Use of Colour in Machine Vision: Colour Representation, Edge Detection with Colour, Segmentation of Colour Space and Colour Constancy

Tin Wai Robert, HUNG
December 1999

Title :

Use of Colour in Machine Vision: Colour Representation, Edge Detection with Colour, Segmentation of Colour Space and Colour Constancy.

Abstract:

This report is a study of the role of colour and its use in machine vision. Intuitively, for low level image processing, colour provides greater discrimination than grey level for separating different homogenous regions in an image. The first part describes the use of colour in edge detection. In current vision systems, extracting object features such as lines and arcs relies on edge detection. The use of colour images (RGB) can offer additional confidence in the existence of an edge element in one plane when it is corroborated by pixels at the location on one or more of the other planes. The second part investigates the problems and techniques associated with colour image segmentation. A spectral segmentation algorithm based on locating the boundaries of each colour cluster in the spectral space is proposed. The third part investigates the use of colour features for object recognition. Colour information also provides a useful cue for object localisation and identification. The major issues that have to be addressed are colour constancy and representation, and also, their connections to segmentation. Finally, a system for locating object surfaces based on a simplified colour constancy and its colour representation is proposed.

ACKNOWLEDGEMENTS

Declaration

I grant powers of discretion to the University Librarian to allow this thesis to be copied, in whole or in part, without further reference to me. This permission covers only single copy made for study purposes, subject to normal conditions of acknowledgement.

Contents

Chapter 1 Introduction

Chapter 2 Colour Information

Chapter 3 Use Of Colour In Edge Detection

Chapter 4 Segmentation Of Colour Space

Chapter 5 Colour Constancy

Chapter 6 Conclusions & Further Work

Chapter 1
Introduction

1.1 Introduction

The motivation for employing colour in a machine vision system is commonly based on the following reasons:

- Colour provides a larger discriminant space:

  A grey level image consists of the achromatic information, where each pixel is represented by a single attribute. Colour adds chromatic information, and each pixel is represented by several attributes, forming a larger attribute space.

- Colour provides a stable cue for object localisation [ODE99] and is one of the intrinsic features for object identification:

  The colour information of an object's surface is local and can be directly obtained from the pixel values, hence the quality of colour cues is distorted less in comparison with geometric cues for changes of viewpoint and reduction of resolution of the images.

Although it is the fact that colour provides an important attribute of an object, there are many difficulties associated with the use of colour to aid the object recognition. These difficulties can be divided into two categories. One is related to the extraction of geometrical features from the image, and the other is related to the extraction and representation of colour features of the surface.

One of the common tasks in machine vision is object recognition. The goals are the identification of an object at a known location or the localisation of an known object. In order to identify an object at the known location, two sets of data are required: one is the object's features, extracted from the image; the other is the model feature of an object, the features of which describe an object. An identification is verified if the extracted features well match to a set of features of an object model under some pre-defined criterion. Therefore feature extraction and representation plays an important role in the overall identification process. Object features can be generally divided into two types: geometric features and surface features; e.g. colour or texture.

Geometric features are commonly extracted by edge detection or other segmentation methods. Edge detection is used to detect discontinuities in the image space. Segmentation partitions the image into a set of homogenous regions. Where the use of colour information has been studied previously, methods focused on building a colour extension to the existing grey-level image processing techniques. For example, in edge detection, Nevatia [Nev77] has extended the Hueckel operator to colour images, and found that the resulting image is visually better. However, there is lack of any quantitative comparison between the performance of the grey-level and colour edge detection. For segmentation, Ohta [Oht85] extended the histogram thresholding technique to cope with colour images, to obtain the colour features of the surface, yielding closed boundaries. The use of colour in segmentation can provide a greater discrimination, but the additional features will increase the complexity of the task.

There are two steps in extraction; first segmentation of the image and then obtaining an accurate description of the segmented regions (surfaces). For the segmentation step, the difficulties associated with the use of colour images is the complexity of high dimensional data. It seems straight forward to measure the colour of a region by measuring the colour value of each pixel within the region. However, the pixel values are a function of the reflected light from the surface, formed from the product of the spectrum of the surface reflectance and the ambient light impinging on it. If the ambient light varies, the sensed pixel values will change accordingly. Such an unreliable measure of the colour feature of any object may result in mis-identification or incorrect localisation. It is desirable to have a mechanism that would enable the machine vision system to adapt to the effect of the changes of the ambient light. This mechanism is commonly called colour constancy and has been widely studied by the researchers investigating human vision. Another issue is that if colour features are used in object identification, what is an appropriate representation for the colour surfaces in order to use it for storing colour information and identification.

This thesis addresses issues mentioned in the last two paragraphs associated with the use of colour in extracting geometric and colour features of an object. It is divided into four main topics: use of colour in edge detection, segmentation of colour images, colour constancy and colour information representation. Those are the key topics in constructing modules of a machine vision system using colour.

1.2 Objectives

There is not a general architecture for a machine vision system and the general vision task has not been defined. An example of a colour machine vision system can be seen in figure 1.0, where processing modules provide colour constancy, edge detection, colour segmentation, and object identification. A colour database supports several of the processing modules, in order to provide colour reference information.

This thesis is an attempt to illustrate some of the problems associated with using colour and experiments with some possible solutions. In particular, it will consider the use of colour for segmentation.

An attempt is made to establish a quantitative measure for comparison of the performance of a colour and a grey level edge detector. This measure is based on the difference between the strength of the grey-level and the colour edgel. Previous researchers on the use of colour in edge detection, have concluded that most of colour edgels are correlated in three colour planes. An approach to combine edges based on the correlation among the three colour planes is considered. Thus, a measure is needed for evaluating the suitability of such a combination method.

