

# COLOR INVARIANT OBJECT AND TEXTURE RECOGNITION

by

Subho Sanjay Chatterjee

Bachelor of Engineering (B.E.)

Victoria Jubilee Technical Institute,

University of Bombay 1992

A THESIS SUBMITTED IN PARTIAL FULFILLMENT  
OF THE REQUIREMENTS FOR THE DEGREE OF  
MASTER OF SCIENCE  
in the School  
of  
Computing Science

© Subho Sanjay Chatterjee 1995

SIMON FRASER UNIVERSITY

August 1995

All rights reserved. This work may not be  
reproduced in whole or in part, by photocopy  
or other means, without the permission of the author.

## APPROVAL

**Name:** Subho Sanjay Chatterjee  
**Degree:** Master of Science  
**Title of thesis:** Color Invariant Object and Texture Recognition

**Examining Committee:** Dr. Jim Delgrande, Computing Science, SFU  
Chair

---

Dr. Brian V. Funt, Senior Supervisor  
Professor, Computing Science, SFU

---

Dr. David Fracchia, Supervisor  
Asst. Professor, Computing Science, SFU

---

Dr. Ze-Nian Li, External Examiner  
Associate Professor, Computing Science, SFU

**Date Approved:** August 15, 1995.

SIMON FRASER UNIVERSITY

**PARTIAL COPYRIGHT LICENSE**

I hereby grant to Simon Fraser University the right to lend my thesis, project or extended essay (the title of which is shown below) to users of the Simon Fraser University Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users. I further agree that permission for multiple copying of this work for scholarly purposes may be granted by me or the Dean of Graduate Studies. It is understood that copying or publication of this work for financial gain shall not be allowed without my written permission.

Title of Thesis/Project/Extended Essay

Color Invariant Object and Texture Recognition.

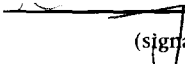
---

---

---

---

Author:

  
(signature)

\_\_\_\_\_  
(name)

August 16, 1995

\_\_\_\_\_  
(date)

# Abstract

A fast, color-based algorithm for recognizing objects and textures viewed under an unknown illuminant is presented. Objects are indexed by six numbers: the angles of the object's color distribution and the angles of the color edge distributions. If  $R$ ,  $G$  and  $B$  denote the 3 color bands of the image of an object (stretched out as vectors) then the color angular index comprises the 3 inter-band angles (one per pair of vectors). Similarly the edge angular index comprises the angles of within band color edges represented as feature vectors. In the general case the distribution of colors, and in turn the angular index, will depend on the color of the illuminant. If, however, the original color bands are transformed by a *sharpening* transform before computing the distribution angles, then the angular index is illuminant independent. The *sharpening* transform in this case refers to linear combinations of the sensor responses which yield narrow band or *sharper* sensors. Indexing using angles calculated post sharpening delivers excellent recognition for a variety of illuminations.

Textures are represented by the six spatial correlation functions of the intra and inter band image color distributions. Treating these correlations as vectors, a new set of angular invariants is computed as the set of angles between the six vectors. The spatial correlation angles are shown to preserve illumination independence under the sharpening transform. Tests show that these angles possess sufficient discrimination to correctly identify a set of textures. The color and edge angular index is used to represent the spatial interactions in color textures and provides a rich representation for color textures. A coarse form of texture classification and segmentation is developed by computing these texture features locally across the image.

# Acknowledgments

My sincere thanks to my supervisor, Dr. Brian Funt, for his help and advice over the last two years. He has shaped much of my understanding and interest in color. This thesis owes a lot to the stimulating discussions with him.

Dr. Dave Fracchia has cheerfully allowed himself to be drawn into color vision and I am grateful for his many comments towards the improvement of this work. I thank Dr. Ze-Nian Li for examining my research and for his valuable suggestions.

Graham Finlayson deserves much of the credit for the ideas in this thesis; it was he who first pointed out the invariants I have used. He has been an invaluable source of encouragement, both as friend and colleague, during my stay here. Thanks go to Mark Drew and Louis Brassard for their interest in my work, to Janet Dueck for the mega-bytes and the nicotine and to Kobus Barnard for his help in the lab.

To Daya Ram Gaur, for volunteering to proof-read my thesis and to Ashish Pimplapure for his suggestions. A special mention for Julie Turcotte, for being a good friend, for the haircuts and the carcinogens.

I am grateful to a close friend, Apurva Desai, for introducing me to Simon Fraser University. Lastly, I am indebted to my parents for their help and support. I would not have been here without their constant encouragement — as Joey the Lips Fagan once said, “Parents are soul”.

# Contents

<b>Abstract</b>	<b>iii</b>
<b>Acknowledgments</b>	<b>iv</b>
<b>1 Introduction</b>	<b>1</b>
<b>2 Literature Review</b>	<b>4</b>
2.1 Introduction . . . . .	4
2.2 Color object recognition . . . . .	5
2.2.1 Swain: color indexing . . . . .	5
2.2.2 Color constant color indexing (CCCI) . . . . .	6
2.2.3 Moment invariants of color histograms . . . . .	8
2.3 Texture Analysis . . . . .	11
2.3.1 Applications . . . . .	12
2.3.2 Use of Color in Textures . . . . .	12
<b>3 Color and Edge angle invariants</b>	<b>16</b>
3.1 Color Distribution Invariants . . . . .	16
3.1.1 Finite Dimensional Models . . . . .	17
3.1.2 Diagonal Models . . . . .	18
3.1.3 Angle Invariants of color distributions . . . . .	19
3.1.4 Data Sets used to test invariance of angles . . . . .	20
3.1.5 Normalized Euclidean Distance as a measure . . . . .	22
3.1.6 Are 3 numbers sufficient ? . . . . .	24

3.2	Angular edge invariants . . . . .	26
3.2.1	Choosing an edge operator . . . . .	28
3.2.2	Edge invariants . . . . .	29
3.3	Combining edges and distributions . . . . .	29
<b>4</b>	<b>Angular color texture descriptors</b>	<b>33</b>
4.1	Introduction . . . . .	33
4.2	Representations of color textures . . . . .	34
4.3	Color texture invariants . . . . .	35
4.3.1	Implementation details . . . . .	39
4.3.2	Datasets . . . . .	41
4.3.3	Performance Analysis . . . . .	42
4.4	Color texture classification . . . . .	47
4.4.1	Local texture invariants . . . . .	49
4.4.2	Invariance of local representations . . . . .	51
4.4.3	Representing the invariant set . . . . .	52
<b>5</b>	<b>Results</b>	<b>53</b>
5.1	Object Recognition Results . . . . .	54
5.1.1	Objects under changing illumination: . . . . .	54
5.1.2	Swain's Database: . . . . .	56
5.1.3	Remarks . . . . .	57
5.2	Color texture recognition results . . . . .	57
5.2.1	Color texture classification . . . . .	60
<b>6</b>	<b>Conclusions</b>	<b>65</b>
6.1	Application based invariant representation . . . . .	66
6.2	Extensions . . . . .	67
6.3	Texture segmentation schemes . . . . .	67
6.4	Concluding remarks . . . . .	68
	<b>Bibliography</b>	<b>74</b>

# List of Tables

3.1	Comparison of invariant distances for color and edge distributions of two images. . . . .	28
5.1	Recognition results for angle invariants . . . . .	55
5.2	Results using different edge operators . . . . .	55
5.3	Comparative results of different recognition algorithms . . . . .	56
5.4	Database of 55 real objects . . . . .	57
5.5	Texture recognition for a database of ten textures . . . . .	59
5.6	Texture recognition with change in illumination and rotation about the optical axis. . . . .	59



# List of Figures

2.1	Gaussian sphere . . . . .	14
3.1	Real objects imaged under varying illumination conditions . . . . .	21
3.2	Camera sensitivities . . . . .	23
3.3	Measure of invariance for color distribution angles . . . . .	24
3.4	Measure of invariance with sharpening . . . . .	25
3.5	Insufficiency of color angle invariants . . . . .	27
3.6	Color edges of the images . . . . .	27
3.7	Edge invariant distances between images . . . . .	30
3.8	Invariant distances between images for invariant set $\mathcal{I}$ . . . . .	31
3.9	Sharpened invariant distances between images for invariant set $\mathcal{I}$ . . . . .	31
4.1	Color texture images . . . . .	36
4.2	Spatial correlation functions of a color texture . . . . .	37
4.3	Spatial correlation functions changed due to: a) rotation in the image plane; b) due to changing illumination . . . . .	38
4.4	Correlation of image color planes . . . . .	40
4.5	A texture under varying illumination and rotations . . . . .	43
4.6	Another color texture under similar illumination and rotation conditions . . . . .	44
4.7	Invariant distances for normalized correlations . . . . .	45
4.8	Invariant distances for correlations defined by Eq 4.1 . . . . .	45
4.9	Invariant distances for correlations using a Fourier transform approach . . . . .	46
4.10	Invariant distances for color-edge angle invariants for natural textures . . . . .	47
4.11	A natural scene with multiple textures . . . . .	48

5.1	Database of natural textures used for texture classification . . . . .	61
5.2	Context images of natural scenes used for texture classification . . . .	62
5.3	a)Invariant images for a context scene; b) Clustering results for the invariants . . . . .	63
5.4	Segmented texture image . . . . .	64

# Chapter 1

## Introduction

“When considering, with reference to color, the objects that constitute the world, it is quickly noted that these fleeting appearances, which so readily appear and disappear through certain angles of these objects, are not accidental but are dependent upon definite laws.”

— Goethe, *“Color Theory”*

Object recognition — the identification of key features from an image which correctly characterize the given object as a member of a set of well known objects — is one of the major goals of computer vision. This raises the question of what constitutes ‘key features’: typically they are measurable object properties such as 2D shape, edges, surface curvature and color which can be extracted from images of the object and *a priori* knowledge about the world. The obvious problem is that objects may be viewed from different positions under various rotations, orientations, changes of scale and lighting conditions. This often gives rise to different features for the same object and makes recognition a hard task. It is desirable therefore that the characterizations be invariant to the set of transformations which give rise to these differences.

The work presented in this thesis is based on the premise that color distributions in

an object are sufficiently characteristic and discriminatory in nature to serve as a basis of identification. Previous work done by Swain [SB91] and Funt and Finlayson [FF95] has shown that color distributions are indeed sufficient for recognizing a wide variety of colorful man-made objects. The use of color distributions renders the characterization invariant to translation and rotation about the optical axis. Distributions normalized by the area of the image are robust to changes in scale. Equally importantly, distributions change gradually with occlusion and rotations about other axes — the only caveat being that they are sensitive to changes in illumination. Swain’s “Color Indexing” [SB91] which does not account for illumination changes thus fares poorly in spectrally different lighting conditions. Funt and Finlayson solve this by using color ratio distributions which are invariants under this transformation. More recently work done by Healey and Slater [HS94] has used moment invariants of color histograms as a key to recognition.

The key motivation here is to find a color invariant descriptor of this distribution space which is compact, reliable and efficiently computable. With this in mind, I use Finlayson et al.’s [FDF94b] result for a generalized diagonal model of color constancy which essentially states that a change in illumination color can be accounted for by an independent scaling of each color band in the image, after the sensors are transformed by a linear combination of the original sensor responses. With this assumption, we can treat each band color distribution as a feature vector in  $N$  dimensions, where  $N$  is the number of pixels in the image. A change in the color of illumination corresponds to a change in length of each feature vector; the angles between these vectors however remain invariant. I index objects on these angle invariants. A more formal derivation of these invariants is described in chapter 3. The algorithm has been tested on a variety of colored objects with impressive results.

I try another variation on the above theme where I investigate the angle invariants of image edges by taking the Laplacian of the image and computing the invariants of the resultant image. Since the color and edge distributions will in general yield different information about the object, I conjecture that using both sets of invariants will improve the performance of the algorithm. Indeed, we will see for Swain’s database, the combination of color and edge invariants exhibits better results than the use of

color distribution invariants alone.

The notion of angle invariants has been extended to color texture recognition. Representing textures as spatial correlation feature vectors, I show that the diagonal model holds for the set of feature vectors. I propose that textures can be represented by the angles between these six feature vectors. While this is a deviation from the norm in terms of how textures are recognized with statistical methods, these angles are surprisingly robust to rotation and changes in color of light and possess sufficient discriminatory power. The algorithm in this case has been formalized in Chapter 4. Results are provided for tests on a variety of colored textures. I also examine the use of the color distribution and edge distribution invariants described above as a representation for textures. We observe that the results are superior to those obtained by using a spatial correlation model.

Taking this a step further, the color and edge angles are locally computed across an image to yield a set of invariant descriptors for multi-textured images. These invariants are indexed by a hashing function which partitions the invariant space equally. Hash buckets with high bin counts are matched to a model database of textures for classifying the various textures present in an image. By computing these invariants for every pixel location in the image, a set of invariants for every image point is obtained. A clustering algorithm applied to these invariants yields a form of coarse image segmentation.

This thesis is organized linearly. In the next chapter, I discuss some of the related work in color object recognition as well as certain relevant aspects of color texture recognition. Chapters 3 and 4 discuss formal aspects of the invariant derivation and cite some results about the performance of the various algorithms to be discussed. Chapter 5 is essentially devoted to the results obtained from the various algorithms and as such can be distinctly separated into two sections — the first dealing with object recognition and the next with texture. I conclude by remarking upon the limitations of the method, improvements which can be applied to the algorithm and possible applications.

# Chapter 2

## Literature Review

“It seemed by then as if there were no facts anywhere in the universe, in his own brain or anyone else’s or just lying loose, which could possibly be brought within his present scope.”

— *Kingsley Amis, “Lucky Jim”*

### 2.1 Introduction

The usefulness of color in identifying objects has been apparent for quite some time; what is less clear is the role color plays in the recognition process. Even a cursory study of the psychophysical literature in the field [WL93] shows a wide disagreement over exactly what benefits humans accrue from color vision. Of course a point to note here is that most of the psychophysical experiments in this regard have been carried out on natural objects with a predominant color. The emphasis has been on a single color as a feature of recognition and the extent to which it helps in pruning a search tree.

My contention here is significantly different in that I say that the distribution of all the colors in an object is characteristic. For example, if we say that an object is red, it could either be an apple or a tomato (given we have a small database of vegetables). If however we say that the object is red with whitish patches, then the

probability of it being an apple is far larger than that of it being a tomato. This assumption holds very strongly in the case of man-made objects which usually have fixed proportions of colors in them.

## 2.2 Color object recognition

### 2.2.1 Swain: color indexing

Swain [SB91] was the first to exploit the notion of uniqueness in color distributions. Each color pixel was represented by its opponent colors which are defined as follows [BB82]:

$$\begin{aligned}rg &= r - g \\by &= 2 * b - r - g \\wb &= r + g + b\end{aligned}$$

Here  $r$ ,  $g$ , and  $b$  represent the red green and blue values respectively. The  $rg$ ,  $by$  and  $wb$  axes are analogous to the opponent color axes used by the human visual system (Lennie and D'Zmura [DL86]). They were used to allow the intensity ( $wb$ ) axis to be more coarsely sampled than the other two, as the intensity axis is more sensitive to lighting variation from shadows and distance to the light source. This distribution was then modeled as a three dimensional histogram of the pixel values. For recognition, the color histogram of a test image was matched with the model histograms in the database. Histogram intersection was defined as a metric :

$$\sum_{r=1}^{N_r} \sum_{g=1}^{N_g} \sum_{b=1}^{N_b} \min(h_{r,g,b}^i, h_{r,g,b}^j) \quad (2.1)$$

where  $h_{r,g,b}^i, h_{r,g,b}^j$  represent the image and model histogram bins respectively,  $\min()$  returns the minimum of two numbers, and  $N_r, N_g, N_b$  represent the number of bins for each color band.

The result of the histogram intersection of a model histogram with an image histogram is the number of pixels in the model which have the corresponding pixels of the same color in the image. To obtain a fractional match, the intersection is normalized

by the number of pixels in the model. This normalized intersection becomes what is known as the *city block metric* when each histogram is normalized by its area.

Such a representation is invariant to rotation of the object about the optical axis, as the color histogram does not encode spatial information about the object. Rotation about a different axis and occlusion change the resultant histogram. However as Swain has shown with various experiments, the histogram changes slowly with respect to these transformations. The scale of the object affects the size of the histogram and thus any match to a database object. The histogram can be normalized by its total area to make it invariant to changes in scale.

Swain tested his algorithm and reported excellent results for Color Indexing. His database consisted of a set of 66 objects, all highly colored and of different shapes and sizes. His test data set comprised 24 objects, taken under the same illuminant, but with different orientations in space.

The shape of the histogram changes with a change in illuminant color, and thus Swain's color indexing does not work under a change in illumination. Swain suggested the use of a color constancy algorithm as a preprocessing step to color indexing; this would solve the problem posed by a changing illumination, and render color indexing useful under spectrally varying conditions.

### 2.2.2 Color constant color indexing (CCCI)

Subsequent to Swain's "Color Indexing", Funt and Finlayson [FF95] worked on a more elegant approach to indexing color invariant pixel distributions. Their approach was based on the observation that the spectral variation of illumination over an object is smooth, i.e. without local discontinuities. In such a case the illumination spectrum can be assumed to remain the same over a local region in the image. Color constant descriptors can be obtained by taking ratios of adjacent pixels in the image. This follows from a hypothesis that a change in the illumination of an object corresponds to an independent scaling in each color band.



