This is to certify that the

dissertation entitled

A Framework for Texture Analysis
Based on Spatial Filtering

presented by

James Michael Coggins

has been accepted towards fulfillment
of the requirements for

Ph. D. degree in Computer Science

Anil Kumar Jain
Major professor

Date 12/6/82

A FRAMEWORK FOR TEXTURE ANALYSIS BASED ON SPATIAL FILTERING

By

James Michael Coggins

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Computer Science

1982

ABSTRACT

A FRAMEWORK FOR TEXTURE ANALYSIS BASED ON SPATIAL FILTERING

By

James Michael Coggins

A texture analysis method motivated by a theory of human visual information processing and based on spatial filtering is defined and evaluated for classification and segmentation of textured images.

The problem of texture analysis is viewed as an attempt to duplicate the texture analysis performance of human vision. This performance is considered to be a consequence of certain information reductions (filtering) performed in early stages of vision which are modelled by a sequence of filters defined in the spatial frequency domain. The implementation of the model results in a sequence of spatial domain filtered images which contain limited spectral information from the original image.

Features which are interpreted as measurements of average local energy are defined and evaluated in texture classification experiments. The energy features are found to outperform power spectral features; this is attributed to the use of phase information in the filtered images. This contrasts with previous studies in which phase was assumed to be unimportant for texture analysis. The channel filtering features are found to be insensitive to global, constant gray level changes, making some preprocessing operations unnecessary. Procedures are demonstrated for determining that two images portray the same texture at different magnifications or orientations.

A method for computing a texture feature vector for a neighborhood about each pixel in the image is defined and demonstrated in texture segmentation experiments. An image is segmented by identifying clusters in the feature space and labelling each pixel by the cluster in which its feature vector lies. A cluster validity statistic is used to determine an appropriate number of clusters. The results indicate that the channel filtering feature space is an appropriate representation of image texture for segmentation.

TABLE OF CONTENTS

List of Tables v

List of Figures viii

Acknowledgements x

1 Introduction

1.1 Computational Vision 1

1.2 Texture 2

1.3 Guidance from Studies of Human Vision 3

1.4 Spatial Filtering 5

1.5 Organization of This Dissertation 6

2 Toward a Definition of "Texture"

2.1 Introduction 8

2.2 Studies of Human Texture Perception 9

2.3 Properties of Texture and of Texture Perception 12

2.4 Comments on the Catalog of Texture Definitions 15

2.5 Summary 16

3 Models and Methods in Texture Analysis

3.1 Introduction 17

3.2 Local Analysis Methods 17

3.3 Global Analysis Methods 21

3.4 Intermediate Analysis Methods 26

3.5 Image Modelling Methods 29

3.6 Summary 31

4 A Texture Analysis Method Based on a Theory of Human Vision

4.1 Introduction 33

4.2 A Critical Re-evaluation of Texture Analysis Methodology . 33

4.3 Spatial Frequency Channels 35

4.4 Constraints on Possible Texture Features 44

4.5 Properties of Channel-Filtered Images 47

4.6 Definition of Texture Features 48

4.7 How Will the Texture Features Be Used? 51

4.8 Summary 52

5 Evaluation of Channel Filtering Features for Texture Classification

5.1 Introduction 53

5.2 Evaluations of the Features on Natural Images 55

5.2.1 Experiment 1: Natural Images, 64x64 Subimages 58

5.2.2 Experiment 2: Natural Images, 32x32 Subimages 62

5.2.3 Experiment 3: Natural Images, 16x16 Subimages 64

5.2.4 Summary 65

5.3 The Effect of Histogram Equalization 65

5.3.1 Experiment 4: Histogram Equalized Images, 64x64 Subimages.69

5.3.2 Experiment 5: Histogram Equalized Images, 32x32 Subimages.72

5.3.3 Summary 75

5.4 Experiment 6: Uniform Gray Level Changes 76

5.5 Experiment 7: Magnification Changes 79

5.6 Experiment 8: Orientation Changes 82

5.7 Experiment 9: Phase Spectrum Changes 86

5.8 Second-Order Statistics and Channel Filtering 92

5.8.1 Experiment 10: Application of the Co-occurrence Method.95

5.8.2 Experiment 11: Application of Channel Filtering 101

5.8.3 Summary 103
5.9 Computational Simplifications 104
5.9.1 Experiment 12: Using Fewer Channels 104
5.9.2 Experiment 13: Ideal Bandpass Channels 105
5.10 Summary 107
6 Evaluation of Channel Filtering Features for Texture Segmentation
6.1 Segmentation 109
6.2 Computing Texture Features for Segmentation 110
6.3 Segmentation Using Feature Images 113
6.4 Evaluating the Segmentations 115
6.5 Segmentation Experiment 1: Dot Textures 116
6.6 Segmentation Experiment 2: Gaussian White Noise 123
6.7 Segmentation Experiment 3: Natural Image Composite 124
6.8 Segmentation Experiment 4: SCRE Image 136
6.9 Summary 136
7 Summary and Conclusions
7.1 Summary 143
7.2 Conclusions 144
7.2.1 Classification 144
7.2.2 Segmentation 146
7.2.3 General 147
7.3 Advantages of the Channel Filtering Method 148
7.4 Disadvantages of the Channel Filtering Method 149
7.5 Suggestions for Further Research 150
Appendix A: A Catalog of Texture Definitions 152
Appendix B: Definition of Channel Filters 156
List of References 158

List of Tables

Table 1: Classification of Eight Natural Image Classes Using 64x64 Subimages 59

Table 2: Classification of Eight Natural Image Classes Using 32x32 Subimages 63

Table 3: Classification of Four Natural Image Classes Using 16x16 Subimages 65

Table 4: Classification of Eight Histogram-Equalized Natural Image Classes Using 64x64 Subimages 70

Table 5: Classification of TILE, ROCK, SAND, PAPE and TILQ, ROCQ, SANQ, PAPQ Using 64x64 Subimages 71

Table 6: Classification of CORK, GRAS, WOOD, SCRE and CORQ, GRAQ, WOOQ, SCRQ Using 64x64 Subimages 72

Table 7: Classification of Eight Histogram-Equalized Image Classes Using 32x32 Subimages 73

Table 8: Classification of TILE, ROCK, SAND, PAPE and TILQ, ROCQ, SANQ, PAPQ Using 32x32 Subimages 74

Table 9: Classification of CORK, GRAS, WOOD, SCRE and CORQ, GRAQ, WOOQ, SCRQ Using 32x32 Subimages 75

Table 10: Classification Results Showing the Effect of Average Gray Level Changes 78

Table 11: Size Change Experiment Using 6 Spatial Frequency Channels 80

Table 12: Size Change Experiment Using 5 Spatial Frequency Channels 82

Table 13: Orientation Experiment Using No Channel Shifting 85

Table 14: Orientation Experiment Using Circular Channel Shifting 86

Table 15: Classification Results on Four Phase-Modified Gaussian White Noise Classes 90

Table 16: Classification Results Illustrating the Effect of Phase Spectrum Changes on Gaussian White Noise Images 91

Table 17: Phase Modification Experiment with Indiscriminable Classes Merged 92

Table 18: Black-Black Displacement Vectors in Four-Dot Micropatterns 96

Table 19: Computation of Expected Number of Black-Black Co-occurrences in 256x256 Four-Dot Images 98

Table 20: Classification of Four-dot Textures Using Co-occurrence Matrices 100

Table 21: Classification of Four-dot Textures by Channel Filtering 101

Table 22: Classification of Four-dot Textures Using Six Spatial Frequency Channels 102

Table 23: Classification of Four-dot Textures Using Four Orientation Channels 103

Table 24: Classification Results on Natural Images Using Eight Channels 106

Table 25: Classification Results on Natural Images Using Ideal Bandpass Channels 107

Table 26: Evaluation of Clustering on Dots Image Using 8x8 Averaging Windows 122

Table 27: Evaluation of Clustering on Dots Image Using 16x16 Averaging Windows 123

Table 28: Evaluation of Clustering on Gaussian White Noise Image Using 8x8 Averaging Windows 131

Table 29: Evaluation of Clustering on Gaussian White Noise Image Using 16x16 Averaging Windows 132

Table 30: Evaluation of Clustering on Composite Natural Image Using 8x8 Averaging Windows 134

Table 31: Evaluation of Clustering on Composite Natural Image Using 16x16 Averaging Windows 135

Table 32: Evaluation of Clustering on SCRE Image Using 8x8 Averaging Windows 140

Table 33: Evaluation of Clustering on SCRE Image Using 16x16 Averaging Windows 141

LIST OF FIGURES

Figure 1: Sinusoidal Gratings 24

Figure 2: Diagram of the Spatial Filtering Procedure 38

Figure 3: A Spatial Frequency Channel Filter 39

Figure 4: An Orientation Channel Filter 40

Figure 5: An Example of Channel Filtering 42

Figure 6: Indiscriminable Textures which Could Be Discriminated Based on the Edge Structure of the Micropatterns 45

Figure 7: Texture Feature Definitions 50

Figure 8: Natural Image Classes 56

Figure 9: Mean Values of Feature 1 in Each Channel for the Eight Natural Image Classes 61

Figure 10: Histogram-Equalized Natural Images 67

Figure 11: Gaussian White Noise Images 77

Figure 12: Bar Images for Orientation Experiment 84

Figure 13: Phase Spectrum Modification Procedure 88

Figure 14: Effect of Phase Modification on GWN1 89

Figure 15: Four-Dot Micropatterns and Corresponding Image Classes 93

Figure 16: Diagram of the Segmentation Procedure 111

Figure 17: Dots Image for Segmentation Experiment 1 117

Figure 18: Segmented Images for Dots Using 8x8 Averaging Windows 119

Figure 19: Segmented Images for Dots Using 16x16 Averaging Windows 120

Figure 20: Gaussian White Noise Image for Segmentation Experiment 2 125

Figure 21: Segmented Images for Gaussian White Noise Using 8x8 Averaging Windows 126

Figure 22: Segmented Images for Gaussian White Noise Using 16x16 Averaging Windows 127

Figure 23: Natural Image Composite for Segmentation Experiment 3 128

Figure 24: Segmented Images for Natural Image Composite Using 8x8 Averaging Windows 129

Figure 25: Segmented Images for Natural Image Composite Using 16x16 Averaging Windows 130

Figure 26: SCRE Image for Segmentation Experiment 4 137

Figure 27: Segmented Images for SCRE Using 8x8 Averaging Windows 138

Figure 28: Segmented Images for SCRE Using 16x16 Averaging Windows 139

ACKNOWLEDGEMENTS

I would like to thank Dr. Anil K. Jain for his guidance, patience, and persistence in his role as advisor, critic, editor, and taskmaster. His careful editing has contributed greatly to the quality of this document and his critical scrutiny of my work has kept me out of a lot of trouble. I am also grateful to Major Arthur Ginsburg, whose filtering model of human vision forms the foundation on which my work is built. In addition to teaching me about the spatial filtering model, psychophysics, and scientific writing, Maj. Ginsburg also served on my doctoral committee. I would also like to thank the other members of the committee, Dr. Richard Dubes, Dr. Carl Page, and Dr. James Zacks for their contributions throughout my research.

Several persons have provided software tools without which some aspects of my work would have been nearly impossible. These include Larry Johnson, George Cross, Ken Dimoff, Steve Smith, Neal Wyse, and Gautam Biswas.

I would also like to thank Dr. Harry Hedges for his support, and Ardeshir Goshtasby, John Piippo, Steve Smith, and Steve Fullenkamp for their various contributions.

Chapter 1

Introduction

1.1 Computational Vision

Computer applications involving visual input include problems in aerial image interpretation [Darling and Joseph, 1968; Kettig and Landgrebe, 1976; Landgrebe, 1981], biomedical image analysis [Bacus, 1976; Pressman, 1976; Hall et al, 1977; Landeweerd and Gelsema, 1977; Mui et al, 1977; Jain et al, 1980; Trussel, 1981], industrial processes [Perkins, 1978; Tennenbaum et al, 1978; Agin, 1980], and robotics and scene analysis [Duda and Hart, 1973; Marr and Nishihara, 1978; Barrow and Tennenbaum, 1981; Stevens, 1980; Aggarwal et al, 1981]. Hardware for acquiring, storing, and displaying images is available, and current computers are capable of performing some image analysis tasks in real-time. The development of algorithms for processing images has proven to be a slow and difficult task [Gurari and Wechsler, 1982]. Image analysis methods have been developed in the fields of pattern recognition [Duda and Hart, 1973; Fu, 1974, 1977; Pavlidis, 1977], image processing [Pratt, 1978; Rosenfeld and Kak, 1981] and artificial intelligence [Winston, 1977; Nilsson, 1980]. The inquiries in these three fields involving approaches to automatic image analysis, interpretation, and understanding are sometimes referred to collectively as "computational vision" [Brady, 1982; Barrow and Tennenbaum, 1981].

In order to cope with the variety and complexity of activities associated with vision, several subproblems within computational vision have been identified including image enhancement, edge detection, segmentation, image registration, and texture analysis. These subproblems have been attacked separately and have tended to evolve into subfields themselves. This has the advantage that many different techniques with different characteristics are available as tools for use in applications. The separation of the subproblems has the disadvantage that similarities in the solutions to different problems may be overlooked and that diverse solutions to the subproblems of vision complicate the construction of unified computational vision systems.

1.2 Texture

Human observers are capable of some image segmentation and discrimination tasks under conditions (such as brief exposure to a test image) which prevent detailed scrutiny of the image. This ability is referred to as "effortless" or "pre-attentive" visual discrimination. When an image does not portray any particular object or form, only certain aspects of the overall pattern of gray level changes in the image is effortlessly perceived. "Texture" sometimes refers to the pattern of gray level variations produced by some image generation procedure. However, different procedures may yield images which are not effortlessly discriminable to human observers. In this thesis, two images which do not portray particular objects or forms will be considered to have the same "texture" if they are not effortlessly discriminable to human observers.

Texture is recognized as being fundamental to the perception of regions and surfaces in images [Brady, 1982; Stevens, 1980]. Textural information can potentially be used by automatic vision systems in region identification, image segmentation and classification tasks. Texture analysis is a major component in discussions of general computer vision systems [Sklansky, 1978; Wechsler, 1980; Barrow and Tennenbaum, 1981; Brady, 1982].

The potential importance of texture for automatic vision systems has inspired many attempts to develop algorithms for texture analysis. Unfortunately, the characterization of texture in terms of human performance does not suggest a simple measurement on images which can duplicate human texture perception. This lack of precise guidance has led to a proliferation of ad hoc texture analysis methods based on diverse mathematical, statistical and heuristic measurements on images. Most texture analysis methods duplicate human performance reasonably well for some image classes, but they fail to duplicate human performance in more general problems.

1.3 Guidance from Studies of Human Vision

Since a texture analysis algorithm is supposed to duplicate the performance of human vision, it seems reasonable to look to vision science for guidance in developing such algorithms. Computational methods need not emulate the human visual system, but knowledge of the limitations and strategies present in human vision could guide the development of useful algorithms.

Some texture analysis algorithms have been guided by results of psychophysical experiments designed to find an upper bound on the complexity of texture perception. The nature of texture is described by specifying a statistical property which indiscriminable textured images appear to have in common. This type of description is intended to constrain the nature of texture by a criterion which is independent of human performance, but the generality or minimality of such a criterion is difficult to establish. Such results have guided many texture analysis studies, but bounds on the complexity of indistinguishable stimuli do not necessarily suggest a particular algorithm which will reproduce human performance.

While algorithms for computational vision need not emulate human vision at the neural level, knowledge of the overall strategies used by the visual system for analyzing images could provide more definite guidance for developing algorithms. We will use one such theory in this dissertation to motivate the development of a new approach to texture analysis.

This theory characterizes early stages of the human visual system as being composed of quasi-independent mechanisms, called channels, which decompose an image into certain bands of spatial frequency and orientation [Ginsburg, 1971]. This decomposition is modelled by a spatial filtering operation using filters defined in the spatial frequency domain [Ginsburg, 1978, 1980b]. The result of the decomposition is a sequence of filtered images in the spatial domain which contain limited spectral information from the original image.

The channel filtering theory is not simply a generalization of human visual behavior; it is an attempt to describe the information content and possible processing strategies in early human vision. The theory implies that many aspects of visual perception are consequences of the information reduction (filtering) performed by the perceptual system [Ginsburg, 1978, 1980b]. Thus, in order to characterize texture perception, it may be most effective to determine first how the human visual system filters the information in an image. The particular decomposition of an image performed in early stages of vision can explain why certain stimuli are discriminable and others are confused.

1.4 Spatial Filtering

The channel filtering theory asserts that the early processing of an image by the human visual system is effectively modelled by a spatial filtering operation. Spatial filtering is a technique which has been useful in several areas of computational vision. Two equivalent implementations of spatial filtering exist; one involves convolution in the spatial domain, and the other involves multiplication in the spatial frequency domain. The use of spatial filtering as a computational tool for image analysis is attractive because of its well-known mathematical basis, because of the availability of efficient algorithms for performing the Fourier transform, and because of the existence of intuitively satisfying interpretations for the Fourier transform and for spatial filtering.

Some uses for spatial filtering in computational vision are two-dimensional extensions of well-known one-dimensional operations

[Papoulis, 1962]. Low-pass filtering has been used to remove noise from an image [Rosenfeld and Kak, 1981]. High-pass filtering has been applied to edge detection, and in fact, a certain filtering approach has been shown to be an "optimal" edge detection method [Shanmugam et al, 1979]. Band-pass filtering has been applied to form and edge detection [Ginsburg, 1976, 1978, 1979a, 1980a, 1980b; Crowley and Parker, 1978; Marr et al, 1979; Marr and Hildreth, 1980]. Spatial filtering was also used in an implementation of a theory of stereo vision which involved operations similar to those required in image registration [Marr and Poggio, 1979]. Several approaches to texture analysis, such as those based on edge detection or template matching can be implemented using spatial filtering operations [Laws, 1980].

In this study, a computational vision system will be defined by specifying its filtering properties in the spatial frequency domain. The filter characteristics to be used are adopted from studies of human vision, but this study will not attempt to determine in detail the correspondence between human vision and the computational vision system. Methods for texture classification and segmentation based on the computational vision system will then be defined and evaluated.

1.5 Organization of This Dissertation

This dissertation will begin by reviewing the results of previous research in texture analysis. Chapter 2 reviews attempts to define or at least to delimit the nature of "texture". This review includes results from vision science and principles derived from human intuition and experience with texture perception. In Chapter 3, texture analysis algorithms proposed in the literature will be reviewed. Chapter 4 will present a new texture analysis method based on spatial filtering. Chapters 5 and 6 present the results of experiments which evaluate the proposed method for classifying and segmenting textured images. The conclusions of this study will be summarized and further research will be suggested in Chapter 7.

Chapter 2

Toward a Definition of "Texture"

2.1 Introduction

In order to use textural information, the computer vision system requires an operational definition of "texture". In the absence of a sufficiently precise general definition, the operational definition of texture is implicitly supplied in the features computed from an image. This feature space then constitutes a model of "texture".

How can an operational definition for texture be constructed? A reasonable first step is to observe the texture analysis performance of the human visual system [Haralick et al, 1973; Barrow and Tennenbaum, 1981]. Texture is a common and important aspect of human visual perception. Everyday experiences with texture analysis and recognition tasks provide an extensive background of intuition regarding the nature of "texture". One common manifestation of this intuition is in the adjectives we use to describe textures. For computational texture analysis, however, the nature of texture must be precisely quantified. Texture analysis has sometimes been approached as the problem of quantifying our intuition and our vocabulary about texture.

Some early attempts to construct automatic texture analysis procedures referenced certain results from vision science which appeared to limit the complexity of human texture perception. Textured areas which did not differ in certain simple statistical measurements were found to be indiscriminable to human observers. These psychophysical results provided a useful upper bound on the complexity of texture in human vision but little guidance concerning what approaches should be effective for computational purposes. Section 2.2 reviews these influential findings and recent extensions and modifications to them.

Several intuitive properties of texture and of texture perception were identified, and image analysis techniques were proposed which seemed to measure aspects of an image which were relevant to those properties. These results provide some valid insights into the nature of texture, though they actually tell more about how textures are perceived than about what texture is or how to compute texture. A summary of these results is presented in Section 2.3.

The intuitive insights into the nature of texture have not produced a characterization from which simple computational procedures can be developed. Researchers attempting to present new texture analysis techniques have found it awkward and difficult to describe the properties of images which their techniques are supposed to measure. Section 2.4 will discuss the Catalog of Texture Definitions in Appendix A.

2.2 Studies of Human Texture Perception

Experiments by Julesz [Julesz, 1962, 1965, 1975; Julesz et al, 1973] have influenced the development of texture analysis methodology.

In these experiments, human observers are presented an image composed of subimages which are generated by different rules. The subject's task is to find the different subimages. The composite image is displayed to the observer for very brief periods (100 ms) in order to prevent the observer from scrutinizing the image to find the different areas. Since the test images typically contained many repetitions of some micropattern and since the segmentation was to be performed without scrutinizing the image, the subimages were considered to have different "textures" and the experiments were interpreted as tests of pre-attentive human texture analysis ability. The tests would demonstrate the complexity of the processing performed in the visual system before detailed analysis involving cognition or memory could begin.

Notice that in such experiments, including [Julesz, 1962, 1965, 1975; Julesz et al, 1973; Pratt et al, 1978; Richards and Polit, 1974], "texture" is treated as an intrinsic property of an image determined by the image generation procedure. Pre-attentive human vision is then tested, and can succeed or fail in the texture discrimination task. The results of these tests were examined to find some properties of the generated images which could predict whether the human visual system could preattentively discriminate different subimages.

One generalization of the results of these preattentive discrimination experiments became known as the "Julesz Conjecture". This conjecture asserted that image areas which do not differ in their second-order gray level joint probability distributions cannot be preattentively discriminated by human observers. The conjecture was seen to imply that computational methods no more complex than the computation of second-order gray level distributions should be sufficient to rival human performance in texture discrimination tasks. Texture analysis researchers used this conjecture as a justification for limiting the complexity of their techniques to fairly simple computations.

Unfortunately, the Julesz Conjecture was wrong. Later studies [Caelli and Julesz, 1978a, 1978b; Julez et al, 1978; Julesz, 1981; Gagalowicz, 1981] have produced many examples of images which have identical second- and even higher-order gray level joint probability distributions but which are visually discriminable.

Revisions to the original Julesz Conjecture have been attempted. One recent attempt [Julesz, 1981] involves the assumption of special geometrical structures called "textons" which, allegedly, are detected in early stages of vision. The textons involve local geometrical structures and include "corner", "closure", and "connectivity". Detection of these special features is assumed to enable discrimination of textures which have identical second-order distributions.

Another recent paper [Gagalowicz, 1981] modifies the original conjecture by suggesting that the actual co-occurrence matrices (an estimate of the second-order gray level distribution; see Section 3.2) derived from the given image should be used to characterize textural properties rather than the underlying probability distribution. The argument is that two discriminable texture fields can be produced from the same probability distribution if the generating process is not ergodic [Papoulis, 1965]. In such a case, the random differences between the images due to the non-ergodicity of the generating procedure can be a textural discriminating factor.

In spite of the problems which have been discovered with texture characterization based on second-order distributions, methods based on the Julesz Conjecture are still used in texture analysis applications. One study [Pratt et al, 1978] attempts to characterize the cases when the conjecture fails and when the conjecture works. They conclude that the Julesz conjecture is a reasonable approximation to human performance in many texture discrimination problems.

2.3 Properties of Texture and of Texture Perception

Introspective observation of human texture perception has provided some insight into the nature of texture. These observations do not constitute a definition of texture; instead, they serve to guide and constrain the development of computational texture analysis methods.

The use of constraints in the absence of precise definitions is a basic approach in artificial intelligence (AI) research. As computational vision problems (such as texture analysis) have been found to resist attempts to define fundamental terms, the AI approach of finding and exploiting constraints has been adopted [Zucker, 1981].

The intuitive guidance for texture analysis methods can be summarized in four principles (cf. [Beck, 1980]).

1. Texture is a property of areas; the texture of a point is undefined. Thus, operations on a random sampling of pixels would not be an appropriate texture analysis procedure since texture does not exist in context-free intensities. Texture is a global property of a region, but the region under consideration can be a fairly small subimage of a larger scene. Analysis of textures over small areas

occurs when one attempts to find a boundary between two textured areas [Haralick, 1975; Thompson, 1977].

2. Texture involves the spatial distribution of gray levels throughout a region. This spatial property could be characterized by statistical features computed over regions or by maintaining position (or relative position) information throughout processing. A one-dimensional gray level histogram is not, by itself, an appropriate texture analysis tool because it captures no spatial information. Two-dimensional histograms, or co-occurrence matrices (Section 3.2), are more reasonable texture analysis tools because a spatial parameter is incorporated in the histogram computation.

The minimum size of an image which adequately characterizes a texture depends on several factors including the apparent size of "objects" in the image, the resolution of the image acquisition system, and the size of the "operators" used in the texture analysis procedure. The perception of texture depends on the assumption of a frame of reference [Haralick, 1979] governing the sizes of gray level changes which are to be considered significant and the spatial scale which is to be operated upon. Since texture information can be found at many different scales, recent papers have advocated multi-level or hierarchical descriptions of regions for texture analysis [McCormick and Jayaramamurthy, 1975; Crowley and Parker, 1978; Ehrich and Foith, 1978; Tomita, 1981; Zucker and Kant, 1981]. Alternatively, the frame(s) of reference can be adopted from human performance [Ginsburg and Coggins, 1981].

3. Texture is perceived in image regions which contain many equally significant gray level changes. This implies that the gray

levels in a "textured" region have low redundancy and high information content [Resnikoff, 1981]. A textured region can be created by inserting large numbers of "objects" such as edges, regions of constant gray level, or micropatterns [Hall et al, 1977]. Alternatively, texture can be created by removing all "enumerable objects" [Richards and Polit, 1974]. In either case, a "texture" is perceived when significant individual forms are not present.

4. Textural properties of regions are invariant through moderate changes in overall brightness, orientation, and size (as in magnification/shrinkage). While these changes are observable, an original texture and a modified version of the same texture are recognizable as being samples of the same texture. (Some interesting points about textural invariance are made in [Modestino et al, 1981; Ginsburg and Coggins, 1981].)

Another potential source of guidance is in the perceived qualities of texture. Textural qualities are typically expressed by adjectives such as coarse, streaked, sharp, irregular, fine, cellular, rippled, directional, etc. One texture study [Tamura et al, 1978] investigated rank correlations between human evaluations of natural textures and statistical features designed specifically to quantify textural adjectives. Six perceptual dimensions of texture were specified, as follows: coarseness (coarse vs. fine), contrast (high contrast vs. low contrast), directionality (directional vs. non-directional), line-likeness (line-like vs. blob-like), regularity (regular vs. irregular) and roughness (rough vs. smooth). In psychophysical experiments, subjects evaluated 16 natural textures from [Brodatz, 1966] along the six dimensions. Correlations among these six features indicated that the features were not independent. In fact, the features cluster into two groups [Ginsburg and Coggins, 1981]. One group (coarseness, contrast, and roughness) appears to be affected by the apparent size of the texture while the other group (directionality, line-likeness, and regularity) appears to be affected by directional dependencies in the image. The clustering of the subjective features suggests that more general or fundamental descriptions of textured images may exist. Statistical features designed to duplicate the human evaluations of image textures along the six intuitive dimensions yielded poor results.

2.4 Comments on the Catalog of Texture Definitions

Appendix A contains a collection of attempts by texture analysis researchers to define "texture". It should be noted that many papers dealing with methods for quantifying visual texture do not even attempt to define the concept of "texture". The selections in the appendix are typical of the definitions which do appear.

Usually, attempts to define "texture" are either constructed specifically for a particular texture analysis approach [Appendix A, items 1, 3, 4], or they are so general that they are of little practical value [item 2]. Some "definitions" simply characterize "texture" by certain aspects of the human perception of textures [items 4-6].

Texture can be characterized as a global property [items 5, 7], a local property [item 2], a random phenomenon [item 3], a non-random phenomenon [item 9], a property of a region which remains when all "objects" are removed [item 6], a property of a region when many "objects" are present [item 5], a property determined by structural arrangements of objects [items 1, 4, 5, 9], a property determined by statistical distributions [items 3, 7].

Unfortunately, these attempts to define texture have not suggested simple computational texture measurements [item 10].

2.5 Summary

Texture is a commonly perceived quality of regions and surfaces, but no precise, general definition of texture exists. Texture analysis research has been guided by some results concerning the complexity of texture analysis in human vision and by intuitive properties of texture. These results serve to constrain the nature of texture, but not to define texture or to suggest what computational approaches might be appropriate for texture analysis. An attempt to directly quantify the textural qualities perceived by human observers yielded poor results. Due to the lack of more precise guidance, awkward and even contradictory "operational definitions" of texture have appeared in the literature.

Chapter 3

Models and Methods in Texture Analysis

3.1 Introduction

The process of analysis or measurement of texture is inseparable from the process of creating an operational definition of texture. The features computed from an image and the decision procedures applied to the features constitute a working definition of "texture". Existing reviews of texture analysis methods [Hawkins, 1969; Rosenfeld and Troy, 1974; Sklansky, 1978; Haralick, 1979; Wechsler, 1980] concentrate on the mathematical or computational techniques on which the methods are based. This review will organize the methods according to the nature of the operational definition of "texture" implied by the methods.

3.2 Local Analysis Methods

The unifying aspect of texture analysis methods reviewed in this section is a dependence on small groups of neighboring pixels. This local information may then be averaged or accumulated over a region for use in characterizing the texture of the region.