Obtaining the boundaries of an object by partitioning the image into a set of regions is studied. The methods used in this thesis are based on segmenting the colour space, such that the colours from a homogenous region should have pixels of similar colour based on their closeness in colour space. Two approaches for the segmentation of colour space are considered: one signal-based and one model-based segmentation. The signal-based approach considers the image as a set of random data. This technique will exploit the assumption that each subset of colours is distributed in a unimodal manner, with the use of a merging method to group the colours accordingly. The model-based approach considers the image colours corresponding to the colours in a database, that contains the modelled colours of certain object surfaces. A measure is required to define the similarity between an image colour and a model colour.
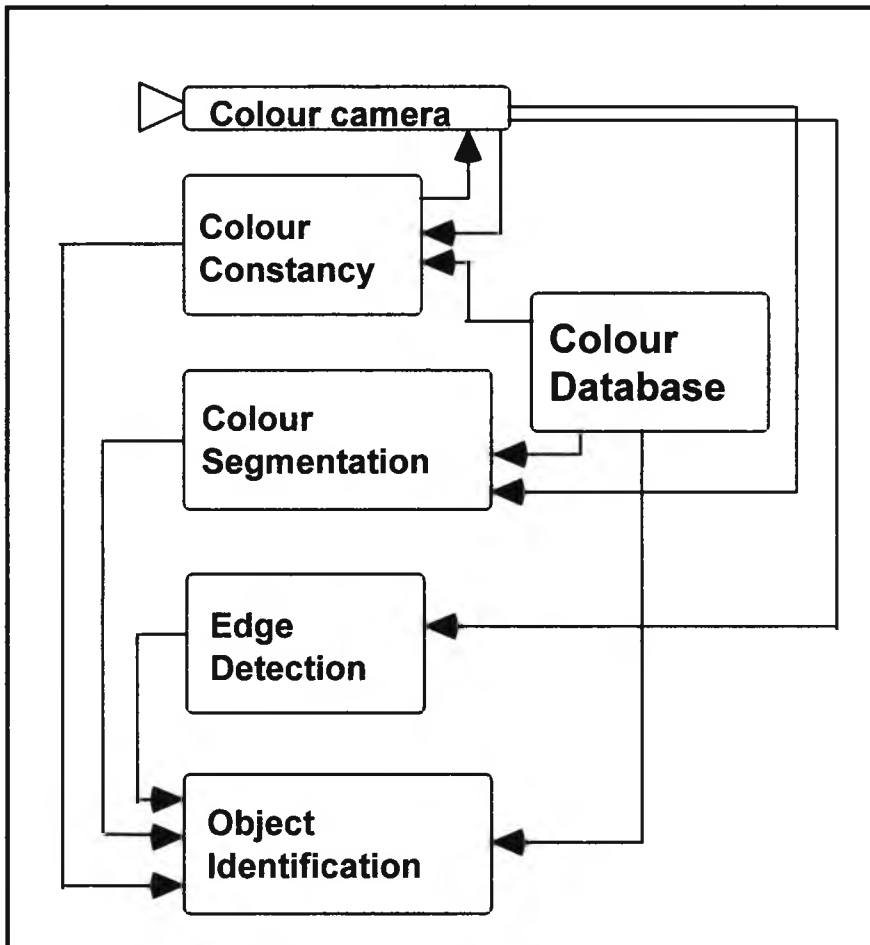
Figure 1.0 Example of a colour machine vision system for object identification.

Issues associated with the extraction of colour (surface) features are: Colour constancy and colour information representation.

It is desirable to have a machine vision system equipped with a colour constancy algorithm to adapt to different lighting conditions. Because changes of lighting result in changes of the measured colour of the observed surfaces, a review of existing colour constancy algorithms is made to identify common features. A colour constancy algorithm based on these common features will be developed.

One important issue in object recognition is the representation of features for identification. In this report, an investigation is made to find common features in different colour information representations for the object or colour recognition system.

1.3 Outline

This report is organised as self-contained chapters. The reader can read each one independently; in general, all of them contain relevant background knowledge. Additionally, chapter 2 provides a more general background on the use of colour in machine vision such as image formation, colour co-ordination systems, colour information representation and the proposed colour database which is exploited in chapters 4 and 5.

Chapter 2 introduces the image formation, colour co-ordination system and reviews previous research into the uses of colour in machine vision, from low level image processing to colour object or surface recognition. Section 2.3 demonstrates two common features that can be found in the representation of colour surfaces for object recognition or inspection. A proposed representation of a colour surface can be seen in section 2.4 and is called 'memory colour'. This is a key part in this report because the colour segmentation and the colour constancy algorithm are based on the use of 'memory colour'. Finally a new data structure called the Colour Information LookUp Table (CILUT) is used to accommodate the 'memory colour'.

Chapter 3 introduces the evolutionary development of a colour edge detector , high-lighting the difference between the performance of a colour and a grey level edge detector, the corroboration of an edgel in three colour planes and the method of combining colour edgels. A quantitative measure of this difference is defined in section 3.3. A mathematical and statistical approach is used to investigate whether there is any significant performance difference, and these results can be seen in section 3.3.3. Based on the conclusions from previous research on colour edge detection and the observation of my own experiments, it found that the edgel from three colour planes are highly correlated at the same and neighbouring locations. A method combining colour edgels is proposed and described in section 3.4. Finally, a set of experiments are carried out to evaluate the performance of this method and the results can be seen in section 3.4.2.

Chapter 4 introduces some alternative approaches for colour image segmentation. The segmentation of a colour image based on segmenting the colour space is studied extensively. An algorithm for signal-based segmentation is constructed and based on the assumption that each subset of

colours is distributed in a unimodal manner in colour space. Details of the algorithm and its resulting segmentation are described in section 4.4-6 and 4.8 respectively. A model-driven segmentation based on the classification of image colours with a set of modelled colours from the database is also described. This database is constructed to accommodate the 'memory colour' which is described in chapter 2. The detail of the algorithm and the experimental results are described in section 4.7 and 4.9 respectively.