This was first hypothesized by von Kries in 1904 [WS82] to explain the phenomenon of color constancy in humans on the basis of chromatic adaptation. Chromatic adaptation is the phenomenon by which our eyes adjust to the ambient illumination in a scene over time. The von Kries adapted responses to a surface  $S^x(\lambda)$  in sensor channel  $k$  can be written as

$$d_k^x = \frac{\int S^x(\lambda)E(\lambda)R_k(\lambda)d\lambda}{\int E(\lambda)R_k(\lambda)d\lambda} \quad (2.2)$$

where  $E(\lambda)$  is the incident illumination and  $R_k(\lambda)$  is the sensor response for a sensor class  $k$ . von Kries conjectures that for any given illuminant  $E(\lambda)$ ,  $d_k^x$  will remain constant. This constant is often computed by assuming the presence of a white (uniform) reflector in every scene of the image. In reality, surfaces under varying illuminants are only approximately modeled by the von Kries invariants [WB82].

Under this assumption of independent scaling, given two pixel locations  $x_1$  and  $x_2$  we can write the ratio invariants as:

$$\rho_k^{x_1} = \delta_k \varrho_k^{x_1} \quad \rho_k^{x_2} = \delta_k \varrho_k^{x_2} \quad \Rightarrow \quad \frac{\rho_k^{x_1}}{\rho_k^{x_2}} = \frac{\varrho_k^{x_1}}{\varrho_k^{x_2}} \quad (2.3)$$

where  $\rho_k^{x_1}$  and  $\varrho_k^{x_1}$  represent the responses under two different illuminations  $E(\lambda)$  and  $\hat{E}(\lambda)$  at location  $x^1$  and

$$\delta_k = \frac{\int \hat{E}(\lambda)R_k(\lambda)d\lambda}{\int E(\lambda)R_k(\lambda)d\lambda}$$

These ratios are illumination invariant and are computed for each of the three bands. 3-dimensional histograms created on the basis of these ratio invariants (called *ratio histograms*) are therefore illumination independent descriptors of the color distribution of a scene. Ratios of course encode color edge information about a scene (ratios of a patch with constant color will be unity, and not contribute information about the color content of the patch). Thus they encode substantially different information about a scene in comparison to Swain's "Color Indexing". These ratios are efficiently computed by taking the derivative of the log of the image.

Funt and Finlayson [FF95] use a histogram intersection metric similar to the one defined by Swain for histogram matching. The histogram creation process is complex, involving a probability analysis of the ratio space. This follows from the observation

that some ratios are more likely than others. The ratio histograms therefore have a bin distribution based on non-uniform sampling of the ratio space. Probability distributions for the ratio space are computed and used to generate optimal bin distributions for the ratio histograms. For the histogram intersection, in a slight deviation from Swain, the trivial ratio bin  $(1, 1, 1)$  is neglected for purposes of intersection.

Funt and Finlayson show that ratio histograms provide a rich representation for color-based recognition tasks by successfully recognizing a wide variety of objects. Results of CCCI on Swain's database are comparable to Color Indexing.

Amongst the various advantages of using ratios as a representational framework is the fact that they encode spatial information about the object by identifying *color edges* in the scene. Problems arise due to the imperfection of the diagonal model as a vehicle for color constancy, in addition to the problem of ratios at low pixel values, where the noise in an image can introduce significant errors into the computation. Finlayson et al. [FDF94b] have subsequently shown that the diagonal model holds for color constancy if the sensors are transformed by a linear transformation to narrow-band sensors. In the case of narrow-band sensors, intuitively, there is little interaction between the sensor channels, and an independent scaling proves to be a good model of color constancy. Indeed, they have shown that under certain assumptions of surfaces or illuminants, a diagonal model under a sharpening transformation provides perfect color constancy [FDF94a].

### 2.2.3 Moment invariants of color histograms

Recently, Healey has examined the behavior of higher order moment invariants of color histograms. Using a finite dimensional linear model of color image formations, he shows that the color histograms of a scene under different illuminations are related by an affine transformation.

Under an image formation system, let

$$\rho_k^x = \int_w E(\lambda) S^{x'}(\lambda) R_k(\lambda) d\lambda \quad (2.4)$$

where  $\lambda$  denotes the wavelength,  $E(\lambda)$  denotes the spectral power distribution of the illuminant,  $S^{x'}(\lambda)$  is the spectral reflectance of the surface and  $R_k(\lambda)$  is the sensitivity

of the  $k$ th sensor class. The location  $x'$  on the surface is projected onto location  $x$  on the sensor array.

The spectral reflectance function  $S^{x'}(\lambda)$  of a surface can be approximated at location  $x'$ , by using a set of basis functions:

$$S^{x'}(\lambda) = \sum_{1 \leq j \leq n} \sigma_j^{x'} S_j(\lambda)$$

where  $S_j(\lambda)$  is a set of  $n$  fixed basis functions. If  $\rho^x$  denotes the set of set of sensor measurements at location  $x$  on the sensor array as a column vector  $\rho^x = (\rho_1^x, \rho_2^x, \dots, \rho_n^x)^t$  and  $\sigma^{x'}$  denotes the column vector  $\sigma^{x'} = (\sigma_1^{x'}, \sigma_2^{x'}, \dots, \sigma_n^{x'})^t$  of spectral reflectance coefficients, we can write

$$\rho^x = A \sigma^{x'} \quad (2.5)$$

where  $A$  is the  $n \times n$  matrix with entries

$$A_{kj} = \int E(\lambda) S_j(\lambda) R_k(\lambda) d\lambda$$

We see that the matrix  $A$  depends on the illumination, but not on the location  $x'$  in the image. Images taken under a different illumination  $\bar{E}(\lambda)$ , will yield a different matrix  $\bar{A}$ . For any illuminants,  $E(\lambda)$  and  $\bar{E}(\lambda)$  such that  $A$  and  $\bar{A}$  are non-singular, we can write

$$\bar{\rho}^x = M \rho^x \quad (2.6)$$

where  $M = \bar{A} A^{-1}$ . If  $H$  and  $\bar{H}$  are  $n$ -dimensional histograms that represent the distribution of color pixels in images  $\rho^x$  and  $\bar{\rho}^x$ , then from Eq 2.6 they are related as

$$\bar{H}(M\rho) = H(\rho) \quad (2.7)$$

Thus for a change in illuminants, the color histograms are related by an affine transformation.

Taubin and Cooper [TC92] have recently developed efficient algorithms for the computation of vectors of affine moment (or algebraic) invariants of functions. These vectors are invariant to affine transformations of the function and as such can be used for recognition in the presence of illumination changes. Let us define a centered

moment of a color histogram  $H(\rho)$  as

$$A^\alpha = \frac{1}{|H(\rho)|} \int (\rho - H')^\alpha H(\rho) d\rho \quad (2.8)$$

where  $H' = [1/|H(\rho)|] \int \rho H(\rho) d\rho$  is the mean and  $|H(\rho)| = \int H(\rho) d\rho$ . The notation  $\rho^\alpha$  denotes the monomial  $\rho_1^{\alpha_1} \rho_2^{\alpha_2} \rho_3^{\alpha_3}$  for a vector of non-negative integers  $\alpha = (\alpha_1, \alpha_2, \alpha_3)$  called a multi-index. The size of a multi-index is  $|\alpha| = \alpha_1 + \alpha_2 + \alpha_3$ . Vectors and matrices of these centered moments can be defined by a lexicographical ordering of these centered moments for a given size of index  $\alpha$ .

If the matrix  $M$  as defined in Eq 2.6 is orthogonal corresponding to a rotation in color space, then certain moment-matrix eigenvalues are invariants and do not depend on  $M$ . For affine transformations, invariants cannot be directly computed. An affine transformation is realized by a change of coordinates followed by an orthogonal transformation. As a result of this, Healey uses higher order moments (at least 3rd and 4th order) to calculate affine invariants for object recognition. These moment invariants encode higher order global shape information about the histogram and are used for recognition. Healey uses a set of six invariants in his work [HS94] to recognize objects. The computation of the moment invariants is efficient and can be done in  $30N$  multiplications, where  $N$  is the size of the image. Recognition is based on the Euclidean distance between the six invariants. This is far quicker than indexing on color or ratio histograms.

The initial impetus for the work done in this thesis derived from this result of Healey. Under the assumption of a diagonal model of color constancy, the affine transformation  $M$  reduces to a scaling function  $D$  (where  $D$  is a diagonal matrix). Moment invariants in this case can be easily computed and lower order moments which have to be neglected by Healey can be used for recognition. I argue that often higher order moments do not yield sufficient discriminatory invariants for recognition and that lower order moments contain more useful information about the image scene. As we will see later, the algorithm presented in this thesis proves more effective than Healey's in all the recognition tasks that it is used for.

## 2.3 Texture Analysis

Textures are studied in the computer vision community from a wide variety of perspectives depending upon motivation and application. This has led to a variety of different definitions for textures. Coggins [Cog82] has compiled a catalogue of texture definitions in the computer vision literature and I cite a few examples here.

- “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”. [TMY78]
- “The image texture we consider is non figurative 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 scalar 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.” [Har79]
- “The notion of texture seems to depend upon three ingredients: (i) some local ‘order’ is repeated over a region which is large in comparison to the order’s size, (ii) the order consists in the nonrandom arrangement of elementary parts, and (iii) the parts are roughly uniform entities having approximately the same dimension everywhere within the textured region.” [Haw69]

As we can see from the above, the definition of texture is formulated by different people depending upon the particular application and that there is no generally agreed upon definition. Image texture, defined as a function of the spatial variation in pixel intensities, is useful in a variety of applications and has been a subject of intense study by many researchers. One immediate application of image texture is the recognition of image regions using texture properties. Texture in this sense forms an important visual cue in identifying various types of homogeneous regions. This is known as *texture classification*. The goal of texture classification is to produce a classification

map of the input image where each uniform textured region is identified with the texture class to which it belongs.

### 2.3.1 Applications

Texture analysis methods are used in a variety of application domains [TJ93]. In inspection problems, texture plays a limited role in the automated defect detection in textiles and in the automated inspection of carpet wear and automobile paints. In medical image analysis, textures are used in the automated extraction of features from the image which are then used for a variety of classification tasks such as distinguishing abnormal tissue from normal tissue. Depending upon the application, extracted features capture morphological properties or color properties. Textural properties have been used in the classification of pulmonary disease. Interstitial fibrosis can be clearly identified from lesions in X-rays by texture variations. Sutton and Hall have used isotropic contrast measures, a directional contrast measure and a Fourier domain energy sampling to distinguish normal lungs from diseased ones. Texture analysis has been extensively used in remote sensing for classification of homogeneous regions with different kinds of terrains.

### 2.3.2 Use of Color in Textures

Multi-spectral data contains a large amount of information about a scene which can be used to advantage in recognition tasks. Color, while an important and defining quality of texture in many cases, has largely been neglected by the vision community. Some color texture models which capture the spatial interaction between the red, green and blue bands of a color image and are used for image synthesis are discussed in [GMTL86]. Color random field models for spatial interactions between generalized neighbor sets have been used for segmentation of natural scenes by Healey [HP95]. Healey has also used moment invariants of spatial correlation functions in order to model 3-dimensional color textures for geometry invariant recognition [KH94]. Scharcanski et al. [SHS92] have used image edges at different scales to discriminate between color textures. For each of these models, the color texture representation depends on

the scene illumination. Healey [HW95] has recently proposed a method of color texture recognition based upon the spatial correlations between color bands which is invariant to the incident illumination.

I will discuss some of Healey's work related to color texture recognition in greater detail as much of my work in color textures derives directly from his results.

### Geometry invariant recognition of color textures

Kondepudy and Healey [KH94] have derived a 3-dimensional model of color textures based upon the spatial interactions amongst the color bands of an image. They use algebraic moment invariants (described in an earlier section) to select a best match from the model database to the image. The model is then fitted to the image to verify the match and identify orientation angles of the 3-dimensional texture in space.

They define a correlation function for a pair of images  $I_i(x, y)$  and  $I_j(x, y)$  as

$$R_{ij}(n, m) = E\{[I_i(x, y) - \bar{I}_i][I_j(x + n, y + m) - \bar{I}_j]\} \quad (2.9)$$

where  $\bar{I}_i$  and  $\bar{I}_j$  are the means of the images, and  $E$  denotes the expectation. For a color image, there can be six correlation functions  $R_{rr}, R_{rg}, R_{rb}, R_{gg}, R_{gb},$  and  $R_{bb}$  covering all the possible autocorrelations and cross correlations.

To show that 3-dimensional geometric transformations of the texture produce corresponding transformations in these correlation functions, they develop a 3-dimensional model of image formation. Under an arbitrary position in space, a point on a textured planar patch can be represented by a point on the Gaussian sphere which represents its normal. This point is specified by an azimuth angle  $\theta$  and a polar angle  $\phi$  as shown in Fig 2.1.

Based on this model, Kondepudy derives a relationship between the correlation functions for an arbitrary rigid motion of the body as :

$$R_{ij}(\bar{n}, \theta, \phi, \alpha, k) = R_{ijN}(M_1 \bar{n}) \quad (2.10)$$

where  $R_{ijN}(\bar{n})$  represents the color correlation of an image of the plane when it is perpendicular to the optical axis.  $\alpha$  indicates the rotation of the object in the image

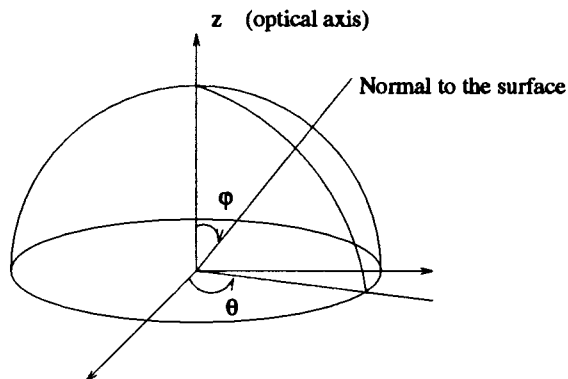


Figure 2.1: Gaussian sphere

plane, while  $k$  indicates a magnification or scaling factor to account for distance of the texture from the image plane.  $\bar{n}$  is the surface normal of the point and  $M_1$  is the transformation defining the rigid motion of the body. Under a geometric transform of the texture surface, the image coordinates are transformed as:

$$\hat{x} = M_1 \bar{x} + D$$

where  $\hat{x}$  is the transform of the image coordinate  $\bar{x} = (x, y)$ ,  $M_1$  includes effects of rotation and orientation and  $D$  is the translation vector. Given two images of texture planes, the transformation matrix is calculated from the images for each of the correlation functions and a fit is ascertained by a minimization procedure.

As this process is computationally intensive, Kondepudy and Healey use moment invariants as described before for selecting a best match to the set of texture models being recognized. Higher order moment invariants (up to 7th and 8th order moments) are calculated for each of the six correlation functions. A total of 54 moment invariants is used to match each texture. Only the first ranked match is selected for the minimization procedure; a successful identification is returned if the square of the difference error between the fitted model and the original image is below a chosen threshold. While excellent results are reported in identification of color textures, the model remains sensitive to illumination change as the structure of the correlation matrix changes with a change in the incident illumination.



### **Illumination invariant color texture recognition**

Work on illumination invariant texture representations has been done by Healey and Wang [HW95] using spatial correlation models to represent textures. These spatial correlation equations are described in Eq 2.9. A change in illumination transforms the structure of the spatial correlations. They show that a linear transformation maps the correlations for a change in illumination. Stretching the components of the correlation function out as a vector, he derives a set of orthonormal basis functions by a singular value decomposition of the six correlation vectors. For a change in illumination, it can be determined how well the new correlations can be represented using the basis vectors corresponding to the original correlation vectors. A distance function which calculates the squared fitted error is used as a matching function for a set of database textures. The biggest restriction here is that the textures are assumed to remain at the same orientation. A change in the orientation of a texture will change the structure of the correlation vector quite differently and the distance function will compute an incorrect match for the object.

# Chapter 3

## Color and Edge angle invariants

“Its chief merit is its simplicity — a simplicity so pure, so profound, in a word, so *simple*, that no other word will fitly describe it.”

— *Lewis Carroll, “Divisions and Digressions”*

### 3.1 Color Distribution Invariants

The light reflected from a surface depends on the spectral properties of the surface reflectance and the illumination incident on the surface. In the case of Lambertian surfaces, this light is simply the product of spectral distribution of the light source with the spectral reflectance of the surface. I restrict this discussion to Lambertian surfaces. In an imaging system, light reflected from a surface falls onto a planar array of sensors in the camera. Each location  $x$  on the array has  $k$  classes of sensors. The value  $\rho_k^x$  of each sensor output is given by the integral of its response function multiplied by the incoming color signal:

$$\rho_k^x = \int_w S^{x'}(\lambda) E^{x'}(\lambda) R_k(\lambda) d\lambda \quad (3.1)$$

where  $\lambda$  is the wavelength,  $R_k$  is the response function of the  $k$ th sensor class,  $E^{x'}$  is the incident illumination and  $S^{x'}$  is the surface reflectance function at location  $x'$  on the surface which is projected onto location  $x$  on the sensor array. I further assume

here that the illumination does not vary spectrally over the given surface, and so drop the index  $x'$  from  $E(\lambda)$ .