This approach is typical of many "statistical" texture analysis procedures. Among these are the gray level run length method

[Galloway, 1975], the gray level difference method [Weszka et al, 1976] and the gray level co-occurrence method [Haralick et al, 1973]. All three methods involve counting the occurrences of a simple local property over the entire region. In each case, the local property is easy to identify and each instance of the property involves only a few pixels. Since the local property depends on the gray levels of the pixels, accumulations are stored in a matrix whose size depends on the number of gray levels available in the quantization procedure. Statistics are then computed from the matrix for use as texture features.

In the gray level run length method, the local property is the number of linearly adjacent pixels with a specific gray level. The matrix Rd(i,j) gives the number of runs of length j of pixels with gray level i in direction d. Four such matrices can be computed, for d=0, 45, 90, and 135 degrees.

In the gray level difference method, the local property is the absolute difference between the gray levels of pixels at a specified displacement d=(dx,dy) from each other. The matrix Gd(i) gives the number of times that a gray level difference of i occurs between pixels at displacement d from each other.

In the gray level co-occurrence method, the local property consists of the gray levels of two pixels with a given displacement d=(dx,dy). The matrix Hd(i,j) gives the number of occurrences of a pixel with gray level i at displacement d from a pixel with gray level j.

The displacement vectors for the difference and co-occurrence methods could be any vectors which can occur within the image. In practice, however, fairly small vectors ($|d|<10$) are used. Since the co-occurrence matrix is sensitive to gray level changes over a certain distance, the use of small displacement vectors involves an assumption that texture exists in the gray level distributions in local areas of an image.

In the co-occurrence method especially, the matrices for several displacement vectors can be added together to remove certain orientation dependencies. For example, Haralick uses a definition which involves the sum of the co-occurrence matrices for vectors (dx,dy) and (-dx,-dy) [Haralick, 1979]. This particular definition results in a symmetric co-occurrence matrix.

In comparative evaluation studies [Weszka et al, 1976; Conners and Harlow, 1980a] involving these texture analysis methods, the co-occurrence method was found to be superior, with the gray level difference method a close second. These comparative studies have influenced the frequent use of the co-occurrence method in applications.

Some of the problems associated with local analysis methods are as follows:

(1) The choice of the displacement vectors is critical.

(2) The visual significance of the features computed from the matrices is sometimes difficult if not impossible to understand. This difficulty has led to attempts to capture textural adjectives directly in statistical features [Tamura et al, 1978; Conners, 1979].

(3) The local analysis methods are sensitive to "noise" due to their dependence on actual gray level values. In addition, the gray scale resolution determines the size of the accumulation matrices

independent of the image size, making comparison of textural properties of images digitized under different circumstances difficult.

(4) Local analysis methods are insensitive to global aspects of the image such as brightness gradients and directional tendencies.

Another type of local analysis method attempts to characterize the pattern of gray levels encountered in one-dimensional scans of the image. This scan can be analyzed as a one-dimensional function using time series techniques [McCormick and Jayaramamurthy, 1974; Deguchi and Morishita, 1978] or heuristic analysis of extrema in the function [Ehrich and Foith, 1978; Mitchell et al, 1977; Mitchell and Carlton, 1978]. In the time series analysis method, autoregression coefficients are the texture features. This method has been used for synthesis of certain types of streaked textures, but its applicability to less directional or less regular textures is questionable [Haralick, 1979]. Heuristic analysis of the one-dimensional scan involves identifying gray level extrema and associated properties such as the height and width of the peaks and the distance to the next higher peak. The texture features computed from such methods include the average height (contrast) of peaks and the density of peaks of a particular height.

The density of edges in a local region can also be used as a texture feature [Rosenfeld and Thurston, 1971; Rosenfeld et al, 1972; Rosenfeld and Troy, 1974]. A simple local edge operator such as the gradient can be used to identify edge pixels. The gradient computation can be adjusted to be a function of distance to obtain a hierarchical characterization of edge densities in the image.

Edge detection can be expressed as a spatial filtering operation by defining a set of templates in the spatial domain and convolving them with the image. One such template is the Sobel operator [Pratt, 1978]. Laws [1980] uses several templates (including Sobel operators and gradients) to detect edges at different orientations and with different contrasts.

3.3 Global Analysis Methods

Global tendencies in an image such as the average sizes of objects or areas or directional preferences are difficult to capture using local analysis methods. Since such global information seems to be related to texture (in particular, to coarseness and directionality), methods which describe global properties of the image could be useful in texture analysis.

Global size and directional tendency information can be derived from the autocorrelation function [Hayes et al, 1974]. Let $I(x,y)$ be the image function which gives the gray level at position (x,y) for $0 \le x,y \le N-1$ and is 0 otherwise. The values of $I(x,y)$ are integers with $0 \le I(x,y) \le G-1$. The (normalized) autocorrelation function $R(dx,dy)$ is the product of the image function with a shifted [by $d=(dx,dy)$] copy of itself. That is,

$$R(dx,dy) = \frac{1}{R0} \left[\sum_{x=0}^{N-1} \sum_{y=0}^{N-1} I(x,y) * I(x+dx,y+dy) \right]$$

where R0 is a normalizing factor. The maximum value of $R(dx,dy)$ occurs at $R(0,0)$. If the image contains large areas of constant gray level, the autocorrelation function will decrease slowly with distance from (0,0); if the image contains mostly small areas of constant gray level (the texture is "busy") then the autocorrelation function will drop off sharply. If the image is periodic, the autocorrelation function will rise and fall with the same period as the pattern in the image.

The value of R(dx,dy) is related to the co-occurrence matrix Hd(g1,g2) where d=(dx,dy) as follows:

$$R(dx,dy) = \frac{1}{R0} \left[\sum_{g1=0}^{G-1} \sum_{g2=0}^{G-1} g1*g2*Hd(g1,g2) \right]$$

Thus, the autocorrelation function can be derived from the entire ensemble of N**2 co-occurrence matrices of the image, but the converse is not true. For a given displacement, the co-occurrence matrix provides a more detailed characterization of the spatial distribution of gray levels than the autocorrelation at the same displacement.

The Fourier transform has been used more frequently than the autocorrelation function as a texture analysis tool [Lendaris and Stanley, 1970; Bajcsy, 1973; Bajcsy and Lieberman, 1976; Weszka et al, 1976; Conners and Harlow, 1980a; D'Astous and Jernigan, 1981; Eklundh, 1979]. The two-dimensional Discrete Fourier Transform (DFT) of an NxN image $\mathcal{I}(x,y)$ is defined as follows:

$$F(u,v) = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} \mathcal{I}(x,y) \exp\left[-j2\pi \frac{(ux+vy)}{N}\right] .$$

The Fast Fourier Transform algorithm can be used to compute the two-dimensional DFT [Johnson and Jain, 1981]. The Fourier spectrum can also be expressed in magnitude-phase form as

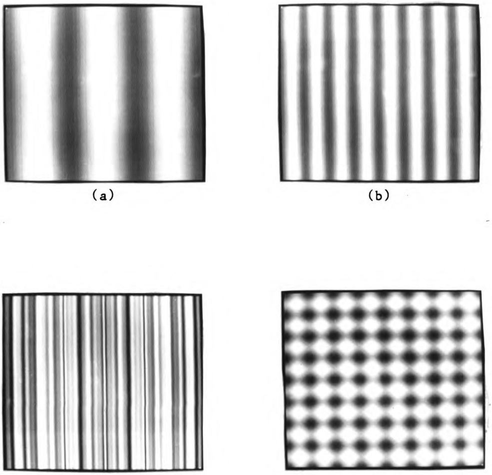
$$F(u,v) = M(u,v)\ \exp[jP(u,v)]\ .$$

An intuitively appealing interpretation of the Fourier transform is based on the representation of an image as a weighted sum of sinusoidal gratings (Figure 1) [Rosenfeld and Kak, 1981; Duda and Hart, 1973]. The parameters u and v determine the frequencies of the horizontal and vertical sine waves in the gratings. The amplitude of the sine waves (the contrast of the grating) is given by $M(u,v)$, and their phase is given by $P(u,v)$.

We note that the autocorrelation function and the power spectrum ($M(u,v)^2$) are a Fourier transform pair. Thus, by the argument given earlier, the ensemble of N^2 possible co-occurrence matrices also determines the magnitude spectrum [Julesz and Caelli, 1979].

As spatial frequency (distance from (0,0) increases, the wavelength of a cycle decreases. Thus, high spatial frequency is associated with small areas and low spatial frequency is associated with large areas in the image. This association means that a coarse texture - one composed of large areas of constant gray level - will have strong low spatial frequency components and a fine texture will have strong high spatial frequency components.

Because of the associations between size and spatial frequency, several texture features have been defined on the power spectrum. The phase spectrum has been largely ignored because of associations between phase and position. Since textures can be identified regardless of the position of particular components in the visual field or the position of a (sufficiently large) window in a uniformly textured plane, phase information has been assumed to be unimportant for texture analysis but



(a) (b) (c) (d)

Figure 1: Sinusoidal Gratings.
(a) 3 cycles per image horizontal (b) 8 cycles per image horizontal
(c) Combination of five spatial frequencies with arbitrarily selected amplitudes. This is an example of a "one-dimensional texture".
(d) 8 cycles per image horizontal and 8 cycles per image vertical

useful for pattern recognition [Richards and Polit, 1974 (Appendix A, item 6); Bajcsy and Lieberman, 1976]. Some recent studies have attempted to reexamine the potential of phase information for texture analysis, but the results have not been encouraging [Eklundh, 1979; Julesz and Caelli, 1979; Zucker and Cavanaugh, 1980; D'Astous and Jernigan, 1981]. It will be argued later that the apparent failure of phase information in texture analysis is due to its improper use.

Two types of features are computed from the power spectrum: spatial frequency energy and orientation energy [Weszka et al, 1976; Haralick, 1979]. Spatial frequency energy features have the form (using polar coordinates)

$$\rho_k = \int_0^{2\pi} \int_{r_k}^{r_k+\Delta r} F^2(r\cos\theta, r\sin\theta)\, dr\, d\theta$$

These features give the total energy in a limited band of spatial frequencies. Orientation energy features have the form (again in polar coordinates)

$$\omega_k = \int_0^{r_{max}} \int_{\theta_k}^{\theta_k+\Delta\theta} F^2(r\cos\theta, r\sin\theta)\, d\theta\, dr$$

These features give the total energy in limited orientation bands, which indicate directional tendencies in the image.

These global analysis methods have several disadvantages.

(1) The shapes and widths of the spatial frequency and orientation bands are free parameters which are arbitrarily specified.

(2) Differences in illumination, contrast, or gray level resolution cause significant changes in the power spectrum. This problem can be alleviated by preprocessing techniques such as histogram

equalization, but this can cause a loss of gray level resolution and image fidelity [Haralick, 1979] and can dramatically change the appearance of the image.

(3) Computation of the Fourier transform or autocorrelation function over irregularly shaped regions (as are commonly encountered) involves some difficulty [Wechsler, 1980].

(4) Every entry in $F(u,v)$ is determined by all of the image function. It is not possible to extract spatially-limited information from $F(u,v)$ by any method short of an inverse Fourier transform. This implies that the spatial frequency domain features cannot be used for texture segmentation without recomputation of the Fourier transform for each subimage to be analyzed [Bajcsy, 1973; Bajcsy and Lieberman, 1976].

3.4 Intermediate Analysis Methods

Several texture analysis methods are based on the relationships among particular types of objects ("primitives") in an image. These methods typically involve definition of primitives, extraction of primitives from a textured image, and characterization of the texture by formal language or heuristic methods.

In the formal languages approach, textures are considered to be language classes for which separate grammars can be inferred. Texture analysis is seen to consist of extracting the primitives and then parsing the image according to the texture grammars. Unfortunately, both of these operations are nontrivial, even in simple test cases.

Alternatively, the primitives can be analyzed by heuristic methods, often involving statistical measurements of the structure of the primitives. For example, "generalized co-occurrence matrices" [Davis et al, 1979; Haralick, 1979] can capture some aspects of the spatial distribution of primitives. A generalized co-occurrence matrix gives the number of times different primitives occur in the image at a particular displacement. The same features used for the gray level co-occurrence method can be computed from generalized co-occurrence matrices.

An orderly review of intermediate analysis methods is obtained by examining them in order of increasing complexity of the primitives.

One method based on formal language techniques uses 9x9 pixel windows as the primitive [Lu and Fu, 1978a, 1978b]. The texture is characterized by a stochastic tree grammar which specifies the assignment of gray levels to the windows. Error-correcting parsing methods are used to eliminate noise, and functions to compute the "distance" between languages are used to measure differences between texture classes.

Another simple primitive is an area of constant or nearly constant gray level [Tsuji and Tomita, 1973; Tomita and Yachida, 1973; Tomita, et al, 1973]. Features such as size, gray level, curvature and directionality are computed to characterize the regions. Textures are identified by multiple modes in the histograms of feature values. Labelling each region by the mode in which it appears provides a first approximation to the textured regions. Generalized co-occurrences or split-and-merge techniques can then be used to refine the segmentation.

One obvious disadvantage of this approach is that its complexity increases rapidly with the fineness of the texture and with the gray scale resolution due to the large number of primitive areas which must be evaluated.

An edge can also be used as a primitive. One recent texture analysis method attempts to characterize repetitive patterns of edge pixels in an image [Vilnrotter et al, 1981] using an "edge repetition array" constructed from horizontal and vertical scans of the image. The array is then analyzed to find repetitive edge structures. In another approach, texture is characterized by computing generalized co-occurrences of edge pixels [Rosenfeld, 1979]. A spatial frequency domain form extraction method [Crowley and Parker, 1978] which identifies edges whose contrast exceeds a threshold is also suggested for texture analysis.

A more complex type of primitive involves geometric shapes such as lines, curves, angles, open polygons, and closed polygons. These primitives are difficult to use, but in some cases complex geometrical texture patterns can be synthesized from simple grammars and fairly complex primitives [Carlucci, 1972; Siromoney et al, 1972].

A theoretical scheme for an intermediate-level structural analysis of texture is presented in [Zucker, 1976]. In this theory, an observed texture is considered a transformation of an "ideal texture" which is typically a polygonal tesselation of the plane. The ideal texture governs the placement of primitives, and transformation rules determine the relationships between primitives (such as superposition, adjacency, etc.). This theory is an idealized characterization of "structural" texture analysis, but there seem to have been no practical spinoffs of the theory.

Other intermediate-level analysis methods have been suggested based on combinations of statistical and structural methods [Tomita, 1981; Conners and Harlow, 1980b].

3.5 Image Modelling Methods

This class of methods for characterizing texture assumes that the textured image is a realization of some stochastic process which is governed by a few parameters. Texture analysis is viewed as a parameter estimation problem; given an image $I(x,y)$, estimate the parameters of the assumed random process so that the probability of obtaining $I(x,y)$ is maximized. The estimates of the parameters serve as texture features for classification problems. The estimates can also be used to synthesize other images with similar (in the sense of the model) texture.

One type of image synthesis model is based on stochastic tesselations of the plane [Schachter et al, 1978a; Ahuja and Rosenfeld, 1981]. These random mosaic models produce polygonal regions in the image plane. Conceptually, the models produce a textured image either by generating random lines through the image or by generating random points and growing regions around them. In [Modestino et al, 1981], gray scale random textured images are generated which are reminiscent of the random mosaic images. This model allows control of gray level correlations between adjacent regions as well as control of the polygonal generation process. A log-likelihood texture analysis and segmentation procedure based on the synthesis method is presented and demonstrated on compound images composed of subimages generated by the model. The segmentation results are good. Further experiments reveal some difficulties in generalizing the method to discriminate natural textures.

A different, very flexible image generation model is developed in [Pratt et al, 1978; Gagalowicz, 1978; Pratt et al, 1981]. In this model, an image is considered to be the output of a homogeneous spatial operator responding to noise input. The noise input supplies the randomness in the texture and the spatial operator supplies local structure. Characterization of a texture then requires specification of a noise distribution and a spatial operator.

Several statistical models for image generation and image modelling are reviewed in [Kashyap, 1980; Garber and Sawchuk, 1981; Chellappa and Kashyap, 1981a, 1981b].

One class of well-known texture models is based on the Markov random field [Rosenblatt and Slepian, 1962; Besag, 1974; Hassner and Sklansky, 1978; Cross and Jain, 1981; Chellappa and Kashyap, 1981a, 1981b; Garber and Sawchuk, 1981; Schmitt and Massaloux, 1981]. One-dimensional Markov random field models have been used to synthesize images for psychophysical experiments [Julesz 1962, 1965] and for theoretical comparisons of texture analysis methods [Conners and Harlow, 1980a]. In one recent study [Cross and Jain, 1981], parameters of a Markov process were fitted to natural textures from [Brodatz, 1966]. The parameters were then used in the model to generate random images. The results are typical of the fundamental problem with model-based approaches to characterizing textures: while some successful model-based parameterizations can be found, and the model can be used to randomly generate some images which appear similar to a prototype, in general, natural images are not bound to conform to the restrictions of a particular model. This fundamental problem severely limits the potential for model-based approaches in general texture analysis problems.

3.6 Summary

A critical review of texture analysis methods has been presented. The large number and the variety of texture analysis methods is caused in part by the lack of a general definition of texture and by the difficulties encountered when existing methods have been applied to general texture analysis problems. Many methods have been developed on an ad hoc basis guided by intuition.

The few comparative evaluations which exist conclude that the co-occurrence method is the best statistical approach to texture. The co-occurrence matrix is widely used in applications and as a basis for more sophisticated techniques including generalized co-occurrences, visually interpretable texture features, and combined statistical-structural approaches to texture.

Features from the Fourier transform of the image have been suggested, but their performance has been poor. The comparative studies conclude that the statistical features provide better results than power spectral features, and attempts to develop texture features from the phase spectrum have been unsuccessful.

The performance of existing texture analysis methods depends on the data. Acceptable performance can be obtained for some specific problems, but the approaches lack generality. Many approaches have disadvantages such as insensitivity to some aspects of texture or susceptibility to irrelevant gray level variations ("noise").

Chapter 4

A Texture Analysis Method Based on a Theory of Human Vision

4.1 Introduction

In this chapter, the channel filtering theory of early human vision will be used to motivate a new approach to texture analysis. First, however, we pause to explain why and how this theory of human vision will be used to guide the development of a texture analysis method. Then, the channel filtering theory and its potential significance for computational vision will be briefly presented. A feature space for texture analysis will then be developed and its plausibility will be examined.

4.2 A Critical Re-evaluation of Texture Analysis Methodology

Computational vision methods are not evaluated for use in applications based on whether the methods accurately emulate the mechanisms (or sometimes even the performance) of human vision. Methods are sometimes evaluated by correlations with human performance, but correlation does not imply equivalence or causality. High correlations with human performance might be obtained using processes very different from those present in human vision. But if correlation with human performance is desirable, it seems reasonable to base the computational techniques on whatever is known about human vision. Insights into human vision can serve to motivate computational texture analysis procedures.

How should information concerning human vision be used to guide computational vision research? Ideally, we would know the overall strategies used in human vision to analyze an image. We could then implement a vision system which might well rival human performance on many visual tasks. But the computer vision system would not necessarily emulate the human visual system at the neural level. The operation of low-level components of a complex system may give no useful information about the overall strategies used in the system. (This is discussed at length by Hofstadter [1979].) Once the overall strategies in vision are identified, their implementation can be tailored to whatever devices or constraints are relevant.

In some approaches to computational vision, a technique which appears to reproduce or explain human performance is developed, then neural mechanisms are postulated which correspond to some aspects of the computational methods [Julesz et al, 1973; Julesz, 1981; Marr et al, 1979; Marr and Hildreth, 1980; Marr, 1980; Hildreth, 1980]. The use of vision science to construct post hoc justifications for existing methods is not effective since the value of a proposed technique for computational vision is determined by correlations between observed and desired performance. Correspondences between a proposed computational method and hypothesized mechanisms of human vision are ultimately irrelevant.

In this study, the opposite approach is used. Information from vision science will be used to guide the development of a new method for texture analysis, not to bolster or to validate an existing method. A theory which attempts to describe the possible processing strategies in early human vision will be used to guide the development of the computational procedure.

4.3 Spatial Frequency Channels

In the late 1960's, researchers found that the threshold visibility of sinusoidal gratings (Figure 1) depends on the spatial frequency of the gratings [Campbell and Robson, 1968; Pantle and Sekuler, 1968; Blakemore and Campbell, 1969; Campbell and Maffei, 1970]. The potential power of Fourier analysis as a tool for studying human vision was noted immediately [Campbell and Robson, 1968]. Attempts were made to determine how well spatial frequency domain analysis of visual stimuli correlated with actual human performance. One key issue involved whether the visual system combines spatial frequency information linearly, and thus whether visual analysis of complex objects could be expressed easily in terms of the spatial frequencies in the complex stimuli. Another important issue was whether the analysis of a stimulus by the visual system involves several independent mechanisms (called channels) which analyze different aspects of the stimulus. It was hypothesized that the channels might have a convenient spatial frequency domain representation. This hypothesis sparked additional activity in which neurological and psychophysical experiments were interpreted as providing evidence for or against various single-channel or multi-channel models of early stages of human visual processing. For reviews of this activity, see [Graham, 1981; Ginsburg, 1978].

These results were unified and extended in [Ginsburg, 1978, 1980b]. An implementation of the multi-channel theory was used to illustrate the action of spatial frequency domain filters on various images containing complex forms, visual illusions, multistable images, and certain visual textures. In addition, the contrast sensitivity of abnormal visual systems was found to have unique properties in the spatial frequency domain. The channels in the visual system are modelled by filters defined in the spatial frequency domain, but a critical feature of this particular channel analysis is the assumption that phase information is an essential part of the internal representation of visual stimuli. Thus, the decomposition of an image is modelled by a sequence of spatial domain filtered images [Ginsburg, 1971].

These studies did not develop algorithms for automatic image analysis, but other papers [e.g. Ginsburg, 1973, 1979a] suggested that filtering in spatial frequency channels might be useful for machine pattern recognition, including texture analysis.

The assumption of a channel decomposition of an image suggests two computational implementations [Hall, 1972; Nathan, 1970; Ginsburg, 1979b]. In one, a series of point spread functions, or templates, is convolved with the image to obtain a series of distorted (filtered) versions of the original image. Alternatively, the convolution can be performed in the spatial frequency domain by applying an inverse Fourier transform to the product of the filter transfer function (the

Fourier transform of the point spread function) and the Fourier transform of the image. This process is illustrated in Figure 2. The resulting sequence of filtered images is exactly the same as those produced by spatial convolution. The choice of the method is largely a matter of convenience and computational efficiency. For this study, the spatial frequency domain method will be used due to the availability of fast algorithms for the discrete Fourier transform [Johnson and Jain, 1981] and due to the intuitive value of defining the filters in the spatial frequency domain [Ginsburg and Coggins, 1981; Graham, 1981].

The next issue is to decide the shapes, sizes, locations, and number of the channels. The filtering properties of the human visual system are not precisely known, but psychophysical and neurological data can be used to guide the selection of the filter parameters. This study will use filters whose parameters are within the constraints specified in [Ginsburg, 1978]. The transfer functions for spatial frequency channels are defined by a Gaussian function (on a log scale); their center frequencies are one octave apart and their width is between 1 and 2 octaves. The number of spatial frequency channels used depends on the size of the image (see Appendix B for details). Figure 3 shows two representations of a spatial frequency channel filter in the spatial frequency domain (u-v plane). The height of the surface above the u-v plane (Figure 3a) and the intensity of the gray levels (Figure 3b) represent the filter amplitudes, $|F(u,v)|$, which are between 0 and 1. Four orientation channels are implemented with center orientations directed horizontally, vertically, and along the two diagonals (see Appendix B for details). Figure 4 shows two

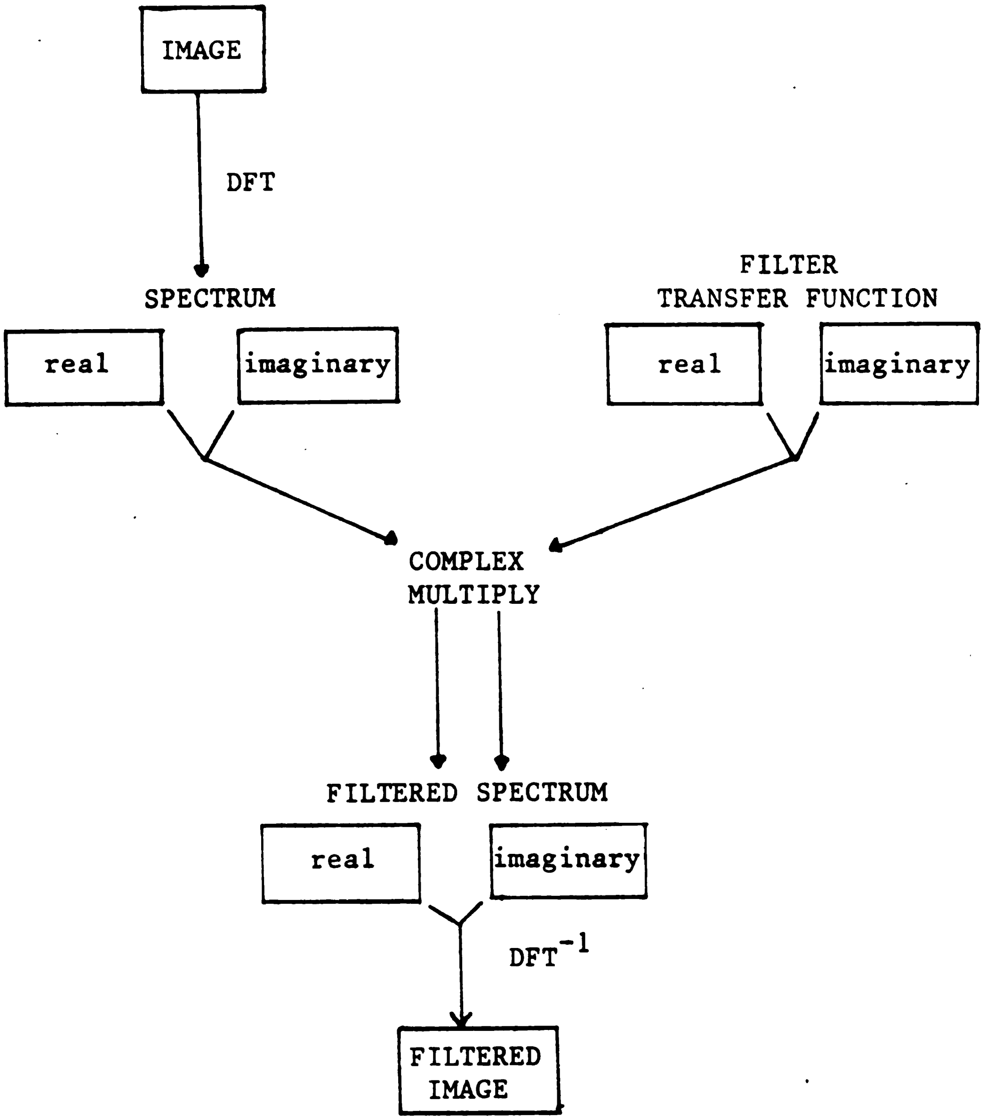
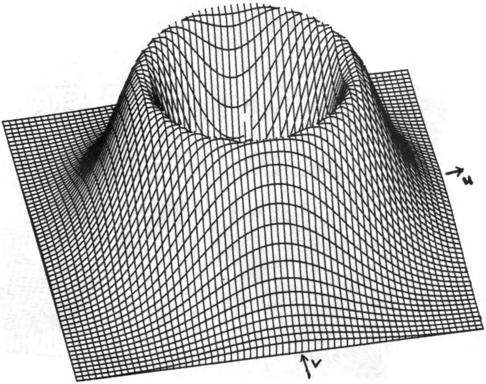
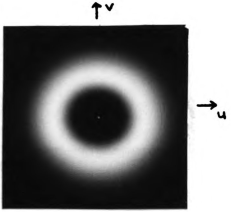


Figure 2: Diagram of the Spatial Filtering Procedure.
This procedure is repeated once for each channel.

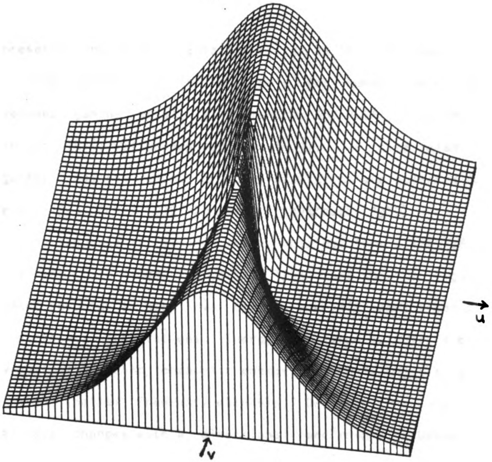


(a)

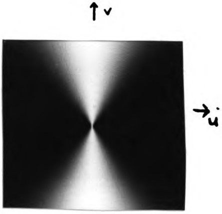


(b)

Figure 3: A Spatial Frequency Filter.
(a) A transect plot portraying the amplitude of the real part of the filter transfer function in the spatial frequency domain. The zero-frequency component is in the center of the region.
(b) An image representation of the same filter transfer function.



(a)



(b)

Figure 4: An Orientation Channel Filter.
(a) A transect plot portraying the amplitude of the real part of the filter transfer function in the spatial frequency domain. The zero-frequency component is in the center of the region.
(b) An image representation of the same filter transfer function.

representations of an orientation filter in the u-v plane.