Chapter 5 introduces the problem of colour constancy, and is divided into two parts: the first considers what form of transformation is used in the algorithm to discount the effect of the illuminant (the transformation of colour constancy); secondly, how the parameters of the desired transformation are computed (the algorithm of colour constancy). Following a review of some existing algorithms, it is shown how they can be incorporated into 'memory colour'. A detailed illustration of the use of 'memory colour' in some existing algorithms is described in section 5.3. A proposed algorithm can be seen in section 5.4. It is a colour matching algorithm, because the algorithm generates a match between the observed colours and the 'memory colours' and then calculates the required transformation for that match. The results in section 5.7 show the algorithm performing effectively and accurately on the test images.

Finally conclusions are made in chapter 6 together with some suggestions of further work on colour edge detection and colour constancy.

Chapter 2
Colour Information

"...For if we are capable of knowing what is where in the world, our brains must somehow be capable of representing this information- in all its profusion of COLOUR..."
David Marr 1982

Abstract:

This chapter provides the background of the use of colour in low and high level vision. Concepts from image formation to the colour database are introduced. Different colour spaces and their use in vision systems will be explained. On the review of colour object recognition, colour provides a stable cue for localisation and is a generic feature for identification. Therefore, what is the necessary colour information that should be included in any system that is to cope with various vision tasks? A colour database is introduced, which is a data structure based on a lookup table and the element within the structure is called colour memory. The colour database is exploited in the work described in chapters 4 & 5.

2.1 Introduction

This chapter is divided into two parts: first to study different colour representations and their use in machine vision; second to study the use of colour information in high level vision, such as colour recognition.

In the past few decades, research on the use of colour in machine vision has mostly concentrated on the use of different colour systems for pixel representation. The purpose is to evaluate certain colour features for their suitability to some low level image processing, such as edge detection or region segmentation. Most of the approaches are based on statistical and empirical methods, since the image data is considered as a set of patterns. The trichromatic representation (red, green and blue) is most commonly used to aid the understanding of physical properties of the surface reflection and illumination, because it is directly related to the physical process of image formation. The various colour systems are described in section 2.2.1-7.

Colour information is a useful attribute of objects for object recognition, because it is invariant to object pose. Also the geometrical information present in an image or a sequence of images may not be sufficient or accurate enough for object localisation and identification, due to the inadequacy of the information, which is caused by the presence of noise, inconsistent viewing orientation and occlusion. Extraction of geometrical information from images requires operations such as region segmentation, edge detection, edge linking, etc. On the other hand, colour information of regions can be directly extracted from the pixels, because colour values of the pixels directly reflect the physical properties of the surface at that location. Whether the recognition of an object is based on geometrical or colour information, there must be some kind of knowledge-base or database in the system, which contains the information about the object.

The question arises: what is sufficient colour information in the database and image required to accomplish the task of recognition. There are two common and essential features that can be found in colour (surface) information for object or colour recognition in the literature. Section 2.3

overviews the features that have been used by various researchers. For the work described here, a generic colour information for each surface colour is proposed and named *memory colour*[1]. Also a data structure called the *colour information lookup table (CILUT)* which will accommodate the two basic parts of memory colour is proposed for increasing the efficiency of the manipulation of the colour information. Moreover the memory colour is exploited in conjunction with a colour image segmentation and a colour constancy algorithm, which are described in chapters 4 and 5 respectively.
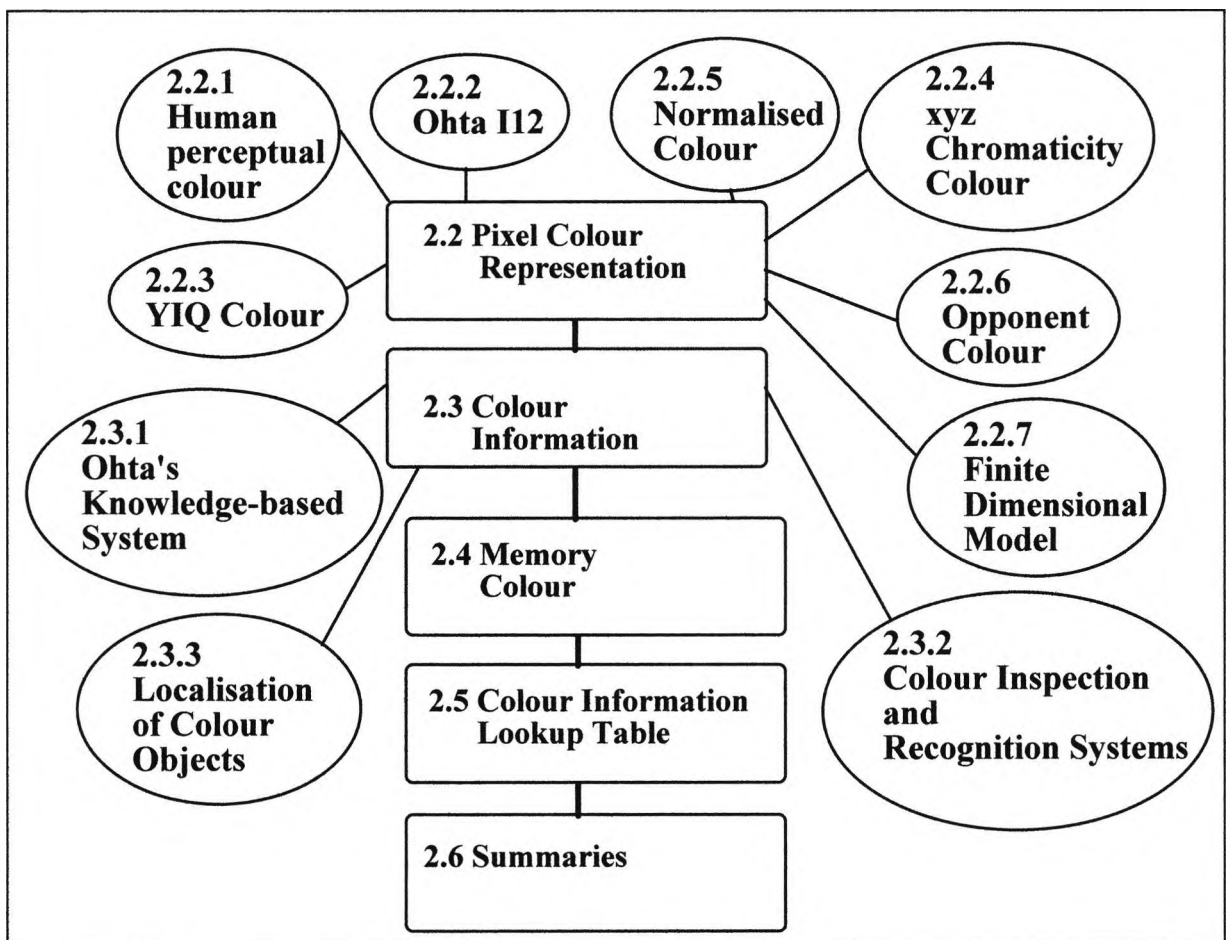
Figure 2.0 The structure for chapter 2.