Consider the same surface illuminated by another illuminant  $\hat{E}$ , which is spectrally different from  $E$ . In this case the sensor response is given by :

$$\tilde{\rho}_k^x = \int_w S^{x'}(\lambda) \hat{E}(\lambda) R_k(\lambda) d\lambda \quad (3.2)$$

The human visual system has three classes of photo receptor cells, called *cones*, which are used to sense the color signal entering the eye. The sensor outputs from Eq 3.1 can therefore be represented by an ordered triple as  $p = \{\rho_1, \rho_2, \rho_3\}$  for each location  $x$ .

### 3.1.1 Finite Dimensional Models

A response to visual stimuli is band-limited by the frequency response of the sensor and can be represented by a discrete summation of a sampling of the color signal [SW64]. Cone sensitivities are usually considered to be significant between 380 nm to 770 nm. A sampling interval of 10 nm is common practice in the color science community [Kri47, Nic57]. Within this discretization framework, we can rewrite Eq 3.1 as

$$\rho_k^x = \sum_{i=1}^n S^{x'}(\lambda_i) E(\lambda_i) R_k(\lambda_i) \quad (3.3)$$

For almost all real illuminants and reflectances, Eq 3.3 approximates Eq 3.1 with small errors [SSS92].

The above equation can be further simplified by the use of finite dimensional models. Judd et al. [JMW64] measured 605 daylight illuminants and showed that they can be modeled well by a set of three basis functions.

$$E(\lambda) \approx \sum_{i=1}^{D(E)} \epsilon_i E_i(\lambda) \quad (3.4)$$

where  $E_i(\lambda)$  is a basis vector and  $\epsilon_i$  is the weight vector for the surface.

Parkkinen et al [PHJ89] measured the spectra of Munsell colors. Vrhel [VGI94] and Maloney [Mal86] evaluated various surface spectral reflectances and concluded that

reflectances could in general be well modeled by a set of 3 to 8 basis vectors [Fin95]

$$S^x(\lambda) \approx \sum_{j=1}^{D(S)} \sigma_j^{x'} S_j(\lambda) \quad (3.5)$$

where  $S_j(\lambda)$  is a basis vector and  $\sigma_j^{x'}$  is the weight vector for the surface  $x'$ . Given this, we can now rewrite Eq 3.3 in terms of a finite dimensional linear model :

$$\rho_k^x \approx \sum_{k=1}^n \sum_{i=1}^{D(E)} \sum_{j=1}^{D(S)} \epsilon_i E_i(\lambda_k) \sigma_j^x S_j(\lambda_k) R_k(\lambda_k) \quad (3.6)$$

Under a change of illumination as represented by Eq 3.2, the triples  $p$  and  $\hat{p}$  are now related by a general linear transformation.

$$\hat{p}^x = \mathcal{M} p^x \quad (3.7)$$

### 3.1.2 Diagonal Models

von Kries (1904) put forward his famous hypothesis on the *law of coefficients* for chromatic adaptation where he suggested chromatic adaptation as a vehicle for human color constancy. The central idea was that the eye adjusts to the ambient illumination in a scene over time and all colors are seen relative to this adapted state. Specifically, the responses under two different stimuli are related by an independent scaling in each sensor class [WS82].

$$\hat{\rho}_k^x = d_k \rho_k^x \quad (3.8)$$

where  $d_k = 1 / \int E(\lambda) R_k(\lambda) d\lambda$ . This modifies Eq 3.7 to a diagonal linear transform

$$\hat{p}^x = \mathcal{D} p^x \quad (3.9)$$

The diagonal matrix  $\mathcal{D}$  contains the von Kries coefficients for the given sensor responses. Note that the same diagonal matrix  $\mathcal{D}$  takes the response vectors at every image location  $x$  between illuminants; that is  $\mathcal{D}$  maps the entire image.

Several quantitative studies have shown the von Kries adaptation to be imperfect in one or more ways (see Wyszecki and Stiles 1982 [WS82] for a detailed study).

The von Kries model can however be quite accurate so long as there exist linear combinations of sensitivities which are narrow-band. This observation forms the basis of Finlayson et al's [FDF94b] spectral sharpening method. They show that if the sensor response functions are first transformed to a more narrow-band (or sharper) sensor basis [FDF94b] then the accuracy of Eq. (3.9) is improved. Indeed a von Kries type model is quite adequate for all sensor sets [FDF94a]. Von Kries plus sharpening can be written as

$$\mathcal{T} \hat{p}^x \approx \mathcal{D} \mathcal{T} p^x \quad (3.10)$$

where  $\mathcal{T}$  denotes the sharpening transformation of the original sensor response functions.

Let  $\psi^x = \mathcal{T} p^x$  where  $\psi^x$  is called the sharpened response. Let  $W$  be a  $N \times 3$  matrix representing the set of sharpened sensor responses for a 3 sensor imaging environment with  $N$  array elements. From Eq. (3.10), we can then write

$$\widehat{W} \approx \mathcal{D} W \quad (3.11)$$

where  $\widehat{W}$  is the set of responses under a different illumination.

### 3.1.3 Angle Invariants of color distributions

Consider  $W$  to be a set of three row vectors in  $N$  dimensions as

$$u_k = \{\psi_k^1, \psi_k^2, \dots, \psi_k^n, \dots, \psi_k^N\} \quad \text{for } 1 \leq k \leq 3 \quad (3.12)$$

where  $\psi_k^n$  is the sharpened response as defined above for sensor class  $k$  at location  $n$ . We see from Eq. (3.11) that a change in illumination corresponds to a change in length for each of these three vectors<sup>1</sup>. Upon normalizing each of these vectors to unit length, the diagonal matrix  $\mathcal{D}$  reduces to the identity matrix.

It follows from the above that the angles between the three  $N$ -dimensional vectors are invariant to changes in the illumination. These angles are computed as

---

<sup>1</sup>The author thanks Graham Finlayson for the idea

$$\phi_{i,j} = \cos^{-1} \left( \frac{\sum_{x=1}^N \psi_i^x \psi_j^x}{|u_i| |u_j|} \right) \quad \text{for } 1 \leq i < j \leq 3 \quad (3.13)$$

where  $|u_i|$  and  $|u_j|$  are the lengths of the band vectors. The  $\phi_{i,j}$  form the color distribution descriptors. These can be computed with  $6N$  multiplications from the sharpened responses.

This idea is the crux of the thesis. Intuitively, consider a colored surface under a blue light. When the incident light is spectrally different, for example, a red light, the relative magnitudes of the 3 color bands will be scaled by 3 independent numbers. Thus the red band will gain in magnitude while the blue band will have a lower magnitude. This simply means that the red band vector will become longer, while the blue band vector becomes shorter. The angles between the vectors will remain the same. Let the scaling factor (or the diagonal elements of  $\mathcal{D}$ ) be  $\{d_1, d_2, d_3\}$ . From Eqs 3.12 and 3.13, we see that the angles for the new illumination can be written as

$$\phi_{i,j} = \cos^{-1} \left( \frac{\sum_{x=1}^N d_i \psi_i^x d_j \psi_j^x}{|d_i u_i| |d_j u_j|} \right) \quad \text{for } 1 \leq i < j \leq 3 \quad (3.14)$$

From the above equation, it is obvious that the  $d_i$  and  $d_j$  in the numerator can be taken outside the summation. The scaling factors cancel out of the equation, and therefore, a change in illumination does not affect the angles.

It is instructive to note here that the angles possess no discrimination with regards to changes in scale; thus we cannot identify a smaller object relative to an identical larger one.

### 3.1.4 Data Sets used to test invariance of angles

To test whether the angular invariants possess sufficient discrimination, I imaged a set of thirteen objects under varying lighting conditions and at different orientations. The database comprised highly colored man-made objects such as cereal boxes, books, t-shirts and sweaters. A few of the objects are shown in Figure 3.1. The set of objects were imaged three times; lighting conditions were varied across each set of images, but were identical within the set. No attempt was made to keep each object at a



Figure 3.1: Real objects imaged under varying illumination conditions

fixed orientation — I was only careful to preserve a major portion of a selected face of each object in the three images. The light source used was a halogen lamp; the other two sets of images were generated by using an orange and a blue filter. Spectral characteristics of these filters can be found in Barnard's thesis [Bar95].

A Sony 3-CCD DXC-930 color camera and a Parallax 24-bit frame grabber card attached to a SUN SparcLX formed the main components of the imaging system. The spectral sensitivities of the camera responses are shown in Figure 3.2. Each set of sensor responses is from 380 nm to 780 nm with an interval of 4nm. The responses are estimated at the central region of the camera with an aperture of 2.8 and a focal length of 25. Data based sharpening as described by Finlayson et al [FDF94b] was performed in our laboratory by Barnard [Bar95]. The sharpened camera responses are represented by the dotted lines in the figure. The sharpening transform calculated is

$$\mathcal{T} = \begin{bmatrix} 0.9495 & 0.0038 & 0.0099 \\ -0.1182 & 0.9535 & -0.1279 \\ 0.0356 & 0.2380 & 0.9805 \end{bmatrix} \quad (3.15)$$

Visually, we note that not much sharpening seems to have resulted. It seems worthwhile therefore as a first approximation to set the sharpening transform matrix of Eq 3.10 to the identity matrix, i.e. use the original camera responses.

### 3.1.5 Normalized Euclidean Distance as a measure

A normalized Euclidean distance was used to measure the invariance of the angles. Let  $\Phi = \{\phi_{1,2}, \phi_{1,3}, \phi_{2,3}\}^t$  represent the angles as a three dimensional vector. Given two such vectors  $\Phi^i$  and  $\Phi^m$  for the test image and the model respectively, the normalized distance is defined as

$$\delta_{i,m} = \frac{\sum_{k,l=1}^3 (\phi_{k,l}^i - \phi_{k,l}^m)^2}{\sum_{k,l=1}^3 (\phi_{k,l}^m)^2} \quad (3.16)$$



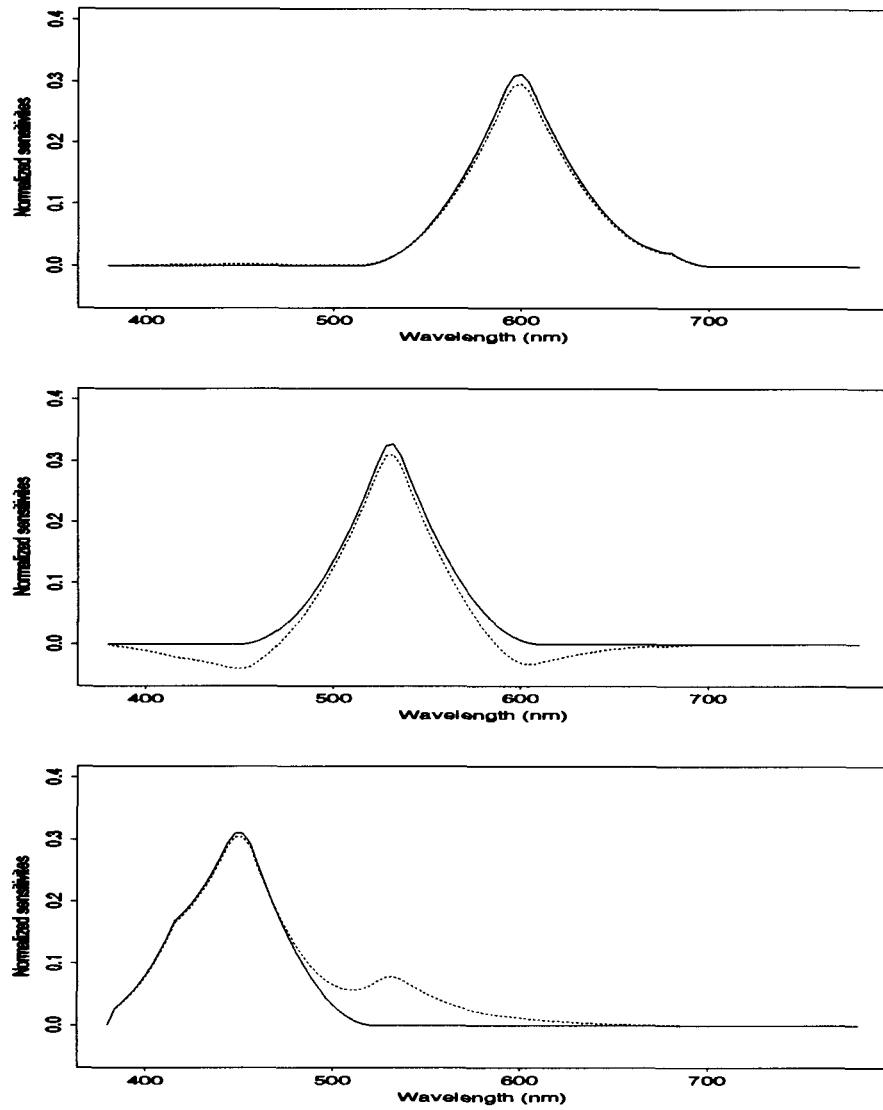


Figure 3.2: Camera sensitivities

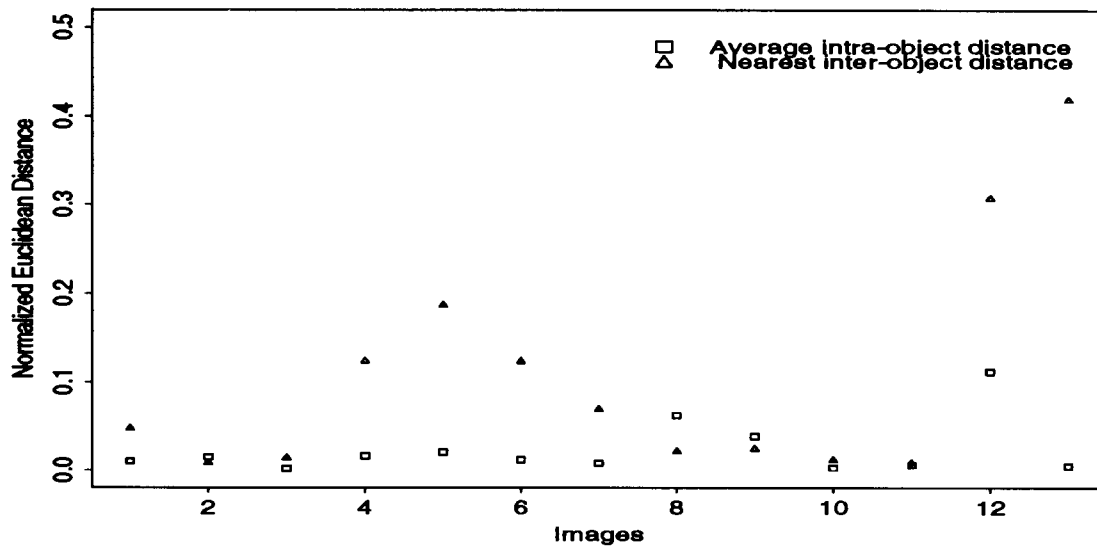


Figure 3.3: Measure of invariance for color distribution angles

### 3.1.6 Are 3 numbers sufficient ?

Angular invariants were calculated for each of the thirty nine images. A distance matrix based on the measure in Eq 3.16 was generated by computing the distance for each image against the others. A model base was chosen as the set of images taken under the halogen lamp. Distances between each model image and the two test images for the same object were averaged ( $\alpha_i$  for each object  $i$ ). For each object, the distances between it's test images and the other model images were averaged ( $\beta_{i,j}$  for each model  $i$  and object  $j$ , where  $j \neq i$ ). Figure 3.3 is a plot of the intra-object distance ( $\alpha_i$ ) and the nearest inter-object distance ( $\text{minimum}\{\beta_{i,j}\}$ ) for each object  $i$ .