Other channel filtering models for human vision have been developed [Sachs et al, 1971; Richards and Polit, 1974; Mostafari and Sakrison, 1976; Wilson and Bergen, 1979]. We note also that some arguments against a Fourier model of vision have appeared [Julesz and Caelli, 1979; Ochs, 1979; Zucker and Cavanaugh, 1980].

The sequence of filtered images obtained from a 128x128 sample of a ceiling tile image [Brodatz, 1966] are shown in Figure 5. Each channel responds to gray level changes over different sized regions or at different orientations. Energy from large objects is displayed in low spatial frequency channels; energy from small objects is displayed in high spatial frequency channels. Orientation channels respond to gray level changes with a directional preference. Since the phase information from the original image is retained in the filtered images, the spatial distribution of the spectral energy in each channel is apparent in the positions of gray level variations in the filtered images. The use of phase information captures in the filtered images the gray level spatial distribution information which is essential to texture.

The specification of the channels completes the definition of the initial information processing stage of a computational vision system. The output of this stage is a series of channel-filtered images, each of which contains limited spectral information from the original image. The next task is to exploit the channel decomposition, reducing the series of filtered images down to a set of texture features which can be used to duplicate some texture analysis capabilities of human vision.

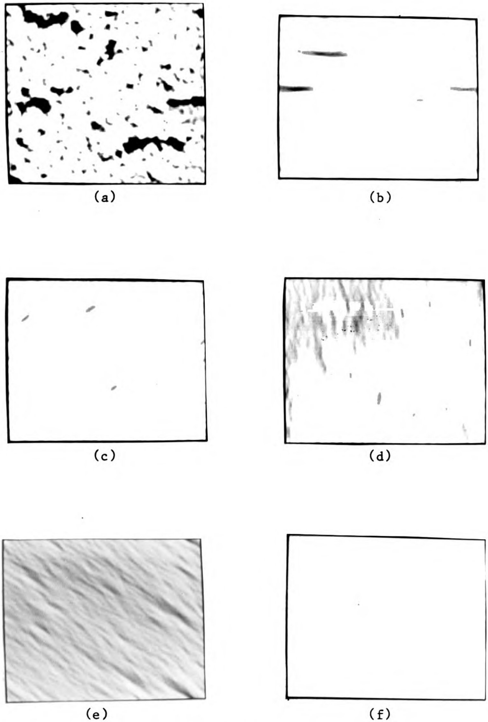
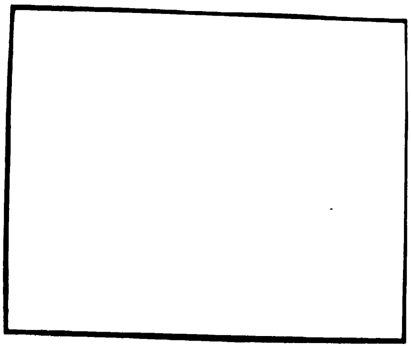
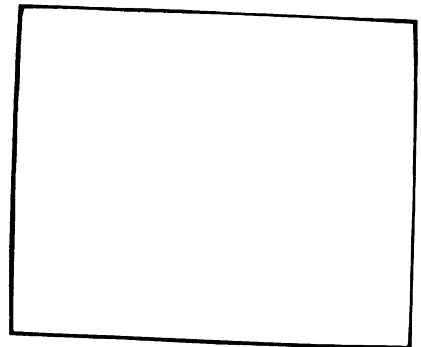


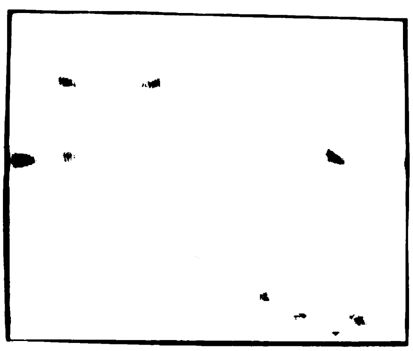
Figure 5: An Example of Channel Filtering.
(a) Original 128x128 TILE image. (b)-(l) filtered images with center frequencies (in cycles per image) and orientations (in degrees) as follows: (b) 0 degrees (c) 45 degrees (d) 90 degrees (e) 135 degrees (f) 1 c/i (g) 2 c/i (h) 4 c/i (i) 8 c/i (j) 16 c/i (k) 32 c/i (l) 64 c/i



(g)

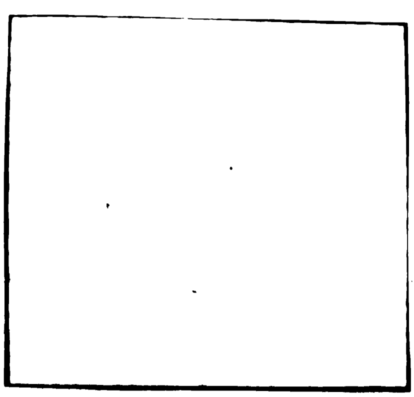


(h)

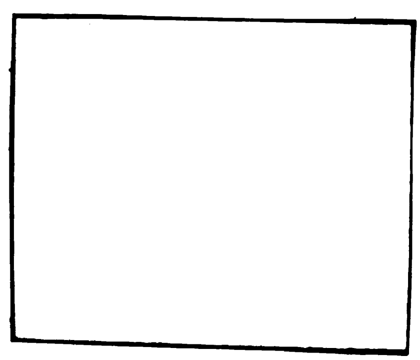


(i)

(j)



(k)



(l)

Figure 5 (cont'd)

4.4 Constraints on Possible Texture Features

Pragmatic requirements constrain the nature of the features we are willing to compute from the filtered images. The features must be very simple since we now have a whole series of images to analyze rather than a single image on which more extensive analysis can be performed. We would like to find some evidence from vision science to suggest that simple features which duplicate human performance could exist.

Several observations suggest that texture is a consequence of crude information reduction in vision which simple computational methods should be sufficient to capture. First, textures can be analyzed and discriminated by the human visual system quickly and effortlessly, but only very simple, crude mechanisms have been found in the early stages of mammalian visual systems [Hubel, 1963; Ginsburg, 1978; Graham, 1981]. Second, texture is perceived most clearly in an image area which has many intensity changes and thus small areas of constant intensity [Crowley and Parker, 1978; Resnikoff, 1981]. Individual edges and their placement are not critical for identifying textures. In fact, very different generation procedures can yield images which are not preattentively discriminable. "Texture" appears, then, to be the result of a crude information reduction which occurs in an image area with no distinctive individual features.

This crude information reduction is evident in several indiscriminable textures presented by Julesz (Figure 6). In these images, a simple check of the edge structure of the micropatterns would immediately yield discrimination of the different regions, yet these

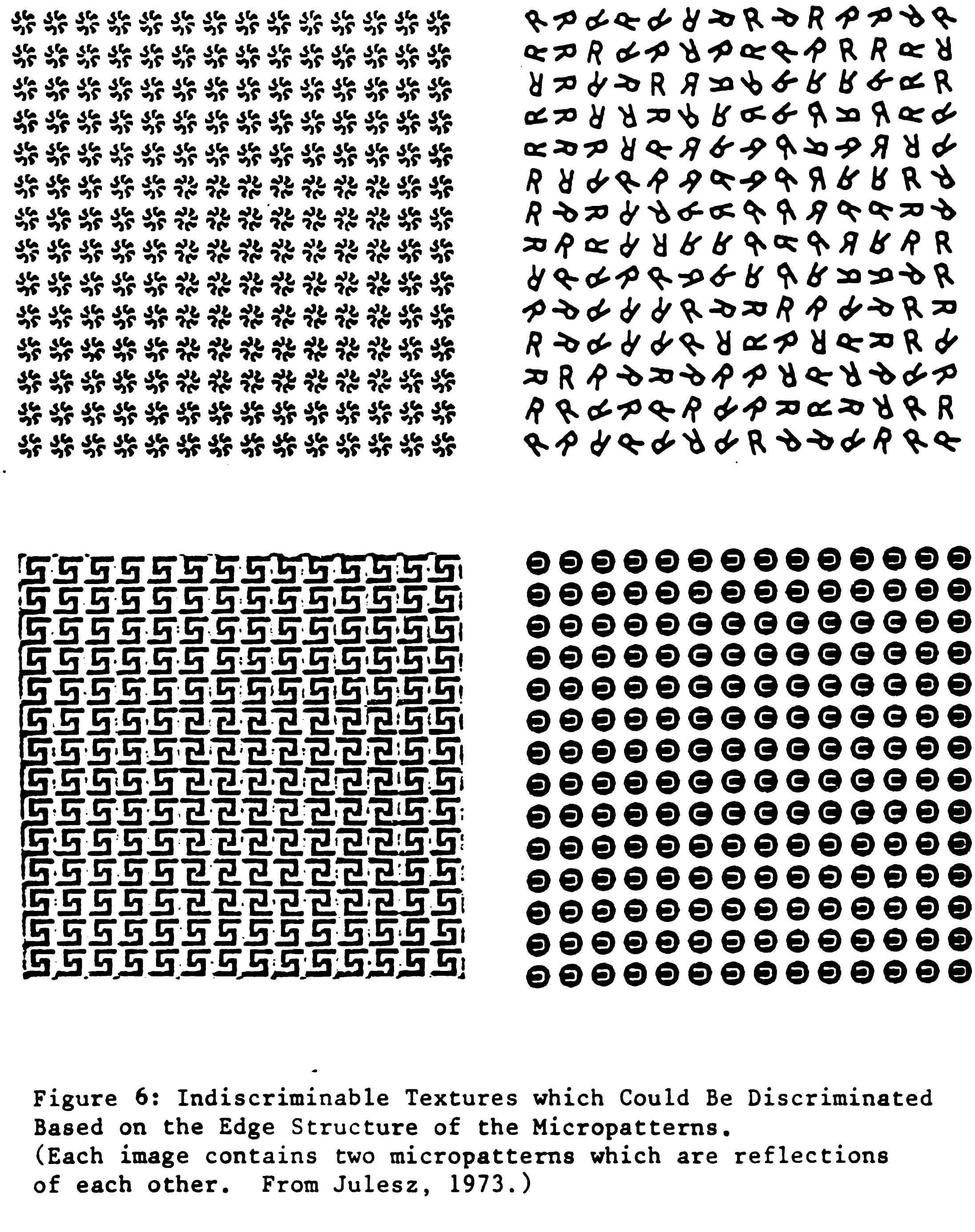


Figure 6: Indiscriminable Textures which Could Be Discriminated Based on the Edge Structure of the Micropatterns.
(Each image contains two micropatterns which are reflections of each other. From Julesz, 1973.)

are among the most difficult patterns to visually discriminate, even given the nature of the difference between the micropatterns. Simple pixel-level operations could be constructed to discriminate the image fields in Figure 6, but the criteria for the discrimination would no longer be "texture". The ability to discriminate micropatterns does not necessarily imply that textures composed of the micropatterns are easy to discriminate [Beck, 1980; Ginsburg, 1978].

Further evidence for the simplicity of human texture perception comes from psychophysical experiments involving "one-dimensional textures" (Figure 1c). In these images, the gray levels in each column are constant and the gray levels along each row are determined by a weighted sum of sinusoidal functions. Psychophysical experiments show that arbitrary one-dimensional textures can be matched with images containing only a few (4) selected spatial frequencies [Richards and Polit, 1974]. Another study shows that human similarity judgements between one-dimensional textures are predicted better by the outputs of four spatial frequency channels than by the actual spatial frequencies present in the textures [Harvey and Gervais, 1981].

An observation made in passing by Julesz et al [1973:Figure 6] provides further evidence that texture is a consequence of simple operations on images. Two patterns generated by a geometrical method involving repetitions of four-dot micropatterns were presented in an image. The generation method insures that the two patterns have identical "second-order statistics", and it was observed that the patterns are not effortlessly discriminable to human observers. Julesz then presented a "blurred" version of the image in which the patterns were found to be discriminable. In addition, it was noted that the blurred patterns have different second-order statistics. The significance of this observation is that "blurring" is an alternative characterization of certain spatial filtering operations. Blurring can enable discrimination using simpler measurements than were required on the original image.

Thus, there is evidence that simple operations on channel-filtered images could be effective for texture analysis.

4.5 Properties of Channel-Filtered Images

In addition to the evidence for the simplicity of "texture" presented in the last section, we can obtain guidance in constructing texture features from certain properties of the channel-filtered images which are consequences of the filter definitions and of the spatial filtering operation. Four properties which have proven useful from preliminary studies are as follows:

1. The mean gray level of the channel-filtered images can be fixed in advance. In the implementation used in this study, the mean gray level of the filtered images is made equal to the mean gray level of the original image by setting the value of the zero-frequency component of the filters to 1, thereby passing the zero-frequency component (average gray level) of the original image unchanged. This property makes possible meaningful displays and comparisons of the channel filtered images. Keeping the average gray level of the filtered images the same simplifies the construction of features which are invariant over global, constant gray level changes.

2. The gray level frequency histograms of the filtered images tend to be symmetrical about the mean gray level due to the symmetrical response of the filters to objects in the original image. Asymmetry in the gray level histograms is caused by the interaction of the responses of spatially close objects.

3. The magnitude of the deviation of the gray level of a pixel in a filtered image from the mean gray level is directly related to the spectral energy contained in a neighborhood of that pixel in the original image. Since phase information is retained in the filtered images, the spatial distribution of the spectral energy passed by the filter is reflected in the gray level distribution in the filtered image. Thus, the spectral energy in different spectral bands arising from small spatial areas can be measured from the sequence of filtered images without recomputing a Fourier transform for each local area of interest.

4. The differences between the channel-filtered images lies in their sensitivities to gray level variations in the original image over regions of different sizes and orientations. The channel decomposition, then, allows separate measurements of local energy over spatial domain neighborhoods of different sizes and orientations.

4.6 Definition of Texture Features

Property 3 above implies that the gray levels in each filtered image can be interpreted as representing the spectral energy arising from local areas of the original image. This interpretation motivates the selection of features for use in texture analysis experiments in this thesis. The features will measure the average local spectral energy in an image by computing the spread of the gray level frequency histogram of the filtered images.

A number of features can be defined to measure the spread of the histograms. We have used eight such features in classification experiments (Figure 7). The first feature is the average absolute deviation from the mean gray level. Features 2-4 are functions of the second through fourth moments of the gray level histogram. Feature 2 is actually the standard deviation of the gray levels. We note that the third moment is not strictly a measure of spread, but it is included for completeness. Functions of the moments are used rather than the actual values of the moments because of computational difficulties encountered in preliminary experiments caused by the large magnitudes of the third and fourth moments. Features 5-8 are heuristic spread measurements which assume that the histogram is symmetric. Property 2 from the previous section states that this assumption is generally reasonable, but can fail in specific cases. The threshold values of .25, .50, .75, and 1.0 used in the feature definitions are arbitrarily selected.

Similar features are investigated by Laws [1980], though the filters in that study are defined as small spatial domain templates.

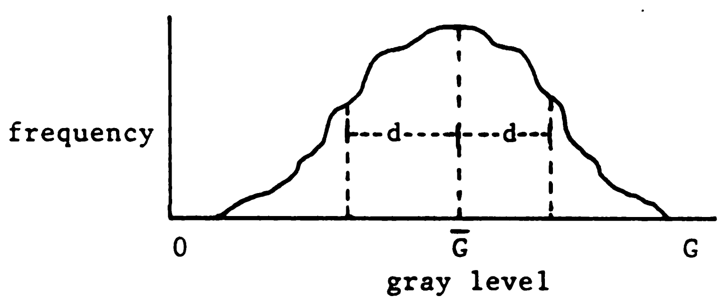
The sequence of filters satisfies the requirement that texture be measured at several different scales. Each gray level in the filtered images is determined by a neighborhood in the original image, so the filtered images are consistent with the spatial property of texture. Statistics of the gray level histograms of filtered images are easy to compute, so they satisfy the pragmatic requirement that channel

$$F1_k = \frac{1}{N^2} \sum_{g=0}^{G} |g-\bar{G}| * f_k(g)$$

$$F2_k = \left[\frac{1}{N^2} \sum_{g=0}^{G} (g-\bar{G})^2 * f_k(g) \right]^{1/2}$$

$$F3_k = \left[\frac{1}{N^2} \sum_{g=0}^{G} (g-\bar{G})^3 * f_k(g) \right]^{1/3}$$

$$F4_k = \left[\frac{1}{N^2} \sum_{g=0}^{G} (g-\bar{G})^4 * f_k(g) \right]^{1/4}$$



$$\text{Let } S_k(d) = \sum_{g=\bar{G}-d}^{\bar{G}+d} f_k(g)$$

$F5_k$ = d such that $S_k(d) = .25*N^2$

$F6_k$ = d such that $S_k(d) = .50*N^2$

$F7_k$ = d such that $S_k(d) = .75*N^2$

$F8_k$ = d such that $S_k(d) = N^2$

Figure 7: Texture Feature Definitions.
$f_k(g)$ denotes the frequency of gray level g ($0 \leq g \leq G$) in the filtered image for channel $\underset{\smile}{k}$. $\bar{G}$ denotes the mean gray level. Image size is NxN pixels.

filtering features be simple. Moreover, the features have reasonable visual interpretations in terms of spectral energy, size, and orientation. In addition, these features can be computed for any subimage without repeating the channel filtering operation; this property will be important for texture segmentation (Chapter 6).

4.7 How Will the Texture Features Be Used?

An image will be represented by a feature vector consisting of the values of one of the texture features computed over each filtered image. The feature vector will map the image into a point in the feature space; this point will represent the image's texture. The distance between points in the feature space will be used as a measure of the textural difference between images.

In most approaches to texture analysis, a number of features are defined and then they are applied in various combinations to classify texture samples [e.g. Weszka et al, 1976; Faugeras and Pratt, 1980; Haralick and Shanmugam, 1973]. This approach is more likely to produce, at some point, "good" classification results than using each feature in isolation. But combining various features whose properties are not well understood leads to problems in interpreting results. In this study, the features will be used separately, expecting that the results for all of them (except perhaps Feature 3) will be similar. Selection of the "best" feature from these will have to be based on computational considerations and compatibility with solutions to other image analysis problems such as form extraction, image registration, or edge detection.

4.8 Summary

Existing texture analysis methods use little guidance from properties of human visual perception. A theory concerning the early information processing in human vision was used to motivate the development of a new feature space for texture analysis. The theory asserts that an image is decomposed by the human visual system into channels which are modelled by spatial filtering operations. Since phase is retained in the channel-filtered images, the spectral energy caused by gray level variations in local spatial areas can be measured from the channel-filtered images. Features which measure this local energy were defined for use in texture analysis studies.

Chapter 5

Evaluation of Channel Filtering Features for Texture Classification

5.1 Introduction

This chapter will present the results of several experiments designed to evaluate the channel filtering features, to compare their performance against some other texture features proposed in the literature, and to illustrate properties of the feature space which may be useful in applications.

The experiments in this chapter are texture classification problems using a supervised learning paradigm. Sets of test images are labelled according to an external criterion which, for these experiments, is a visual evaluation of the image texture. The sets of images are then analyzed by the channel filtering method, resulting in a feature vector for each image. These feature vectors are then input to a classification algorithm, and the estimated error rate is used to evaluate the class separations in the feature space.

Most of the experiments in this chapter use 25 samples of each image class. The samples are extracted as subimages from a large (256x256 pixels) image which represents a single texture class. Nonoverlapping subimages are extracted as far as possible, then the sampling origin is shifted by half the subimage size to complete the set if necessary. The number of samples is fairly small, but since the large images are homogeneous, with no complications such as shadows, receding surfaces, or magnification, this number of samples from throughout the image should adequately characterize the image texture. However, in interpreting the estimated error rates, we must keep in mind that the estimates are based on a small number of samples.

In all of the experiments in this chapter, different texture features are used in separate classification problems. The feature vector for each image consists of the values of one of the texture features F1...F8 defined in Figure 7 computed over all of the channel-filtered images. The dimensionality of that vector depends on the size of the images being analyzed; for 64x64 images, eleven channels (four orientation, seven spatial frequency) are used; for 32x32 images, ten channels (four orientation, six spatial frequency) are used.

For the experiments in this chapter, a Nearest Neighbor Classifier will be used. The decision rule is to classify an unlabelled point into the class of its nearest labelled neighbor in the feature space, where the distances are determined by the Euclidean distance metric. This decision rule results in piecewise linear decision boundaries. The error rate of the classifier will be estimated using the Leave-One-Out rule, which is often used when sample sizes are small. In this method, each point in turn is treated as a test sample with all other points as training samples. The results will be presented as a confusion matrix for each feature. The (i,j)th entry of the matrix gives the number of times a sample from class i was classified as class j.

The performance of the channel filtering features will be compared to the performance of another feature, the total power spectral energy (PSE) in the sequence of filtered spectra. This feature has been used with ideal band-pass channels in previous studies (Section 3.3). This feature uses no phase information, by definition, but it uses the same power spectral information as the channel filtering features. The effect of using phase information in the channel filtering features will be determined by comparing the error rates of PSE and channel filtering features.

The experiments will first establish the performance of the proposed texture features on classification of natural images whose sizes are 64x64, 32x32, and 16x16 pixels. Then the effect of histogram equalization on the natural images will be evaluated. The sensitivity of the features to changes in the average gray level, magnification, and orientation of the textured images will be evaluated. Computational simplifications of the channel filtering procedure using fewer channels and using channels with an ideal band-pass characteristic will also be evaluated. The features will also be applied to a particular set of images to compare their performance to that of the co-occurrence method.

5.2 Evaluations of the Features on Natural Images

In this sequence of experiments, eight images from [Brodatz, 1966] will be used to provide an initial test of the proposed texture analysis method. The data for these experiments consists of subimages of eight 256x256 images illustrated in Figure 8. The images (and their

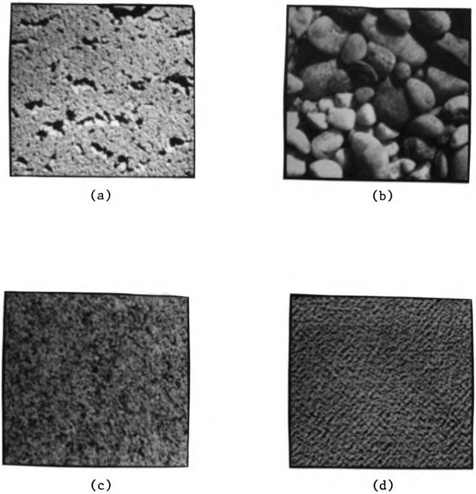
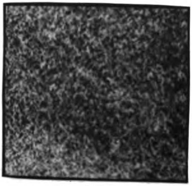
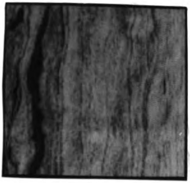


Figure 8: Natural Image Classes.
(a) TILE (b) ROCK (c) SAND (d) PAPE
(e) CORK (f) GRAS (g) WOOD (h) SCRE

(e)



(f)



(g)

(h)

Figure 8: (cont'd)

abbreviations) are as follows: ceiling tile (TILE), beach pebbles (ROCK), beach sand (SAND), handmade paper (PAPE), pressed cork (CORK), grass lawn (GRAS), wood grain (WOOD), and straw screening (SCRE). These images were selected to provide a variety of textural properties and a range of difficulty in discrimination. The images were digitized using a Spatial Data Systems Eyecom image processing system. No preprocessing was applied to the digitized images for the experiments in this section.

5.2.1 Experiment 1: Natural Images, 64x64 subimages

In the first experiment, 25 samples of size 64x64 were extracted from the 256x256 images. Sixteen of the samples were nonoverlapping, and nine more were extracted by shifting the sampling origin by (32,32). The 256x256 images are visually discriminable. The feature space will now be tested on smaller images which could, in principle, provide a more difficult discrimination problem since a smaller texture sample is available for analysis. Each of the 200 samples were filtered and the eight features in Figure 7 were computed on all of the filtered images.

The classification results for all of the features are given in Table 1. The PSE feature provides 69% correct classification while all of the channel filtering features (except feature 3) provide better than 90% correct classification. Feature 3 is the only channel filtering feature which does not measure the spread of the gray level frequency histogram of the filtered images and therefore cannot be easily interpreted as a local energy measure. Feature 3 yields 79.5%

Table 1: Classification of Eight Natural Image Classes Using 64x64 Subimages.

PSE		TILE	ROCK	SAND	PAPE	CORK	GRAS	WOOD	SCRE	
	TILE	19	2	3	1	0	0	0	0	
	ROCK	3	18	1	0	1	0	1	1	
	SAND	4	0	14	2	0	1	4	0	
	PAPE	0	0	1	22	1	1	0	0	69%
	CORK	0	0	0	0	15	10	0	0	accuracy
	GRAS	0	0	0	1	9	13	2	0	
	WOOD	2	0	5	0	0	4	14	0	
	SCRE	0	0	0	0	0	2	0	23	

F1		TILE	ROCK	SAND	PAPE	CORK	GRAS	WOOD	SCRE	
	TILE	22	2	1	0	0	0	0	0	
	ROCK	1	24	0	0	0	0	0	0	
	SAND	0	0	25	0	0	0	0	0	
	PAPE	0	0	0	25	0	0	0	0	98%
	CORK	0	0	0	0	25	0	0	0	accuracy
	GRAS	0	0	0	0	0	25	0	0	
	WOOD	0	0	0	0	0	0	25	0	
	SCRE	0	0	0	0	0	0	0	25	

F2		TILE	ROCK	SAND	PAPE	CORK	GRAS	WOOD	SCRE	
	TILE	22	2	1	0	0	0	0	0	
	ROCK	0	25	0	0	0	0	0	0	
	SAND	0	0	23	0	0	2	0	0	
	PAPE	0	0	0	25	0	0	0	0	97.5%
	CORK	0	0	0	0	25	0	0	0	accuracy
	GRAS	0	0	0	0	0	25	0	0	
	WOOD	0	0	0	0	0	0	25	0	
	SCRE	0	0	0	0	0	0	0	25	

F3		TILE	ROCK	SAND	PAPE	CORK	GRAS	WOOD	SCRE	
	TILE	25	0	0	0	0	0	0	0	
	ROCK	2	16	0	0	1	5	1	0	
	SAND	0	0	16	2	1	0	3	3	
	PAPE	0	0	0	25	0	0	0	0	79.5%
	CORK	0	0	1	0	17	4	1	2	accuracy
	GRAS	0	0	0	2	11	12	0	0	
	WOOD	0	0	1	0	0	0	23	1	
	SCRE	0	0	0	0	0	0	0	25	

F4		TILE	ROCK	SAND	PAPE	CORK	GRAS	WOOD	SCRE	
	TILE	23	1	1	0	0	0	0	0	
	ROCK	0	25	0	0	0	0	0	0	
	SAND	0	0	23	0	0	2	0	0	
	PAPE	0	0	0	25	0	0	0	0	98%
	CORK	0	0	0	0	25	0	0	0	accuracy
	GRAS	0	0	0	0	0	25	0	0	
	WOOD	0	0	0	0	0	0	25	0	
	SCRE	0	0	0	0	0	0	0	25	

Table 1 (cont'd)

F5	TILE	ROCK	SAND	PAPE	CORK	GRAS	WOOD	SCRE	
TILE	21	1	1	0	0	2	0	0	
ROCK	5	20	0	0	0	0	0	0	
SAND	1	0	22	0	0	2	0	0	
PAPE	0	0	0	25	0	0	0	0	92%
CORK	0	0	0	0	24	1	0	0	accuracy
GRAS	0	0	1	0	0	24	0	0	
WOOD	0	0	0	1	0	1	23	0	
SCRE	0	0	0	0	0	0	0	25	

F6	TILE	ROCK	SAND	PAPE	CORK	GRAS	WOOD	SCRE	
TILE	22	1	1	0	0	1	0	0	
ROCK	3	22	0	0	0	0	0	0	
SAND	0	0	24	0	0	1	0	0	
PAPE	0	0	0	25	0	0	0	0	96.5%
CORK	0	0	0	0	25	0	0	0	accuracy
GRAS	0	0	0	0	0	25	0	0	
WOOD	0	0	0	0	0	0	25	0	
SCRE	0	0	0	0	0	0	0	25	

F7	TILE	ROCK	SAND	PAPE	CORK	GRAS	WOOD	SCRE	
TILE	23	1	1	0	0	0	0	0	
ROCK	1	24	0	0	0	0	0	0	
SAND	0	0	25	0	0	0	0	0	
PAPE	0	0	0	25	0	0	0	0	98.5%
CORK	0	0	0	0	25	0	0	0	accuracy
GRAS	0	0	0	0	0	25	0	0	
WOOD	0	0	0	0	0	0	25	0	
SCRE	0	0	0	0	0	0	0	25	

F8	TILE	ROCK	SAND	PAPE	CORK	GRAS	WOOD	SCRE	
TILE	22	0	1	0	0	2	0	0	
ROCK	0	25	0	0	0	0	0	0	
SAND	0	0	19	0	2	4	0	0	
PAPE	0	0	0	25	0	0	0	0	91%
CORK	0	0	2	2	21	0	0	0	accuracy
GRAS	0	0	5	0	0	20	0	0	
WOOD	0	0	0	0	0	0	25	0	
SCRE	0	0	0	0	0	0	0	25	

correct classification in this experiment.