---

[1] The name is originated by Hering 1874 and is used in human vision.

2.2 Image Formation

Thomas Young's theory[1] provided the foundation for the formation of colour images. Any visible colour within the spectrum can be reproduced by a mixture of three principal colour lights with appropriate intensities. It has been determined that certain wavelengths of red, green and blue, when combined with each other in various intensities will produce a wider range of colours than any other combination of three colours. These colours are thus referred to as the primary colours.

Acquisition of a colour image is based on this tristimulus system that contains three sensors, which respond to three different spectral bands. The formation of an image from the light impinging on a surface to the sensor in the camera is shown below (in figure 2.1). The interaction between illumination, surface reflectance and sensors is shown in eqn. 2.1-3 where *r, g* and *b* are the measurements obtained at each location in the image. Technically, the sensor responses to different colour bands can be achieved by using filters to separate into different bands of wavelengths. Figure 2.2 shows the transmittance of three common filters.
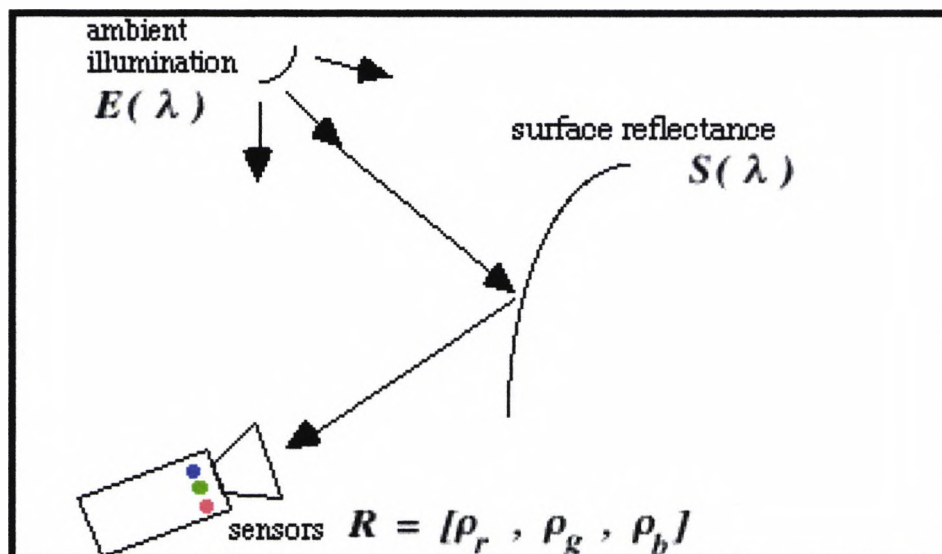
Figure 2.1 Acquisition of image.

---

[1] Details can be found in any book on colour or colour theory. e.g [Gra86]

$E(\lambda)$ is the spectral power distribution of the ambient illumination

$S(\lambda)$ is the surface reflectance function

$\rho_r(\lambda), \rho_g(\lambda)$ and $\rho_b(\lambda)$ are the sensor responses to red, green and blue wavelengths respectively

$$r = \int E(\lambda)S(\lambda)\rho_r(\lambda)d(\lambda)\cdots\cdots(2.1)$$

$$g = \int E(\lambda)S(\lambda)\rho_g(\lambda)d(\lambda)\cdots\cdots(2.2)$$

$$b = \int E(\lambda)S(\lambda)\rho_b(\lambda)d(\lambda)\cdots\cdots(2.3)$$

r, g and b are the measurements of the intensity of the light reflected from the surface with respect to their bands of wavelengths.