Two things are of importance in the graph. First, the  $\alpha_i$ 's should be close to zero as they represent distances between invariants of the same object. Secondly, the distance between the  $\alpha_i$  and the corresponding  $\text{min}\{\beta_{i,j}\}$  should be relatively large for successful discrimination. We see from the graph that in three cases (images 2, 8 and 9 corresponding to the light bulb box, Javex bleach bottle and the Knorr Soup packet) the inter-object distances were smaller than the intra-object distance. Thus

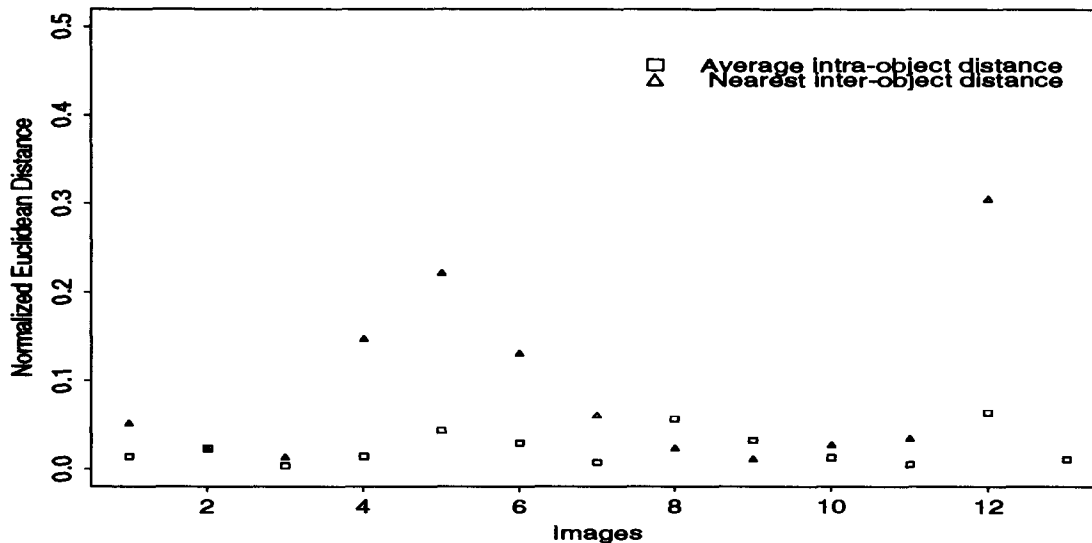


Figure 3.4: Measure of invariance with sharpening

these objects would be incorrectly recognized. Furthermore, image 11 (a shampoo bottle) was found to have poor discrimination. For object 12 (a sweater) the intra-object invariant distance is quite large. The discrimination for this database is fairly good; however one would expect problems as the dataset grows larger. The average intra-object distance was found to be approximately 20% of the nearest inter-object distance.

Using the sharpening transform previously calculated, the invariants were recomputed. The corresponding plot of inter and intra-object distances is shown in Figure 3.4

We see that for image 12, the intra-object distance has reduced significantly. The invariance for image 2 has improved a little; however, inter-object distances are still in the same neighborhood, thus affecting the stability of discrimination. Inter-object distance in case of image 11 has increased, improving discrimination. These observations are supported by the table of results in chapter 5. For image 13, the inter-object distance was too large to fit into the scale used for the plot. We observe that the initial assumption of an identity sharpening is reasonably justified as no significant

improvement is noted. Of course this argument applies only to systems with narrow band sensors. For systems like the human eye where the cone responses are broad band, sharpening is necessary for the diagonal model to work.

## 3.2 Angular edge invariants

Funt and Finlayson [FF95] argue that color ratios possess sufficient discrimination to be able to identify objects from a large database. This seems to follow from the observation that highly colorful objects will have distinctive color edge information which can be used as a cue to recognition. The motivation here lies in the fact that color edge distributions provide information which differs from that provided by the color distributions themselves. We would like to examine whether this information is sufficiently uncorrelated as to overcome the discrimination problems of the original set of invariants. Note that since the objects under consideration in general will not be 2 dimensional surfaces, false edges may be encountered due to shading or surface edges. We can expect that surface edges will remain similar across all images of an object since an important constraint on the algorithm is that a major portion of a chosen surface area be preserved in each image.

Consider this argument in the context of angle invariants. Given an image, we can extract color edges by differentiating the image and compute angle invariants of the edge distributions as described in the previous section. As an illustration, consider the two images shown in Fig 3.5. They are two synthetically generated images, containing equal surface areas of four different Mondriaans with the only difference being that the patches are arranged differently in the two images. The difference between the images is immediately obvious; the method of angular invariants as described above does not distinguish between the two images, as they contain equal amounts of identical color pixels for each surface patch. If we extract the edges from the two images as shown in Fig 3.6 and compute the angular invariants between them, we note that there is a large difference in the results obtained.

Table 3.1 shows the invariants obtained by the two methods and the Euclidean distance between them. In the table, CD refers to the angle invariants of the color

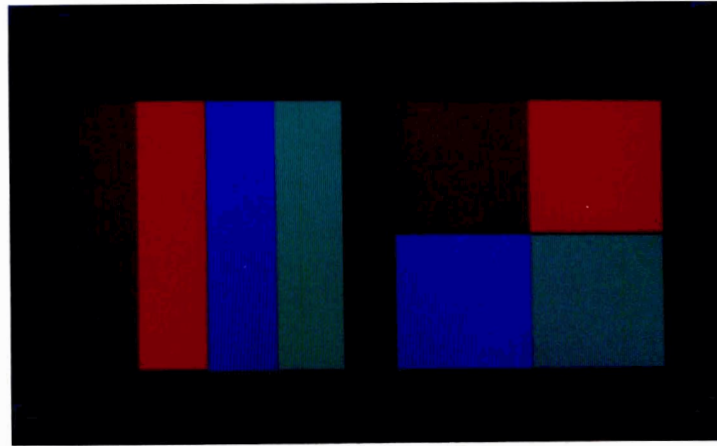


Figure 3.5: Insufficiency of color angle invariants

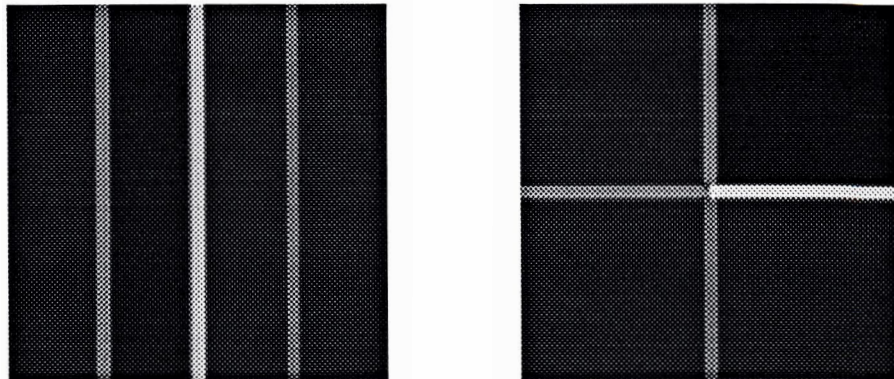


Figure 3.6: Color edges of the images

distribution, while ED refers to the invariants of the edge distributions in the image. Distance refers to the Euclidean distance between the two image invariants of each type.

This is as much an argument for using edge information as showing that just using color distributions is not sufficient in many cases.

Type	Image1			Image2			Distance
C D	57.727	54.003	19.432	57.727	54.003	19.432	0.000
E D	29.177	20.622	11.950	25.911	29.226	24.260	15.370

Table 3.1: Comparison of invariant distances for color and edge distributions of two images.

### 3.2.1 Choosing an edge operator

The Laplacian of a function  $f(x, y)$  is defined as

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \quad (3.17)$$

The Laplacian is the lowest order linear combination of partial derivatives that is rotationally symmetric [Hor86]. Rotationally symmetric operators are attractive because they treat image features in the same way, irrespective of their orientation. It is standard practice to apply a low frequency attenuation filter to the image before applying the Laplacian. This minimizes the effect of noise by preventing high frequency noise edges from being detected. One of the most popular filters used is a Gaussian,

$$h(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{1}{2} \frac{x^2+y^2}{\sigma^2}} \quad (3.18)$$

where  $\sigma$  defines the spread of the Gaussian. This *Laplacian of Gaussian* [GW92] (or *LoG*) operator was also used by Funt and Finlayson in Color Constant Color Indexing [FF95] to differentiate the log of the image. Finlayson also cites other examples of the use of this operator. Hurlbert [Hur89] examined opponent invariants in the human visual system; her model used the *LoG* operator, which she proposes is implemented in humans by the double opponent cells [Fin92].

### 3.2.2 Edge invariants

The angular invariants for the edge distributions can be written from Eq 3.13 as

$$\varphi_{i,j} = \cos^{-1} \left( \frac{\sum_{x=1}^N \chi_i^x \chi_j^x}{|X_i| |X_j|} \right) \quad \text{for } 1 \leq i < j \leq 3 \quad (3.19)$$

where  $\chi_i^x = LoG(\psi_i^y)$ ,  $y \in \mathcal{N}(x)$ .  $\mathcal{N}(x)$  signifies the neighborhood of  $x$  as the *LoG* operator has a support over the neighborhood of the pixel being examined. In our case, it is calculated over a 7x7 pixel region with the given pixel at the center of the mask.  $X_i = \{\chi_i^1, \dots, \chi_i^N\}$  represents the edge feature vector. To verify my hypothesis about the uncorrelated information content between image color distributions and image edge distributions, I repeat my invariance test on the same data set of 13 objects. The *LoG* operator is applied to each image and its angular invariants are calculated. Distances between them are calculated as per Eq 3.16. Intra-object and inter-object distances are calculated and represented as in the earlier case. The plots are shown in Fig 3.7 for the case where the sharpening transform is set to identity. It is obvious from a comparison of the two graphs — Fig 3.3 and Fig 3.7 that the focus of our problems has shifted to other images. For the edge distribution invariants, we note discrimination problems in images 7, 10 and 11 while there are significant improvements in images 2, 8 and 9 which were the main cause of concern in the color distribution invariants.

## 3.3 Combining edges and distributions

Given the results in the preceding section, we would expect the combination of the two sets of invariants to offset the problems of each other. The final set of invariants is represented as the combination of both color and edge distribution angles. Given an image  $i$ , its angle invariants are:

$$\mathcal{I}_i = \{\phi_{1,2}, \phi_{1,3}, \phi_{2,3}, \varphi_{1,2}, \varphi_{1,3}, \varphi_{2,3}\} \quad (3.20)$$

Distances are plotted as before for the set of six invariants for each of the 39 images. Figures 3.8 and 3.9 show the  $\alpha_i$  and  $\min(\beta_{i,j})$  plots for the set of invariants  $\mathcal{I}$  without

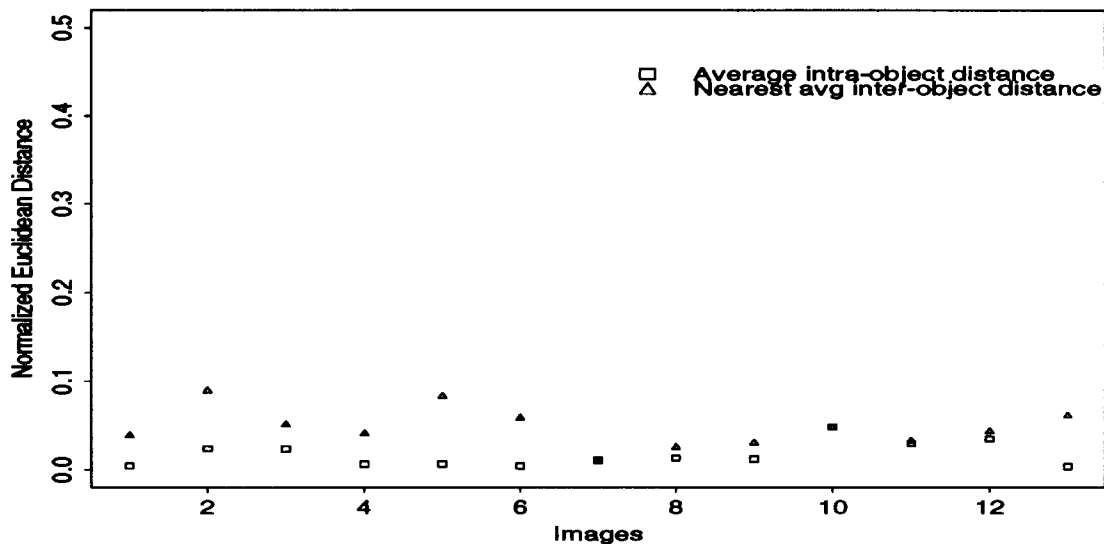


Figure 3.7: Edge invariant distances between images

and with sharpening respectively. In the case where the sharpening transform is unused or set to unity, we note that problems still persist in certain cases. In particular, images 10 and 11 still remain unstable in the discrimination problem. Application of the sharpening transform to the images before the edges and the invariants are calculated shows some improvement in the distances between  $\alpha_i$  and  $\min(\beta_{i,j})$ .

We will see in Chapter 5 that the results obtained are comparable to Color Constant Color Indexing.

In conclusion, it is necessary to point out that these invariants are suitable only for certain classes of objects. Highly textured surfaces in general, are not well modeled by the color distribution invariants, and as Healey [KH94] has shown, color histograms are not sufficient for many textures. This issue is examined in detail in the next chapter, as we look at the usefulness of certain statistical models in the characterization of colored textures. Many natural colored objects (for instance various fruits and vegetables) also do not respond well to the treatment described here. There is too much variation in color distributions and color edges for a stable recognition technique based on the invariants described above. This is not a problem of the algorithm, but rather the



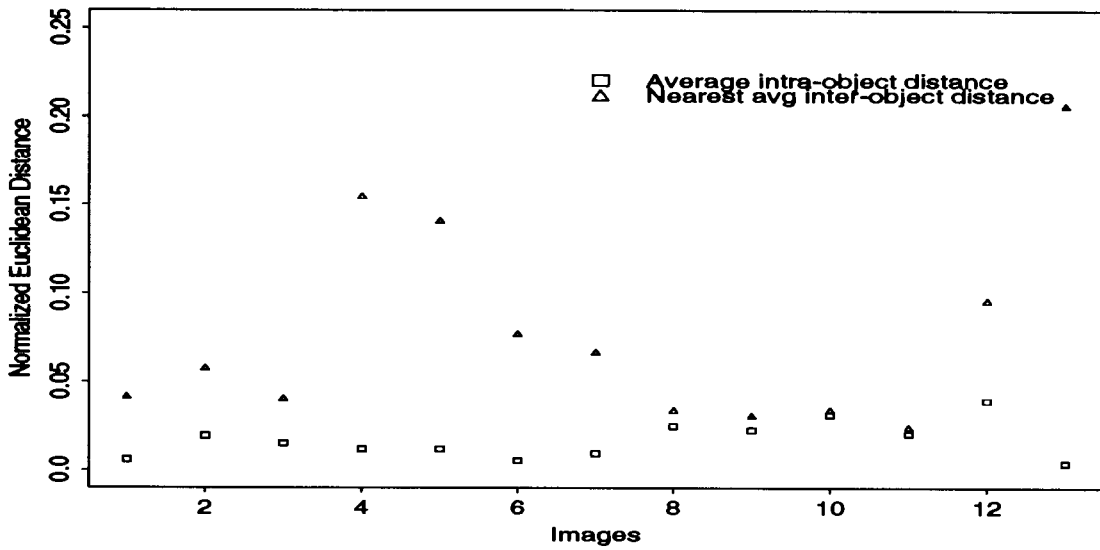


Figure 3.8: Invariant distances between images for invariant set  $\mathcal{I}$

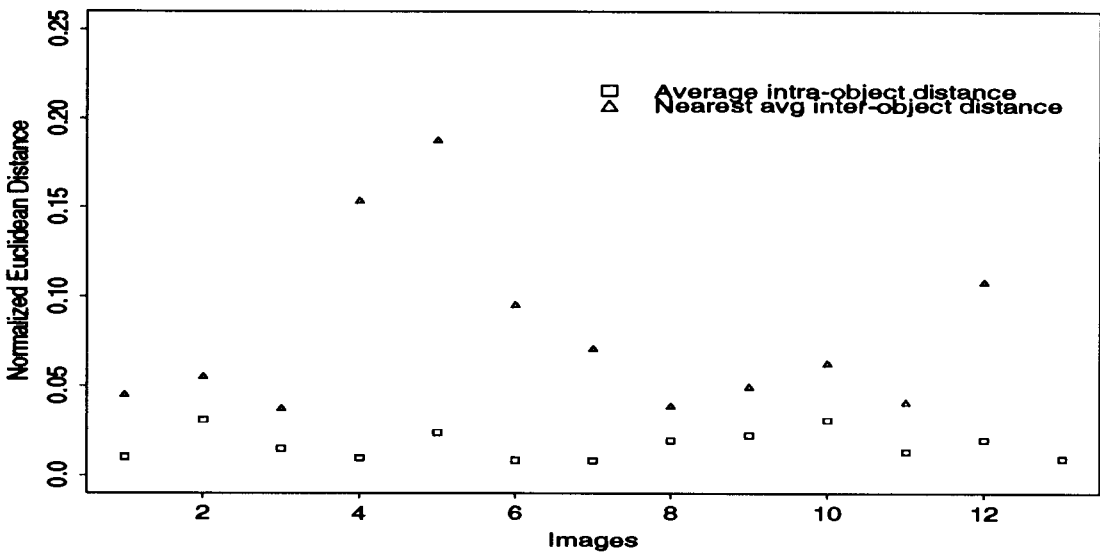


Figure 3.9: Sharpened invariant distances between images for invariant set  $\mathcal{I}$

fact that the feature base used (color distributions) is unstable for these objects. The algorithm is fast — linear in the number of pixels in the image — and as the results in Chapter 5 show, robust in its application over large databases of objects.

