaries, and emergent segmentation." In S. Grossberg (Ed.), The Adaptive Brain II. (Amsterdam: North Holland, 1987), 143-210. (a)

Grossberg, S., "Cortical dynamics of three-dimensional form, color, and brightness perception: I. Monocular theory," Perception & Psychophysics, 41 (1987), 87-116. (b)

Grossberg, S., "Cortical dynamics of three-dimensional form, color, and brightness perception: II. Binocular theory," Perception & Psychophysics, 41 (1987), 117-158. (c)

Havens, W, & Mackworth, A. "Representing knowledge of the visual world," Computer, (October 1983), 90-96.

Hubel, D.H. "Vision and the brain," Los Alamos Science (Los Alamos National Labs) 16 (1980), 15-35.

Intelligent Task Automation Progress Reports and Final Reports. Project conducted at SRI International, Honeywell SRC, and Stanford, under AFOSR DARPA Contract No.: F49620-83-C-0134.

Jain, R., & Grosky, W.I. "Hyper-pyramids for integration of spatial information," Proc. 1987 AAAI Workshop on Spatial Reasoning and Multi-Sensor Fusion (St. Charles, IL., Oct. 5-7, 1987), 72-81.

Julesz, B. "Experiments in the visual perceptual of texture," Scientific Am., 232 (1975), 34-43.

Kohler, W. Gestalt Psychology (New York: Liveright, 1947).

Koffka, K. Principles of Gestalt Psychology (New York: Harcourt, Brace and Co., 1935).

Kosslyn, S. "Aspects of a cognitive neuroscience of mental imagery," Science, 240 (1988), 1621-1626.

Lawton, D.T., & McConnell, C.C. "Perceptual organization using interestingness," Proc. 1987 AAAI Workshop on Spatial Reasoning and Multi-Sensor Fusion, (St. Charles, IL., Oct. 5-7, 1987), 405-419.

Lowe, D. G. Perceptual Organization and Visual Recognition, (Boston, MA: Kluwer, 1985).

Lowe, D. G. "Three-dimensional object recognition from single two-dimensional images," Artificial Intelligence, 31 (1987), 355-395.

Maren, A.J. "Hierarchical region structures: A method for structuring image regions," Proc. of SPIE Conference (Cambridge, MA: September 16-20, 1985), 189-195.

Maren, A.J., & Ali, M. "Hierarchical scene structures to facilitate improved image understanding," Proc. of the First International Conference of Industrial and Engineering Applications of Artificial Intelligence and Expert Systems. (Tullahoma, TN, June 1-3, 1988), 657-667.

Maren, A.J., Minsky, V., & Ali, M. "A multilayer cooperative/competitive method for creating hierarchical structures by clustering maximally-related nodes," Proc. Second IEEE Intl. Conf. on Neural Networks, (San Diego, CA; July 24-27, 1988)

Marr, D. Vision (San Francisco: W.H. Freeman and Co., 1982).

Minsky, V. & Maren, A. "A Multilayered Cooperative/Competitive Neural Network for Segmented Scene Analysis." Jr. of Neural Network Computing, 3 (1990), 14-32.

Olson, R. K. & Attneave, F. "What variables produce similarity grouping?," Am. Journal Psych., 83 (1970), 1-21.

Palmer, S. "The psychology of perceptual organization: A transformational approach." In J. Beck, B. Hope, and A. Rosenfeld (Eds.), Human and Machine Vision (New York: Academic, 1983).

Pentland, A.P. "Representation of natural form," Artificial Intelligence, 28 (1986), 293-331.

Pomerantz, J.R., "Visual form perception: An overview." In E. Schwab, and H. Nusbaum (Eds.), Pattern Recognition by Humans and Machines, Volume 2, Visual Perception. (Academic, 1986), 1-30.

Prinzmetal, W. "Principles of feature integration in visual perception," Perception & Psychophysics, 30 (1981), 330-340.

Rearick, T. C., & Frawley, J. L., The Generalized Cylinder Picture Element, Lockheed-Georgia Company, LG87RR004, May, 1987.

Rosenfeld, A. "Image analysis: problems, progress, and prospects," Pattern Recognition, 17 (1984), 3-12.