For a monochromatic camera, the sensor generally respond to the full range of the visible spectrum.

$$i = \int_{l}^{u} E(\lambda)S(\lambda)\rho(\lambda)d(\lambda)\cdots\cdots(2.4)$$

where u and l are the upper and lower bound of the visible spectrum respectively.
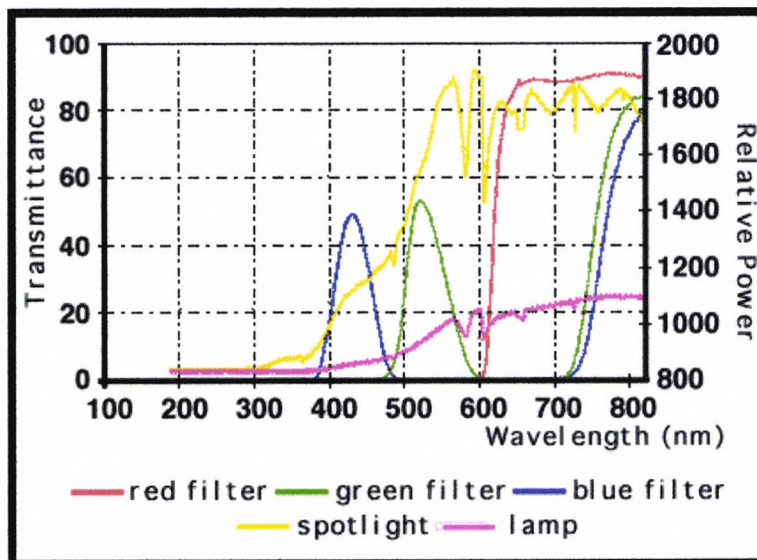
Figure 2.2 Transmittance of different filters and the spectral power density of two different illuminants.

The three mutually independent primary colours constitute an RGB colour system for colour image analysis and display. A vector representation of

colour that allows a colour to be quantitatively represented in the three-dimensional Cartesian space is shown in figure 2.3.
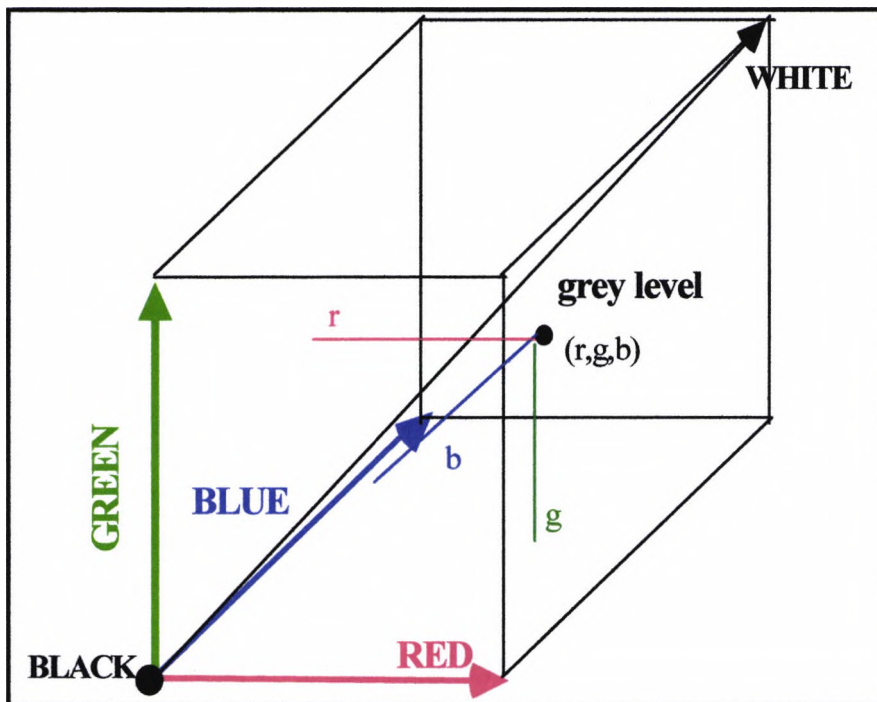
Figure 2.3 RGB colour space and quantitative representation of colour.

Other colour systems are commonly used in colour image analysis, separating the colour signal into chromatic and achromatic parts, for example:

- ISH (intensity, saturation, hue) which is convenient for human colour perception;
- XYZ is used to represent the chromaticity of the colour;
- YIQ is used to efficiently encode colour information in achromatic and chromatic components for television broadcasting signals;
- Ohta computed $I_1I_2I_3$ [Oht85] to decorrelate the RGB signal using the Karhunen-Loeve transformation;
- Opponent colour (r-g, b-y, w-bl) that is employed in human colour vision.

The transformations from RGB to other colour systems and their uses in machine vision are described in sections9 2.2.1-5.

Some early researches [Rob77, Nev77, ASA81, Oht85, Tom86, BH88, GHS90, GBS90] considered the input data to a colour image segmentation algorithm as a set of random signals. The motivation for using different colour co-ordinates is to reduce the correlation of the data or reduce the dimensionality of the colour space, so that it will improve the performance or reduce the complexity of the segmentation algorithm. Moreover, most of the algorithms are based on a statistical approach and ignore the physical properties of the surfaces. Consequently over-segmented images resulted, due to specularity and other physical effects. Since the dichromatic model [Sha85] was introduced to separate reflection components of dielectric materials, researchers [KSK90, BLL90] have investigated algorithms for identifying the specular parts of objects. The model is described in section 2.2.6.

2.2.1 Human Perceptual Colour (ISH)

The ISH (Intensity, Saturation and Hue) system is commonly used to describe colour.