# Chapter 4

## Angular color texture descriptors

“We never sufficiently reflect that a language, strictly speaking, can only be symbolical and figurative, that it can never express things directly, but only as it were, reflectedly. This is especially the case in speaking of qualities which are only imperfectly presented to observation. They are not to be arrested, and yet we find it necessary to describe them; hence we look for all kinds of formulae in order, figuratively at least, to define them.”

— Goethe, *“Color Theory”*

### 4.1 Introduction

While there exists a huge body of work dealing with various aspects of texture analysis, most of it is based on grey scale variations in image patterns. The role of color in the analysis of textures has been generally neglected in the pursuit of obtaining reliable characterizations for textures. Recently, Healey [HW95, KH94, HP95] has shown that color yields useful characterizations of texture both in statistical as well as model-based methods of representation.

In this work, I extend some of the ideas presented by Healey and obtain a set of feature descriptors invariant to the color of illumination and to rotation in the image plane. The color-edge angles described in the previous chapter are used as another

representation for textures, and it is shown that they provide a good framework for texture recognition.

An important problem here is that the images are assumed to be pre-segmented into regions of a single texture. This is significant because image segmentation in general and texture segmentation in particular remain hard problems in computer vision. I show that by computing texture invariants locally across the image, we can account for multiple textures in an image.

## 4.2 Representations of color textures

Common statistical methods of representing tonal variations across an image consist of co-occurrence matrices and spatial correlations (usually autocorrelation functions as grey scale textures are commonly used in the vision community). The use of algebraic moments has also resulted in some interesting results in the field [Tuc94]. I use the spatial correlation function defined by Healey [KH94] to represent texture features. Surface reflectance patterns are considered to be two dimensional random processes. The three color bands of the texture image are considered to be wide sense stationary and can therefore be characterized by a mean and an autocorrelation function. Furthermore, it is assumed that any two color bands are jointly wide sense stationary with each other. For a color image we can define six correlation functions as

$$C_{i,j}(a,b) = \sum_{x,y=1}^n [\rho_i(x,y) - \bar{\rho}_i][\rho_j(x+a,y+b) - \bar{\rho}_j] \quad \text{for } 1 \leq i,j \leq 3 \quad (4.1)$$

where  $\rho_i(x,y)$  represents the pixel value at location  $(x,y)$  on the image plane for the color band  $i$ .  $\rho_j(x+a,y+b)$  correspondingly defines the pixel value at a location offset from  $(x,y)$  by  $(a,b)$  for another color band  $j$  of the same image.  $\bar{\rho}_i$  and  $\bar{\rho}_j$  represent the image means in the respective color bands. The number of pixels in the image is  $n^2 = N$ .

The structure of this correlation function in general will depend upon the spectral properties of the incident illumination. As the signals are assumed to be wide sense stationary, the correlation functions are independent of translations in the image.

Let us consider how these spatial correlation functions represent the textural properties of an image. Fig 4.1 shows some examples of color textures. The corresponding spatial correlation functions of the texture in the top left corner of Fig 4.1 are shown in Fig 4.2. It is clear that the six correlations encode different information about the color spatial interactions in the image. This function is related to the size of the texture primitive (i.e. the fineness of the texture). If the texture is coarse, the correlation function will drop off slowly; otherwise, it will drop off very rapidly. For regular textures, the surface defined by the function will exhibit peaks and valleys. We can observe the effects of rotation and change in illumination upon the surfaces defined by these functions in Fig 4.3. Here the texture has been rotated by  $30^\circ$  in the image plane and in the second case has been imaged with a red filter over the camera. The surface generated by the autocorrelation in the red color band is shown in the figure. We see that the surface has undergone a rotation corresponding to the rotation in the image plane. With the change in illumination, there is a decrease in the overall height of the surface.

### 4.3 Color texture invariants

Given a change in illumination from  $E(\lambda)$  to  $E'(\lambda)$ , the corresponding spatial correlation equation can be written as

$$C'_{i,j}(a,b) = \sum_{x,y=1}^n [\rho'_i(x,y) - \bar{\rho}'_i][\rho'_j(x+a,y+b) - \bar{\rho}'_j] \quad \text{for } 1 \leq i,j \leq 3 \quad (4.2)$$

where  $\rho'$  represents the sensor outputs under the illumination  $E'(\lambda)$ . Applying a sharpening transformation to the image, we can write Eq 4.1 as:

$$C^{\#}_{i,j}(a,b) = \sum_{x,y=1}^n [\phi_i(x,y) - \bar{\phi}_i][\phi_j(x+a,y+b) - \bar{\phi}_j] \quad \text{for } 1 \leq i,j \leq 3 \quad (4.3)$$

where  $\phi(x,y) = \mathcal{T}\rho(x,y)$  as given by Eq 3.10. Applying a similar transformation to Eq 4.2, we can establish a relation between the correlation functions of the sharpened color signal of an image under a changing illumination as:



Figure 4.1: Color texture images

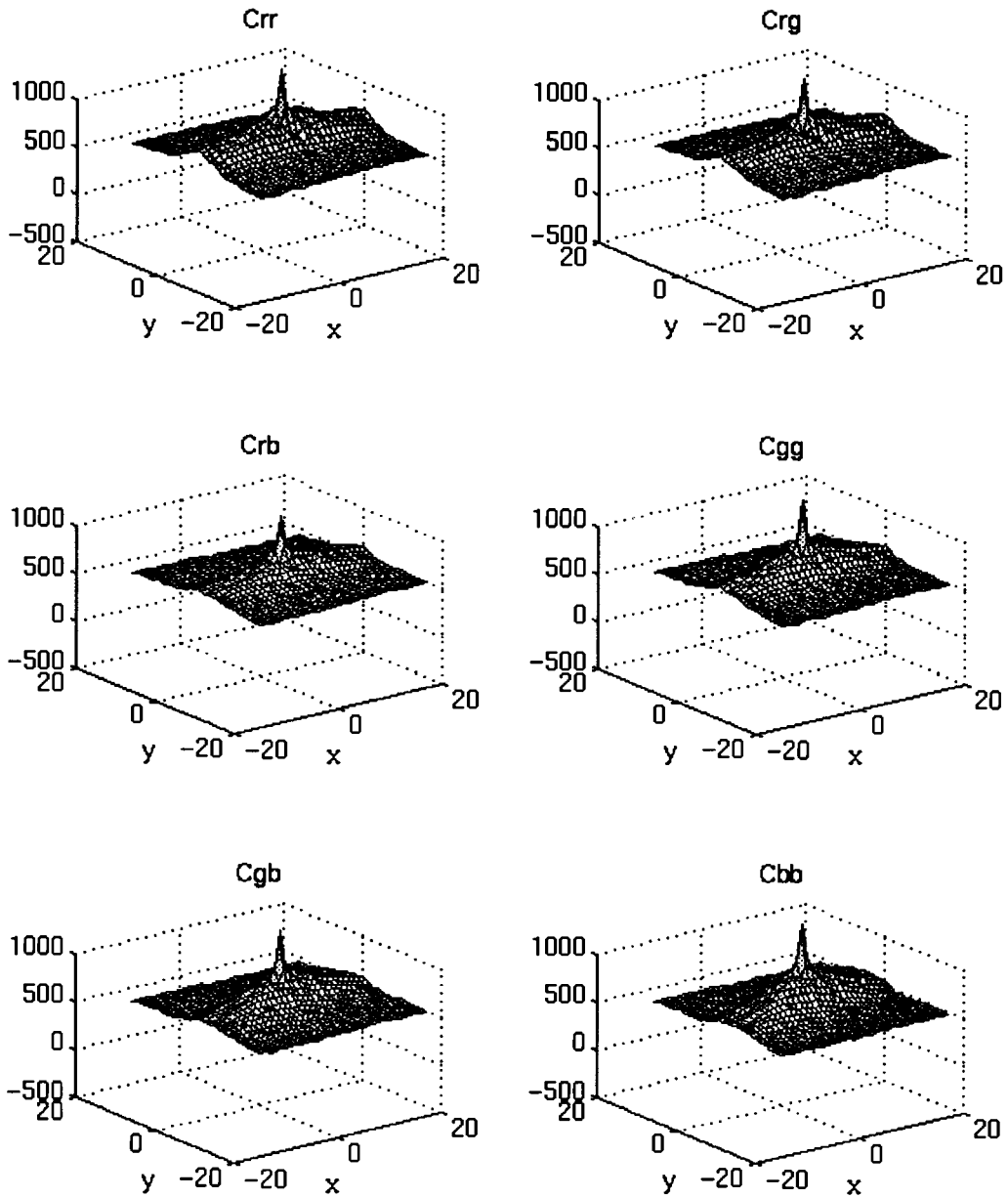


Figure 4.2: Spatial correlation functions of a color texture



Figure 4.3: Spatial correlation functions changed due to: a) rotation in the image plane; b) due to changing illumination

$$C_{i,j}^{\#'}(a, b) = \sum_{x,y=1}^n [d_i \{\phi_i(x, y) - \bar{\phi}_i\}] [d_j \{\phi_j(x + a, y + b) - \bar{\phi}_j\}] \quad \text{for } 1 \leq i, j \leq 3 \quad (4.4)$$

where  $d_i, d_j$  represent the elements of the diagonal matrix of illumination change  $\mathcal{D}$ . Moving these outside the summation, we can write

$$C_{i,j}^{\#'}(a, b) = d_i d_j C_{i,j}^{\#}(a, b) \quad (4.5)$$

For the complete set of six correlation functions, dropping the subscript, we can write  $C^{\#}(a, b) = \{C_{1,1}^{\#}(a, b), C_{1,2}^{\#}(a, b), \dots, C_{3,3}^{\#}(a, b)\}$ . Equation 4.5 can now be written as

$$C^{\#'}(a, b) = \mathcal{D}_2 C^{\#}(a, b) \quad (4.6)$$

where  $\mathcal{D}_2$  represents the diagonal transformations in the case of the spatial correlation functions.

$$\mathcal{D}_2 = \begin{bmatrix} d_1 d_1 & \dots & \dots & \dots & \dots & \dots \\ \dots & d_1 d_2 & \dots & \dots & \dots & \dots \\ \dots & \dots & d_1 d_3 & \dots & \dots & \dots \\ \dots & \dots & \dots & d_2 d_2 & \dots & \dots \\ \dots & \dots & \dots & \dots & d_2 d_3 & \dots \\ \dots & \dots & \dots & \dots & \dots & d_3 d_3 \end{bmatrix}$$



We see therefore that the correlation functions of the sharpened color signal preserve the diagonal transformation between changes in illumination. A rotation of the texture in the image plane corresponds to a similar rotation of the surface defined by the correlation functions. For 3-dimensional textures, Healey [KH94] shows that the correlation surfaces are related by an affine transformation for any arbitrary change in the object's position in space. Texture representations are examined in the context of changes in illumination and rotation about the optical axis.

From a formulation of the problem similar to that described in the previous chapter, angles between the six surfaces defined by the correlation functions can be easily obtained. This gives a set of fifteen invariants with which to describe textures.

$$\vartheta_{i,j} = \frac{\sum_{a,b} C_i^\#(a,b) C_j^\#(a,b)}{|C_i^\#| |C_j^\#|} \quad \text{for } 1 \leq i, j \leq 6 \quad (4.7)$$

where  $C_i^\#$  and  $C_j^\#$  represent elements of the sharpened spatial correlation function set  $C^\#$ . This normalized dot product between the two correlation surfaces is invariant to changes in illumination and rotation of the texture about the optical axis. It is of course not obvious as to how well these angles will perform in terms of discriminating between various textures. This problem is examined in the following sections.

### 4.3.1 Implementation details

An important problem faced in the case of texture analysis is the treatment at image boundaries. If we assume the texture to be terminated at the image boundaries, then the correlation function displays characteristics which reflect that assumption. In particular, there will be a tail-off in the correlation function of Eq. 4.1 as the offsets  $a, b$  increase. Eq. 4.1 is adopted as the first method for computing correlation functions.

Another aspect is that in reality the texture does not suddenly terminate at the image boundaries. In this case, there are two ways to deal with the issue. We can divide the correlation function by the number of pixels under consideration which will offset the effect of the reduced effective image size at the image boundaries. Eq. 4.1

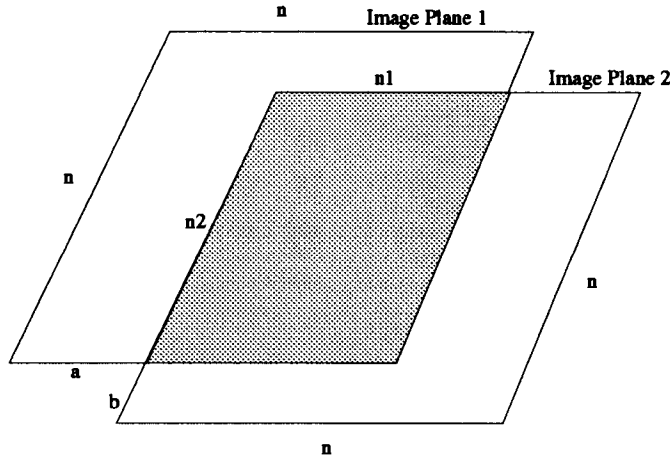


Figure 4.4: Correlation of image color planes

now becomes

$$C_{i,j}^{\#}(a, b) = \frac{\sum_{x,y=1}^{x=n1,y=n2} [\phi_i(x, y) - \bar{\phi}_i][\phi_j(x + a, y + b) - \bar{\phi}_j]}{n1n2} \quad \text{for } 1 \leq i, j \leq 3 \quad (4.8)$$

where  $n1$  and  $n2$  are the number of overlapping pixels under consideration. This is illustrated in Fig 4.4. The shaded portion in the image is the overlapping area, given by  $n1 \times n2$ , where  $n1 = n - a$  and  $n2 = n - b$ . This modified correlation function is the basis of a second implementation.

The other method of dealing with image boundaries is achieved by reflecting the image beyond its boundaries to give an impression of continuity of texture to the algorithm. This can be achieved in various ways — for example, below I outline a Fourier transform approach for computing the correlations which automatically incorporates this into the algorithm.

Consider the number of computations required to generate the surfaces defined by Eq. 4.1. We see that for any significant number of offsets  $a, b$  typically taken over half the image area,  $-n/4 \leq a, b, \leq n/4$ , the number of multiplications required is of the order  $N^2$ . Here  $N$  is the number of pixels in the image and  $n$  is the number of rows and columns of the image (assuming a square image for simplicity of notation), so that  $N = n^2$ .

It is well known that convolution in the spatial domain corresponds to multiplication in the Fourier domain. Correlations are similar to convolution, except that

the correlating (or convolving) signal is not flipped about its axes. The computation of the correlation function in the Fourier domain transforms into a multiplication of the correlated signal transform with the complex conjugate of the correlating signal transform. Let  $R_i(u, v)$  represent the Fourier transform of  $(\phi_i(x, y) - \bar{\phi}_i)$ . For both the continuous and discrete cases, the following theorem holds [GW92]

$$(\phi_i(x, y) - \bar{\phi}_i) \circ (\phi_j(x, y) - \bar{\phi}_j) \Leftrightarrow R_i(u, v) R_j^*(u, v) \quad (4.9)$$

where  $\circ$  represents the correlation function and  $R_j^*(u, v)$  is the complex conjugate of the Fourier transform of  $(\phi_j(x, y) - \bar{\phi}_j)$ . Here  $\phi_i(x, y)$  is the correlated signal while  $\phi_j(x, y)$  is the correlating signal. The computation of the correlation is often more efficiently done in the frequency domain using an FFT algorithm to obtain the forward and inverse transforms. An important point here is that due to the discrete nature of the image, the resultant correlation is what may be called a circular correlation corresponding to the circular convolution which is obtained in the case of the convolution function. This derives from the fact that the signals are treated as periodic in the size of the image by the FFT, and essentially constitutes a repetition of the image beyond its boundaries. This can of course be avoided, as in the case of convolution, by padding the image with a border of zeros corresponding to the image dimensions. The implementation used here retains the periodicity of the image in the Fourier domain due to two reasons: simplicity of implementation and the fact that such a repetition of the image does not always imply a discontinuity of the texture.

### 4.3.2 Datasets

A set of ten textures under varying illumination conditions forms the test bed for the texture invariants. The images were obtained from Glenn Healey<sup>1</sup>, and have been used in his work on illumination invariant descriptors of color textures [HW95]. Some typical textures consisted of fabric patterns, carpets, trees, clouds and other natural scenes. The textures were imaged under nearly white light using a Sony XC-77 CCD camera and a RasterOps TC-PIP frame grabber. Varying illumination was simulated

---

<sup>1</sup>The author is grateful to Glenn Healey, University of California, Irvine, for providing the images.

by the use of 3 color filters (Corion CA-600, CA-550, CA-500). This gave a set of 10 database texture images (under nearly white light) and 30 test images under varying illumination. The images were rotated in the image plane at five different angles ( $30^\circ, 45^\circ, 60^\circ, 90^\circ, 110^\circ$ ). Two such textures under varying illumination and rotation are shown in Figs 4.5 and 4.6. Including the initial set of images I now have a set of 240 images. Invariants were computed for these test sets using both methods outlined in the previous section.