Some insight into the structure of the patterns in the feature space can be obtained by plotting the average, over the 25 subimages, of the feature value in each channel. Figure 9 shows the plot for

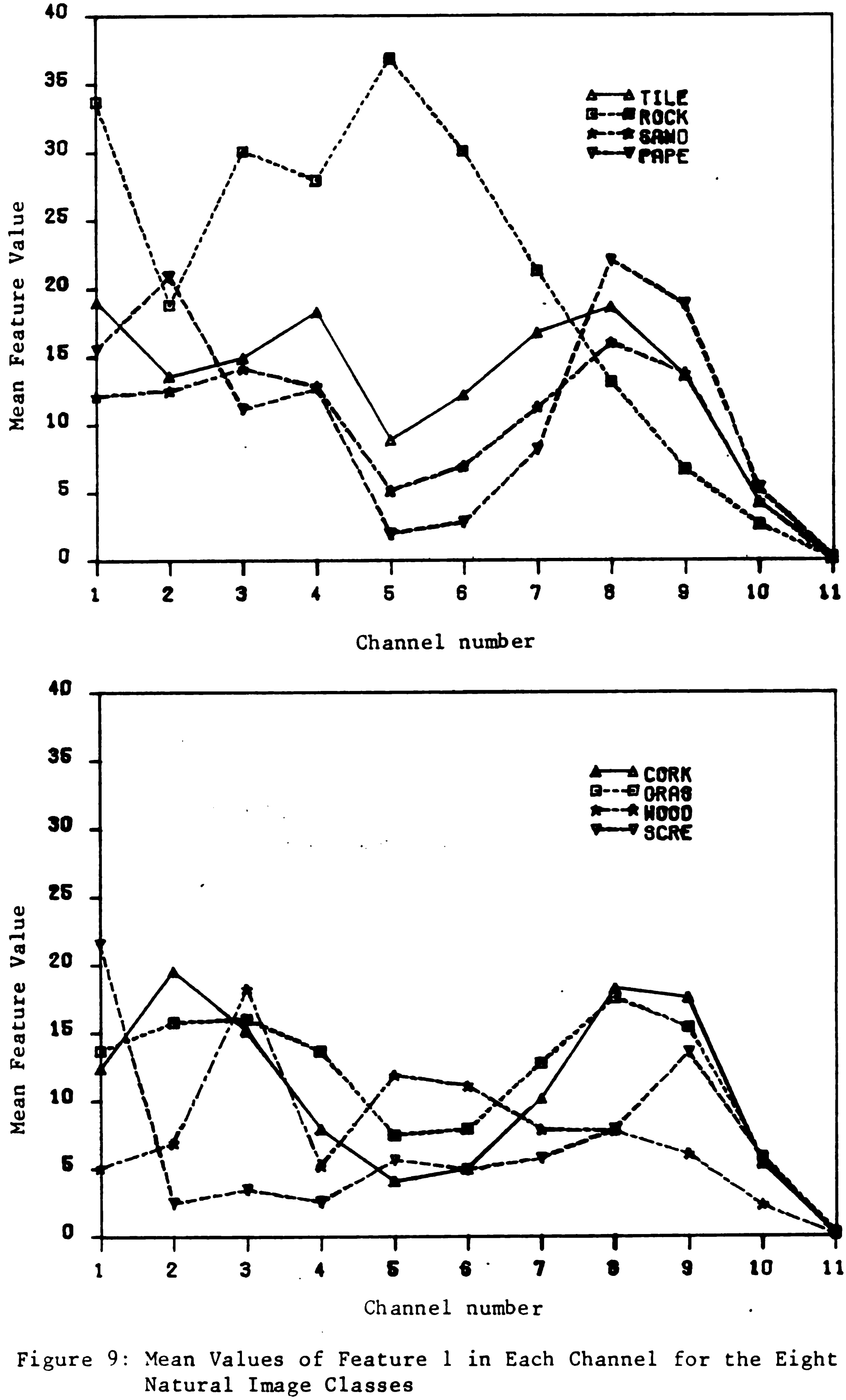


Figure 9: Mean Values of Feature 1 in Each Channel for the Eight Natural Image Classes

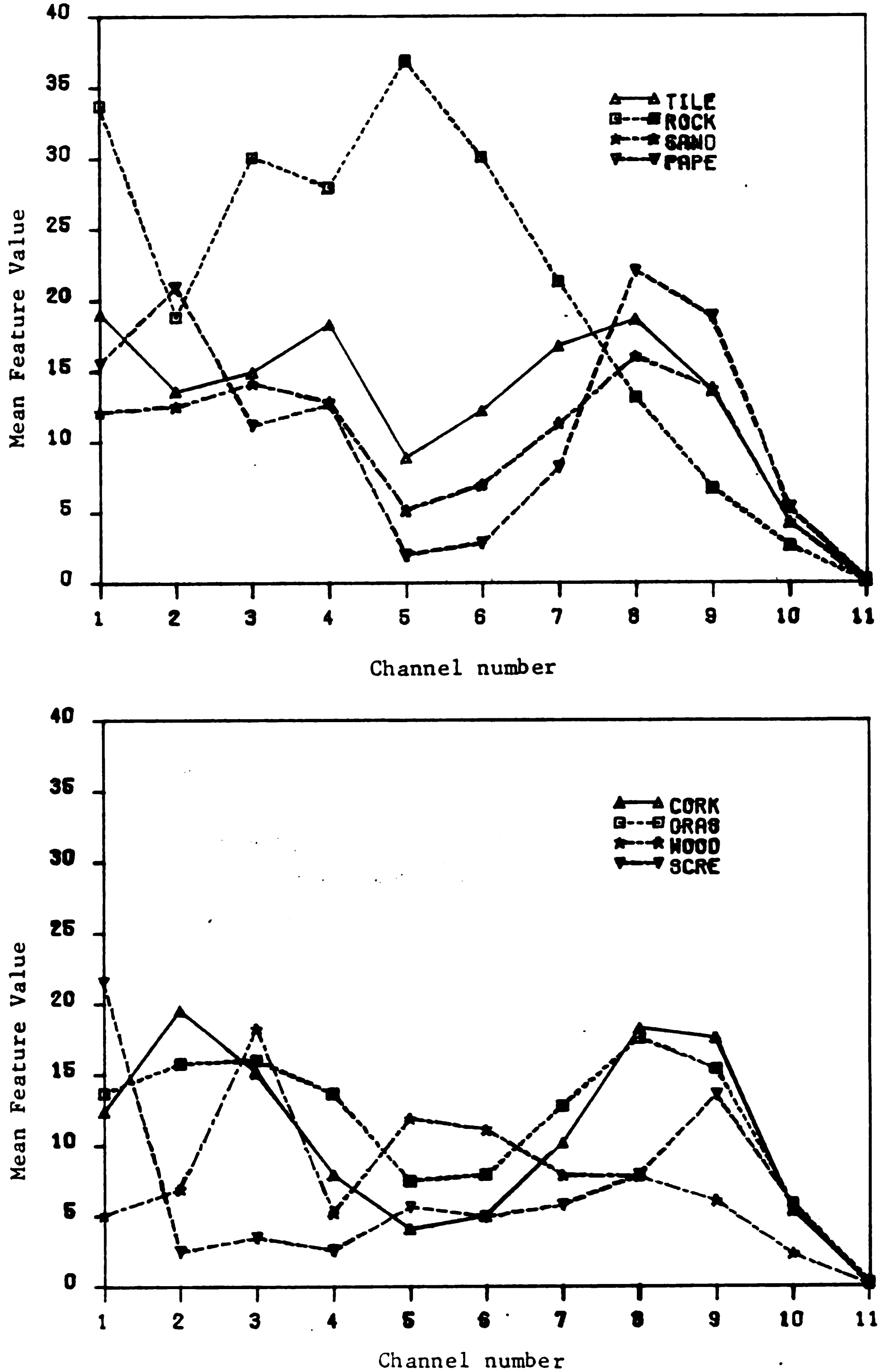


Figure 9: Mean Values of Feature 1 in Each Channel for the Eight Natural Image Classes

feature 1 for the eight natural image classes. The plot of the mean feature values in each channel has unique properties for each image class. For example, ROCK is the only class with high feature values in low frequency channels due to the presence of large objects in the ROCK image. SCRE has a unique pattern in orientation channels; the horizontal channel (number 1) has a high value and the other orientation channels have low values. This reflects the horizontal directional tendency and the regularity of SCRE. In addition, SCRE has low feature values in spatial frequency channels except for the channel centered at 32 cycles per image which captures the periodicity of the bars in the screen. For other texture classes, the description of the mean feature value plot might not be as easy or convenient. These examples show how the values of the channel filtering features can be interpreted as visible image properties.

In the remaining experiments, the results of only the PSE feature and feature 1 will be reported. The channel filtering features (except feature 3) were found in preliminary experiments to have similar performance. Feature 1 is selected for reporting because its behavior was typical of the channel filtering features, because it is a simple feature to compute, and because the feature has its own intuitive interpretation.

5.2.2 Experiment 2: Natural Images, 32x32 subimages

The second experiment involves 32x32 samples of the same eight 256x256 images. The same procedures as in the first experiment were used except that six spatial frequency channels (rather than seven)

Table 2: Classification of Eight Natural Image Classes Using 32x32 Subimages.

PSE	TILE	ROCK	SAND	PAPE	CORK	GRAS	WOOD	SCRE	
TILE	19	0	5	1	0	0	0	0	
ROCK	4	11	1	0	0	1	6	2	
SAND	2	0	19	3	0	0	1	0	
PAPE	1	0	0	23	0	0	0	0	68%
CORK	0	0	0	0	21	2	0	2	accuracy
GRAS	0	0	0	0	4	12	8	1	
WOOD	0	0	4	0	1	11	9	0	
SCRE	0	0	0	0	0	4	0	21	

F1	TILE	ROCK	SAND	PAPE	CORK	GRAS	WOOD	SCRE	
TILE	20	1	2	0	0	2	0	0	
ROCK	0	22	0	0	0	0	3	0	
SAND	1	0	22	0	0	2	0	0	
PAPE	0	0	0	25	0	0	0	0	91%
CORK	0	0	0	0	25	0	0	0	accuracy
GRAS	0	0	6	0	0	19	0	0	
WOOD	0	1	0	0	0	0	24	0	
SCRE	0	0	0	0	0	0	0	25	

were computed. The decreased size of the subimages means that no information beyond the Nyquist frequency of 16 cycles per image width is available, so the highest-frequency channel from the previous experiment (center frequency 64 cycles per image width) is not usable.

The classification results for the PSE feature and feature 1 are given in Table 2. The PSE feature provides 69% correct classification, which is almost the same as for the 64x64 subimages, but the errors occur between different classes. In spite of the reduction in the size of the subimages from those used in the previous experiment, the results show that there is only slight degradation in the performance of the channel filtering features; feature 1 provides 91% correct classification. About half of the errors occur between the GRAS and SAND classes. These results imply that the 32x32 subimages contain enough textural information to enable good classification results on the eight natural image classes. This result increases our confidence in the adequacy of the 64x64 subimage size for characterizing the textural properties of the natural images.

5.2.3 Experiment 3: Natural Images, 16x16 subimages

In this experiment, 50 samples of size 16x16 are extracted from four of the natural images (TILE, ROCK, SAND, PAPE). For subimages of this size, only five spatial frequency channels are useful. Because of this limitation and the small size of the subimages, the number of samples per class was doubled for this experiment.

The classification results for the PSE feature and feature 1 are given in Table 3. In this experiment, the PSE feature achieved 83% correct classification with errors distributed mostly among TILE, ROCK, and SAND. The improved performance of the PSE feature using these small subimages may be due to the larger variance in the average gray levels of small subimages. The increased variance can result in the (accidental) formation of clusters from the same image class which enhance the classification accuracy using the nearest-neighbor classifier. The sensitivity of the PSE feature to changes in the average gray level of an image will be demonstrated in Experiment 6. Channel filtering feature 1 yields 68% correct classification with more than half of the errors occurring between TILE and SAND. In fact, the TILE and SAND classes are practically indistinguishable. The results indicate that 16x16 subimages do not capture enough of these texture patterns to enable accurate nearest-neighbor classification using the channel filtering features.

Table 3: Classification of Four Natural Image Classes
Using 16x16 Subimages.

PSE		TILE	ROCK	SAND	PAPE	
	TILE	39	2	9	0	
	ROCK	7	41	2	0	83%
	SAND	3	1	41	5	accuracy
	PAPE	1	0	4	45	

F1		TILE	ROCK	SAND	PAPE	
	TILE	20	7	22	1	
	ROCK	8	38	4	0	68%
	SAND	16	1	32	1	accuracy
	PAPE	1	0	3	46	

5.2.4 Summary

This sequence of experiments has served as a test of the channel filtering features on a variety of natural textures which were not preprocessed. Images of different sizes were extracted from large images which are visually interpreted as having different textures. The results show that a nearest-neighbor classifier using channel filtering features is able to identify the image classes with 90% or greater accuracy when the subimages are as small as 32x32. An attempt to extend this to 16x16 subimages failed when classes TILE and SAND proved indistinguishable to channel filtering features and error rates between other pairs of classes were also higher.

5.3 The Effect of Histogram Equalization

One common preprocessing technique applied in existing texture analysis studies is histogram equalization [Haralick et al, 1973]. The technique is used in texture studies to standardize the average gray level and contrast of the images, though the method has also been used for contrast enhancement. For some images, histogram equalization can dramatically change the image's appearance while for other images, histogram equalization has little or no effect. The procedure takes an image with G gray levels and an arbitrary gray level frequency histogram and produces an image with G' gray levels (G'<G) and a flattened gray level histogram [Rosenfeld and Kak, 1981].

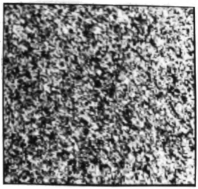
This series of experiments will determine the effect of histogram equalization on the texture of an image as measured by the channel filtering features. The image classes for this experiment will consist of the eight natural image classes from the previous experiments and equalized versions of those images which will be coded TILQ, ROCQ, SANQ, PAPQ, CORQ, GRAQ, WOOQ, and SCRQ (Figure 10). The equalized images are of size 256x256 with 128 gray levels (instead of 256 gray levels in the original images). The histogram equalization procedure is applied to the entire 256x256 image, not on each sample subimage. This serves to limit the effect of histogram equalization on the subimages since without this constraint the procedure could produce widely varying effects on different subimages of the same image class.

Of special importance is the fact that the 256x256 ROCQ image is almost identical to the 256x256 ROCK image. In experiments in which both image classes are used, we will require ROCK and ROCQ to be evaluated as the same texture class, and we will ignore misclassifications between them in computing error rates, leaving a seven-class discrimination problem.

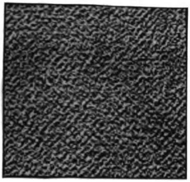
The experiments to be reported evaluate the performance of the feature space on classification problems involving the equalized images

(a)

(b)

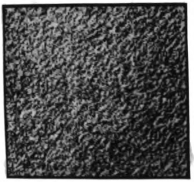


(c)

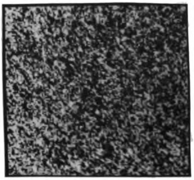


(d)

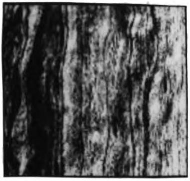
Figure 10: Histogram-Equalized Natural Images
(a) TILQ (b) ROCQ (c) SANQ (d) PAPQ
(e) CORQ (f) GRAQ (g) WOOQ (h) SCRQ



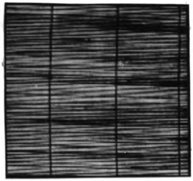
(e)



(f)



(g)



(h)

Figure 10: (cont'd)

alone and combined with the original image classes for subimage sizes 64x64 and 32x32.

5.3.1 Experiment 4: Histogram Equalized Images, 64x64 subimages

From the eight histogram-equalized images, 25 samples of size 64x64 are extracted, each of the 200 samples is filtered and features are computed from the filtered images. Using the nearest-neighbor classifier and estimating the error rate using the leave-one-out rule, the results shown in Table 4 are obtained. The PSE feature gives 63% correct classification. Feature 1 achieves 96% correct classification with all but one of the errors occurring between GRAQ and SANQ. This result indicates that 64x64 subimages are adequate to provide good discrimination of the histogram-equalized textures.

Table 4: Classification of Eight Histogram Equalized Natural Image Classes Using 64x64 Subimages.

PSE	TILQ	ROCQ	SANQ	PAPQ	CORQ	GRAQ	WOOQ	SCRQ	
TILQ	12	1	9	0	0	3	0	0	
ROCQ	3	16	1	0	0	1	4	0	
SANQ	2	0	12	0	1	9	1	0	
PAPQ	0	0	1	23	1	0	0	0	63%
CORQ	0	0	3	1	16	5	0	0	accuracy
GRAQ	1	0	10	0	6	7	1	0	
WOOQ	0	4	1	0	0	4	16	0	
SCRQ	1	0	0	0	0	0	0	24	

F1	TILQ	ROCQ	SANQ	PAPQ	CORQ	GRAQ	WOOQ	SCRQ	
TILQ	25	0	0	0	0	0	0	0	
ROCQ	0	24	0	0	0	0	1	0	
SANQ	0	0	21	0	0	4	0	0	
PAPQ	0	0	0	25	0	0	0	0	96%
CORQ	0	0	0	0	25	0	0	0	accuracy
GRAQ	0	0	3	0	0	22	0	0	
WOOQ	0	0	0	0	0	0	25	0	
SCRQ	0	0	0	0	0	0	0	25	

Another important question is whether the equalized images are perceived by the channel filtering features to be different from the

Table 5: Classification of TILE, ROCK, SAND, PAPE and TILQ, ROCQ, SANQ, PAPQ Using 64x64 Subimages. (percent accuracy reflects merging of ROCK and ROCQ)

PSE	TILE	ROCK	SAND	PAPE	TILQ	ROCQ	SANQ	PAPQ	
TILE	19	2	3	1	0	0	0	0	
ROCK	3	8	1	0	0	13	0	0	
SAND	5	0	17	3	0	0	0	0	
PAPE	0	0	2	22	1	0	0	0	75.5%
TILQ	2	0	0	1	14	0	8	0	accuracy
ROCQ	3	10	1	0	2	9	0	0	
SANQ	0	0	0	4	5	0	15	1	
PAPQ	0	0	0	0	0	0	1	24	

F1	TILE	ROCK	SAND	PAPE	TILQ	ROCQ	SANQ	PAPQ	
TILE	22	1	1	0	0	1	0	0	
ROCK	0	2	0	0	0	23	0	0	
SAND	0	0	25	0	0	0	0	0	
PAPE	0	0	0	25	0	0	0	0	98.5%
TILQ	0	0	0	0	25	0	0	0	accuracy
ROCQ	0	21	0	0	0	4	0	0	
SANQ	0	0	0	0	0	0	25	0	
PAPQ	0	0	0	0	0	0	0	25	

original images. Table 5 shows the results obtained by taking 64x64 subimages of TILE, ROCK, SAND, PAPE and their equalized versions. Ignoring misclassifications between classes ROCK and ROCQ as explained earlier, we find that the PSE feature yields 75.5% correct classification, and feature 1 yields 98.5% correct classification. Table 6 shows the results obtained by taking 64x64 subimages of CORK, GRAS, WOOD, and SCRE and their equalized counterparts. The PSE feature gives 68.5% classification accuracy while the channel filtering feature 1 yields 100% correct classifications. This result shows that except for the ROCK-ROCQ, histogram equalized images are perceived by the channel filtering features as having different textures from the original classes.

Table 6: Classification of CORK, GRAS, WOOD, SCRE and CORQ, GRAQ, WOOQ, SCRQ Using 64x64 Subimages.

PSE		CORK	GRAS	WOOD	SCRE	CORQ	GRAQ	WOOQ	SCRQ	
	CORK	13	9	0	0	3	0	0	0	
	GRAS	8	12	3	0	0	2	0	0	
	WOOD	0	4	19	0	0	1	1	0	
	SCRE	0	2	0	23	0	0	0	0	68.5%
	CORQ	2	0	0	0	17	6	0	0	accuracy
	GRAQ	4	2	0	0	7	11	1	0	
	WOOQ	0	2	6	0	0	2	15	0	
	SCRQ	0	0	0	0	0	0	0	25	

F1		CORK	GRAS	WOOD	SCRE	CORQ	GRAQ	WOOQ	SCRQ	
	CORK	25	0	0	0	0	0	0	0	
	GRAS	0	25	0	0	0	0	0	0	
	WOOD	0	0	25	0	0	0	0	0	
	SCRE	0	0	0	25	0	0	0	0	100%
	CORQ	0	0	0	0	25	0	0	0	accuracy
	GRAQ	0	0	0	0	0	25	0	0	
	WOOQ	0	0	0	0	0	0	25	0	
	SCRQ	0	0	0	0	0	0	0	25	

5.3.2 Experiment 5: Histogram Equalized Images, 32x32 subimages

The experiments of the previous section were repeated using 32x32 subimages. In the experiment using all eight histogram equalized image classes (Table 7), the PSE feature yields 64% correct classification. Feature 1 yields 80% correct classification, but 24 of the 40 misclassifications occur between SANQ and GRAQ, indicating that those classes are indistinguishable to this feature. Using 32x32 subimages from the original textures, only 6 misclassifications occurred between SAND and GRAS. This result indicates that the histogram equalization procedure confuses the differences which exist between the SAND and GRAS images, so that the histogram equalized versions cannot be discriminated by the channel filtering procedure using 32x32 subimages.

Table 7: Classification of Eight Histogram-Equalized Image Classes Using 32x32 Subimages.

PSE	TILQ	ROCQ	SANQ	PAPQ	CORQ	GRAQ	WOOQ	SCRQ	
TILQ	14	0	6	0	0	5	0	0	
ROCQ	2	13	1	0	0	2	5	2	
SANQ	4	0	14	1	1	5	0	0	
PAPQ	0	0	1	23	1	0	0	0	64%
CORQ	0	0	2	4	17	2	0	0	accuracy
GRAQ	2	3	8	1	1	9	1	0	
WOOQ	0	4	1	0	0	4	16	0	
SCRQ	1	1	0	0	0	1	0	22	

F1	TILQ	ROCQ	SANQ	PAPQ	CORQ	GRAQ	WOOQ	SCRQ	
TILQ	21	0	1	0	0	3	0	0	
ROCQ	0	20	0	0	0	0	5	0	
SANQ	1	0	12	0	0	12	0	0	
PAPQ	0	0	0	25	0	0	0	0	80%
CORQ	0	0	0	0	25	0	0	0	accuracy
GRAQ	2	0	12	0	0	11	0	0	
WOOQ	0	4	0	0	0	0	21	0	
SCRQ	0	0	0	0	0	0	0	25	

The classification results between the original and equalized versions of the image classes yielded results similar to those obtained

Table 8: Classification of TILE, ROCK, SAND, PAPE and TILQ, ROCQ, SANQ, PAPQ Using 32x32 Subimages. (percent accuracy reflects merging of ROCK and ROCQ)

PSE	TILE	ROCK	SAND	PAPE	TILQ	ROCQ	SANQ	PAPQ	
TILE	19	0	5	1	0	0	0	0	
ROCK	2	6	0	0	1	16	0	0	
SAND	2	0	20	3	0	0	0	0	
PAPE	1	0	1	23	0	0	0	0	80.5%
TILQ	0	0	0	0	15	0	10	0	accuracy
ROCQ	2	12	1	0	2	8	0	0	
SANQ	0	0	1	0	4	0	19	1	
PAPQ	0	0	0	0	0	0	2	23	

F1	TILE	ROCK	SAND	PAPE	TILQ	ROCQ	SANQ	PAPQ	
TILE	21	0	3	0	1	0	0	0	
ROCK	0	3	0	0	0	22	0	0	
SAND	2	0	23	0	0	0	0	0	
PAPE	0	0	0	25	0	0	0	0	95%
TILQ	1	0	0	0	22	0	2	0	accuracy
ROCQ	0	21	0	0	0	4	0	0	
SANQ	0	0	0	0	1	0	24	0	
PAPQ	0	0	0	0	0	0	0	25	

in the 64x64 experiment reported earlier. In the experiment using TILE, ROCK, SAND, PAPE and their equalized versions, and ignoring misclassifications between ROCK and ROCQ (Table 8) the PSE feature yields 80.5% correct classification and feature 1 yields 95% correct classification. Table 9 gives the results of the experiment using 32x32 samples from CORK, GRAS, WOOD, SCRE and their equalized counterparts. The PSE feature yields 69% correct classification results. Feature 1 yields 97% correct classification results with all misclassifications occurring between WOOD and WOOQ. Using 64x64 subimages, no misclassifications occurred between WOOD and WOOQ, suggesting that the channel filtering features have some difficulty discriminating these classes.

Table 9: Classification of CORK, GRAS, WOOD, SCRE and CORQ, GRAQ, WOOQ, SCRQ Using 32x32 Subimages.

PSE	CORK	GRAS	WOOD	SCRE	CORQ	GRAQ	WOOQ	SCRQ	
CORK	20	2	0	2	0	1	0	0	
GRAS	4	11	0	1	0	1	0	0	
WOOD	1	10	11	0	0	1	2	0	
SCRE	0	4	0	21	0	0	0	0	69%
CORQ	3	2	1	0	10	1	0	0	accuracy
GRAQ	3	7	2	1	1	10	1	0	
WOOQ	0	0	5	0	0	4	16	0	
SCRQ	0	0	0	1	0	1	0	23	

F1	CORK	GRAS	WOOD	SCRE	CORQ	GRAQ	WOOQ	SCRQ	
CORK	25	0	0	0	0	0	0	0	
GRAS	0	25	0	0	0	0	0	0	
WOOD	0	0	22	0	0	0	3	0	
SCRE	0	0	0	25	0	0	0	0	97%
CORQ	0	0	0	0	25	0	0	0	accuracy
GRAQ	0	0	0	0	0	25	0	0	
WOOQ	0	0	3	0	0	0	22	0	
SCRQ	0	0	0	0	0	0	0	25	

5.3.3 Summary

This section has presented experiments which evaluate the ability of the channel filtering features to discriminate histogram-equalized images from each other and from the corresponding unequalized images. The results show that histogram-equalized images are perceived by the channel filtering features to have different textures from the original images. In addition, histogram equalization was found to confuse some discriminable (unequalized) image classes. These observations imply that histogram equalization should be used with caution as a preprocessing procedure when channel filtering features are used for texture analysis.

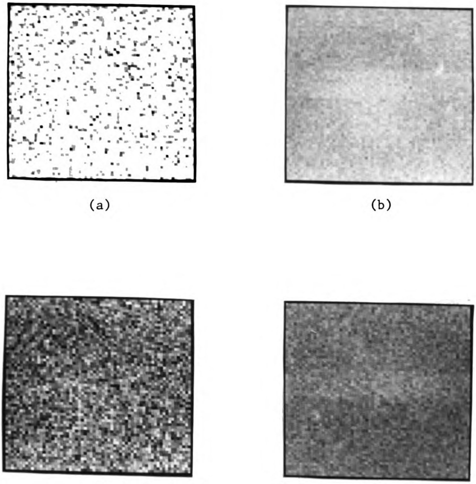
The performance of the PSE feature continues to be poor relative to the channel filtering features.

5.4 Experiment 6: Uniform Gray Level Changes

The conclusion of the previous section suggests that histogram equalization should not be used to preprocess images for the channel fiiltering features. Without preprocessing, however, irrelevant gray level variations can interfere with texture analysis unless the features are insensitive to such variations. Since the channel filtering features are defined relative to the average gray level of the image, we might expect them to be insensitive to changes in the average gray level. This experiment will test this hypothesis.

The data to be used for the following experiment consists of four artificial image classes in which the gray level at each pixel is an independent random variable from a Gaussian distribution. The means and variances are selected to enable the effects of changes in the mean and changes in the variance to be determined. The image classes are illustrated and the generation parameters are given in Figure 11. The parameter values used insure that three standard deviations about the mean gray level are within the allowed gray level range of 0 to 255. Gray levels generated outside that range were rejected.

Twenty-five subimages of size 64x64 for each class were used in this experiment. The results for the PSE feature and channel filtering feature 1 are given in Table 10. The PSE feature discriminates all four classes perfectly, indicating that the PSE feature is sensitive to both the variance and the average gray level of the images. This implies that the PSE feature is not a good texture measurement since average gray level changes should not affect the image texture unless



(a) (b)

(c) (d)

Figure 11: Gaussian White Noise Images
(a) GWN1: mean 160, sd 30 (b) GWN2: mean 160, sd 10
(c) GWN3: mean 96, sd 30 (d) GWN4: mean 96, sd 10

Table 10: Classification Results Showing the Effect of Average Gray Level Changes.

PSE	GWN1	GWN2	GWN3	GWN4
GWN1	25	0	0	0
GWN2	0	25	0	0
GWN3	0	0	25	0
GWN4	0	0	0	25

F1	GWN1	GWN2	GWN3	GWN4
GWN1	17	0	8	0
GWN2	0	9	0	16
GWN3	10	0	15	0
GWN4	0	10	0	15

the change is so severe that the gray level distribution is modified by limitations of the image acquisition system. In the evaluations by the channel filtering features, the classes which have different means but identical variances are confused. Feature 1 misclassifies 44% of the subimages, and all of the misclassifications are between classes with different means and identical variances. Classes with different variances are discriminated with 100% accuracy by the channel filtering features.