Intensity   ≈ Luminance or brightness
Saturation   ≈ Purity
Hue   ≈ Dominant wavelength

Figure 2.4 shows how colour is represented in ISH space: the intensity is the vertical distance from B to the colour, saturation is the perpendicular distance from the line BW to the colour, hue is the clockwise angle from the reference colour (red is used in this case). For example, if the length of line BW is 1.0, the maximum saturation is 1.0 and the hue is from 0° to 360°. If the colour green is registered, its intensity is 0.5; saturation is 1; hue is 120°. Ohta [Oht85] transformed the RGB values of a pixel to ISH (the formula is shown in eqn.2.5-7), then image segmentation was applied to the transformed data. Massen & Böttcher [MB89] and Hung et al [HEC91] used the ISH system for interpretation of regions for a manufacturing product inspection task. There are several reasons for

employing the ISH system: it decorrelates the colour data [Oht85, MB89, HEC91] and reduces the dimensionality of the colour space by ignoring one the colour attributes and providing an object interpretation [Oht85] which is related to human perception.
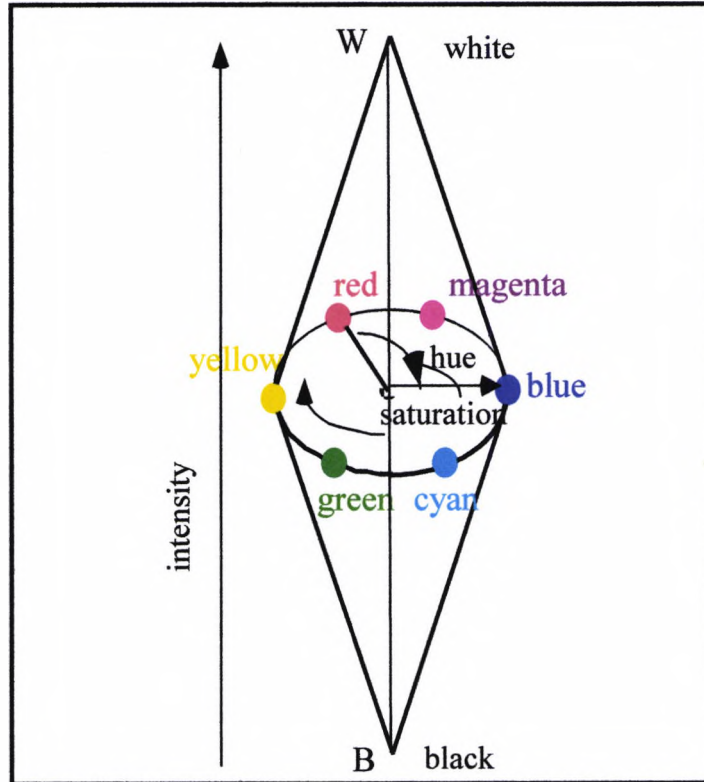
Figure 2.4 (I)Intensity, (S)Saturation, (H)Hue colour Space.

$$I = r + g + b \cdots\cdots (2.5)$$

$$r_n = \frac{r}{I}, g_n = \frac{g}{I}, b_n = \frac{b}{I}$$

$$S = 1 - 3 * \min(r_n, g_n, b_n) \cdots\cdots (2.6)$$

$$H = \arctan(\frac{\sqrt{3} * (g - b)}{2 * r - g - b}) \cdots\cdots (2.7)$$

2.2.2 Ohta $I_1I_2I_3$

Ohta [Oht85] attempted to derive a set of colour features that provide a large discriminating power for histogram thresholding of each colour attribute, using a technique based on the Karhunen-Loeve transformation. Interestingly the effective set of colour features found in each case is approximately the same. The transformations from RGB to Ohta's colour features are shown in eqn. 2.8-11.

$$I_1 = \frac{(R+G+B)}{3}\cdots\cdots(2.8)$$

$$I_2 = (R-B)\cdots\cdots(2.9)$$

$$I_3 = \frac{2G-R-B}{2}\cdots\cdots(2.10)$$

Combining the above three equations and describing them in matrix form

$$\begin{bmatrix} I_1 \\ I_2 \\ I_3 \end{bmatrix} = \begin{bmatrix} 0.33 & 0.33 & 0.33 \\ 1.0 & -1.0 & 0.0 \\ -0.5 & 1 & -0.5 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}\cdots\cdots(2.11)$$

This set of colour features is not necessarily applicable to all images because only a limited number of images were used in his experiments. However, this set has the identical formulation for the opponent colours which are employed in the human visual system (see eqn. 2.19-21).

2.2.3 YIQ Colour

The YIQ colour system is a standard for Television broadcasting; Y is considered as the achromatic part of the signal, Q and I are used to describe the chromatic parts. The transformation is shown in eqn. 2.12. It is similar to other colour systems that separate the colour signal into achromatic and

chromatic parts; the difference is that only a simple linear transformation is required.

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.500 & -0.230 & -0.270 \\ 0.202 & -0.500 & 0.298 \end{bmatrix} \begin{bmatrix} r \\ g \\ b \end{bmatrix} \cdots\cdots(2.12)$$

2.2.4 xyz Chromaticity Colour

This is a system that is used to evaluate the quality of colour [Mac81] (chromaticity). The X, Y and Z values represent the amounts of the standard primaries (established by CIE, Commission Internationale de l'Échairage in 1931) needed to match a given colour. Instead of using X, Y and Z for colour specification, more easily understood information is given if the colour is specified in terms of x and y. Therefore it is used to reduce the dimensionality of the colour space and hence the complexity of colour segmentation algorithm is reduced [ASA81]. Nevatia's [Nev77] colour edge detector was used to separate the colour signal into chromatic and luminance parts. The chromatic part is represented by x and y, and luminance part is the summation of X, Y and Z.

$$x = \frac{X}{X + Y + Z} \cdots\cdots(2.13)$$

$$y = \frac{Y}{X + Y + Z} \cdots\cdots(2.14)$$

$$z = \frac{Z}{X + Y + Z} \cdots\cdots(2.15)$$

Also, $x + y + z = 1 \cdots\cdots(2.16)$

2.2.5 Normalised Colour

Healey [Hea92] proposed a colour system to overcome the problem of over-segmented images due to the physical effects, such as the surface shape and viewing geometry. He considered that the variation of pixel colours from a homogenous surface is only in their achromatic attributes. Therefore, using a colour system that factors out the achromatic attribute helps reduce the occurrence of over-segmented regions. It is similar to the xyz chromaticity system; the difference is that this system is more relaxed in the number of colour attributes and the achromatic part is based on the Euclidean distance of S in eqn. 2.17. The formula can be seen in eqn. 2.17-18.

$$S = (S_1, S_2, S_3, \cdots\cdots S_n) \cdots\cdots (2.17)$$

where S_i are the measured sensor values for each pixel.