### 4.3.3 Performance Analysis

The distance function defined in Chapter 3 is used to measure the invariance of the texture descriptors for the given data set of texture images.

$$\delta_{i,m} = \frac{\sum_{k=1}^{15} (\varphi_k^i - \varphi_k^m)^2}{\sum_{k=1}^{15} (\varphi_k^m)^2} \quad (4.10)$$

where  $\varphi_k^i$  is the  $k^{\text{th}}$  angle of the  $i^{\text{th}}$  texture being matched with the model texture  $m$ . The set of images under nearly white light with no rotation is chosen as the model set. A graph is plotted showing the average distance between the invariants of each object ( $\alpha_i$  for each object  $i$ ). The average distances between a model texture  $i$  and test image invariants of other textures  $j$  are calculated ( $\beta_{i,j}$ ). The smallest of these distances ( $\min\{\beta_{i,j}\}$ ) is plotted on the same graph to indicate the clustering characteristics of the data. The plot is shown in Fig 4.7. In this case the invariants have been calculated using the normalized correlation computation of Eq 4.8.

The second graph (fig 4.8) is a similar plot where the invariants have been calculated from Eq 4.1. We note that both graphs have similar characteristics. Image 7 has poor discrimination due to its low inter-texture distance. The intra-texture invariant differences are quite significant and affect the recognition process as well. Images 3 and 4 from the graph show good discrimination, but instability in the invariant space.

The invariants computed using the Fourier transform approach are shown in Fig 4.9. Here we have a significant departure from the previous set of graphs where we note that results for image 7 have improved substantially, while images 2 and 10

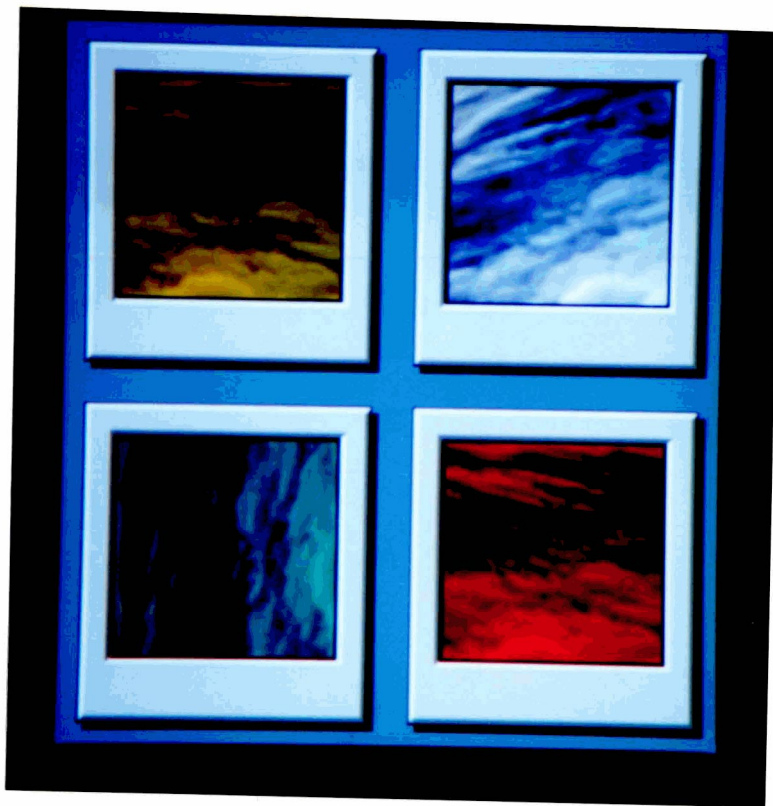


Figure 4.5: A texture under varying illumination and rotations

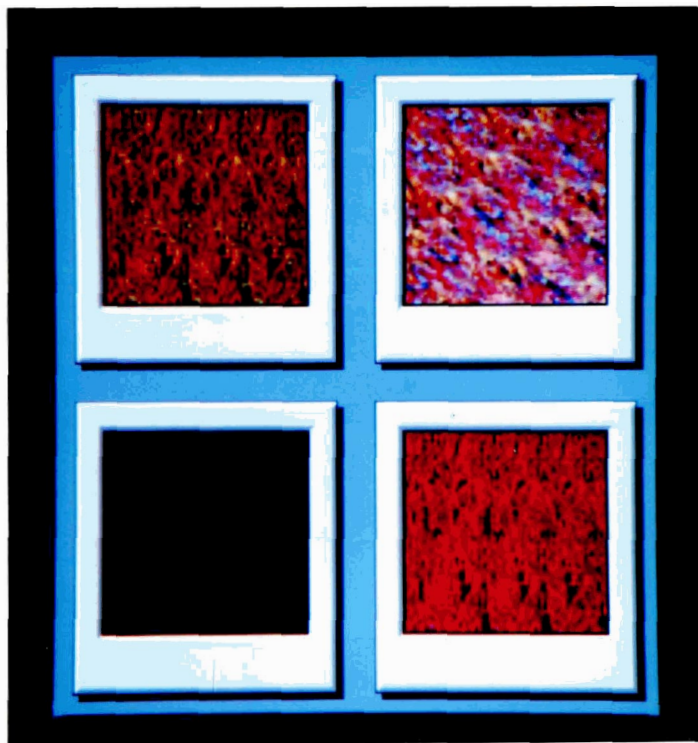


Figure 4.6: Another color texture under similar illumination and rotation conditions

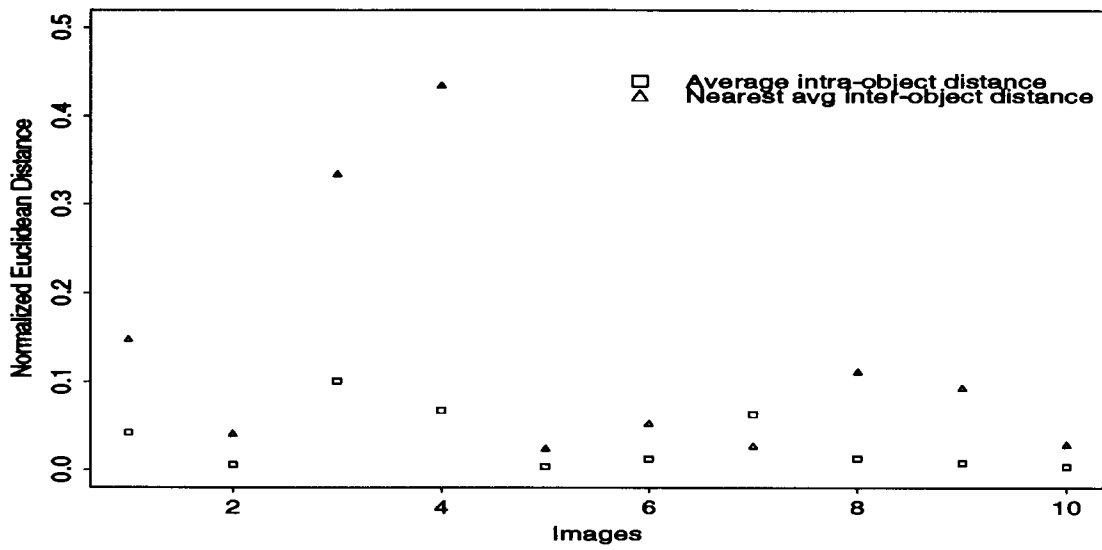


Figure 4.7: Invariant distances for normalized correlations

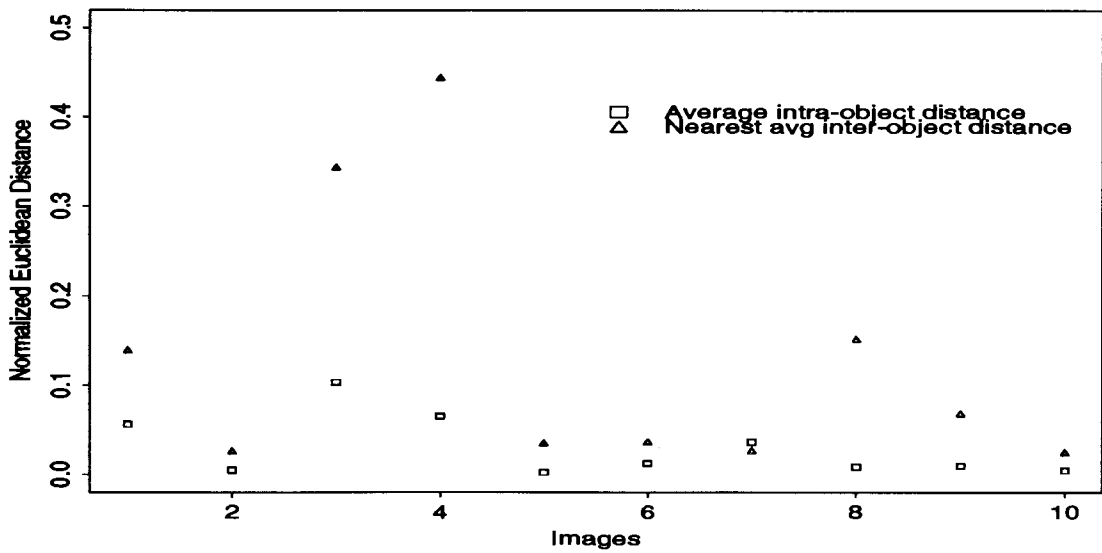


Figure 4.8: Invariant distances for correlations defined by Eq 4.1

now show poor clustering. This can be attributed to the assumption of periodicity of the image which influences the Fourier transform method of correlations.

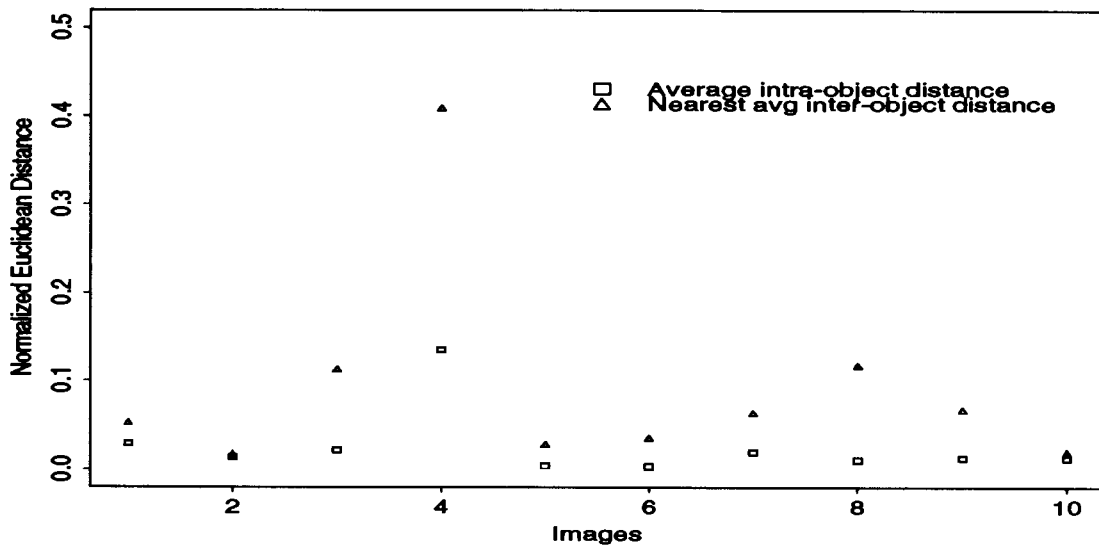


Figure 4.9: Invariant distances for correlations using a Fourier transform approach

Finally, note that the color and edge angle invariants discussed in section 3.3 also incorporate spatial information about the image. We therefore use the six angles to index color textures. The results obtained are remarkably good. An invariant plot as, in the previous cases, is displayed in Fig 4.10. We see that the color-edge invariants have low intra-texture distances with large inter-texture distances. This provides for a stable representation with good discrimination. In fact, as we shall see in Chapter 5, the performance of these invariants is superior to any of the algorithms discussed above.



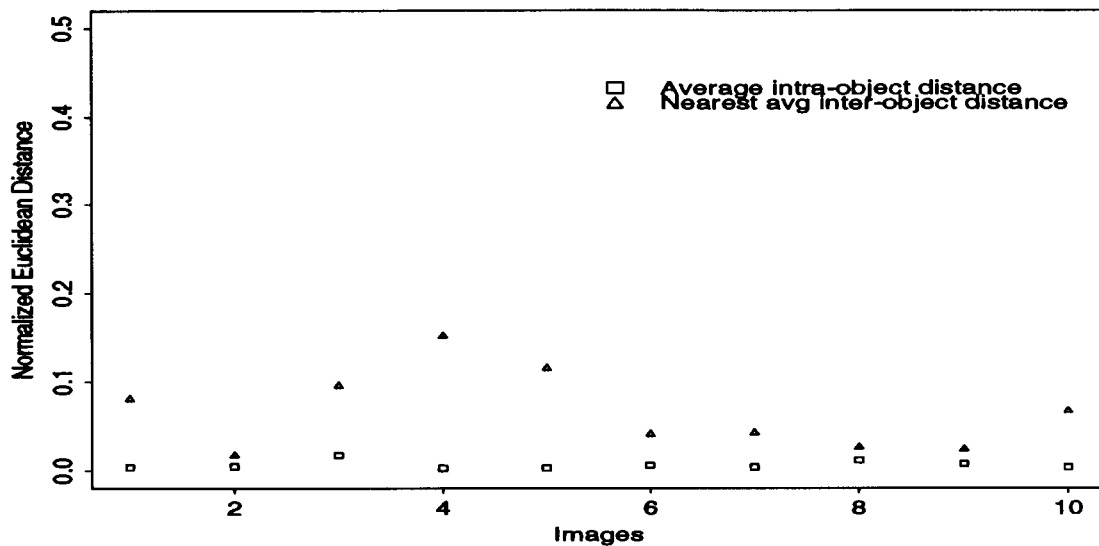


Figure 4.10: Invariant distances for color-edge angle invariants for natural textures

## 4.4 Color texture classification

It is clear from the above section that the proposed texture invariants possess sufficiently stable clustering characteristics for recognizing a wide variety of colorful textures. A primary assumption for the algorithm to work correctly is that only one texture be present in the image under consideration. Such an assumption can be a serious drawback when we wish to identify images which contain multiple textures. Spatial interactions in the image are no longer well modeled by the correlation functions as the properties of one texture begin to interfere with another. Pre-segmentation of the image becomes a necessary part of the recognition process. Segmentation however is a non-trivial process, and such an assumption weakens the viability of the algorithm. When many textures are present in an image (consider Fig 4.11), texture classification is often a desirable property of image analysis. By texture classification, I mean here the correct recognition of each of the distinct textures in the image based on a model database of textures.

The logical step in classifying textures based on angular invariants is to compute the texture characteristics locally across the image. Such local characteristics can be



Figure 4.11: A natural scene with multiple textures

computed by defining a window of interest within which the texture characteristics are computed. This window is then translated across the image, giving rise to a set of invariants for each window position. It is important to keep in mind that I am not at this stage concerned with the segmentation of the image into its component textures — only with the identification of the various textures present in the image. These invariants can then be indexed on some selected criteria; for invariants with high bin counts the database can be searched for the closest model invariants.

One of the significant problems faced in this approach is the selection of a proper window size. The window must be large enough to capture the spatial interactions which characterize the texture; it must also be small enough not to overlap more than one texture. As we keep translating the window, it will eventually move from one texture region in the image to another. At these texture boundaries the window will straddle at least two textures so the invariants computed will be significantly different from those computed previously. These invariants will however have a low bin count as we do not expect the textures to repeat frequently. Since we are interested only in recognizing those invariants which have high bin counts, invariants computed at texture edges will be neglected in the recognition process.

#### 4.4.1 Local texture invariants

I examine two sets of local invariants in this section, based on the experiments with the spatial operators described above. First I consider a localized texture feature based on sharpened spatial correlations as described in Eq 4.3 and examine whether the representation remains stable under varying window sizes. Another method of computing local texture invariants using the color-edge distribution angles is described. Experiments show that the color-edge angles preserve invariance far better than the spatial correlation angles for small window sizes.

### Spatial correlation based local texture representations

I define a window function as

$$W(x, y) = \begin{cases} 1 & \text{for } u_x \leq x \leq v_x, \quad u_y \leq y \leq v_y \\ 0 & \text{otherwise} \end{cases} \quad (4.11)$$

where  $(v_x - u_x)(v_y - u_y) \ll N$  defines the size of the window and  $N$  is the size of the image under consideration. The function  $W(x, y)$  can be translated by applying an offset to its  $x$  and  $y$  limits. If  $u_{x,m} = u_x + m$  where  $m$  is the offset, we can rewrite Eq 4.11 as

$$W_{m,n}(x, y) = \begin{cases} 1 & \text{for } u_{x,m} \leq x \leq v_{x,m}, \quad u_{y,n} \leq y \leq v_{y,n} \\ 0 & \text{otherwise} \end{cases} \quad (4.12)$$

where  $m, n$  specifies the translation of the window in the  $x$  and  $y$  dimensions. Portions of the image can now be isolated by applying the window function to the image under consideration with the desired offset.