This experiment has demonstrated that the channel filtering features are insensitive to moderate variations in the average gray levels of the regions being analyzed. The features were found to be sensitive to differences in the gray level variance as might be expected since the variance determines the spectral energy, which is measured by the channel filtering features. Changes in the mean affect only the zero-frequency component and are eliminated by computing the features relative to the mean gray level of the filtered images. This result also verifies that the discrimination performance observed in earlier experiments using images which were not preprocessed were not due to average gray level differences between the image classes. The

PSE feature was found to be sensitive to both the mean and the variance of the gray level distribution. This enabled the PSE feature to discriminate all four Gaussian white noise classes when properties of human texture perception indicate that certain classes should be confused.

5.5 Experiment 7: Magnification Changes

This experiment will attempt to determine whether and how changes in the apparent size (magnification) of a textured image can be identified by the channel filtering method. The objective is to take a textured image and a magnified version of the same image and determine from the channel filtering features that the two images are from the same texture class.

The data for this experiment are taken from the TILE, ROCK, SAND and PAPE image classes used previously. Twenty-five subimages of size 32x32 from each class will be compared with twenty-five subimages of size 64x64 generated by expanding each pixel from the 32x32 subimages into a 2x2 pixel block. This means that the 64x64 images are 2X magnifications of the 32x32 subimages. The magnified image classes will be abbreviated TI2X, RO2X, SA2X, and PA2X.

The difference between images which is of interest in this experiment is not simply a change in image size. For example, a 64x64 image and another 64x64 image which portrays a magnified 32x32 area provide different coverage of the texture pattern, so the variance of the features may differ for the two classes. This difference could enable the nearest-neighbor classifier to discriminate the classes even

Table 11: Size Change Experiment Using 6 Spatial Frequency Channels.

PSE	TILE	ROCK	SAND	PAPE	TI2X	RO2X	SA2X	PA2X
TILE	18	0	4	3	0	0	0	0
ROCK	3	22	0	0	0	0	0	0
SAND	2	0	20	3	0	0	0	0
PAPE	1	0	2	22	0	0	0	0
TI2X	0	0	0	0	18	0	4	3
RO2X	0	1	0	0	3	21	0	0
SA2X	0	0	0	0	2	0	20	3
PA2X	0	0	0	0	1	0	3	21

F1	TILE	ROCK	SAND	PAPE	TI2X	RO2X	SA2X	PA2X
TILE	16	0	8	0	1	0	0	0
ROCK	0	11	0	0	0	14	0	0
SAND	5	0	20	0	0	0	0	0
PAPE	0	0	0	25	0	0	0	0
TI2X	1	0	0	0	17	0	7	0
RO2X	0	14	0	0	0	11	0	0
SA2X	0	0	0	0	5	0	20	0
PA2X	0	0	0	0	0	0	0	25

though they portray the same texture. Thus, we will require the classes to provide equivalent coverage of the texture pattern. This restriction makes the fundamental frequency of an object in the textured image appear in the same channel at both magnification levels. In the spatial frequency domain, the effect of this magnification is to add a high frequency channel which contains noise caused by the 2x2 pattern of the magnified pixels.

Another critical factor is the number of spatial frequency channels to use in the classifier. Orientation channels cannot be used; since they contain information from all spatial frequencies, they can provide discriminating information between image classes which portray the same texture. Clearly, the number of channnels to use for this problem must be less than or equal to the number of channels available from the smaller subimage size, in this case six. Table 11 shows that using six spatial frequency channels, all of the image classes are discriminable. This discrimination is based on the spatial frequency information available in the sixth channel. Channel 6 has a center frequency of 32 cycles per image, which is beyond the Nyquist frequency for 32x32 images. Thus, only a fraction of the spatial frequency information for channel 6 is present in the image. For 64x64 images, the center frequency of channel 6 is the Nyquist frequency, so more spatial frequency information is available in the channel.

The results after eliminating channel 6 from consideration, leaving five spatial frequency channels for the classifier, are shown in Table 12. The PSE feature provides overall classification accuracy of 81%. Using channel filtering features, the image classes which portray the same texture are completely confused in spite of the difference in magnification. Merging the image classes which portray the same textures yields a four class discrimination problem with 50 samples per class. In this four-class problem, feature 1 achieves

Table 12: Size Change Experiment Using 5 Spatial Frequency Channels.
(percent accuracy reflects merging of classes which portray the same texture)

PSE	TILE	ROCK	SAND	PAPE	TI2X	RO2X	SA2X	PA2X	
TILE	18	0	6	1	0	0	0	0	
ROCK	3	22	0	0	0	0	0	0	
SAND	2	0	20	3	0	0	0	0	
PAPE	1	0	3	21	0	0	0	0	81%
TI2X	0	0	0	0	18	0	6	1	accuracy
RO2X	0	1	0	0	3	21	0	0	
SA2X	0	0	0	0	2	0	20	3	
PA2X	0	0	0	0	1	0	3	21	

F1	TILE	ROCK	SAND	PAPE	TI2X	RO2X	SA2X	PA2X	
TILE	2	0	2	0	20	0	1	0	
ROCK	0	0	0	0	0	25	0	0	
SAND	2	0	2	0	0	0	21	0	
PAPE	0	0	0	10	0	0	0	15	95.5%
TI2X	22	0	0	0	2	0	1	0	accuracy
RO2X	0	25	0	0	0	0	0	0	
SA2X	2	0	20	0	1	0	2	0	
PA2X	0	0	0	16	0	0	0	9	

95.5% classification accuracy.

This experiment has demonstrated the ability of the channel filtering features to identify textures through magnification changes. Certain restrictions were imposed for processing images to test whether the images might portray the same texture at different magnifications. These restrictions were justified by the interpretations of the spatial frequency filters. Further research may provide insights which will relax these restrictions and enable simpler, more natural procedures for identifying texture through magnification changes.

5.6 Experiment 8: Orientation Changes

This experiment will test whether the channel filtering method can determine when two images portray the same texture at different orientations. The method to be used involves determining the minimum distance between patterns using any cyclical shift of the feature values from orientation channels. This minimum distance gives the best fit of one texture to another through the orientations tested.

The data for this experiment consists of four 256x256 images containing regular bar patterns at horizontal (Z000), vertical (Z090) and diagonal (Z045 and Z135) orientations (Figure 12). The bars are about 6 pixels wide. The results of nearest-neighbor classification on 25 samples per class using only the four orientation channels and without any shifting are shown in Table 13. The classes are perfectly discriminated by both the PSE feature and feature F1. Using the minimum distance between patterns produced by a circular shift of the orientation features, the results in Table 14 are obtained. This table shows that the circular shift enables both features to confuse the diagonal bar patterns and to confuse the horizontal/vertical bar patterns. The features are still able to separate the diagonal bar patterns from the horizontal/vertical bar patterns. The results indicate that the diagonal bar patterns are somehow different from the horizontal and vertical patterns. This may be attributed to the slightly different bar width caused by the rectangular quantization

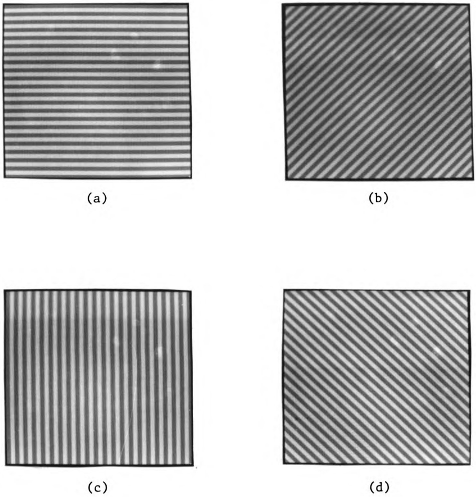


Figure 12: Bar Images for Orientation Experiment
(a) 0-degree bar pattern (b) 45-degree bar pattern
(c) 90-degree bar pattern (d) 135-degree bar pattern

Table 13: Orientation Experiment Using No Channel Shifting.

PSE	Z000	Z045	Z090	Z135
Z000	25	0	0	0
Z045	0	25	0	0
Z090	0	0	25	0
Z135	0	0	0	25

F1	Z000	Z045	Z090	Z135
Z000	25	0	0	0
Z045	0	25	0	0
Z090	0	0	25	0
Z135	0	0	0	25

grid and to the different bar lengths in the diagonal orientations due to the square image shape.

In this experiment, the PSE feature provides the same results as the channel filtering features. This is because the power spectral energy of all the subimages is concentrated in a single orientation channel which is different for each class. Since the differences between these image classes is apparent in the power spectrum, the PSE feature is able to duplicate the performance of the channel filtering features.

This experiment shows that a circular treatment of orientation channels can easily detect orientation differences of 90 degrees, but that due to the square shape of the images and the quantization grid,

Table 14: Orientation Experiment Using Circular Channel Shifting.

PSE		Z000	Z045	Z090	Z135
	Z000	0	0	18	7
	Z045	0	0	0	25
	Z090	0	0	18	7
	Z135	0	0	0	25

F1		Z000	Z045	Z090	Z135
	Z000	0	0	18	7
	Z045	0	0	0	25
	Z090	0	0	18	7
	Z135	0	0	0	25

45 degree orientation differences are more difficult to detect. The perceived difference between diagonal and horizontal/vertical channels might be overcome by computing features only within a circular subimage.

5.7 Experiment 9: Phase Spectrum Changes

In this experiment, the phase spectra of the Gaussian white noise subimages (Section 5.4) will be modified to determine the effect of such changes on the textural evaluations by the PSE feature and the channel filtering features. Only the phase spectra will be modified; the modified images will have the same power spectra as the Gaussian white noise images.

Determining the nature of the modifications which could be made was a difficult task. Since each entry in the phase spectrum affects the entire image, even slight perturbations can have drastic effects [Oppenheim and Lim, 1981]. The most troublesome constraint in generating the images for this experiment was that the phase-modified images had to remain in the gray scale 0 to 255. Otherwise, the phase-modified images would have to be scaled, and scaling changes the power spectrum of the image.

The method which was finally found satisfactory was to change the phase angles of entries near the zero-frequency component in the Fourier transform as illustrated in Figure 13. Due to the small number of entries changed and the low energy associated with those entries, the effect on the space domain image is minimal. The entries which are changed affect spatial frequency channels centered at 1, 2, 4, and 8 cycles per image. The orientation channels are also slightly affected.

Figure 14 shows the result of the changes in GWN1 (mean 160, standard deviation 30) which yield a new phase-modified image labelled PHM1. The figure also shows the original and modified filtered images for two spatial frequency channels. The effect of the phase modification is to increase the regularity of the filtered images, resulting in lightened or darkened patches at the center of each side and a barely visible diamond shape in the PHM1 image. Since only the lowest spatial frequency channels are affected by the modifications, the changes apparent in the PHM1 image are large in size. Due to the low energy in the affected spectral region, however, the contrast of the changes is small.

v

0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	π	π	π	π	π	π	π	π	$-\pi$	$-\pi$	$-\pi$	$-\pi$	$-\pi$	$-\pi$	$-\pi$	0
0	π	0	0	0	0	0	0	0	0	0	0	0	0	0	π	0
0	π	0	π	π	π	π	π	π	$-\pi$	$-\pi$	$-\pi$	$-\pi$	$-\pi$	0	$-\pi$	0
0	π	0	π	0	0	0	0	0	0	0	0	0	$-\pi$	0	$-\pi$	0
0	π	0	π	0	π	π	π	π	$-\pi$	$-\pi$	$-\pi$	0	$-\pi$	0	$-\pi$	0
0	π	0	π	0	π	0	0	0	0	0	$-\pi$	0	$-\pi$	0	$-\pi$	0
0	π	0	π	0	π	0	π	π	$-\pi$	0	$-\pi$	0	$-\pi$	0	$-\pi$	0
0	π	0	π	0	π	0	π	X	$-\pi$	0	$-\pi$	0	$-\pi$	0	$-\pi$	0
0	π	0	π	0	π	0	π	$-\pi$	$-\pi$	0	$-\pi$	0	$-\pi$	0	$-\pi$	0
0	π	0	π	0	π	0	0	0	0	0	$-\pi$	0	$-\pi$	0	$-\pi$	0
0	π	0	π	0	π	π	π	$-\pi$	$-\pi$	$-\pi$	$-\pi$	0	$-\pi$	0	$-\pi$	0
0	π	0	π	0	0	0	0	0	0	0	0	0	$-\pi$	0	$-\pi$	0
0	π	0	π	π	π	π	π	$-\pi$	$-\pi$	$-\pi$	$-\pi$	$-\pi$	$-\pi$	0	$-\pi$	0
0	π	0	0	0	0	0	0	0	0	0	0	0	0	0	$-\pi$	0
0	π	π	π	π	π	π	π	$-\pi$	$-\pi$	$-\pi$	$-\pi$	$-\pi$	$-\pi$	$-\pi$	$-\pi$	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

u

Figure 13: Phase Spectrum Modification Procedure.
Phase spectrum entries around the zero-frequency component (marked with an X) were replaced with the values shown above.

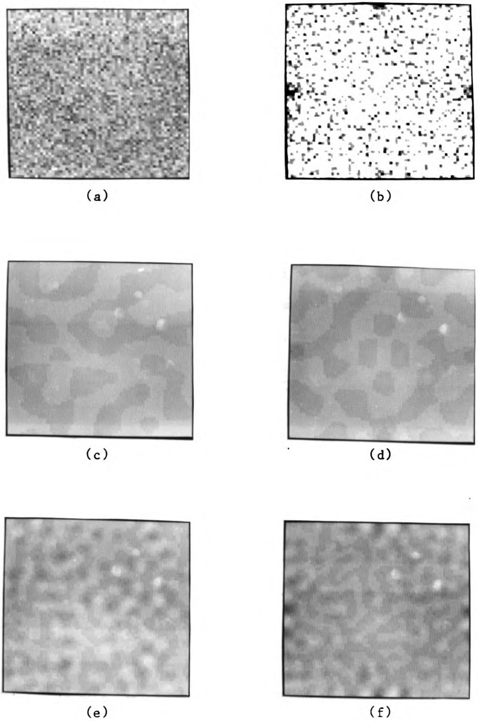


Figure 14: Effect of Phase Modifications on GWN1.
(a) GWN1. (b) PHM1. (c) Filtered image of GWN1 using the spatial frequency channel centered at 4 cyles per image (c/i).
(d) Filtered image of PHM1 using 4 c/i filter.
(e) Filtered image of GWN1 using 8 c/i filter.
(f) Filtered image of PHM1 using 16 c/i filter.

Table 15: Classification Results on Four Phase-Modified Gaussian White Noise Classes.

PSE	PHM1	PHM2	PHM3	PHM4
PHM1	25	0	0	0
PHM2	0	25	0	0
PHM3	0	0	25	0
PHM4	0	0	0	25

F1	PHM1	PHM2	PHM3	PHM4
PHM1	18	0	7	0
PHM2	0	9	0	16
PHM3	9	0	16	0
PHM4	0	14	0	11

In the classification experiment, the phase modification was performed separately on each of the 25 samples of the Gaussian white noise images. The classification results are given in Table 15. The PSE feature obtains perfect discrimination of the four classes. Channel filtering feature 1 does not discriminate image classes which came from Gaussian white noise images with identical means. This is an additional demonstration of the invariance of the channel filtering representation of texture to changes in average gray level. This result will be used to allow us to merge pairs of the phase modified classes in the next experiment.

The effect of the phase modification on the texture features will be demonstrated by using the Gaussian white noise classes and the phase-modified Gaussian white noise classes together in a classification experiment. The results of this experiment are shown in Table 16. The PSE feature confuses each Gaussian white noise class with its phase-modified version. This can be attributed to the fact that the PSE feature does not use phase, thus the differences which exist between the image classes are ignored. The results from channel filtering feature 1 are more clearly illustrated in Table 17, where the

Table 16: Classification Results Illustrating the Effect of Phase Specturum Changes on Gaussian White Noise Images.

PSE	GWN1	GWN2	GWN3	GWN4	PHM1	PHM2	PHM3	PHM4
GWN1	0	0	0	0	25	0	0	0
GWN2	0	3	0	0	0	22	0	0
GWN3	0	0	0	0	0	0	25	0
GWN4	0	0	0	0	0	0	0	25
PHM1	25	0	0	0	0	0	0	0
PHM2	0	22	0	0	0	3	0	0
PHM3	0	0	25	0	0	0	0	0
PHM4	0	0	0	23	0	0	0	2

F1	GWN1	GWN2	GWN3	GWN4	PHM1	PHM2	PHM3	PHM4
GWN1	15	0	7	0	3	0	0	0
GWN2	0	4	0	8	0	11	0	2
GWN3	9	0	13	0	0	0	3	0
GWN4	0	5	0	9	0	3	0	8
PHM1	3	0	0	0	16	0	6	0
PHM2	0	10	0	1	0	4	0	10
PHM3	0	0	3	0	8	0	14	0
PHM4	0	3	0	5	0	8	0	9

image classes which have been found indiscriminable to feature 1 in previous experiments are merged. Taking these identities into account leaves a four-class discrimination problem with 50 samples per class. In this four-class problem, feature 1 separates the classes with different variances. In addition, the phase-modified class with standard deviation 30 is separated from the unmodified class with standard deviation 30.

This experiment has illustrated the effects of a small phase spectrum change on textural evaluations using the channel filtering

Table 17: Phase Modification Experiment with Indiscriminable Classes Merged.

F1	GWN1 GWN3	PHM1 PHM3	GWN2 GWN4	PHM2 PHM4
GWN1,3	44	6	0	0
PHM1,3	6	44	0	0
GWN2,4	0	0	26	24
PHM2,4	0	0	19	31

method. The PSE feature is completely insensitive to phase changes, by definition. This was borne out in the results of this experiment. However, the phase modified images were discriminated from the original image classes by the channel filtering feature for the classes with standard deviation 30. Due to the lower energy in the images with standard deviation 10, the effect of the phase change in the image is decreased. The smaller effect is not detectable using feature F1.

5.8 Second-Order Statistics and Channel Filtering

In this section, the results of experiments designed to compare the performance of the co-occurrence and channel filtering methods will be presented. Four image classes will be generated using a procedure based on the four-disk method of [Caelli and Julesz, 1978a, 1978b]. Pairs of 9x9 micropatterns are generated by a geometrical method such that two images constructed of randomly rotated copies of the micropatterns will have identical second-order gray level distributions. In this experiment, however, the micropatterns will not be rotated through random angles; instead, rotations through random multiples of 90 degrees will be used. This results in images with slightly different second-order statistics and thus slightly different power spectra. The micropatterns to be used and the corresponding

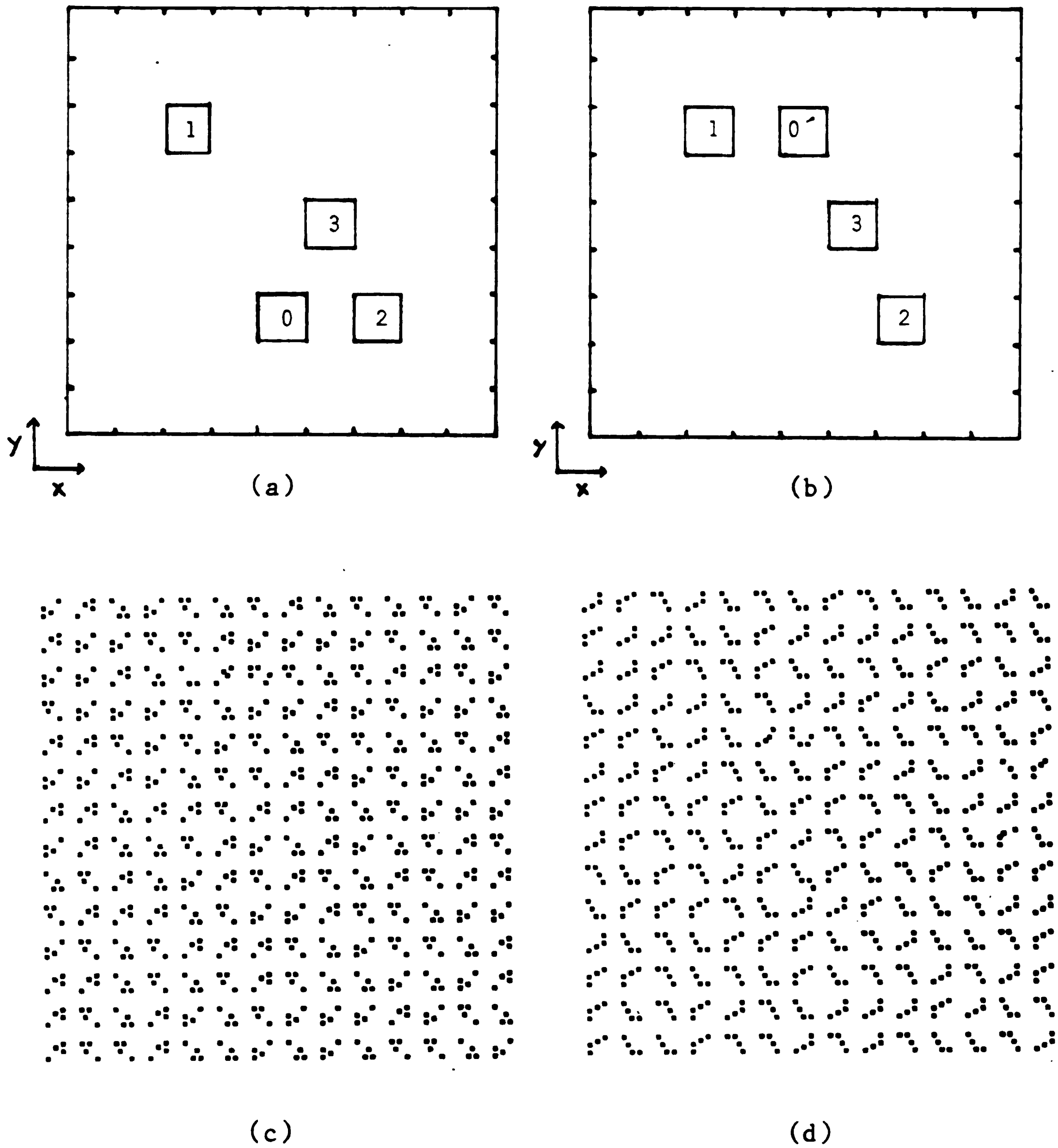
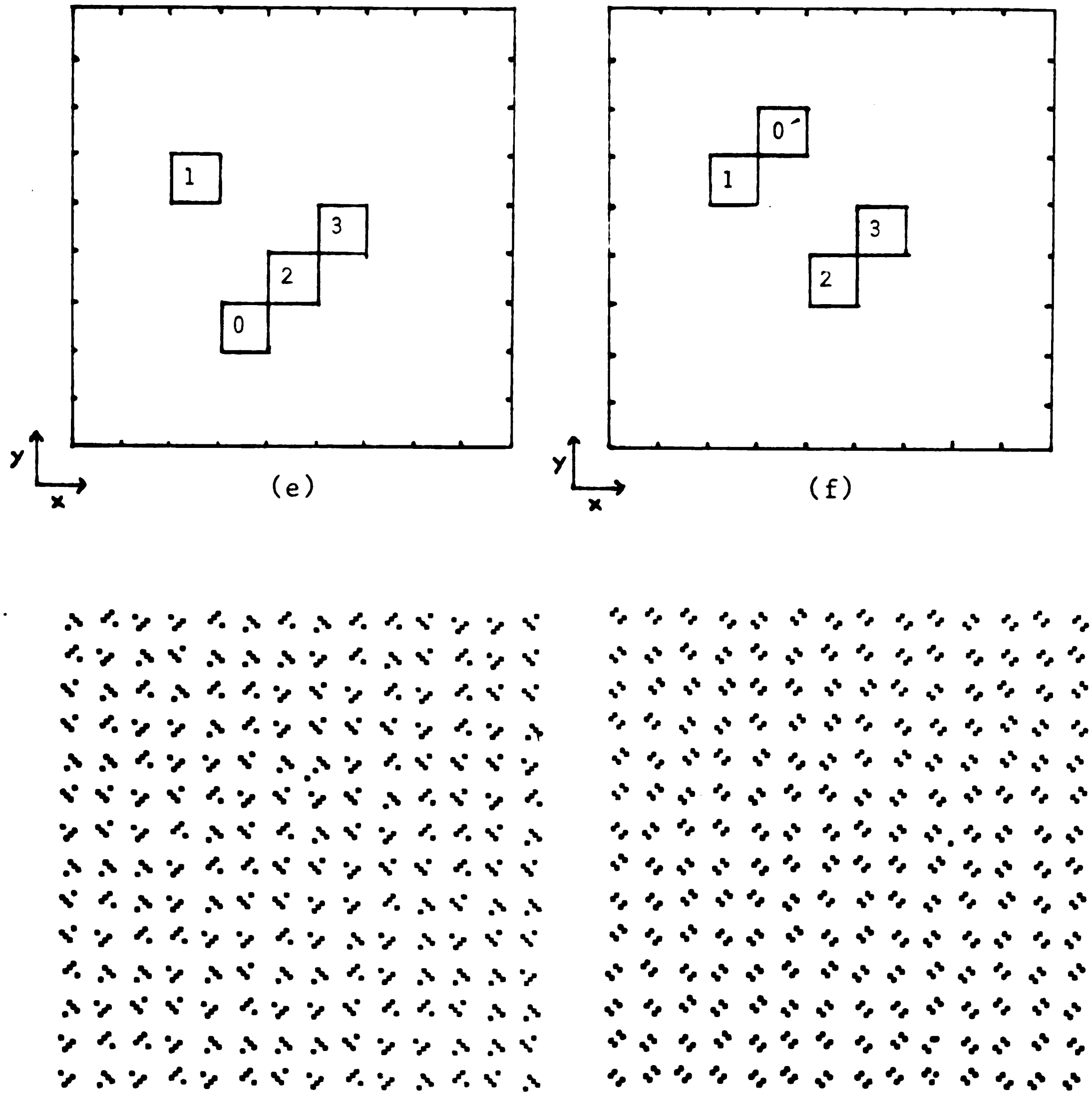


Figure 15: Four-Dot Micropatterns and Corresponding Image Classes. (The squares in the micropattern diagrams correspond to black pixels. The labels correspond to the usage in Table 18.)
(a) Micropattern 1A (b) Micropattern 1B (c) Part of image JU1A
(d) Part of image JU1B (e) Micropattern 2A (f) Micropattern 2B
(g) Part of image JU2A (h) Part of image JU2B



(e) (f)

(g) (h)

Figure 15 (cont'd)

image classes are illustrated in Figure 15. The image classes are visually discriminable.

5.8.1 Experiment 10: Application of the Co-occurrence Method

The definition of the co-occurrence matrix [Haralick, 1979] involves the selection of a displacement vector. Since there exist N^2 possible displacement vectors for an NxN image, some guidance in selecting the displacement vector is required. The choice of the displacement vector is critical. The displacement vectors to be used in this experiment will be chosen by a (suboptimal) strategy which takes advantage of known differences between the micropatterns.

The black pixels in the micropatterns in Figure 15 are numbered to demonstrate the similarities between micropatterns of each pair. The placement of only one dot (labelled 0 and 0') differs in the micropatterns of each pair. This implies that there exists a black-black co-occurrence which involves a different displacement

Table 18: Black-Black Displacement Vectors in Four-Dot Micropatterns.
(arrows indicate vectors which discriminate the micropatterns)

Micropattern 1A			Micropattern 1B	
Points	Displacement		Points	Displacement
0-1	(-2, 4)		0'-1	(-2, 0)
0-2	(2, 0)		0'-2	(2,-4)
0-3	(1, 2)	<********>	0'-2	(1,-2)
1-2	(4,-4)		1-2	(4,-4)
1-3	(3,-2)		1-3	(3,-2)
2-3	(-1,-2)		2-3	(-1,-2)

Micropattern 2A			Micropattern 2B	
Points	Displacement		Points	Displacement
0-1	(-1, 3)		0'-1	(-1,-1)
0-2	(1, 1)		0'-2	(1,-3)
0-3	(2, 2)	<********>	0'-3	(2,-2)
1-2	(2,-2)		1-2	(2,-2)
1-3	(3,-1)		1-3	(3,-1)
2-3	(1, 1)		2-3	(1, 1)

vector in the two micropatterns. This displacement vector will be used to discriminate the image classes. (Note: due to symmetry in the co-occurrence matrix definition, vectors (dx,dy) and (-dx,-dy) are considered to be identical.) Table 18 shows the displacement vectors between all pairs of black pixels in each micropattern. This information can be used to select an appropriate displacement vector to distinguish the micropatterns. For example, the (1,1) displacement vector can be used to discriminate the micropattern pairs. Micropatterns 2A and 2B contain black-black co-occurrences at displacement (1,1) but micropatterns 1A and 1B do not. Similarly, the (1,2) vector can discriminate the micropatterns of the first pair, and the (2,2) vector can discriminate the micropatterns of the second pair. These three displacement vectors, selected specifically to capture the differences between micropatterns, will be used to compute co-occurrence matrices for the textures.

For these displacement vectors, there are no interactions between different copies of the micropatterns for black-black co-occurrences. Thus, we can compute the expected number of black-black co-occurrences in the image from rotations of individual micropatterns as shown in Table 19. The number of micropatterns occurring in the 256x256 image is known (784), so we can estimate the number of micropatterns occurring at each of the four allowed orientations (196). By counting the number of black-black co-occurrences in each micropattern at each orientation, we can then compute for each displacement vector the

Table 19: Computation of Expected Number of Black-Black Co-occurrences in 256x256 Four-dot Images.