The normalised colour $\hat{S}$ is

$$\hat{S} = \frac{S_i}{\sqrt{\sum_{i=1}^{n} s_i^2}} = (\hat{S}_1, \hat{S}_2, \hat{S}_3, \cdots\cdots \hat{S}_n) \cdots\cdots (2.18)$$

2.2.6 Opponent Colour

Hering [Her64] introduced the idea of opponent colour, suggesting that redness and greenness, or yellowness and blueness are never simultaneously evident in any colour but appear to be mutually exclusive. The opponent combination consists of three components named red-green, blue-yellow and white-black. The two different sets of formulae for the transformation from RGB to opponent colour are shown in eqn. 2.19-21 and eqn. 2.22-24. Eqn. 2.19-20 are used in Swain's colour indexing [Swa90], ignoring the white-black component, so that the matching parameters are reduced; the remaining colour features are considered to be invariant to changes of illumination intensity. Eqn. 2.19-21 are only calculating the difference between R and G. The correct formulation for a

colour feature invariant to the changes of illuminant intensity should be the ratio of R and G, and is explained by Finlayson [Fin92].

$$rg = R - G \cdots\cdots (2.19)$$

$$yb = \frac{2B - R - G}{2} \cdots\cdots (2.20)$$

$$bw = \frac{R + G + B}{3} \cdots\cdots (2.21)$$
$$rg = \log(R) - \log(G) \cdots\cdots (2.22)$$

$$yb = \log(R) - \log(B) \cdots\cdots (2.23)$$

$$wbl = \alpha \log(R) + \beta \log(G) + \gamma \log(B) \cdots\cdots (2.24)$$

2.2.7 Finite Dimensional Model

The finite dimensional model is commonly used to represent the surface reflection function; it approximates the surface reflection function more accurately. The advantage is that it can simply express the surface reflection functions in terms of the coefficients of the basis functions and then linear algebra can be used for the analysis. Most researches are trying to find a sufficient number of basis functions to accurately model the common surface reflection function and illumination function (see section 5.2 in chapter 5).

2.2.8 Dichromatic Reflection Model

This is a simple mathematical model of reflection [Sha85] that describes the reflection as two separate components, one is the interface (specular) component and the other is the body (diffuse) component Using this model, it is possible to establish the relation between the two components of each pixel [KSK90]. Eqn. 2.25 and figure 2.5 illustrate the model with an example of dielectric material.
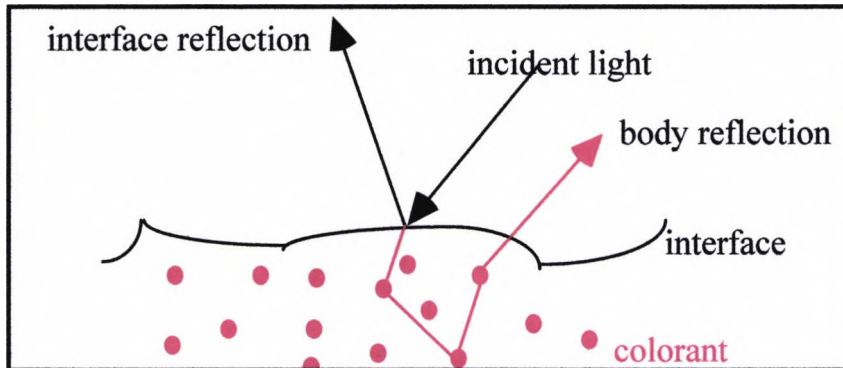
Figure 2.5 Reflected light from a dielectric material

The dichromatic reflection model states :

$$L(\lambda, i, e, g) = L_f(\lambda, i, e, g) + L_b(\lambda, i, e, g) \cdots\cdots (2.25)$$
$$= m_f(i, e, g) c_f(\lambda) + m_b(i, e, g) c_b(\lambda) \cdots\cdots (2.26)$$

where L is the total radiance of the reflected light,

L_f and L_b are the radiance of the light reflected at the interface,

and at the surface body respectively,

i and e are the incident angle of light and the angle of the reflected light to

the surface normal, and g is the angle between the incident and reflected ligh

2.3 Colour Information for Object Recognition

Colour is a powerful cue for the recognition of colour objects in a scene, since the colour is more stable against rotation and size of the objects [Swa90, BE92] than geometrical information. Colour information also helps when objects have no definite shape or are occluded. If the colour database is built for recognition, what essential data should be included? Reviews of the literature [Oht85, Ber87, MB89, Swa90, HEC91, MMK93, BD97, DL98, WE98, SHS99] on colour recognition, object colour and colour surface inspection, in general, show that colour information used in their systems consists of two essential features: the first is the representation or description of the colour; the second is an operational or algorithmic feature. The first feature refers to the colour description of a surface in the database or knowledge base. The second feature is used to