The localized spatial correlation function is now given by

$$\mathcal{LC}_{i,j}^{\#m,n}(a, b) = \sum_{x=u_{x,m}, y=u_{y,n}}^{x=v_{x,m}, y=v_{y,n}} [\varrho_i^{m,n}(x, y) - \bar{\varrho}_i^{m,n}][\varrho_j^{m,n}(x + a, y + b) - \bar{\varrho}_j^{m,n}] \quad (4.13)$$

where  $\varrho_i^{m,n}(x, y) = \phi_i(x, y)W_{m,n}(x, y)$  is the region of interest of the image and  $\bar{\varrho}_i^{m,n}$  is the mean of the sharpened image within the window of interest. Local texture invariants are computed using Eq 4.7 from the localized sharpened spatial correlations described above.

### Color-edge distribution based local operators

Color edge angle invariants provide a rich representation for a variety of textures. They can be computed locally across the image by applying the window function described in Eq 4.12 to the color image and to the Laplacian of the Gaussian of the image. This will yield a set of six angles for each window position. The angles can be computed by Eqs 3.13 and 3.19, where the dimensions of the feature vectors  $\phi$  and  $\chi$  are given by the area under the window function  $W_{m,n}(x, y)$ . Another advantage of

computing these angles locally is that the normalization for changing illumination is done locally, and thus the method can account for varying illumination across the image. The assumption underlying varying illumination is that the incident illumination can change spectrally over the texture surface, but such a change is gradual, and over a small region of the image the illumination remains fixed in its spectral distribution. Of course, the same can be said of the local spatial correlation invariants indicated above.

#### 4.4.2 Invariance of local representations

It is important to examine to what extent textures can be localized by the techniques explained in the previous section. A simple analysis is carried out to investigate the reliability of the local invariants for differing sizes of the window function. I obtain representations for the 10 textures used earlier, by setting the window to the size of the image. These form the model invariant set. While progressively reducing the window size, the local invariants are calculated and compared to the model set. The error introduced by the window causes a shift in the invariants, which is measured using the normalized Euclidean distance function described earlier. These errors are averaged over the set of textures for a given window size. I also averaged the number of correct matches to the database over the set of textures for each window size.

It was observed that as the window size decreased the representations became more susceptible to error. This is to be expected as a smaller window is unable to encapsulate all the spatial interactions in the texture. However in the case of the spatial correlation based local invariants, the results were extremely poor, even for relatively large window sizes: a 32x32 window over a 64x64 image yielded no better than a 60% recognition rate. The color-edge distributions on the other hand proved remarkably stable even at window sizes of 16x16 pixels for the same set of textures. The investigation of texture classification was therefore carried out using the color-edge local angle invariants.

### 4.4.3 Representing the invariant set

The problem of classification in the context of these local invariants is essentially a clustering problem: how many clusters of invariants can be obtained and which database texture invariants are they closest to? Unfortunately, clustering algorithms usually require as an input the number of clusters to be generated. This is obviously unknown for an image about which we have no *a priori* knowledge. Secondly, invariants for windows which contain texture edges will bias the clustering algorithm, as these invariants do not represent a texture and hence should be neglected. An alternative is to index these invariants by partitioning the invariant space using a hashing function. Hash buckets with high counts are matched to the model database, while hash buckets with low counts, which represent texture edge information, are neglected. Thus we can obtain a reliable characterization of the textures present in the image. Results using a set of outdoor scenes show good classification for the images considered.

# Chapter 5

## Results

“A man who holds a cat by its tail learns things he could learn no other way”

— *Mark Twain.*

The algorithms described in Chapters 3 and 4 were implemented and tested on a wide variety of data. Recognition results from these algorithms are described in this part of the thesis. Comparative results, where available from other recognition techniques, have been included. First I describe the performance of the object recognition algorithms, comparing it with Funt and Finlayson’s [FF95] Color Constant Color Indexing and Healey’s [HS94] global color constancy algorithm. Note that the performance of the algorithm described in Chapter 3 is better than Healey’s in all cases, and almost as good as color constant color indexing. Next, the texture recognition algorithm is evaluated for a set of natural textures. Results are fairly good under a wide change in illumination and rotation of the textures.

## 5.1 Object Recognition Results

Results are given for the algorithms explained in Chapter 3 — the color angle invariants, the edge angles invariants and the color-edge angle invariant method. Color constant color indexing and global color constancy are used for a relative evaluation of the performance of the system. Two data sets are used. The first set is the one described in Chapter 3 which was imaged in our laboratory. The second set is Swain's original database of 65 images which was used in his experiments on Color Indexing.

### 5.1.1 Objects under changing illumination:

A set of thirteen objects comprised of a t-shirt, a sweater, a book, some cereal boxes, and a few colored plastic bottles was imaged; some of these are shown in Fig 3.1. The objects were imaged under three sets of lighting conditions. A studio halogen lamp was used to obtain the first set of images. The other two sets were imaged by using a broad band blue filter and an orange filter placed over the camera, with the same halogen lamp used to illuminate the images. In each case, the orientation of the objects was changed in space so as to present a rotated and a slightly occluded, and in some cases (t-shirt and sweater), even a deformed view of the object. One of the objects was then imaged at a highly reduced scale to observe the effects of change in scale on the algorithm. The rank assigned to an object is its position in the sorted list of model objects matched to the image. Thus, a rank of 1 indicates the object was correctly recognized, while a rank of 2 or greater indicates the position of the correct match in the set of database objects. Table 5.1 shows the results of running the three algorithms on the dataset. The first row is the result of using only color distributions in the image. The second row is the set of results obtained by indexing on the angles of the edge distributions alone. Here the edges were produced by using a 7x7 Laplacian of Gaussian mask on the image. Finally the last row contains the results of using a combination of the two sets of angles. We note that the performance of the color distributions alone is the poorest, while the other two methods yield the same recognition rates (though the objects which are incorrectly identified are different in each case). This was predicted in Chapter 3, where we noted that the information



<i>Algorithm</i>	<i>Rankings</i>			
	<i>1</i>	<i>2</i>	<i>3</i>	<i>&gt; 3</i>
Color angles	20	3	3	0
Edge angles	20	5	1	0
Color and Edge angles	20	5	1	0

Table 5.1: Recognition results for angle invariants

content of the two sets of angles is uncorrelated and thus a combination of both could be used to advantage. These are results obtained by setting the sharpening transformation to unity.

Next, the performance of edge invariants is examined, using a variety of gradient operators. Table 5.2 shows the performance of operators such as the sum of gradients, a simple Laplacian and a 7x7 Laplacian of Gaussian (as used in the first table). We see that the Laplacian of the Gaussian does indeed provide the best results.

<i>Algorithm</i>	<i>Rankings</i>			
	<i>1</i>	<i>2</i>	<i>3</i>	<i>&gt; 3</i>
Sum of gradients	20	4	0	2
3x3 Laplacian	17	3	1	5
Laplacian of Gaussian	20	5	1	0

Table 5.2: Results using different edge operators

Finally the results shown in Table 5.1 are recomputed by applying the sharpening transform used in Chapter 3. Recognition results are shown in Table 5.3. Also included in the table are results obtained by running CCCI<sup>1</sup>. The last row is obtained by an implementation of Healey's Global Color constancy algorithm as outlined in [HS94]. We observe that CCCI provides the best recognition results, although the color and edge angles algorithm is nearly as good, with just one match ranked 3rd. Healey's algorithm provides the worst results; indeed we are surprised by its performance which

---

<sup>1</sup>The author thanks Graham Finlayson for running CCCI on the images

<i>Algorithm</i>	<i>Rankings</i>			
	<i>1</i>	<i>2</i>	<i>3</i>	<i>&gt; 3</i>
Color angles	21	4	1	0
Edge angles	20	6	0	0
Color and Edge angles	23	2	1	0
CCCI	24	2	0	0
Healey	11	6	6	3

Table 5.3: Comparative results of different recognition algorithms

is very poor given the small database size. We note that both the color angles and the edge angles have benefited, albeit slightly, by the use of the sharpening transform.

### 5.1.2 Swain's Database:

Swain's database is pruned from 65 objects to 55 in order to eliminate images with saturated pixels in them, as these pixels do not yield useful color information. The images are of a variety of man-made objects such as t-shirts, jerseys, boxes, plastic bottles and various packaged goods such as soup cans, paper towels. The images have been taken under a whitish illumination. Unfortunately all the images in the model as well as the test set have been taken under the same illumination, so we are unable to verify the effects of changing illumination on this database. Changes in orientation and some deformation are however present in the test database relative to the model images. For the test then, we have 55 model images and 24 test images from a subset of the same objects. As before, the three algorithms evaluated are shown in Table 5.4. As seen in the previous dataset, Healey's algorithm performs poorly. It gets noticeably worse as the database size increases with 7 objects matched with a greater than 3 rank. The color distribution angles and edge angles also perform poorly with just 16 and 17 objects recognized correctly in each case. The combination of both however, performs very well, as seen in row 3 of the table, where we note that the match performance is comparable with CCCI. Because these images have been taken with respect to a single illuminant, sharpening is not required. The performance of

<i>Algorithm</i>	<i>Rankings</i>			
	<i>1</i>	<i>2</i>	<i>3</i>	<i>&gt; 3</i>
Color angles	16	5	2	1
Edge angles	17	3	3	1
Color and Edge angles	21	2	1	0
CCCI	22	2	0	0
Healey	7	7	3	7

Table 5.4: Database of 55 real objects

the algorithm is excellent given the large database of objects and the fact that I index on just six numbers.

### 5.1.3 Remarks

For the color and edge distribution angle invariants to work well, the objects need to have a large set of colors in them. Since the algorithm is invariant to changes in scale, it is essentially the number of distinct colors and their respective proportions which provide the richness of representation. For example, if an object has only two colors in fairly equal proportions, the color distribution angles are no longer independent of each other. The  $N$  dimensional feature vector, where  $N$  is the size of the image, collapses to a two dimensional space and the third angle can be computed as the sum of the other two angles. This is not a serious restriction on the domain of the algorithm, however, as man-made objects are often multicolored. Secondly most color constancy algorithms need a large number of colors in the image to perform well so this requirement is present in most color based object recognition schemes.

## 5.2 Color texture recognition results

Using the model described in Chapter 4 for representing color textures, I attempt to recognize a number of natural textures based on the angular invariants of the spatial correlation functions. It was shown in Chapter 4 that these angles are illumination

invariant and are robust to rotation in the image plane. To evaluate the performance of the algorithms described in section 4.3, a set of ten natural textures is used as a model base. These textures have been described in section 4.3.2. They comprise a variety of natural surfaces such as sand, clouds, trees and some man-made textures such as fabrics and carpets. The model base of images has been imaged under nearly white light. For a test set, the textures were imaged with a set of 3 color filters. The filters used have a narrow band response in the blue, green and red regions of the visual spectra. A set of five rotations were applied to the set of 40 images. Rotation angles were arbitrarily chosen to be  $30^\circ$ ,  $45^\circ$ ,  $60^\circ$ ,  $90^\circ$  and  $110^\circ$ .

Three algorithms based on differing computations of spatial correlation functions were described in Chapter 4. In the first algorithm, the correlation function was computed as specified by Eq 4.1. In the second case, the correlation function was normalized by the actual overlapping image area of the two color planes being correlated. Finally, a fast implementation of spatial correlation was achieved using Fourier transforms. As noted in the section 3.3, the color and edge distribution algorithm also encodes spatial interaction in the image. I therefore use this as a fourth approach in texture recognition. Finally, I also match textures on the color distribution angle invariants algorithm to see how much the color distributions alone can characterize the set of textures. In all cases, the camera responses were narrow band and the sharpening transform has been set to unity.

First, I evaluate the performance of the algorithms relative to Healey's [HW95] algorithm. Healey's method, however, is only invariant to changing illumination and does not account for rotation of the texture. I therefore use a test set of 30 textures which are unrotated relative to the model textures. Table 5.5 shows the results of the various algorithms. The ranking scheme described in the previous section is used to evaluate the performance of the various methods. Results of Healey's algorithm are taken from [HW95]. It was felt unnecessary to implement Healey's algorithm here because I use the images he utilized in his experiments.

Healey's algorithm works very well, correctly recognizing all the textures. Similarly the normalized spatial correlation function invariants are also successful in recognizing

<i>Algorithm</i>	<i>Rankings</i>			
	<i>1</i>	<i>2</i>	<i>3</i>	<i>&gt; 3</i>
Spatial Correlation angles	29	1	0	0
Normalized spatial correlation angles	30	0	0	0
Fourier correlation angles	29	0	1	0
Color and edge angles	30	0	0	0
Color angles	25	4	1	0
Healey	30	0	0	0

Table 5.5: Texture recognition for a database of ten textures

all the textures. The Fourier transform correlation invariants and the spatial correlation function angles correctly recognize all but one of the textures. Surprisingly, we note that the color and edge angles are also successful in correctly recognizing all textures. The color distribution angles fare comparatively poorly, with five textures incorrectly recognized. Even then, all the textures are matched within the first three places.

Tests are now run with the set of 230 images obtained by rotating the textures in the image plane. The database remains the same set of ten textures as in the previous case. Results for the various algorithms are shown in Table 5.6. Firstly, we notice that

<i>Algorithm</i>	<i>Rankings</i>			
	<i>1</i>	<i>2</i>	<i>3</i>	<i>&gt; 3</i>
Spatial Correlation angles	190	37	3	0
Normalized spatial correlation angles	195	29	6	0
Fourier correlation angles	190	16	13	11
Color and edge angles	224	6	0	0
Color angles	124	45	29	32

Table 5.6: Texture recognition with change in illumination and rotation about the optical axis.

using color distributions alone fails; almost half the textures are incorrectly recognized. Most surprisingly we note that the color and edge angle distributions have the best

results, with all but six of the images correctly recognized. The invariants based on the spatial correlations perform reasonably well, with most of the texture images being placed within the first two rankings. We observe that the Fourier transform computation of the correlations has a slightly degraded performance relative to the other spatial correlation invariants. This is probably due to the circular correlation effect described in Chapter 4.

### 5.2.1 Color texture classification

Color texture classification, achieved by a local computation of color-edge invariants across the image, yields excellent results for a wide variety of outdoor scenes. A model base of a wide variety of natural textures comprised of different types of grass, rocks, clouds, sand, trees and water was created from a set of images obtained from the VisTex database at the Vision and Modeling group, MIT Media Lab.

A database of 12 natural textures was chosen, some of which are shown in Fig 5.1. Context scenes such as the ones shown in Fig 4.11 and Fig 5.2 were subjected to the texture classification algorithm described in section 4.4. The context images were either 512x512 pixels or 640x480 pixels. A window size of 32x32 pixels was chosen for purposes of obtaining the local texture invariants. All classification results were obtained for non-overlapping windows. In all cases, component textures were correctly identified. By computing these invariants for every pixel in one image, a set of six images was generated: one for each texture invariant, in a manner similar to Tuceryan [Tuc94]. These are shown in Fig 5.3, for the image in Fig 5.2. We observe from the invariant images that the original image can be easily segmented into its two component textures by using the texture invariants. The number of buckets created by the hashing function can be fed to a clustering algorithm, such as K-means in order to obtain a segmentation of the image. This has been tested with a clustering algorithm based on a least-sum-of-squares criterion using competitive learning [UA94]; results are shown in Fig 5.3. We see that the clustering is poor in this case; whereas the segmentation obtained by using the last invariant image (Fig 5.4) is far superior. We predict that good segmentation can be achieved by a more informed approach.