Number of black-black co-occurrences at each orientation | Expected no. black-black co-occurrences in image

DISPLACEMENT (1,1)=(-1,-1)

Micropattern	Rotation angle 0	90	180	270	total	Expected no.
1A	0	0	0	0	0	0*196 = 0
1B	0	0	0	0	0	0*196 = 0
2A	2	0	2	0	4	4*196 = 784
2B	2	0	2	0	4	4*196 = 784

DISPLACEMENT (1,2)=(-1,-2)

Micropattern	Rotation angle 0	90	180	270	total	Expected no.
1A	1	0	1	0	2	2*196 = 392
1B	0	0	0	0	0	0*196 = 0
2A	0	0	0	0	0	0*196 = 0
2B	0	0	0	0	0	0*196 = 0

DISPLACEMENT (2,2)=(-2,-2)

Micropattern	Rotation angle 0	90	180	270	total	Expected no.
1A	0	0	0	0	0	0*196 = 0
1B	0	0	0	0	0	0*196 = 0
2A	1	1	1	1	4	4*196 = 784
2B	0	2	0	2	4	4*196 = 784

expected number of black-black co-occurrences in the entire image.

The result of this computation shows that the expected number of black-black co-occurrences at displacement (1,2) is different for the first pair of images (Figure 15a). Thus, the number of black-black co-occurrences at displacement (1,2) should be able to discriminate image classes 1A and 1B. The other vectors do not provide discriminating information between 1A and 1B. In the second image pair (Figure 15b), the expected numbers of black-black co-occurrences are equal for all three displacement vectors. Even the (2,2) displacement vector, which could discriminate the micropatterns, cannot discriminate images 2A and 2B. The vectors which differ in the second pair of micropatterns have the same length, but their orientations are 90 degrees apart. Since the micropatterns are rotated by random multiples of 90 degrees, the difference between the micropatterns is confused in the images. These results use only black-black co-occurrences to help select reasonable displacement vectors for use in the co-occurrence method. The three vectors will now be used to compute co-occurrence matrices for use in classification problems.

The performance of the co-occurrence matrix method will now be demonstrated on 25 64x64 subimages from each class. Since the images are binary, the co-occurrence matrices are of size 2x2, and since the definition of the co-occurrence matrix to be implemented [Haralick, 1979] imposes symmetry on the matrix, the co-occurrence matrix involves three different values: the numbers of black-black, black-white, and white-white co-occurrences in the image. These numbers are used as features. The nearest-neighbor classification results are given in Table 20. The (1,1) displacement vector discriminates the two pairs but does not discriminate between textures in the same pair. The (1,2) vector can identify texture 1A, but the other classes are confused. The (2,2) vector discriminates between the pairs, but the discrimination within each pair is poor; the (2,2) vector was originally intended to discriminate the second image pair. However, since the black-black co-occurrences are known not to contribute to the discrimination of these classes, the discrimination which does occur can be attributed to white-black and white-white co-occurrences. The results using all three co-occurrence matrices combined are also shown.

Table 20: Classification of Four-dot Textures Using Co-occurrence Matrices.

(1,1)	JU1A	JU1B	JU2A	JU2B
JU1A	2	23	0	0
JU1B	2	23	0	0
JU2A	0	0	1	24
JU2B	0	0	9	16

(1,2)	JU1A	JU1B	JU2A	JU2B
JU1A	25	0	0	0
JU1B	0	1	4	20
JU2A	0	2	4	19
JU2B	0	2	0	23

(1,3)	JU1A	JU1B	JU2A	JU2B
JU1A	1	24	0	0
JU1B	7	18	0	0
JU2A	0	0	18	7
JU2B	0	0	9	16

ALL	JU1A	JU1B	JU2A	JU2B
JU1A	25	0	0	0
JU1B	0	25	0	0
JU2A	0	0	13	12
JU2B	0	0	7	18

This result uses nine features; the three values in each of the three co-occurrence matrices. The images of the first pair are identified, but the second pair is not well separated.

Table 21: Classification of Four-dot Textures by Channel Filtering.

PSE	JU1A	JU1B	JU2A	JU2B
JU1A	11	14	0	0
JU1B	17	8	0	0
JU2A	0	0	17	8
JU2B	0	0	9	16

F1	JU1A	JU1B	JU2A	JU2B
JU1A	25	0	0	0
JU1B	0	25	0	0
JU2A	0	0	25	0
JU2B	0	0	0	25

5.8.2 Experiment 11: Application of Channel Filtering

The channel filtering approach was applied to the same samples of the four four-disk textures. The classification results are shown in Table 21. The PSE feature discriminates the different pairs, but cannot discriminate the textures within each pair. Feature 1 achieves 100% correct classification on all four classes. In this experiment, it is of some interest to determine whether the discrimination of the texture classes is due to the action of orientation channels, spatial frequency channels, or both. Since these images are composed of dots, it might seem that no orientation or size selectivity exists between the image classes. Table 22 gives the classification results using only the six spatial frequency channels, and Table 23 gives the results using only the four orientation channels. The results show that both orientation and spatial frequency information contribute to the discrimination of the classes. These contributions can be understood by considering the interpretation of spatial filtering as a moving average operation. Each gray level in a filtered image is the result

Table 22: Classification of Four-dot Textures Using Six Spatial Frequency Channels.

PSE		JU1A	JU1B	JU2A	JU2B
	JU1A	12	13	0	0
	JU1B	13	12	0	0
	JU2A	0	0	21	4
	JU2B	0	0	7	18

F1		JU1A	JU1B	JU2A	JU2B
	JU1A	24	1	0	0
	JU1B	0	25	0	0
	JU2A	0	0	24	1
	JU2B	0	0	1	24

of an averaging operation over an area of the original image defined by the filter point spread function. The result of this averaging can be expressed for these particular images as a consequence of the local dot density. Variations in the dot density in local areas of different sizes and orientations provides the discriminating information between the image classes. Since the number of dots in a local area determines the average gray level of the area (which is a component of the power spectrum), this interpretation is consistent with the characterization of the channel filtering features as "average local energy" measures.

Table 23: Classification of Four-dot Textures Using Four Orientation Channels.

PSE	JU1A	JU1B	JU2A	JU2B
JU1A	13	12	0	0
JU1B	13	12	0	0
JU2A	0	0	18	7
JU2B	0	0	9	16

F1	JU1A	JU1B	JU2A	JU2B
JU1A	25	0	0	0
JU1B	3	22	0	0
JU2A	0	0	25	0
JU2B	0	0	0	25

5.8.3 Summary

This series of experiments compared the performance of the co-occurrence method to the channel fitering method for discriminating a particular set of textured images. A procedure for selecting displacement vectors for co-occurrence matrices was applied. The displacements were chosen to detect known differences in the micropattern structures. For one image pair, the combined co-occurrence results were satisfactory; for the other, the co-occurrence matrices performed poorly. The method for selecting the displacement vectors was not optimal or exhaustive, but the results demonstrate that the selection of displacement vectors is critical and that the performance of a particular displacement vector cannot be reliably predicted from its performance on micropatterns. In these experiments, only certain black-black co-occurrences could be investigated in detail. Some results indicated that the white-white and black-white co-occurrences might also contribute to class discriminations. The problem of selecting suitable displacement vectors is a major disadvantage to the use of the co-occurrence method.

Channel filtering feature F1 yielded 100% classification results. Both spatial frequency and orientation channels were found to contribute to this result. The performance of the PSE feature indicates that the different image pairs have significantly different power spectra, but the power spectra of the images within each pair are not discriminable. These observations imply that the performance of feature F1 is not due to power spectral differences; it is the use of phase information which enables discrimination of all of the classes.

5.9 Computational Simplifications

In this section, two computational simplifications of the channel filtering approach will be tested to determine if such simplifications cause any significant degradation in classification performance.

5.9.1 Experiment 12: Using Fewer Channels

One possible simplification is to use fewer channels. Since texture is perceived in regions containing small areas of constant gray level, the low-frequency channels, which respond to gray level variations over large areas, might be eliminated without degrading performance. In addition, the highest spatial frequency channel does not capture any spectral information which is not available in other channels.

In this experiment, the two lowest spatial frequency channels and the highest spatial frequency channel will be eliminated. The 64x64 images will be filtered by four orientation channels and four spatial frequency channels (center frequencies 4, 8, 16, and 32 cycles per image). The results of using these eight channels on the eight natural image classes are given in Table 24. The 65.5% correct performance of the PSE feature is slightly worse than its performance using all 11 channels. Feature 1 provides 98% correct classification, the same as with all of the channels present.

Using only eight spatial frequency channels does not appear to degrade the performance of the channel filtering features. The minimum number of channels required for texture discrimination depends on the particular set of textures.

5.9.2 Experiment 13: Ideal Band-Pass Channels

The filtering operation using Gaussian filters requires that the filter amplitudes and the image spectrum be multiplied. If the filter

Table 24: Classification Results on Natural Images Using Eight Channels.

PSE	TILE	ROCK	SAND	PAPE	CORK	GRAS	WOOD	SCRE	
TILE	17	2	4	1	0	0	1	0	
ROCK	3	15	1	0	2	0	0	4	
SAND	4	0	12	2	0	0	7	0	
PAPE	0	0	2	22	1	0	0	0	65.5%
CORK	0	0	0	0	16	9	0	0	accuracy
GRAS	0	1	1	0	8	12	2	1	
WOOD	2	0	5	1	0	4	13	0	
SCRE	0	0	0	0	0	1	0	24	

F1	TILE	ROCK	SAND	PAPE	CORK	GRAS	WOOD	SCRE	
TILE	22	1	2	0	0	0	0	0	
ROCK	0	25	0	0	0	0	0	0	
SAND	1	0	24	0	0	0	0	0	
PAPE	0	0	0	25	0	0	0	0	98%
CORK	0	0	0	0	25	0	0	0	accuracy
GRAS	0	0	0	0	0	25	0	0	
WOOD	0	0	0	0	0	0	25	0	
SCRE	0	0	0	0	0	0	0	25	

amplitudes were all either zero or one, the multiplications could be avoided. In this experiment, channels with an ideal band-pass characteristic will be applied to the natural image classes.

The spatial frequency channels to be used are the same as those defined for power spectral energy features in [Weszka et al, 1976]. The channels pass all spatial frequencies in nonoverlapping bands whose lower and upper bounds, in cycles per image, are as follows: [2,4], [4,8], [8,16] and [16,32]. The orientation channels to be used pass all spectral information within 45 degrees of horizontal, vertical, and both diagonal orientations. The use of ideal band-pass filters causes additional ripples in the filter responses due to the abrupt filter cutoffs in the spatial frequency domain. The response is still symmetrical about the mean gray level, however. In [Marr and Hildreth, 1979] these ripples were found to interfere with edge and form detection in filtered images. Table 25 shows that the ideal band-pass

Table 25: Classification Results on Natural Images Using Ideal Bandpass Channels.

PSE	TILE	ROCK	SAND	PAPE	CORK	GRAS	WOOD	SCRE	
TILE	21	1	3	0	0	0	0	0	
ROCK	4	20	0	0	0	0	0	1	
SAND	4	0	19	0	0	0	2	0	
PAPE	0	0	2	22	0	1	0	0	77%
CORK	0	0	0	0	19	6	0	0	accuracy
GRAS	0	0	2	0	9	14	0	0	
WOOD	1	0	6	0	1	2	15	0	
SCRE	0	0	0	0	0	1	0	24	

F1	TILE	ROCK	SAND	PAPE	CORK	GRAS	WOOD	SCRE	
TILE	23	1	1	0	0	0	0	0	
ROCK	1	24	0	0	0	0	0	0	
SAND	1	0	24	0	0	0	0	0	
PAPE	0	0	0	25	0	0	0	0	98%
CORK	0	0	0	0	25	0	0	0	accuracy
GRAS	0	0	0	0	0	25	0	0	
WOOD	0	0	0	0	0	0	25	0	
SCRE	0	0	0	0	0	0	0	25	

filters do not seem to degrade texture analysis performance; feature 1 provides 98% correct classification results. The PSE feature achieves 77% correct results, which is better than the 69% performance recorded using Gaussian channels. The PSE feature gave similar performance in another study using ideal band-pass filters. [Weszka et al, 1976].

5.10 Summary

This chapter has presented the results of experiments designed to demonstrate properties of the channel filtering feature space. The performance of the method on natural and artificial image classes, on images which differ in average gray level, magnification, orientation, and phase spectra was examined. Some computational simplifications of the channel filtering method were also presented and found not to seriously degrade the performance of the features. Throughout the experiments, the power spectral energy in each channel was used as a feature to demonstrate the effect of using phase information in the channel filtering features. The channel filtering features provide superior performance. A comparison of the channel filtering method with the co-occurrence method on one set of images demonstrated that even when the displacement vectors for the co-occurrence matrix are chosen specifically to capture differences between the micropatterns in a textured image, the channel filtering method outperforms the co-occurrence matrix method.

Chapter 6

Evaluation of Channel Filtering Features for Texture Segmentation

6.1 Segmentation

The previous chapter evaluated the channel filtering features in texture classification problems. In more realistic situations, the images being analyzed contain an unknown number of textured regions which must be identified, thus segmenting the image. This texture segmentation problem is different from a general image segmentation problem in which the objective is to identify objects in a scene.

Segmenting an image into textured regions is more difficult than classifying textured images in three ways. First, the number of classes is specified in advance in classification problems whereas in segmentation problems the number of classes is unknown. This requires a segmentation algorithm to include some means for determining the actual (or at least an appropriate) number of classes from the data. Second, the objects being classified in classification problems are subimages. In segmentation problems, the objects to be classified are individual pixels. The number of classifications required in segmentation is very large. For example, segmenting a 128x128 image requires each of the 16,384 pixels to be classified. This large number of classifications will require some simplifications in the methods to be used. Third, in classification experiments the subimages being classified are known to contain a single texture, but in segmentation experiments, the neighborhoods of some pixels will involve more than one texture.

The segmentation procedure in this chapter will require a determination of the texture in a neigborhood about each pixel. A channel filtering feature will be used for this purpose. A feature vector will be computed for each pixel in the image. The number of texture classes present will be determined from the feature vectors, and all of the pixels in the image will be classified. Figure 16 shows a block diagram of the segmentation procedure.

6.2 Computing Texture Features for Segmentation

Ideally, a "resolution-preserving textural transform" [Haralick, 1975] would be applied to the image which would replace the gray levels by texture feature values computed over some neighborhood about each pixel. Such an ideal is often not practically attainable, but a reasonable solution exists using the channel filtering approach. By applying two operations to the filtered images, the gray level at each pixel of the filtered images can be replaced by a new gray level which is related to the value of channel filtering feature 1. In the first step, we replace each gray level as follows:

$$Ik'(x,y) = 2 * |\bar{G} - Ik(x,y)|$$

where $Ik(.,.)$ is the kth filtered image and $\bar{G}$ is the mean gray level of the filtered images. This step produces an image, $Ik'(.,.)$, in which the gray levels are related to the absolute deviation from the

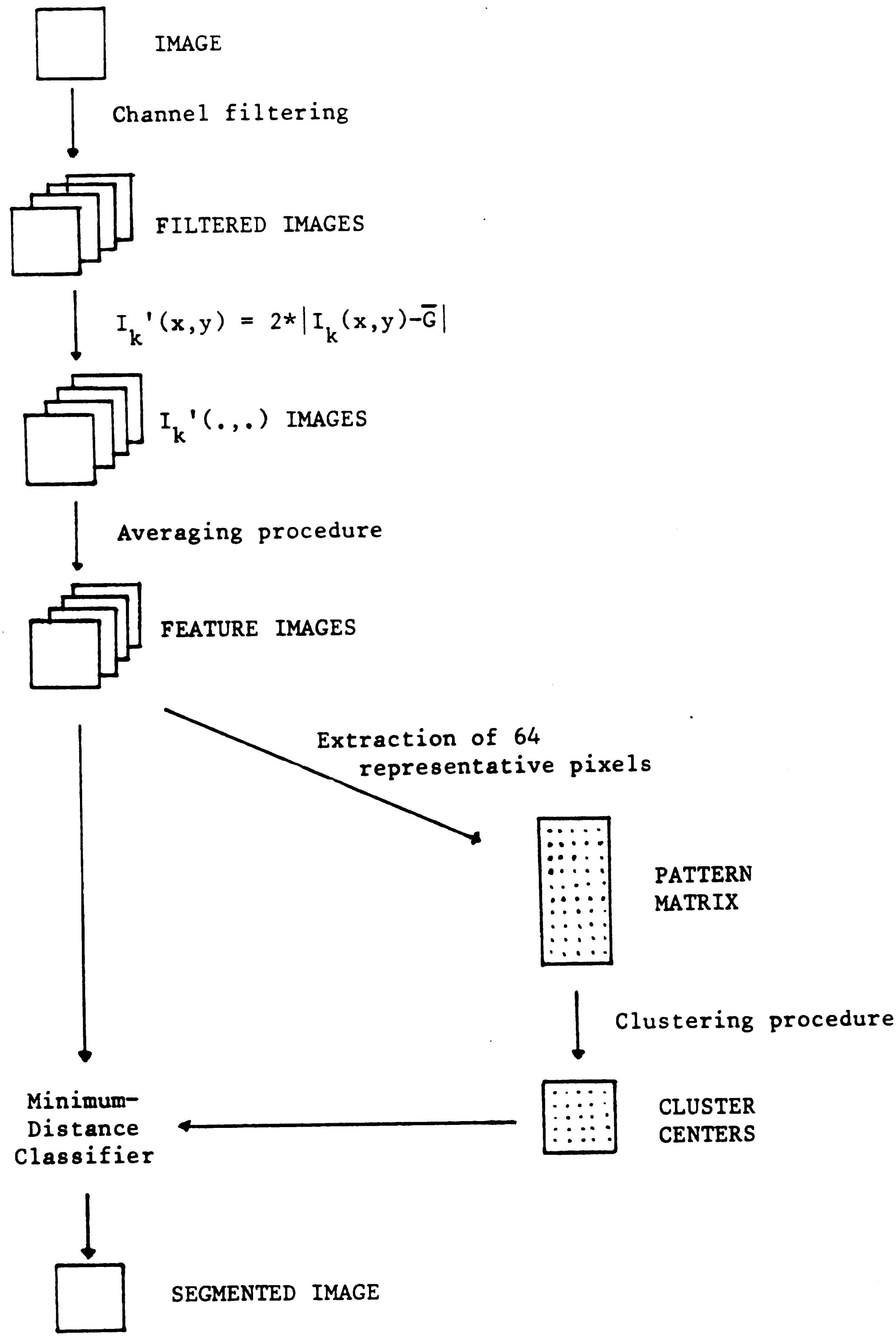


Figure 16: Diagram of the Segmentation Procedure.

mean gray level. Simply taking the absolute deviation would result in an image with about half the original number of gray levels. Since gray levels are integer values, the absolute deviation is scaled to increase the precision of the averaging step which occurs next. In the experiments, $\bar{G}/2$ was added to the pixel values to enhance the visibility of the $\mathcal{I}k'(.,.)$ image.

The second step involves computing a moving average over the $\mathcal{I}k'(.,.)$ images. In the definition of channel filtering feature 1, the absolute deviation from the mean is averaged over the entire subimage. For segmentation experiments, the averaging is performed over small neighborhoods about each pixel. Therefore, we wish to replace each pixel in $\mathcal{I}k'(.,.)$ with the average gray level in a neighborhood about the pixel. This can be accomplished either by a convolution in the spatial domain or by a filtering operation in the spatial frequency domain. The window which defines the neighborhood must be large enough to capture an adequate texture sample, but not so large that transitions between different textures are blurred over a large area. The result of the averaging is a "feature image" in which the gray level at each pixel is a measure of the texture present at the corresponding location of the original image.

The experiments in this chapter will use 8x8 and 16x16 square windows in computing the feature images. In these experiments, the spatial domain averaging method is used.

The feature images defined here correspond to the "texture energy planes" of Laws [1980]. In that study, the filters were defined as spatial domain templates which were convolved with the original image. The filtered images had an average gray level of zero, resulting in a slightly different computational procedure for producing the feature images. The present study differs from Laws [1980] in the use of spatial frequency domain filters rather than spatial domain templates, in the decomposition of the image by isolating bands of spatial frequency and orientation rather than by detection of "edges", "spots", and "rings", and in the use of results from vision science to guide the development of the computational methods.

Note that the computation of the feature images does not require each neighborhood to be filtered separately. This contrasts with some earlier approaches to texture segmentation in which the entire computational procedure is repeated for each neighborhood considered [Haralick, 1975; Bajcsy, 1973; Bajcsy and Lieberman, 1976].

6.3 Segmentation Using Feature Images

The feature images provide an evaluation of the texture in small neighborhoods about each pixel in the image. Corresponding to each pixel is a feature vector in which the number of features equals the number of channels used. The texture segmentation problem has now been transformed into the feature space. The next problem is to assess the structure of the patterns in the feature space. This assessment will be made by applying a clustering algorithm [Anderberg, 1973; Everitt, 1974]. Previous uses of clustering in segmentation problems are discussed in [Mitchell and Carlton, 1978; Schachter et al, 1978b; Rosenfeld, 1981; Coleman, 1979; Davis and Mitiche, 1982]

The clustering algorithm to be used, called CLUSTER [Dubes and Jain, 1976], is a partitional clustering procedure which attempts to mimimize a squared-error criterion. CLUSTER was selected because the clusters it produces are not limited to hierarchical relationships and because this algorithm does not require user-specified parameters. CLUSTER provides partitionings of the data with the number of clusters going from 2 to a user-specified bound (8 in this study). Each partitioning of the data corresponds to a segmentation of the given image. It will be necessary to evaluate these clusterings to determine which ones are appropriate representations of the data, and several statistics are provided by CLUSTER for this purpose.

Clustering algorithms are computationally very demanding. Most of them are designed to cluster only a few hundred points. To segment a 128x128 image using clustering alone would require the algorithm to process 16,384 points. To reduce computational requirements, only 64 pixels spaced 16 rows and 16 columns apart will be clustered. This simplification assumes that the textured regions to be segmented will not be irregularly shaped or very small in area. Additionally, only eight channels, the four orientation channels and spatial frequency channels 3-6 (center frequencies 4, 8, 16, and 32 cycles per image) will be used. These eight channels were found in section 5.9.1 to provide good classification results. This provides a sample of 64 points in 8 dimensions for clustering. The cluster centers will be used to define a minimum-distance classifier to classify the remaining pixels. The classification results will be displayed as a segmented image in which gray levels denote the cluster labels assigned to each pixel.

Note that the segmentation is completely determined by the structure of the points in the feature space according to the clustering algorithm. This differs from interpretation-guided segmentation [Barrow and Tennenbaum, 1981] where semantic information is used to guide the segmentation process.

6.4 Evaluating the Segmentations

The number of segments obtained in an image depends on the clustering algorithm. A problem facing users of any clustering algorithm is the question of cluster validity [Dubes and Jain, 1979]. Since any clustering algorithm will produce clusters regardless of the distribution of points in the feature space, the user must determine whether the partitioning obtained is a consequence of structure in the points or an artifact of the clustering algorithm. A related problem is to determine the actual number of clusters present in the data.

CLUSTER provides several statistics which can be used to qualitatively assess the validity of a clustering. One statistic, which measures the spread of a cluster in the feature space is the average within-cluster distance which is defined for cluster k as follows:

$$CLAVGD(k) = \left[1/N(k) * \sum_{i=1}^{N(k)} \sum_{j=1}^{D} [X(i,j) - C(k,j)]^2 \right]^{1/2}$$

where N(k) is the number of points in cluster k, C(k,.) is the cluster center for cluster k, D is the dimensionality of the feature space, and X(i,j) is the value of the jth feature for the ith pattern.

A measure of the "validity" of cluster k which takes into account the compactness and the isolation of the cluster is defined as

$$S(k) = \frac{\min\limits_{\substack{\ell \\ \ell \neq k}} \sum_{j=1}^{D} [C(k,j) - C(\ell,j)]^2}{CLAVGD(k)} .$$

Large values of S(k) indicate compact, well-isolated clusters. This statistic will be used to determine the acceptability of a clustering. An acceptable clustering is one in which the value of S(k) for all clusters exceeds a threshold. We have empirically determined a threshold of 1.70 for S(k) by tuning the threshold in preliminary experiments to yield approximately the number of clusters perceived by preattentive human vision. This particular value is only an approximation; we found a threshold of 2.0 to reject too many reasonable clustering solutions and a threshold of 1.5 to accept too many clusterings.

The threshold value requires that for each cluster the minimum distance to another cluster center be at least 1.70 times the average within-cluster distance. The clusterings which are accepted by this criterion will be ranked by the value of the average of S(k) over all clusters weighted by the number of points in the clusters. Clusterings with higher average S(k) values will be preferred.

6.5 Segmentation Experiment 1: Dot Textures

The data for this experiment consists of a 128x128 binary image illustrated in Figure 17. The left half of the image is a regular dot pattern in which the dots are separated by 3 pixels from their

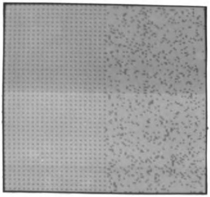
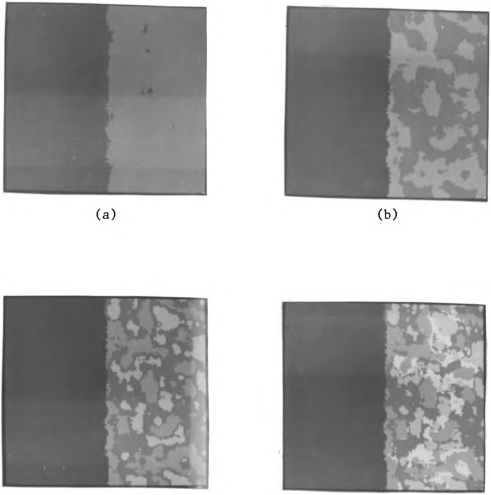


Figure 17: Dots Image for Segmentation Experiment 1.

horizontal and vertical neighbors. The right half of the image is a random dot pattern in which the probability that each pixel is black is independent of the other pixels and is approximately equal to 1/9. (The exact probability is the number of dots in the regular texture divided by half the number of pixels in the image, which results in identical average gray values for both textures). This image was segmented using 8x8 and 16x16 windows in the averaging step for computing the feature images.

The segmented images produced using 8x8 windows are shown in Figure 18, and the segmented images produced using 16x16 windows are shown in Figure 19. The figures show only the segmentations for 2, 3, 4, and 5 clusters because the gray levels used for labelling each segment become too difficult to see when more classes are present. Note that the actual gray levels in the segmented images have no significance other than to distinguish the regions.

The two-cluster segmentations for both window sizes accurately distinguish the regions of different texture; over 98% of the pixels are correctly labelled. The 8x8 segmentation contains a few small areas in the random texture which are classified with the regular texture. These misclassifications do not appear in the 16x16 segmentation because the probability that a region of random dots will resemble a regular texture is lower for larger regions. The segmentation results corresponding to more than two clusters break up the random texture into irregularly shaped regions. The regular texture is not subdivided. None of the segmented regions lie across the boundary between the textures.



(a) (b)

(c) (d)

Figure 18: Segmented Images for Dots Using 8x8 Averaging Windows. (a) 2-cluster solution (b) 3-cluster solution (c) 4-cluster solution (d) 5-cluster solution

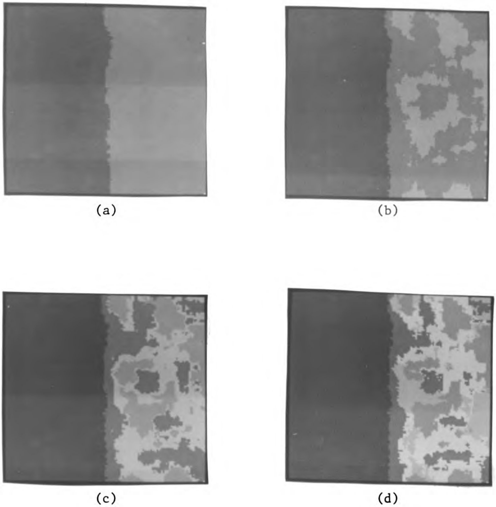


Figure 19: Segmented Images for Dots Using 16x16 Averaging Windows. (a) 2-cluster solution (b) 3-cluster solution (c) 4-cluster solution (d) 5-cluster solution

We now need to determine which of the segmentations are reasonable using the validity test defined earlier. The number of points in each cluster, the value of S(k) for each cluster, and the weighted average values of S(k) are shown in Table 26 for the 8x8-averaged feature images and in Table 27 for the 16x16-averaged feature images. In both

Table 26: Evaluation of Clustering on Dots Image
Using 8x8 Averaging Windows.
(* indicates accepted clustering solutions)

	k	N(k)	S(k)	
2 CLUSTERS				
	1	32	6.93	
	2	32	2.31	
	AVG		4.62	*
3 CLUSTERS				
	1	32	6.17	
	2	19	1.41	
	3	13	1.24	
	AVG		3.76	
4 CLUSTERS				
	1	32	6.08	
	2	20	1.48	
	3	8	1.15	
	4	4	2.20	
	AVG		3.78	
5 CLUSTERS				
	1	32	6.08	
	2	13	1.05	
	3	8	1.14	
	4	4	2.16	
	5	7	0.975	
	AVG		3.64	
6 CLUSTERS				
	1	32	5.75	
	2	7	1.53	
	3	3	1.19	
	4	11	1.24	
	5	6	1.07	
	6	5	1.47	
	AVG		3.53	

	k	N(k)	S(k)
7 CLUSTERS			
	1	32	5.75
	2	3	2.19
	3	3	1.19
	4	9	1.34
	5	9	1.13
	6	5	1.31
	7	3	1.53
	AVG		3.55
8 CLUSTERS			
	1	32	5.75
	2	3	1.98
	3	3	1.19
	4	9	1.41
	5	3	1.66
	6	5	1.31
	7	3	1.53
	8	6	1.29
	AVG		3.59

cases, only the two-cluster solution is accepted. The tables also show that the S(k) value for cluster 1, which corresponds to the regular texture, is always very high. Since the cluster of points from the regular texture is always compact and well-isolated, the regular texture is never subdivided in the segmented images.