Rosenfeld, A. "Divide-and-conquer techniques for biological vision." [Extended abstract.] Proc. AAAI Symposium on Physical and Biological Approaches to Computational Vision (Stanford, CA: March 22-24, 1988).

Rosenfeld, A. "Recognizing unexpected objects: A proposed approach." [Extended abstract.] Proc. of AAAI Spring Symposium Series on Physical and Biological Approaches to Computational Vision (Stanford, CA: March 22-24, 1988, 49-49a.

Rosenfeld, A., & Klette, R. "Degree of adjacency or surroundedness," Pattern Recognition, 18 (1985), 169-177.

Salton, G., & McGill, M.J. Introduction to Modern Information Retrieval (New York: McGraw-Hill Co., 1983).

Samet, H., "Hierarchical Representations of Collections of Small Rectangles," ACM Computing Surveys, 20 (December 1988), 271- 309.

Sejnowski, T.J., Koch, C., & Churchland, P.S., "Computational Neuroscience," Science, 241 (1988), 1299-1306.

Treisman, A., "Features and objects in visual processing", Sci. Am., 255 (Nov. 1986), 114-125.

Treisman, A., & Paterson, R. "Emergent features, attention, and object perception, J. Psych: Human Perception & Performance, 10 (1984), 12-31.

Tsotsos, John K., "Knowledge and the visual process: Content, form, and use," Pattern Rec., 17, (1984), 13-27.

Vernon, M. A Further Study of Visual Perception (Cambridge: Cambridge U. Press, 1952).

von Grunau, M.W., Wiggin, S., & Reed, M. "The local character of perspective organization," Perception & Psychophysics, 35, (1984), 319-324.

Walters, D.K.W. "The use of Natural Contraints in Image Segmentation", SPIE Applications of Artificial Intelligence II 548, (1985), 27-34

Wertheimer, M. "Principles of perceptual organization." In D. C. Beardslee (Ed.), Readings in Perception, (New York: Van Nostrand, 1958), 115-135.

Witkin, A. P., & Tenenbaum, J. M. "On the role of structure in machine vision." In J. Beck, B. Hope, and A. Rosenfeld (Eds.), Human and Machine Vision (New York: Academic, 1983).

Zucker, S. W. "The diversity of perceptual grouping." In M. A. Arbib & A.R. Hanson (Eds.), Vision, Brain, and Cooperative Computation (Cambridge, MA: MIT, 1987).

VITA

My professional goal is to research issues in artificial intelligence. My interest in this area stems from a belief that artificial intelligence and robotics will provide the next major advancements in computer technology. In order to achieve my goal, I have been working on a Master of Science degree in computer science with an emphasis in artificial intelligence and computer vision at The University of Tennessee Space Institute (UTSI). My GPA is 3.95.

I have been employed as a graduate research assistant at UTSI researching a relatively new area in artificial intelligence, neural networks, and their application to problems in computer vision. Publications in this area include: a journal article in the Journal of Neural Network Computing, (Winter 1990), entitled "Representing the perceptual organization of segmented images through hierarchical scene structures", and a paper in the Proceedings of the Second IEEE International Conference on Neural Networks, (San Diego, CA; July 24-27, 1988), entitled "A multilayer cooperative/competitive method for creating hierarchical structures by clustering maximally-related nodes". Awards include a third place award in the ACM Mid-Southeast Conference (Gatlinburg, TN; November 10-11, 1988) for presentation of a paper entitled "A multilayer cooperative/competitive method for identifying the most perceptually distinctive clusters in an image".

Prior to attending UTSI, I worked for six years in various computer related positions and developed an academic background in computer science. The positions I held included computer operator, programmer, systems analyst and manager of information systems services. The most recent employment was at Baptist Hospital in Nashville, TN, as a manager of information systems services. I developed an academic background in computer science by attending computer science courses at The Ohio State University in Columbus, OH. and The Middle Tennessee State University in Murfreesboro, TN. My GPA at these institution is 3.95. I received a Bachelor of Arts degree in 1978 with a major in Linguistics from The University of Michigan, Ann Arbor, MI.

I was born and grew up not to far from Ann Arbor, in Birmingham, MI, a suburb of Detroit. I am 32 years old and currently enjoy living in Nashville, TN with my husband and son.