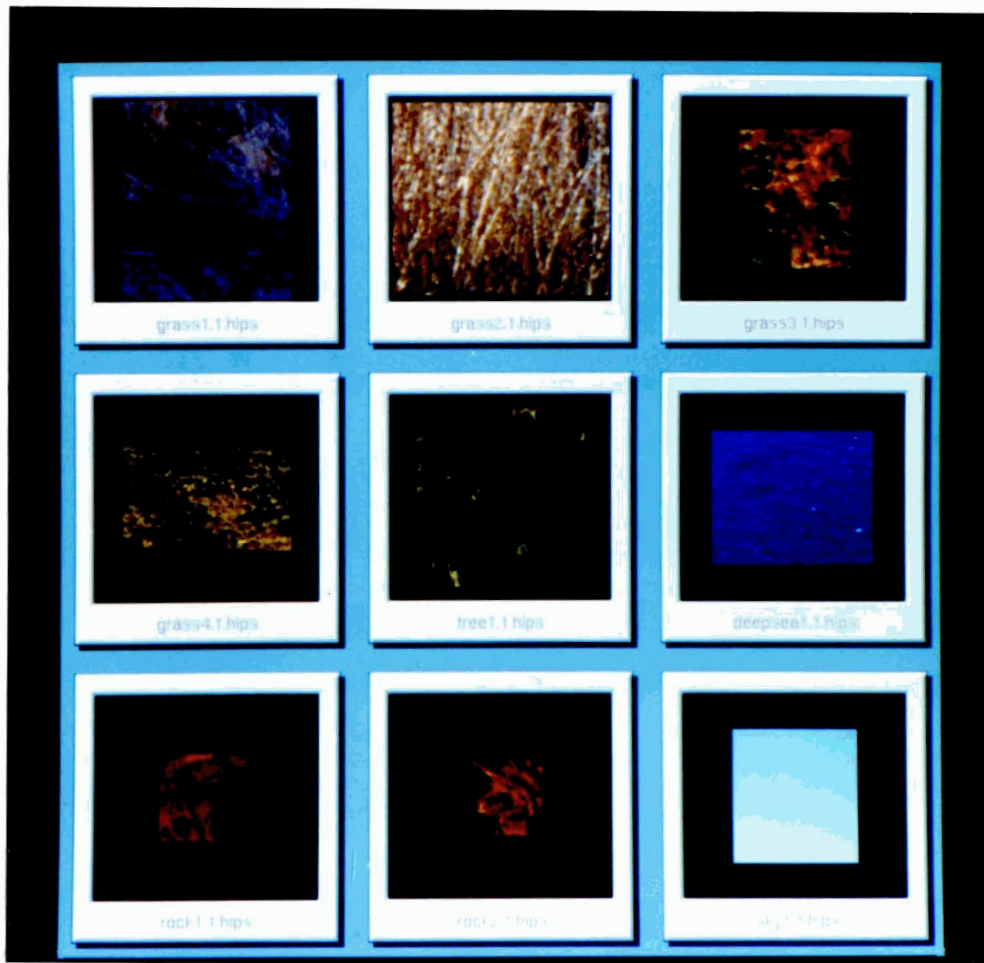


Figure 5.1: Database of natural textures used for texture classification



Figure 5.2: Context images of natural scenes used for texture classification



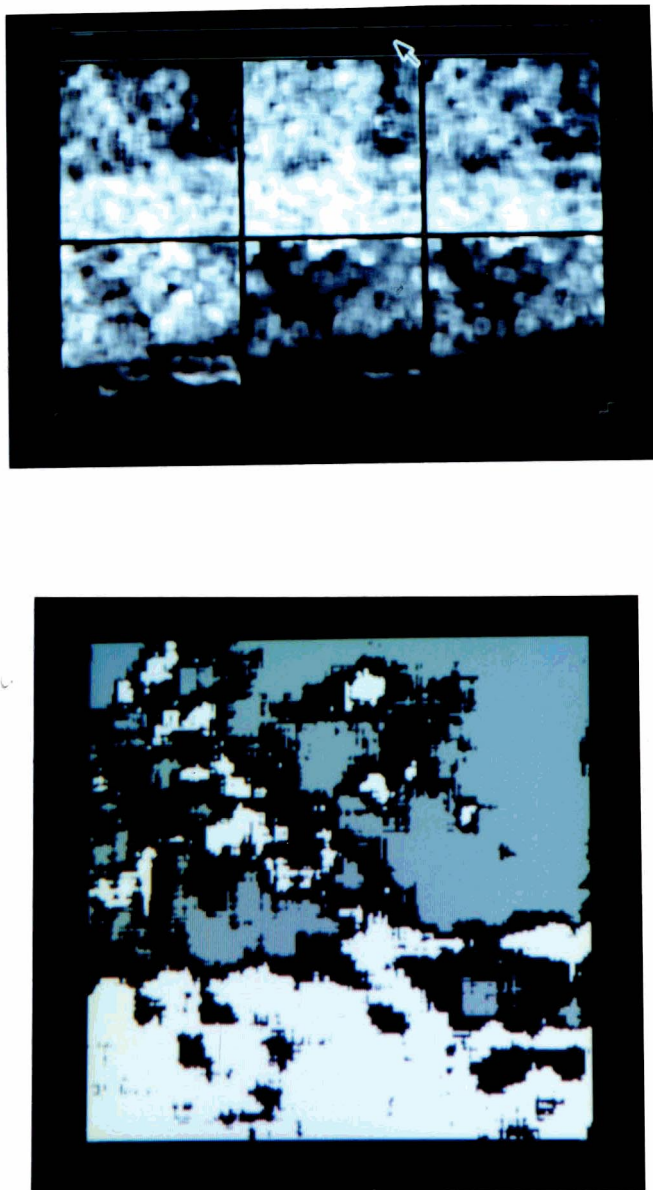


Figure 5.3: a) Invariant images for a context scene; b) Clustering results for the invariants



Figure 5.4: Segmented texture image

# Chapter 6

## Conclusions

Dirk stepped boldly through, . . . and then announced that he saw exactly how it worked, that it was obviously to do with the unreal numbers that lay between minimum quantum distances and defined the fractal contours of the enfolded Universe and he was only astonished at himself for not having thought of it himself.

— *Douglas Adams, "Dirk Gently's Holistic Detective Agency"*

I have presented a new algorithm for object and texture recognition based on the color distributions and spatial color variations in an image. This algorithm has been extended to texture classification and preliminary results for texture segmentation are promising. The work can be extended in various ways. A careful data analysis of the angles shows a lower variance in the edge angles over the entire data set of images than for the color distribution angles. The application of a weighting function to the edge angle invariants is likely to improve the match results. Careful sharpening will also improve results under varying illumination conditions. I consider these issues in the following sections. While results for texture recognition show that angle invariants work reasonably well for spatial correlation representations of textures, I have found

an excellent representational scheme for textures in the framework of color-edge angle invariants. This method proves stable under relatively small window sizes, and thus textures can be classified and segmented on the basis of these invariants.

## 6.1 Application based invariant representation

For the data sets I have tested for object recognition, the variance of the edge angles has been smaller than the variance of the color distribution angles. Thus, given a database, good weighting functions can be found to optimize the representation of the invariant space. This of course assumes that the database is relatively static, with few objects being added, or that the objects added are similar to existing objects in the database so as to avoid a re-computation of the weighting function and updating the invariant database. Similarly, application specific problems, such as ambient lighting conditions or noise, can be accounted for in various ways. If the lighting variations are known or can be controlled, for a given set of objects, database sharpening can be optimized to yield the best possible results for a diagonal model of color constancy [FDF94b].

Limitations of the method have already been pointed out before. A large set of colors is an important requirement for objects if they are to be reliably identified. Single colored or bi-colored objects are unstable in their representations. Edge distributions are limited in their availability in such cases, and the edge angles contribute very little to the representation as well. Limitations of color indexing have been explored by various people (Niblack et al. [NB93], Pentland et al. [PPS94], Gong [GZ94]) and the general conclusion is that color indexing needs to be combined with some sort of textural and geometric properties to reduce the number of false positives. I have, however, shown that edge angles provide a good representation for texture recognition. Spatial interaction about a scene is thus well encoded in the invariants. Whether the representation is rich enough to index a very large database of textures has not been dealt with here.

## 6.2 Extensions

We can observe that the angles represent second order statistical information about the image. In the case of texture representations, it seems likely that higher order information may be useful for the generation of a larger and richer set of invariants. Such higher order information could be obtained by the use of higher order differential operators (I have used a Laplacian, but other higher order differencing operators may be used as well). Alternatively, higher order algebraic moments of the sharpened and normalized data set can be considered for purposes of indexing. As a case in point, the recovery of 3-dimensional textural patches as described by Kondepudy and Healey [KH94] can now be achieved under varying illumination by considering spatial correlations of the sharpened and normalized data.

Another possibility is the consideration of inter-color band edges. Healey's [HW95, HS94] correlation functions model the inter-band spatial interaction by cross-correlating the band images. A different approach would be to consider the application of an inter-band Laplacian mask to the color image. This can be done using a Laplacian of Gaussian mask by using the center of the mask in one band and the support of the mask in the other color band. This will lead to the use of inter-band edge information. In the case of white or uniform (grey) colors, the LoG mask will sum to zero, thus indicating the absence of an inter-band edge. Presence of a color will cause a value to be present for the inter-band interaction. While I have not implemented this, it seems to promise a rich source of spatial color information about a scene. Indirect evidence of the usefulness of inter-band spatial interactions is available from Healey's spatial cross-correlation functions.

## 6.3 Texture segmentation schemes

Methodologies as suggested above can be used to extend the set of invariants in various ways. It is interesting to consider how this affects texture classification and texture segmentation. Most texture segmentation schemes rely on model based methods, such as Markov random field models, for a compact support of representation of

textural features [HP95]. This becomes important as most segmentation schemes depend upon the local computation of textural features. Healey's Markov random field segmentation scheme, for example, is able to use window sizes of up to  $4 \times 4$  pixels for identifying texture features. The possibility of small windows leads to a finer segmentation of the scene. Such a necessity is of course application dependent. By their definition, small window sizes will be unable to capture more global spatial interactions in the image and if these are of importance in segmentation, some other transformation schemes might be considered as a preprocessing step. In our case, a richer set of invariants may enable us to use smaller windows for segmentation. Other applications of local computations of textures can be in the field of texture indexing — where we attempt to recognize objects by the distributions of various textures present in the object. As a preliminary step, I have tried indexing the database of objects, described as the first data set in Chapter 5, using a set of localized spatial correlation angles with reasonably good results. However, such a database is probably not suitable for such a representation — natural scenes with a richer set of textures would perhaps provide for a more interesting case study of texture indexing.

## 6.4 Concluding remarks

A new algorithm for color object and texture recognition, based on just six numbers, has been presented. The six invariants used are the normalized dot products of the inter-band color and edge distributions of the image. They represent second-order statistical information about the scene, and are similar to lower order moments. The representation has been shown to be rich, with very good discriminatory power. Indexing on six numbers is extremely fast, as opposed to indexing on over 4000 numbers in the case of Color Constant Color Indexing. Recognition results are of the same order of CCCI, under varying conditions of illumination, rotation and changes of scale.

The six invariants have been shown to provide a rich representation for color textures. Textures are recognized under rotations about the optical axis, under changes in illumination and over a range of scales. The computation is linear in the size of the image; spatial correlation based texture features are, on the other hand, quadratic in

the image size. This is a distinct computational advantage over the use of spatial correlations. Matching is also very quick, involving only the computation of the distance between the six invariants of a test image and a model database of textures. The six invariants are stable when computed over a small support, and have been successfully used for texture classification and segmentation for images of natural outdoor scenes.

# Bibliography

- [Bar95] J. Barnard. Computational color constancy: Taking theory into practice. Master's thesis, Simon Fraser University, Vancouver, August 1995.
- [BB82] D. H. Ballard and C. M. Brown. *Computer Vision*. Prentice Hall, New York, 1982.
- [Cog82] J. M. Coggins. *A Framework for Texture Analysis based on Spatial Filtering*. PhD thesis, Computer Science Department, Michigan State University, East Lansing, MI, 1982.
- [DL86] M. D'Zmura and P. Lennie. Mechanisms of color constancy. *J. Opt. Soc. Am. A*, 3:1662–1672, 1986.
- [FDF94a] G.D. Finlayson, M.S. Drew, and B.V. Funt. Color constancy: Generalized diagonal transforms suffice. *J. Opt. Soc. Am. A*, 11:3011–3020, 1994.
- [FDF94b] G.D. Finlayson, M.S. Drew, and B.V. Funt. Spectral sharpening: Sensor transformations for improved color constancy. *J. Opt. Soc. Am. A*, 11(5):1553–1563, May 1994.
- [FF95] G.D. Finlayson and B.V. Funt. Color constant color indexing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(5):522–528, May 1995.
- [Fin92] G. Finlayson. Color object recognition. Master's thesis, Simon Fraser University, Vancouver, April 1992.



- [Fin95] G. Finlayson. *Coefficient Color Constancy*. PhD thesis, Simon Fraser University, Vancouver, April 1995.
- [GMTL86] A. Gagalowicz, S. D. Ma, and C. Tournier Lasserre. Efficient models for color textures. In *Proceedings of the 8th International Conference on Pattern Recognition*, pages 412–414, 1986.
- [GW92] R. C. Gonzalez and R. E. Woods. *Digital Image Processing*. Addison-Wesley Publishing Company, Mass., 3 edition, 1992.
- [GZ94] Y. Gong and H. Zhang et al. An image database system with content capturing and fast image indexing abilities. In *Proceedings of the IEEE Internatinal Conference on Multimedia Computing and Systems*, pages 121–130, May 1994.
- [Har79] R. M. Haralick. Statistical and structural approaches to texture. *Proc. IEEE*, 67:786–804, 1979.
- [Haw69] J. K. Hawkins. Textural properties for pattern recognition. In B. Lipkin and A. Rosenfeld, editors, *Picture Processing and Psychopictorics*. Academic Press, New York, 1969.
- [Hor86] B. K. P. Horn. *Robot Vision*. The MIT Press, Cambridge, Mass., 1986.
- [HP95] G. Healey and D. Panjwani. Results using markov random field models for the segmentation of color images of natural scenes. In *Proceedings of the 5th International Conference on Computer Vision, MIT, Cambridge*, pages 714–719, 1995.
- [HS94] G. Healey and D. Slater. Global color constancy: recognition of objects by use of illumination invariant properties of color distributions. *Journal of the Optical Society of America, A*, 11(11):3003–3010, November 1994.
- [Hur89] A.C. Hurlbert. *The Computation of Color (PhD Thesis)*. MIT Artificial Intelligence Laboratory, 1989.

- [HW95] G. Healey and L. Wang. The illumination-invariant recognition of color textures. In *proceedings of the 5th International Conference on Computer Vision, MIT, Cambridge*, pages 128–133, 1995.
- [JMW64] D.B. Judd, D.L. MacAdam, and G. Wyszecki. Spectral distribution of typical daylight as a function of correlated color temperature. *J. Opt. Soc. Am.*, 54:1031–1040, August 1964.
- [KH94] R. Kondepudy and G. Healey. Use of invariants for recognition of three-dimensional color textures. *Journal of the Optical Society of America, A.*, 11(11):3037–3049, November 1994.
- [Kri47] E.L. Krinov. Spectral reflectance properties of natural formations. *Technical Translation TT-439, National Research Council of Canada*, 1947.
- [Mal86] L.T. Maloney. Evaluation of linear models of surface spectral reflectance with small numbers of parameters. *J. Opt. Soc. Am. A*, 3:1673–1683, 1986.
- [NB93] W. Niblack and R. Barber et al. The qibc project: Querying images by content using color, texture and shape. In *Storage and Retrieval for Image and Video Databases I*, volume 1908 of *SPIE Proceedings Series*, February 1993.
- [Nic57] D. Nickerson. Spectrophotometric data for a collection of munsell samples, 1957.
- [PHJ89] J.P.S. Parkkinen, J. Hallikainen, and T. Jaaskelainen. Characteristic spectra of munsell colors. *J. Opt. Soc. Am. A*, 6:318–322, 1989.
- [PPS94] A. Pentland, R. W. Picard, and S. Sclaroff et al. Photobook:tools for content-based manipulation of image databases. In *Storage and Retrieval for Image and Video Databases II*, volume 2185 of *SPIE Proceedings Series*, February 1994.

- [SB91] M.J. Swain and D.H.. Ballard. Color indexing. *International Journal of Computer Vision*, 7(11):11–32, 1991.
- [SHS92] J. Scharcanski, J. K. Hovis, and H. C. Shen. Color texture representation using multiscale feature boundaries. *SPIE, Visual Communications and Image Processing*, 1818:156–165, 1992.
- [SSS92] B. Smith, C. Spiekermann, and R. Sember. Numerical methods for colorimetric calculations: Sampling density requirements. *COLOR research and application*, 17(6):394–401, 1992.
- [SW64] C. E. Shannon and W. Weaver. *The Mathematical Theory of Communication*. The University of Illinois Press, Urbana, IL, 1964.
- [TC92] G. Taubin and D. Cooper. Object recognition based on moment (or algebraic) invariants. In J. Mundy and A. Zisserman, editors, *Geometric Invariance in Computer Vision*, pages 375–397. MIT Press, Cambridge, Mass., 1992.
- [TJ93] M. Tuceryan and A. K. Jain. Texture analysis. In C. H. Chen, L. F. Pau, and P. S. P. Wang, editors, *Handbook of Computer Vision and Pattern Recognition*, pages 235–276. World Scientific Publishing Co., Singapore, 1993.
- [TMY78] H. Tamura, S. Mori, and Y. Yamawaki. Textural features corresponding to visual perception. *IEEE Trans. Syst. Man Cybern.* , pages 460–473, 1978.
- [Tuc94] M. Tuceryan. Moment based texture segmentation. *Pattern Recognition letters*, 15:659–668, July 1994.
- [UA94] T. Uchiyama and M. A. Arbib. Color image segmentation using competitive learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(12), December 1994.

- [VGI94] M.J. Vrhel, R. Gershon, and L.S. Iwan. Measurement and analysis of object reflectance spectra. *COLOR research and application*, 19(1):4–9, 1994.
- [WB82] G. West and M.H. Brill. Necessary and sufficient conditions for von Kries chromatic adaption to give colour constancy. *J. Math. Biol.*, 15:249–258, 1982.
- [WL93] L. H. Wurm and G. E. Legge et al. Color improves object recognition in normal and low vision. *Journal of Experimental Psychology*, 19(4):899–911, 1993.
- [WS82] G. Wyszecki and W.S. Stiles. *Color Science: Concepts and Methods, Quantitative Data and Formulas*. Wiley, New York, 2nd edition, 1982.