Table 27: Evaluation of Clustering on Dots Image
Using 16x16 Averaging Windows.
(* indicates accepted clustering solutions)

k	N(k)	S(k)	
2 CLUSTERS			
1	32	9.39	
2	32	3.79	
AVG		6.59	*
3 CLUSTERS			
1	32	8.86	
2	22	1.33	
3	10	1.62	
AVG		5.14	
4 CLUSTERS			
1	32	8.27	
2	11	1.22	
3	9	1.39	
4	12	1.28	
AVG		4.78	
5 CLUSTERS			
1	32	8.05	
2	8	1.39	
3	8	1.53	
4	6	1.56	
5	10	1.26	
AVG		4.73	
6 CLUSTERS			
1	32	8.06	
2	7	1.45	
3	7	1.63	
4	6	1.58	
5	7	1.14	
6	5	2.97	
AVG		4.87	

k	N(k)	S(k)
7 CLUSTERS		
1	32	8.06
2	7	1.39
3	8	1.58
4	5	1.84
5	2	1.77
6	7	1.71
7	3	1.35
AVG		4.83
8 CLUSTERS		
1	32	8.06
2	7	1.59
3	8	1.58
4	5	1.93
5	2	1.87
6	6	1.97
7	2	1.91
8	2	1.91
AVG		4.91

6.6 Segmentation Experiment 2: Gaussian White Noise

This experiment will apply the segmentation procedure to a 128x128 image which is visually perceived to contain a single homogeneous texture. The gray level at each pixel is generated independently using a Gaussian distribution with mean 128 and standard deviation 30. The image is illustrated in Figure 20. The same filtering and segmentation procedure as in the previous experiment was applied. The segmented images from 8x8-averaged feature images are shown in Figure 21 and the results from 16x16-averaged feature images are shown in Figure 22. In both cases, the images are segmented into irregular patches. The patches for the 8x8 case are generally smaller than those for the 16x16 case.

The statistics for evaluating the clusterings are given in Table 28 for 8x8 averaging and in Table 29 for 16x16 averaging. None of the clusterings pass the threshold test. In fact, none of the individual clusters appear to be valid. Since all multiple-class solutions are rejected, we conclude that the image contains a single texture.

6.7 Segmentation Experiment 3: Natural Image Composite

Figure 23 shows the 128x128 image to be used for this experiment. The image is composed of four natural texture classes: WOOD, PAPE, SCRE and SAND. The results of segmentation based on 8x8 averaging for the feature images is shown in Figure 24, and the results for 16x16 averaging are shown in Figure 25.

Different results are obtained by using different averaging windows. The two-cluster solution using 8x8 averaging groups the WOOD and SCRE areas in one cluster and the PAPE and SAND areas in the other. This clustering appears to be determined by the irregularity of PAPE and SAND and the directional tendencies of WOOD and SCRE. Using 16x16 averaging, the SCRE texture is alone in one cluster in the two-class

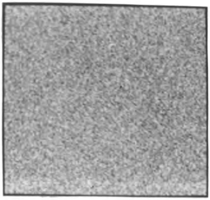
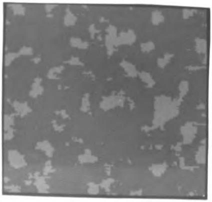
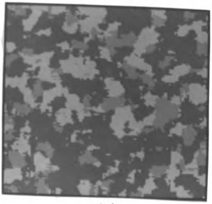


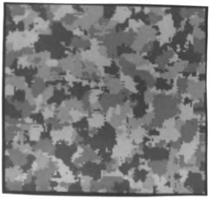
Figure 20: Gaussian White Noise Image for Segmentation Experiment 2.



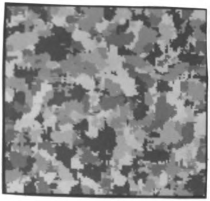
(a)



(b)



(c)



(d)

Figure 21: Segmented Images for Gaussian White Noise Using 8x8 Averaging Windows. (a) 2-cluster solution (b) 3-cluster solution (c) 4-cluster solution (d) 5-cluster solution

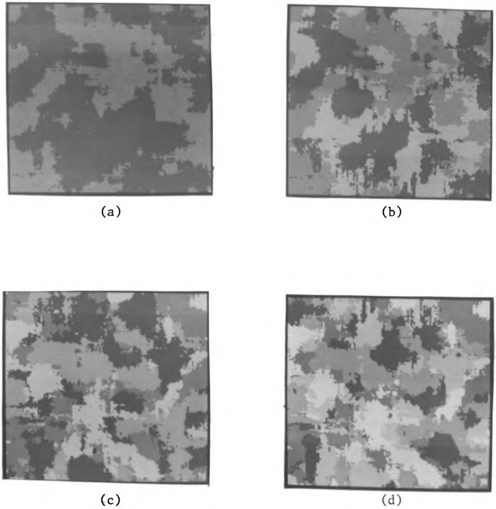


Figure 22: Segmented Images for Gaussian White Noise Using 16x16 Averaging Windows. (a) 2-cluster solution (b) 3-cluster solution (c) 4-cluster solution (d) 5-cluster solution

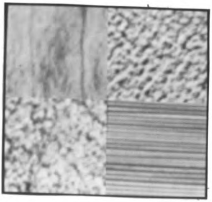
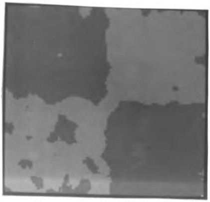
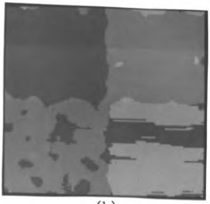


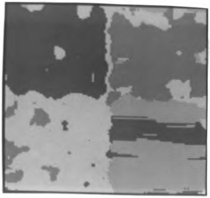
Figure 23: Natural Image Composite for Segmentation Experiment 3.



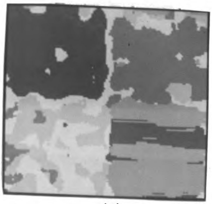
(a)



(b)



(c)



(d)

Figure 24: Segmented Images for Natural Image Composite Using 8x8 Averaging Windows. (a) 2-cluster solution (b) 3-cluster solution (c) 4-cluster solution (d) 5-cluster solution

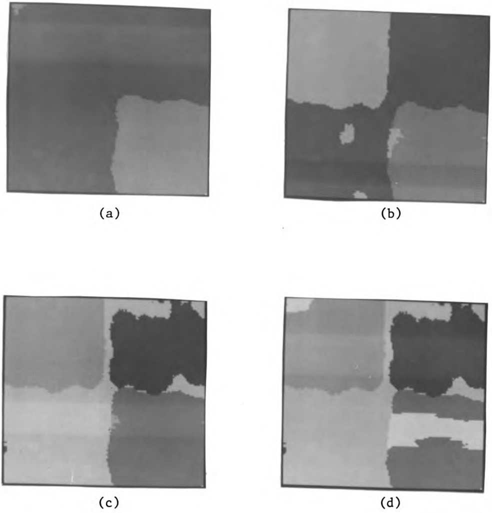


Figure 25: Segmented Images for Natural Image Composite Using 16x16 Averaging Windows. (a) 2-cluster solution (b) 3-cluster solution (c) 4-cluster solution (d) 5-cluster solution

Table 28: Evaluation of Clustering on Gaussian White Noise Image Using 8x8 Averaging Windows.

k	N(k)	S(k)
2 CLUSTERS		
1	55	1.26
2	9	1.12
AVG		1.24
3 CLUSTERS		
1	38	0.831
2	8	1.22
3	18	0.939
AVG		0.91
4 CLUSTERS		
1	14	1.21
2	16	1.11
3	18	1.05
4	16	1.07
AVG		1.11
5 CLUSTERS		
1	13	1.18
2	10	1.04
3	17	1.15
4	10	1.01
5	14	1.13
AVG		1.11
6 CLUSTERS		
1	12	1.08
2	4	1.64
3	7	1.29
4	8	1.04
5	23	1.06
6	10	0.946
AVG		1.10

k	N(k)	S(k)
7 CLUSTERS		
1	10	1.13
2	5	1.66
3	7	1.07
4	7	1.16
5	15	0.923
6	7	1.02
7	13	1.03
AVG		1.09
8 CLUSTERS		
1	9	1.29
2	4	1.44
3	6	1.53
4	7	1.01
5	13	1.18
6	9	1.15
7	6	1.06
8	10	1.09
AVG		1.20

segmentation. The regularity and strong orientation dependence appear to determine this clustering. The segmentations with three and four clusters are similar for the two window sizes. The 8x8 window tends to break up the image into smaller areas and to misclassify more pixels than the 16x16 window. Edges between textures are more sharply identified using the 8x8 window, but the 8x8 window is more likely to

Table 29: Evaluation of Clustering on Gaussian White Noise Image Using 16x16 Averaging Windows.

k	N(k)	S(k)
2 CLUSTERS		
1	39	1.12
2	25	0.957
AVG		1.06
3 CLUSTERS		
1	24	0.951
2	19	1.04
3	21	0.879
AVG		0.954
4 CLUSTERS		
1	21	1.06
2	15	1.00
3	17	0.972
4	11	1.03
AVG		1.02
5 CLUSTERS		
1	13	1.20
2	12	1.10
3	15	1.08
4	13	1.24
5	11	0.973
AVG		1.12
6 CLUSTERS		
1	11	1.25
2	14	1.11
3	11	1.24
4	11	1.34
5	10	1.24
6	7	1.22
AVG		1.23

k	N(k)	S(k)
7 CLUSTERS		
1	11	1.25
2	11	1.34
3	11	1.24
4	11	1.34
5	10	1.24
6	3	1.28
7	7	1.20
AVG		1.27
8 CLUSTERS		
1	10	1.30
2	8	1.10
3	9	1.37
4	11	1.36
5	9	1.21
6	3	1.28
7	6	1.23
8	8	1.13
AVG		1.25

incorrectly subdivide a homogeneous texture area. One significant error made using 8x8 averaging is the misclassification of a low-contrast area of the SCRE panel as WOOD. The same area appears as a separate segment in the 3 and 4 class segmentations using 16x16 averaging.

The statistics for evaluating the clusterings are shown in Table 30 for 8x8 averaging and in Table 31 for 16x16 averaging. The three-cluster solution is the only one which passes the threshold test in the 8x8 case. With 16x16 averaging, the solutions for 3, 4, and 5 clusters are found to be acceptable. Using the weighted average of

Table 30: Evaluation of Clustering on Composite Natural Image Using 8x8 Averaging Windows.
(* indicates accepted clustering solutions)

k	N(k)	S(k)
2 CLUSTERS		
1	35	1.34
2	29	1.58
AVG		1.45
3 CLUSTERS		
1	24	2.11
2	28	1.75
3	12	2.22
AVG		1.97 *
4 CLUSTERS		
1	21	1.61
2	16	1.39
3	12	2.19
4	15	1.46
AVG		1.63
5 CLUSTERS		
1	21	1.46
2	15	1.35
3	12	2.17
4	7	1.06
5	9	1.03
AVG		1.46
6 CLUSTERS		
1	14	1.21
2	7	1.47
3	12	2.13
4	14	1.11
5	10	1.28
6	7	1.33
AVG		1.41

k	N(k)	S(k)
7 CLUSTERS		
1	14	1.21
2	6	1.52
3	12	2.08
4	10	1.27
5	4	1.32
6	7	1.33
7	11	1.32
AVG		1.45
8 CLUSTERS		
1	14	1.21
2	6	1.52
3	12	2.09
4	5	1.37
5	4	1.32
6	7	1.33
7	11	1.43
8	5	1.76
AVG		1.52

S(k) to rank the solutions, we find the three-cluster solution to be the best, followed in order by the 4- and 5-cluster solutions. In the three-cluster solution, the SAND and PAPE regions are merged in a single cluster. The irregularity and the contrast of these areas are similar, so the preference for this solution is plausible. In the three-class segmentation using 16x16 averaging windows, over 95% of the

Table 31: Evaluation of Clustering on Composite Natural Image Using 16x16 Averaging Windows.
(* indicates accepted clustering solutions)

k	N(k)	S(k)	
2 CLUSTERS			
1	48	1.44	
2	16	2.26	
AVG		1.65	
3 CLUSTERS			
1	31	2.18	
2	16	2.51	
3	17	2.95	
AVG		2.47	*
4 CLUSTERS			
1	13	1.92	
2	16	2.40	
3	16	2.59	
4	19	1.89	
AVG		2.20	*
5 CLUSTERS			
1	13	1.92	
2	12	2.22	
3	15	2.05	
4	19	1.89	
5	5	2.76	
AVG		2.06	*
6 CLUSTERS			
1	5	2.53	
2	12	2.22	
3	15	2.05	
4	15	1.61	
5	12	1.56	
6	5	2.76	
AVG		1.98	
7 CLUSTERS			
1	5	2.53	
2	12	2.22	
3	15	2.05	
4	6	1.17	
5	12	1.55	
6	9	1.09	
7	5	2.76	
AVG		1.86	
8 CLUSTERS			
1	5	2.53	
2	12	2.22	
3	15	2.05	
4	6	1.23	
5	12	1.53	
6	6	1.43	
7	3	1.40	
8	5	2.76	
AVG		1.91	

pixels are correctly labelled (assuming the SAND and PAPE panels to be the same class). The 8x8 window segmentation using 8x8 windows classifies about 86% of the pixels correctly.

The four-cluster solution, which is the second choice for 16x16 averaging, reasonably captures the four image types used; the segmentation based on an 8x8 averaging window correctly labels 81% of the pixels while using a 16x16 window, 91% of the pixels are correctly labelled.

6.8 Segmentation Experiment 4: SCRE

Figure 26 shows the 128x128 SCRE natural image which will be segmented using the same procedures as the previous experiments. The image contains a single texture class, but unlike the Gaussian white noise image, the SCRE image contains some internal structure. The segmented images using 8x8 averaging windows are presented in Figure 27 and the results using 16x16 averaging windows are presented in Figure 28. The cluster evaluation statistics are shown in Table 32 for 8x8 averaging and in Table 33 for 16x16 averaging. None of the clustering solutions passes the threshold test, but the values of S(k) are larger than those obtained with Gaussian white noise. There appear to be emerging clusters, but the presence of some non-isolated clusters in each solution causes all of the clusterings to be rejected.

6.9 Summary

This chapter presented an algorithm for computing a textural transform in which the texture of a neighborhood about each pixel is represented by gray level values in a series of feature images. The feature images are computed from filtered images using a two-step procedure in which the second step can be implemented as another filtering operation. The resulting textural transform was evaluated by applying a texture segmentation procedure in four experiments. The segmentation procedure imposed no restrictions on the nature of the

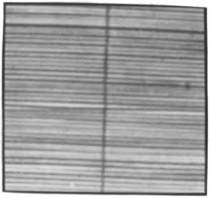
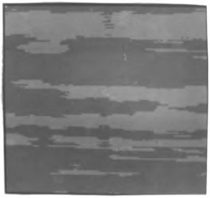
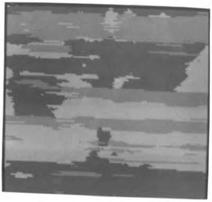


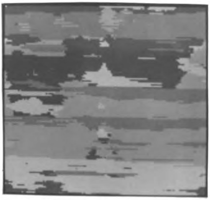
Figure 26: SCRE Image for Segmentation Experiment 4



(a)



(b)



(c)

(d)

Figure 27: Segmented Images for SCRE Using 8x8 Averaging Windows. (a) 2-cluster solution (b) 3-cluster solution (c) 4-cluster solution (d) 5-cluster solution

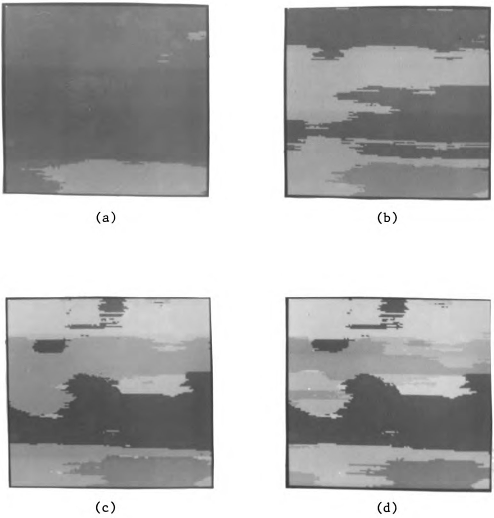


Figure 28: Segmented Images for SCRE Using 16x16 Averaging Windows. (a) 2-cluster solution (b) 3-cluster solution (c) 4-cluster solution (d) 5-cluster solution

Table 32: Evaluation of Clustering on SCRE Image Using 8x8 Averaging Windows.

k	N(k)	S(k)
2 CLUSTERS		
1	44	1.45
2	20	1.60
AVG		1.50
3 CLUSTERS		
1	27	1.24
2	15	2.02
3	22	1.57
AVG		1.54
4 CLUSTERS		
1	18	1.46
2	14	1.88
3	19	1.89
4	13	1.33
AVG		1.65
5 CLUSTERS		
1	19	1.37
2	14	1.85
3	17	1.70
4	5	2.70
5	9	1.83
AVG		1.73
6 CLUSTERS		
1	9	1.74
2	14	1.76
3	17	1.70
4	5	2.54
5	13	1.68
6	6	1.71
AVG		1.78

k	N(k)	S(k)
7 CLUSTERS		
1	9	1.74
2	12	1.82
3	7	1.43
4	5	2.54
5	13	1.71
6	6	1.71
7	12	1.14
AVG		1.66
8 CLUSTERS		
1	9	1.74
2	6	1.37
3	5	2.03
4	5	2.46
5	11	1.54
6	6	1.71
7	10	1.95
8	12	1.51
AVG		1.74

segmented regions such as a minimum size or connectivity requirement. A clustering algorithm was applied to some representative pixels, and the remaining pixels were assigned to the clusters by a minimum distance classifier.

The segmented images indicate that the feature space is a reasonable representation of the textures in an image. This conclusion

Table 33: Evaluation of Clustering on SCRE Image Using 16x16 Averaging Windows.

k	N(k)	S(k)
2 CLUSTERS		
1	57	1.65
2	7	1.81
AVG		1.66
3 CLUSTERS		
1	32	1.34
2	6	1.99
3	26	1.34
AVG		1.40
4 CLUSTERS		
1	18	1.49
2	4	3.99
3	25	1.36
4	17	1.59
AVG		1.62
5 CLUSTERS		
1	18	1.53
2	4	3.50
3	7	1.83
4	19	1.60
5	16	1.64
AVG		1.73
6 CLUSTERS		
1	14	1.86
2	4	3.63
3	7	1.77
4	17	1.50
5	6	1.28
6	16	1.73
AVG		1.79

k	N(k)	S(k)
7 CLUSTERS		
1	13	1.73
2	4	3.10
3	6	1.99
4	12	1.21
5	4	2.41
6	16	1.72
7	9	1.25
AVG		1.72
8 CLUSTERS		
1	13	1.73
2	4	3.10
3	4	2.50
4	11	1.29
5	4	2.41
6	16	1.72
7	6	1.44
8	6	1.70
AVG		1.80

was further confirmed by a statistical evaluation of the clusterings. The evaluation identified reasonable clusterings by requiring all clusters to be compact and isolated as measured by a test statistic. Acceptable clusterings were then ranked by an average isolation statistic. The segmentations corresponding to the preferred clusterings contained reasonable numbers of clusters and the corresponding segmentations generally agree with the image generation method and visual segmentations of the images.

The experiments used 8x8 and 16x16 windows for computing the feature images. Both gave satisfactory segmentation results, though these window sizes are smaller than the subimages used in the classification experiments. This suggests that a minimum distance classification algorithm might be more appropriate than a nearest neighbor algorithm for the channel filtering feature space.

Chapter 7

Summary and Conclusions

7.1 Summary

Texture is characterized by preattentive human visual performance of segmentation and classification tasks. This form of perception has many potential applications in computer vision systems, but the characterization of "texture" in terms of human performance does not lead to simple, effective algorithms for texture analysis. No alternative definition of texture independent of human performance has provided sufficiently precise guidance to enable development of general texture analysis algorithms.

Since texture analysis algorithms attempt to model preattentive human vision, insights from vision science and from intuition have been used to guide the development of algorithms. The lack of precise guidance from a definition of texture and the lack of generality in existing algorithms has resulted in a profusion of diverse approaches to texture analysis.

A recent theory concerning the early information processing strategies in human vision was used to motivate a new approach to texture analysis. A feature space which measures average local energy was defined from filtered images and used in texture classification problems. The features were applied to artificial images, natural images, and preprocessed natural images. The performance of the feature space in various classification tasks was investigated and compared with the power spectral and co-occurrence methods.

A method for computing a texture feature vector over a small neighborhood about each pixel in an image was developed for texture segmentation. A clustering algorithm was applied to the feature vectors, and a segmented image was produced by labelling each pixel according to the cluster in which the pixel's feature vector lies.

Since each clustering solution corresponds to a segmented image, a means for selecting acceptable clusterings is required. A cluster validity statistic was used to determine which, if any, of a set of clusterings provided an acceptable representation of the data. A related statistic was used to rank the accepted solutions. Four texture segmentation experiments were performed using both composite and homogeneous images.

7.2 Conclusions

7.2.1 Classification

The channel filtering features were used to classify subimages of eight natural textures. The results using 64x64 and 32x32 subimages were good. The features performed poorly when presented 16x16 subimages. These results indicate that the features are suitable for classification of natural images based on 32x32 or larger subimages.

The effect of histogram equalization on the evaluation of natural textures by the channel filtering features was investigated in a series of experiments. We found that almost all of the histogram equalized images were perceived as having different textures from the original classes. Histogram equalization was also observed to confuse image classes which were originally separable. The results suggest that histogram equalization should be used carefully, if at all, with channel filtering features.

The channel filtering features were found to be insensitive to global, constant gray level changes. This implies that histogram equalization to remove differences in average brightness is unnecessary.

A procedure for determining whether two images portray the same texture at different magnifications was developed. The procedure was tested for 2X magnifications only. Certain aspects of the procedure might be simplified by using a different classification algorithm, but the results obtained were good.

A procedure for detecting orientation differences was developed which gave excellent performance for 90 degree orientation changes but was less reliable for 45 degree orientation changes due to fundamental differences in 45 degree rotated patterns caused by the rectangular quantization grid and the square image shape.

The effect of phase modification was investigated. The experiment demonstrated that images with identical power spectra could be discriminated by channel filtering features based on differences in the phase spectra.

An experiment comparing the co-occurrence method with the channel filtering method demonstrated that selection of displacement vectors for co-occurrence matrices is a serious problem. Even displacement vectors selected specifically to discriminate micropatterns can fail to discriminate images composed of the micropatterns. On the other hand, the channel filtering features were applied in the same way as in other experiments and produced excellent classification accuracy.

Experiments to test computational simplifications of the channel filtering procedure showed that the method is "robust" in that the classification accuracy is not degraded by using ideal band-pass channels or by eliminating certain channels.

7.2.2 Segmentation

Four experiments were performed to test the utility of the channel filtering approach for texture segmentation. In the first experiment, an image composed of a regular dot pattern and a random dot pattern were segmented. The statistical evaluation of the clustering solutions indicated that two clusters existed in the feature space. The labellings of the pixels corresponded closely to the actual image areas.

Segmentation of a Gaussian white noise image yielded irregularly shaped, non-contiguous segments throughout the image. The cluster validity statistic found no acceptable clustering solution, so the image was correctly identified as portraying a single texture.

A composite image composed of four natural textures was segmented. The statistical evaluation indicated that the best solution involved three clusters. This result was considered reasonable due to the visual similarity of two of the actual texture classes. The second best solution involved four clusters and the segmentation generally corresponded to the actual textures in the image.

Another experiment on a homogeneous natural image resulted in a correct identification of the image as a homogeneous texture. Emerging clusters were found, but they were not well-defined enough to pass the statistical validity test.

7.2.3 General

The channel filtering feature space has been evaluated on a variety of texture classification and segmentation problems. The results indicate that the feature space is a good representation of texture for these problems. Equivalently, the feature space has been validated as a model for preattentive human vision on a variety of stimuli. The investigation of this feature space has involved tests in which visually discriminable image classes were expected to be discriminated by the features and in which visually indiscriminable image classes generated in different ways were expected to be confused by the features.

The classification and segmentation procedures used to evaluate the feature space did not involve sophisticated heuristics; the results are consequences of the structure of the feature space. Some results, in particular the magnification experiment and the results of the segmentation experiments, suggest that a minimum distance classifier may be more appropriate than a nearest neighbor classifier for the channel filtering feature space. Further investigation of the feature space may suggest that a different clustering procedure may be more appropriate for texture segmentation.

The performance of the channel filtering feature demonstrates that a global filtering model of vision is capable of reasonably approximating human preattentive vision. This contrasts to results of previous studies which discounted the value of phase information for texture analysis and which suggested that the Fourier transform is not an appropriate tool for texture analysis. This study shows that phase information provides critical information on the distribution of spectral energy through the image plane. The use of phase information enables spatially local analysis of different spectral energy bands. This information was measured by the channel filtering features and provided good results.

7.3 Advantages of the Channel Filtering Method

1. The channel filtering features have a satisfying intuitive basis. Unlike some ad hoc procedures, the significance of the channel filtering features are explainable in terms of image properties such as contrast, edge density, local energy, and directionality. The explanations are a consequence of the intuitive interpretations of the channels and of the spatial filtering operation.

2. The channel filtering method provides classification results superior to those obtained by the power spectral method. The performance is attributed to the utilization of phase information in the channel filtering procedure.

3. The channel filtering method has no critical parameters similar to the choice of displacement vectors for the co-occurrence method.

4. A method for computing features for texture segmentation exists which does not require recomputation of the entire filtering procedure over every subimage of interest. This means that only a simple feature computation needs to be repeated to evaluate the textures in different regions. In fact, the texture features for every pixel can be computed by a procedure easily amenable to parallel execution.

5. Simple procedures for classification and for segmentation produce good results. This indicates that the results are due to the structure of the feature space and not to a clever classification or segmentation procedure.

6. There are many opportunities for parallelism in computing the filtered images and the texture features. This could enable fast hardware implementations of the channel filtering procedure to be constructed with a high degree of modularity and at fairly low cost.

7.4 Disadvantages of the Channel Filtering Method

1. The computational and storage requirements for sequential, digital implementation of the channel filtering approach are too

demanding for many applications. Other implementation methods using parallel hardware or optical/digital methods may make fast implementations possible. Among the current difficulties are the number of Fourier transforms, inverse transforms, and image multiplications required in the channel filtering procedure. These operations are very demanding when implemented sequentially.

2. A simple method for generating images which have specified characteristics in the texture feature space is not known. Such a procedure would enable further validation of the feature space by enabling generation of images which are "close" to a given image. This capability would enable a quantitative determination of how well the feature space models human preattentive perception. The major obstacles to development of such a method are the overlapping of the channels in the spatial frequency domain and the dependence of the feature values on phase spectrum information.

7.5 Suggestions for Further Research

1. Segmentation results have suggested that the minimum distance classification algorithm is more appropriate than the nearest-neighbor algorithm for the channel filtering feature space. Further research is needed to confirm this hypothesis and to investigate the applicability of other decision rules.

2. Different clustering algorithms for use in the segmentation procedure should be investigated. CLUSTER is known to perform poorly in certain situations, and statistics similar to those used to evaluate the clusterings in this study can be defined for other algorithms.

3. Further investigation of the performance of the segmentation procedure is needed. The averaging windows used to compute the feature images could be implemented as another filtering operation, perhaps using a sequence of filters as in the original channel filtering procedure. The validity of the feature space for images which involve distinct forms should also be investigated.

4. The use of channel filtering for receding surfaces should be investigated. Since the apparent size of the texture on a receding surface shrinks progressively with distance, the spectral information from the surface should move steadily toward high-frequency channels. This would cause the feature vectors for pixels depicting the surface to be strung out though the feature space. A clustering procedure which can detect long, stringy clusters, such as the single-link algorithm, might be useful for analyzing surface structure.

5. Since the feature space developed in chapter 4 is based on a theory of human visual information processing, it may be useful to determine how faithfully the feature space reproduces human texture vision. Psychophysical experiments could help to guide the selection of classification and clustering algorithms which would duplicate human performance.

6. Spatial frequency domain filtering has been used to provide at least partial solutions for several computational vision problems, now including texture analysis. Revision of procedures for solving other computational vision problems using channel filtering as a framework could enable development of unified computational vision systems.

APPENDICES

Appendix A: A Catalog of Texture Definitions

1. "We may regard texture as what constitutes a macroscopic region. Its structure is simply attributed to the repetitive patterns in which elements or primitives are arranged according to a "placement rule"." [Tamura et al, 1978]

2. "We suggest the following operational definition of 'texture'. A region in an image has a constant texture if a set of local statistics or other local properties of the picture function are constant, slowly varying, or approximately periodic." [Sklansky, 1978]

3. "A texture will be considered to be a random field X(n,m) where n and m are integers." (They later specialize to Markov random fields.) [Conners and Harlow, 1980a]

4. "The image texture we consider is nonfigurative and cellular.... An image texture is described by the number and types of its (tonal) primitives and the spatial organization or layout of its (tonal) primitives... a fundamental characteristic of texture: it cannot be analyzed without a frame of reference of tonal primitive being stated or implied. For any smooth gray-tone surface, there exists a scale such that when the surface is examined, it has no texture. Then as resolution increases, it takes on a fine texture and then a coarse texture." [Haralick, 1979]

5. "Image texture refers to the visual sensation that one receives about the structure arrangement of an image region. The

textural properties of a scene are often qualitatively described as coarse, grainy, striated, or rough. Computer simulations have been done in which overlapping particles have been placed at random locations. When only a few particles are present, the visual impression is of discrete, countable objects. However, when the number of objects is increased, the visual impression is of texture rather than countable objects." [Hall et al, 1977]

6. "Texture is defined for our purposes as an attribute of a field having no components that appear enumerable. The phase relations between the components are thus not apparent. Nor should the field contain an obvious gradient. The intent of this definition is to direct the attention of the observer to the global properties of the display - i.e. its overall "coarseness", "bumpiness", or "fineness". Physically, nonenumerable (aperiodic) patterns are generated by stochastic as opposed to deterministic processes. Perceptually, however, the set of all patterns without obvious enumerable components will include many deterministic (and even periodic) textures. Because our criterion for enumerability was subjective rather than objective, many of our patterns actually contained repetitive elements which were not immediately obvious but could be identified when the observer specifically looked for these components. To further minimize very obvious enumerable periodic components of the patterns, all displays contained only components whose spatial frequencies were proportional to prime numbers. We would like to stress that the above constraints imposed on our textures are all designed to minimize the importance of phase information - a variable we consider more important for pattern recognition than for texture perception." [Richards and Polit, 1974]

7. "Images composed of numerous binary pixels which are only weakly correlated with their neighbors form one extreme in a continuum of scene types whose opposite extreme is exemplified by simple line drawings against a uniform background. In the latter case the image pixels are highly redundant and the image consequently carries little information in the sense of Shannon, whereas in the former case, redundancy is low and the information content is high: normally so high that the vision system must selectively disregard information in order to process the scene. The limiting extremes of this continuum of images are the uniform ("constant") image all of whose pixels are identical, and the random image, whose pixel arrangement is completely determined by a probability distribution. The random image appears to be homogeneous (and therefore completely redundant) in a global sense; that is, different subregions large enough to contain many pixels convey an equivalent subjective impression. This type of homogeneity, which is different from a locally highly redundant image, is usually called "texture", and is characteristic of the structure of the probability distribution of the random image." [Resnikoff, 1981]

8. "Texture is an apparently paradoxical notion. On the one hand, it is commonly used in the early processing of visual information, especially for practical classification purposes. On the other hand, no one has succeeded in producing a commonly accepted definition of texture. The resolution of this paradox, we feel, will depend on a richer, more developed model for early visual information processing, a central aspect of which will be representational systems at many different levels of abstraction. These levels will most

probably include actual intensities at the bottom and will progress through edge and orientation descriptors to surface, and perhaps volumetric, descriptors. Given these multi-level structures, it seems clear that they should be included in the definition of, and in the computation of, texture descriptors." [Zucker and Kant, 1981]

9. "The notion of texture appears to depend upon three ingredients: (1) some local 'order' is repeated over a region which is large in comparison to the order's size, (2) the order consists in the nonrandom arrangement of elementary parts, and (3) the parts are roughly uniform entities having approximately the same dimensions everywhere within the textured region." [Hawkins, 1969]

10. "Although these descriptions of texture seem perceptually reasonable, they do not immediately lead to simple quantitative textural measures in the sense that the description of edge discontinuity leads to the quantitative definition of an edge in terms of its location, slope angle, and height." [Pratt, 1978]

APPENDIX B: Definition of Channel Filters

The transfer function of a 2Nx2N spatial frequency channel filter $F_k(u,v)$, where $-N+1 <= u,v <= N$, is defined as follows:

$$\text{Imag} [F_k(u,v)] = 0 \quad \text{for all } u,v,k,$$

$$\text{Real} [F_k(0,0)] = 1 \quad \text{for all } k,$$

and for $(u,v) \neq (0,0)$,

$$\text{Real} [F_k(u,v)] = \exp \left[-.5 * \frac{[\ln(D(u,v)) - \ln(\mu_k)]^2}{\sigma^2} \right]$$

where $D(u,v) = [u^2 + v^2]^{1/2}$, $\sigma = .275$, and $\mu_k = 2^{k-1}$.

For a 2Nx2N image, we will use spatial frequency channels for $k=1,\ldots(\log_2 N + 1)$. This definition yields a series of filters one octave apart whose widths are slightly more than one octave on each side of the center frequency. The value of σ was selected to produce filters whose widths are within the constraints for human visual filters given in [Ginsburg, 1978].

The transfer function for a 2Nx2N orientation channel filter Gk(u,v) where $-N+1 <= u,v <= N$ is defined as follows:

$$\text{Imag}\ [\ Gk(u,v)\ ] = 0 \quad \text{for all } u,v,k,$$

$$\text{Real}\ [\ Gk(0,0)\ ] = 1 \quad \text{for all } k,$$

and for $(u,v) \neq (0,0)$,

$$\text{Real}\ [\ Gk(u,v)\ ] = \exp\left[-.5 * \frac{a_k^2}{\sigma^2} \right]$$

where $a_k = \text{Min}\ \{\ |\mu_k - \arctan(v/u)|\ ,\ |\ (\mu_k - 180) - \arctan(v,u)\ |\ \}$, $\sigma = 17.8533$, $1 <= k <= 4$, and the values of μ_k are given in the table below for each value of k. The value of σ is chosen to produce the same overlap between channels as in the spatial frequency channels. Note that the orientation channels defined here are slightly wider then the 30 degree wide channels proposed for human vision in [Ginsburg, 1978].

Values of μ_k (in degrees) for each k

k	μ_k	
1	0	(horizontal)
2	45	
3	90	
4	135	

List of References

List of References

Agin, G. J., 1980, "Computer Vision Systems for Industrial Inspection and Assembly," Computer, 13: 11-20.

Aggarwal, J. K., L. S. Davis and W. N. Martin, 1981, "Correspondence Processes in Dynamic Scene Analysis," Proceedings of the IEEE, 69: 562-572.

Ahuja, N. and A. Rosenfeld, 1981, "Mosaic Models for Textures," IEEE Transactions on Pattern Analysis and Machine Intelligence, 3: 1-11.

Anderberg, M. R., 1973, Cluster Analysis for Applications, New York: Academic Press.

Bacus, J. W., 1976, "A Whitening Transformation for Two-Color Blood Cell Images," Pattern Recognition, 8: 53-60.

Bajcsy, R., 1973, "Computer Identification of Visual Surfaces," Computer Graphics and Image Processing, 2: 118-130.

Bajcsy, R. and L. Lieberman, 1976, "Texture Gradient as a Depth Cue," Computer Graphics and Image Processing, 5: 52-67.

Barrow, H. G. and J. M. Tennenbaum, 1981, "Computational Vision," Proceedings of the IEEE, 69: 572-595.

Beck, J., 1980, "Texture Segmentation," University of Maryland, Computer Science Technical Report TR-971.

Besag, J., 1974, "Spatial Interaction and the Statistical Analysis of Lattice Systems," Journal of the Royal Statistical Society, B36: 192-236.

Blakemore, C. and F. W. Campbell, 1969, "On the Existence of Neurons in the Human Visual System Selectively Sensitive to the Orientation and Size of Retinal Images," Journal of Physiology, 203: 237-260.

Brady, M., 1982, "Computational Approaches to Image Understanding," Computing Surveys, 14: 3-71.

Brodatz, P., 1966, Textures: A Photographic Album for Artists and Designers, New York: Dover.

Caelli, T. M. and B. Julesz, 1978a, "On Perceptual Analyzers Underlying Visual Texture Discrimination - Part 1," Biological Cybernetics, 28: 167-175.

Caelli, T. M. and B. Julesz, 1978b, "On Perceptual Analyzers Underlying Visual Texture Discrimination - Part 2," Biological Cybernetics, 29: 201-214.

Campbell, F. W. and L. Maffei, 1970, "Electrophysiological Evidence for the Existence of Orientation and Size Detectors in the Human Visual System," Journal of Physiology, 207: 635-652.

Campbell, F. W. and J. G. Robson, 1968, "Application of Fourier Analysis to the Visibility of Gratings," Journal of Physiology, 197: 551-566.

Carlucci, L., 1972, "A Formal System for Texture Languages," Pattern Recognition, 4: 53-72.

Chellappa, R. and R. L. Kashyap, 1981a, "On the Correlation Structure of Random Field Models of Images and Textures," Proceedings of the IEEE Conference on Pattern Recognition and Image Processing, Dallas. 574-576.

Chellappa, R. and R. L. Kashyap, 1981b, "Synthetic Generation and Estimation in Random Field Models of Images," Proceedings of the IEEE Conference on Pattern Recognition and Image Processing, Dallas. 577-582.

Coleman, G. B. and H. C. Andrews, 1979, "Image Segmentation by Clustering," Proceedings of the IEEE, 67: 773-785.

Conners, R. W., 1979, "Towards a Set of Statistical Features Which Measure Visually Perceivable Qualities of Textures," Proceedings of the IEEE Conference on Pattern Recognition and Image Processing, Chicago. 382-390.

Conners, R. W. and C. A. Harlow, 1980a, "A Theoretical Comparison of Texture Algorithms," IEEE Transactions on Pattern Analysis and Machine Intelligence, 2: 204-222.

Conners, R. W. and C. A. Harlow, 1980b, "Toward a Structural Texture Analyzer Based on Statistical Methods," Computer Graphics and Image Processing, 12: 224-256.

Cross, G. R. and A. K. Jain, 1981, "Markov Random Field Texture Models," Proceedings of the IEEE Conference on Pattern Recognition and Image Processing, Dallas. 597-602.

Crowley, J. and A. Parker, 1978, "The Analysis, Synthesis and Evaluation of Local Measures for Discrimination and Segmentation of Textured Regions," Proceedings of the IEEE Conference on Pattern

Recognition and Image Processing, Chicago. 372-378.

Darling, E. M. and R. D. Joseph, 1968, "Pattern Recognition from Satellite Altitudes," IEEE Transactions on Systems, Man and Cybernetics, 4: 38-47.

D'Astous, F. and M. Jernigan, 1981, "Phase Information in Texture Feature Extraction," Proceedings of the International Conference on Cybernetics and Society, Atlanta. 182-186.

Davis, L., S. Johns and J. Aggarwal, 1979, "Texture Analysis Using Generalized Co-occurrence Matrices," IEEE Transactions on Pattern Analysis and Machine Intelligence, 1: 251-259.

Davis, L. and A. Mitiche, 1982, "MITES: A Model-Driven Iterative Texture Segmentation Procedure," Computer Graphics and Image Processing, 19: 95-110.

Deguchi, K. and I. Morishita, 1978, "Texture Characterization and Texture-Based Image Partitioning Using Two-dimensional Linear Estimation Techniques," IEEE Transactions on Information Theory, 27: 783-745.

Dubes, R. C. and A. K. Jain, 1976, "Clustering Techniques: The User's Dilemma," Pattern Recognition, 8: 247-260.

Dubes, R. C. and A. K. Jain, 1979, "Validity Studies in Clustering Methodologies," Pattern Recognition, 11: 235-254.

Duda, R. O. and P. E. Hart, 1973, Pattern Classification and Scene Analysis, New York: Wiley.

Ehrich, R. and J. P. Foith, 1976, "Representation of Random Waveforms by Relational Trees," IEEE Transactions on Computers, 25, no. 7: 725-736.

Ehrich, R. and J. P. Foith, 1978, "A View of Texture Topology and Texture Description," Computer Graphics and Image Processing, 8: 174-202.

Eklundh, J., 1979, "On the Use of Fourier Phase Features for Texture Discrimination," Computer Graphics and Image Processing, 9: 199-201.

Everitt, B., 1974, Cluster Analysis, London: Heinemann Educational Books.

Faugeras, O. D. and W. K. Pratt, 1980, "Decorrelation Methods of Texture Feature Extraction," IEEE Transactions on Pattern Analysis and Machine Intelligence, 2: 323-332.

Fu, K. S., 1974, Syntactic Methods in Pattern Recognition, New York: Academic Press.

Fu, K. S., 1977, Syntactic Pattern Recognition Applications, New York: Springer-Verlag.

Gagalowicz, A., 1978, "Analysis of Texture Using a Stochastic Model," Proceedings of the Fourth International Conference on Pattern Recognition, Tokyo. 541-544.

Gagalowicz, A., 1981, "A New Method for Texture Fields Synthesis: Some Applications to the Study of Human Vision," IEEE Transactions on Pattern Analysis and Machine Intelligence, 2: 520-533.

Galloway, M., 1974, "Texture Analysis Using Gray Level Run Lengths," Computer Graphics and Image Processing, 4(2): 172-179.

Garber, D. and A. Sawchuk, 1981, "Computational Models for Texture Analysis and Synthesis," Proceedings of the DARPA Image Understanding Workshop, Washington, D. C. 69-88.

Ginsburg, A. P., 1973, "Pattern Recognition Techniques Suggested from Psychological Correlates of a Model of the Human Visual System," Proceedings of the IEEE National Aerospace Electronics Conference, 1972, Dayton, OH: 309-316.

Ginsburg, A. P., 1976, "The Perception of Visual Form: A Two-Dimensional Filter Analysis" in Information Processing in the Visual System, Proceedings of the IV Symposium on Sensory System Physiology, I. P. Pavlov Institute of Physiology, Leningrad, U. S. S. R., V. D. Glezer, editor. 46-51.

Ginsburg, A. P., 1978, Visual Information Processing Based on Spatial Filters Constrained by Biological Data, Dissertation for Ph. D. University of Cambridge, England, 1977. Also published as Air Force Aerospace Medical Research Laboratory Technical Report AMRL-TR-78-129, December, 1978.

Ginsburg, A. P., 1979a, "Visual Perception Based on Spatial Filtering Constrained by Biological Data," Proceedings of the International Conference on Cybernetics and Society, Denver. 453-457.

Ginsburg, A. P., 1979b, "Spatial Filtering and Mechanisms of Perception," Proceedings of the Tenth Annual Pittsburg Conference on Modelling and Simulation, Pittsburg. 185-192.

Ginsburg, A. P., 1980a, "Visual Perception Based on Biological Filtering of Spatial Information," Proceedings of the International Conference on Cybernetics and Society, Cambridge, MA. 424-428.

Ginsburg, A. P., 1980b, "Specifying Relevant Information for Image Evaluation and Display Design: An Explanation of How We See Certain Objects," Proceedings of the Society for Industrial Design, 21: 219-227.

Ginsburg, A. P. and J. M. Coggins, 1981, "Texture Analysis Based on Filtering Properties of the Human Visual System," Proceedings of the International Conference on Cybernetics and Society, Atlanta. 112-117.

Graham, N., 1981, "The Visual System Does a Crude Fourier Analysis of Patterns, in Mathematical Psychology and Psychophysiology, Stephen Grossberg editor. SIAM and AMS Proceedings Series. 1-16.

Gurari, E. M. and H. Wechsler, 1982, "On the Difficulties Involved in Segmentation of Pictures," IEEE Transactions on Pattern Analysis and Machine Intelligence, 4: 304-306.

Hall, E. L., 1972, "A Comparison of Computations for Spatial Frequency Filtering," Proceedings of the IEEE, 60: 887-891.

Hall, E. L., B. K. Rouge and R. P. Kruger, 1977, "Automated Chest X-Ray Analysis," SPIE Applications of Optics in Medicine and Biology, 89: 109-118.

Haralick, R. M., 1975, "A Resolution-Preserving Textural Transform for Images," Proceedings of the IEEE Conference on Computer Graphics, Pattern Recognition, and Data Structures, Los Angeles. 51-54.

Haralick, R. M., 1979, "Statistical and Structural Approaches to Texture," Proceedings of the IEEE, 67(5): 786-804.

Haralick, R. M. and K. Shanmugam, 1973, "Computer Classification of Reservoir Sandstones," IEEE Transactions on Geoscience Electronics, 11: 171-177.

Haralick, R. M., K. Shanmugam and I. Dinstein, 1973, "Textural Features for Image Classification," IEEE Transactions on Systems, Man and Cybernetics, 3: 610-621.

Harvey, L. O. and M. J. Gervais, 1981, "Internal Representation of Visual Texture and the Basis for the Judgement Similarity," Journal of Experimental Psychology: Human Perception and Performance, 7: 741-753.

Hassner, M. and J. Sklansky, 1978, "Markov Random Fields as Models of Digitized Image Texture," Proceedings of the IEEE Conference on Pattern Recognition and Image Processing, Chicago. 346-351.

Hawkins, J. K., 1969, "Textural Properties for Pattern Recognition," in Picture Processing and Psychopictorics, B. Lipkin and A. Rosenfeld, editors, New York: Academic Press.

Hayes, K. C., A. N. Shah and A. Rosenfeld, 1974, "Texture Coarseness: Further Experiments," IEEE Transactions on Systems, Man and Cybernetics, 4: 467-472.

Hildreth, E. C., 1980, "Implementation of a Theory of Edge Detection," MIT Artificial Intelligence Laboratory AI-TR-579.

Hofstadter, D. H., 1979, Godel, Escher, Bach: An Eternal Golden Braid, New York: Basic.

Hubel, D. H., 1963, "The Visual Cortex of the Brain," Scientific American, 209(November 1963): 54-62.

Jain, A. K., S. P. Smith and E. Backer, 1980, "Segmentation of Muscle Cell Pictures: A Preliminary Study," IEEE Transactions on Pattern Analysis and Machine Intelligence, 2: 232-242.

Johnson, L. R. and A. K. Jain, 1981, "An Efficient 2-Dimensional FFT Algorithm," IEEE Transactions on Pattern Analysis and Machine Intelligence, 3: 698-701.

Julesz, B., 1962, "Visual Pattern Discrimination," IRE Transactions on Information Theory, 8: 840-892.

Julesz, B., 1965, "Texture and Visual Perception," Scientific American, 212(February 1965): 38-54.

Julesz, B., 1975, "Experiments in the Visual Perception of Texture," Scientific American, 232(April 1975): 34-43.

Julesz, B., 1981, "Textons, the Elements of Texture Perception, and Their Interactions," Nature, 290(12 March 1981): 91-97.

Julesz, B. and T. Caelli, 1979, "On the Limits of Fourier Decompositions in Visual Texture Perception," Perception, 8: 69-73.

Julesz, B., E. N. Gilbert, L. A. Shepp and H. L. Frisch, 1973, "Inability of Humans to Discriminate Between Visual Textures That Agree in Second-Order Statistics - Revisited," Perception, 2: 391-405.

Julesz, B., E. N. Gilbert and J. D. Victor, 1978, "Visual Discrimination of Textures with Identical Third-Order Statistics," Biological Cybernetics, 31: 137-140.

Kashyap, R. L., 1980, "Univariate and Multivariate Random Field Models for Images," Computer Graphics and Image Processing, 12: 257-270.

Kettig, R. and D. Landgrebe, 1976, "Classification of Multispectral Image Data by Extraction and Classifiction of Homogeneous Objects," IEEE Transactions on Geoscience Electronics, 14, no. 1: 19-26.

Landeweerd, G. H. and E. S. Gelsema, 1977, "The Use of Nuclear Texture Parameters in the Automatic Analysis of Leucocytes," Pattern Recognition, 9: 57-61.

Landgrebe, D. A., 1981, "Analysis Technology for Land Remote Sensing," Proceedings of the IEEE, 69: 628-642.

Laws, K. I., 1980, "Textured Image Segmentation," University of Southern California Image Processing Institute TR-940.

Lendaris, G. and G. Stanley, 1970, "Diffraction Pattern Sampling for Automatic Pattern Recognition," Proceedings of the IEEE, 58: 198-216.

Lu, S. Y. and K. S. Fu, 1978a, "A Syntactic Approach to Texture Analysis," Computer Graphics and Image Processing, 7: 303-330.

Lu, S. Y. and K. S. Fu, 1978b, "Stochastic Tree Grammar Inference for Texture Synthesis and Discrimination," Proceedings of the IEEE Conference on Pattern Recognition and Image Processing, Chicago. 340-345.

Marr, D., 1980, "Visual Information Processing: The Structure and Creation of Visual Representations," Philosophical Transactions of the Royal Society of London, B290: 199-217.

Marr, D. and E. Hildreth, 1980, "Theory of Edge Detection," Proceedings of the Royal Society of London, B207: 187-217.

Marr, D. and M. K. Nishihara, 1978, "Visual Information Processing: AI and the Sensorium of Sight," Technology Review, 81: 28-49.

Marr, D. and T. Poggio, 1979, "A Computational Theory of Stereo Vision," Philosophical Transactions of the Royal Society of London, B204: 301-328.

Marr, D., S. Ullman and T. Poggio, 1979, "Bandpass Channels, Zero-Crossings, and Early Visual Information Processing," Journal of the Optical Society of America, 69: 914-916.

McCormick, B. H. and S. N. Jayaramamurthy, 1974, "Time Series Model for Texture Synthesis," International Journal of Computer and Information Science, 3(4): 329-343.

McCormick, B. H. and S. N. Jayaramamurthy, 1975, "A Decision Theory Method for the Analysis of Texture," International Journal of Computer and Information Science, 4(1): 1-38.

Mitchell, O. R. and S. C. Carlton, 1978, "Image Segmentation Using a Local Extrema Texture Measure," Pattern Recognition, 10: 205-210.

Mitchell, O. R., C. Myers and W. Boyne, 1977, "A Max-Min Measure for Image Texture Analysis," IEEE Transactions on Computers, 25: 408-414.

Modestino, J. W., R. W. Fries and A. L. Vickers, 1981, "Texture Discrimination Based Upon an Assumed Stochastic Texture Model," IEEE Transactions on Pattern Analysis and Machine Intelligence, 3: 557-580.

Mostafavi, H. and D. Sakrison, 1976, "Structure and Properties of a Single Channel in the Human Visual System," Vision Research, 16:

957-968.

Mui, J. K., K. S. Fu and J. W. Bacus, 1977, "Automatic Classification of Blood Cell Neutrophils," Journal of Histochemistry and Cytochemistry, 25: 633-640.

Nathan, R., 1970, "Spatial Frequency Filtering," in Picture Processing and Psychopictorics, B. Lipkin and A. Rosenfeld, editors. 151-163.

Nilsson, N. J., 1980, Principles of Artificial Intelligence, Palo Alto:Tioga.

Ochs, A. L., 1979, "Is Fourier Analysis Performed by the Visual System or by the Visual Investigator?" Journal of the Optical Society of America, 69: 95-98.

Oppenheim, A. V. and J. S. Lim, 1981, "The Importance of Phase in Signals," Proceedings of the IEEE, 69: 529-541.

Pantle, A. and R. Sekuler, 1968, "Size-Detecting Mechanisms in Human Vision," Science, 162: 1146-1148.

Papoulis, A., 1962, The Fourier Integral and Its Applications, New York: McGraw-Hill.

Papoulis, A., 1965, Probability, Random Variables and Stochastic Processes, New York: McGraw-Hill.

Pavlidis, T., 1977, Structural Pattern Recognition, New York: Springer-Verlag.

Perkins, W. A., 1978, "A Model-Based Vision System for Industrial Parts," IEEE Transactions on Computers, 27: 126-143.

Pratt, W. K., 1978, Digital Image Processing, New York: Wiley.

Pratt, W. K., O. D. Faugeras and A. Gagalowicz, 1978, "Visual Discrimination of Stochastic Texture Fields," IEEE Transactions on Systems, Man and Cybernetics, 8: 796-804.

Pratt, W. K., O. D. Faugeras and A. Gagalowicz, 1981, "Applications of Stochastic Texture Field Models to Image Processing," Proceedings of the IEEE, 69: 542-551.

Pressman, N. J., 1976, "Markovian Analysis of Cervical Cell Images," Journal of Histochemistry and Cytochemistry, 24(1): 138-144.

Resnikoff, H. L., 1981, "Selective Omission of Information and Machine Intellegence," Presented at the Machine Intelligence and Perception Symposium at the Annual Meeting of the American Association for the Advancement of Science, Toronto.

Richards, W. and A. Polit, 1974, "Texture Matching," Kybernetic, 16: 155-162.

Rosenblatt, M. and D. Slepian, 1962, "Nth Order Markov Chains with Any Set of N Variables Independent," Journal of the Society for Industrial and Applied Mathematics, 10: 537-549.

Rosenfeld, A., 1979, "Some Recent Developments in Texture Analysis," Proceedings of the IEEE Conference on Pattern Recognition and Image Processing, Chicago. 618-622.

Rosenfeld, A., 1981, "Image Pattern Recognition," Proceedings of the IEEE, 69: 596-605.

Rosenfeld, A. and A. C. Kak, 1981, Digital Picture Processing, second edition, New York: Academic Press.

Rosenfeld, A. and M. Thurston, 1971, "Edge and Curve Detection for Visual Scene Analysis," IEEE Transactions on Computers, 20: 562-569.

Rosenfeld, A., M. Thurston and Y. Lee, 1972, "Edge and Curve Detection: Further Experiments," IEEE Transactions on Computers, 21(7): 677-715.

Rosenfeld, A. and E. Troy, 1974, "Visual Texture Analysis," Computer Graphics and Image Processing, 4: 172-179.

Sachs, M. B., J. Nachmias and J. G. Robson, 1971, "Spatial Frequency Channels in Human Vision," Journal of the Optical Society of America, 61: 1176-1186.

Schachter, B., A. Rosenfeld and L. S. Davis, 1978a, "Random Mosaic Models for Textures," IEEE Transactions on Systems, Man and Cybernetics, 8: 694-702.

Schachter, B., A. Rosenfeld and L. S. Davis, 1978b, "Some Experiments in Image Segmentation by Clustering of Local Feature Values," Pattern Recognition, 11: 19-28.

Schmitt, F. and D. Massaloux, 1981, "Texture Synthesis Using a Markov Model," Proceedings of the IEEE Conference on Pattern Recognition and Image Processing, Dallas. 593-596.

Shanmugam, K. S., F. M. Dickey and J. A. Green, 1979, "An Optimal Frequency Domain Filter for Edge Detection in Digital Pictures," IEEE Transactions on Pattern Analysis and Machine Intelligence, 1: 37-49.

Siromoney, G., R. Siromoney and K. Krithivasan, 1972, "Abstract Families of Matrices and Picture Languages," Computer Graphics and Image Processing, 1: 284-307.

Sklansky, J., 1978, "Image Segmentation and Feature Extraction," IEEE Transactions on Systems, Man and Cybernetics, 8: 237-247.

Stevens, K., 1980, "Surface Perception from Local Analysis of Texture and Contour," MIT Artificial Intelligence Lab TR-512.

Tamura, H., S. Mori and Y. Yamawaki, 1978, "Textural Features Corresponding to Visual Perception," IEEE Transactions on Systems, Man and Cybernetics, 8(6): 460-473.

Tennenbaum, J. M., et al, 1978, "Prospects for Industrial Vision," SRI International Technical Report TR-175.

Thompson, W., 1977, "Textural Boundary Analysis," IEEE Transactions on Computers, 26: 272-276.

Tomita, F., 1981, "Hierarchical Description of Textures," Proceedings of the International Joint Conference on Artificial Intelligence, Vancouver. 728-733.

Tomita, F., M. Yachida and S. Tsuji, 1973, "Detection of Homogeneous Regions by Structural Analysis," Proceedings of the International Joint Conference on Artificial Intelligence, Palo Alto. 564-571.

Tomita, F. and M. Yachida, 1973, "A Structural Analyzer for a Class of Textures," Computer Graphics and Image Processing, 2: 216-231.

Trussel, H. J., 1981, "Processing of X-Ray Images," Proceedings of the IEEE, 69: 615-627.

Tsuji, S. and F. Tomita, 1973, "A Structural Analyzer for a Class of Textures," Computer Graphics and Image Processing, 2: 216-231.

Vilnrotter, F., R. Nevatia and K. Price, 1981, "Structural Analysis of Natural Textures," Proceedings of the DARPA Image Understanding Workshop, Washington, D. C. 61-68.

Wechsler, H., 1980, "Texture Analysis - A Survey," Signal Processing, 2: 271-282.

Weszka, J., C. Dyer and A. Rosenfeld, 1976, "A Comparative Study of Texture Measures for Terrain Classification," IEEE Transactions on Systems, Man and Cybernetics, 6(4): 269-285.

Wilson, H. R. and J. R. Bergen, 1979, "A Four-Mechanism Model for Threshold Spatial Vision," Vision Research, 19: 19-32.

Winston, P. H., 1977, Artificial Intelligence, Reading, MA: Addison-Wesley.

Zucker, S. W., 1976, "Toward a Model of Texture," Computer Graphics and Image Processing, 5(2): 190-202.

Zucker, S. W., 1981, "Computer Vision and Human Perception: An Essay on the Discovery of Constraints," Proceedings of the IEEE Conference on Pattern Recognition and Image Processing, Dallas. 1102-1116.

Zucker, S. W. and P. Cavanaugh, 1980, "Constructive Texture Perception: Orientation Anisotopies in Discrimination," McGill University Dept. of Electrical Engineering. TR 80-8.

Zucker, S. W. and K. Kant, 1981, "Multiple-Level Representations for Texture Discrimination," Proceedings of the IEEE Conference on Pattern Recognition and Image Processing, Dallas. 609-614.