evaluate the difference between colour descriptions of extracted regions from the image and the colour in the database. In figure 2.6, the contents of $S_1, S_2, S_3, C_1, C_2, C_3, \cdots C_n$ are the descriptive features of the colour information and the algorithmic feature is used to determine which arrow $A_1, A_2, A_3, \ldots or\, A_3$ is appropriate to map $S_i \rightarrow C_j$ (e.g. for $S_2 \xrightarrow{A_1} C_1$ the arrow indicates the match between S_2 and C_1). Although the contents of the two features vary from one researcher to another, they are essential in a colour recognition and inspection system.
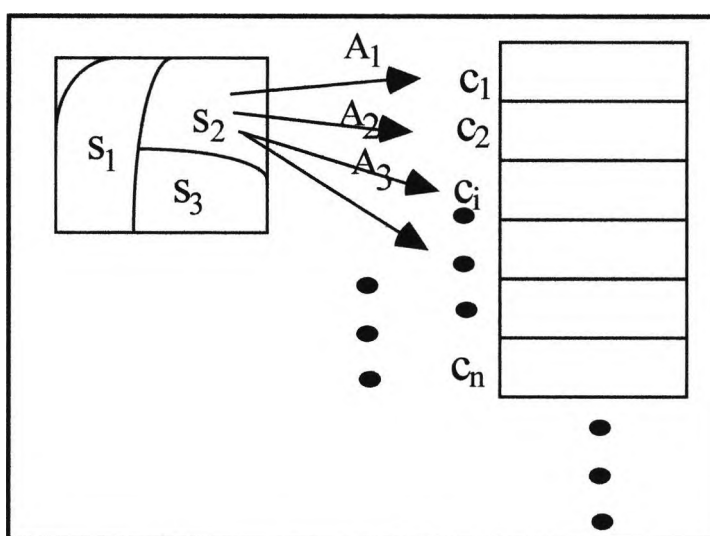
Figure 2.6 Illustrate the relation between image colours and a colour database.

2.3.1 Ohta's Knowledge-Based System

The early work of colour object recognition in machine vision systems can be found in Ohta [Oht85]. In his knowledge-based system, the colour information of objects is characterised by different colour features. The number of features for each object is varied and depends on the type of objects. The objects are mainly found in the outdoor scenes, depicting roads, trees, and buildings. First of all, the colour of a surface is categorised into nine different properties which are dark, bright, shining, grey, vivid, red, green, blue and yellow. For the definition of the colour property, he set two or four thresholds on different attributes such as the intensity, saturation, hue, red, green, blue or contour contrast. In order to evaluate the

similarity between a region extracted from the image to one of the nine properties, a fuzzy truth-value is computed with respect to each colour property. Three types of fuzzy truth functions are used according to the type of colour property. The three functions are shown in figure 2.7a-c. It is not straightforward to define which are the descriptive parts, but if we consider the quantitative colour information of nine colour properties rather than the qualitative description of the surface, then the descriptive part is the colour attribute. The algorithmic part is the type of fuzzy truth functions.
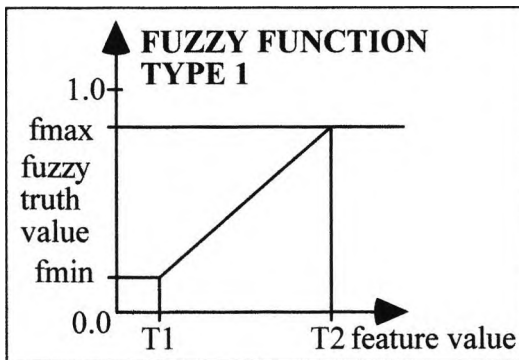
Figure 2.7a Fuzzy truth function type 1.
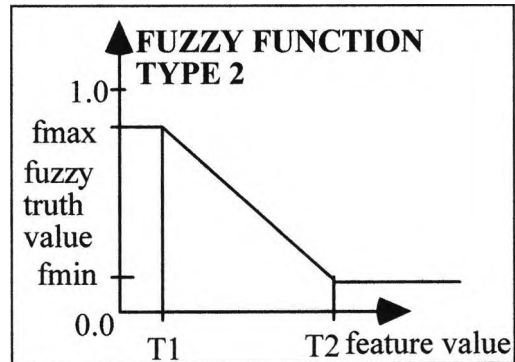
Figure 2.7b Fuzzy truth function type 2.
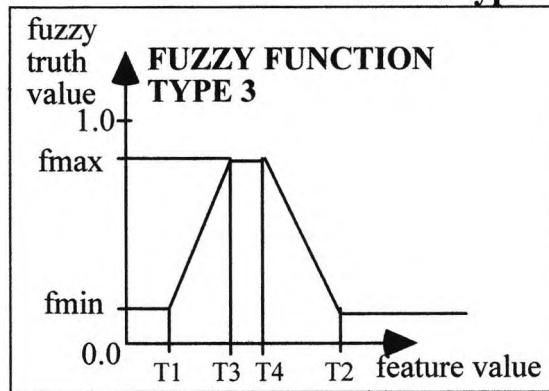
Figure 2.7c Fuzzy truth function 3

The knowledge-base system has been tested on several outdoor scenes and the results have shown that the objects have been successfully identified. There is no detail of how to determine the thresholds and choices of the fuzzy functions in the knowledge-based. Such information is important, because the evaluation of the colour properties of the objects is based on

these parameters and an understanding of how the object model can be interpreted by these functions.

2.3.2 Colour Recognition and Inspection Systems

Berry [Ber87] used colour information extracted from images for recognising objects that have the same shape. He examined which colour feature is the most stable under changes of illumination intensity and a measure based on the fluctuation of the colour features under different illuminations. The most stable colour features are chosen for the representation of the colour surface in the database. A heuristic minimum distance is used as the criterion for matching between colour features extracted from images and the defined colour in the database. The descriptive feature is the most stable colour feature, which is found in the initial experiments. The algorithmic feature is a distance metric between the colour extracted from the each region to the colours in the database.

Massen and Böttcher [MB89] implemented a vision system for real-time inspection of products. They modelled the colours of surfaces by showing typical colour samples to the system, building a lookup table based classifier, after finding the convex hull of each colour set in colour space. The lookup table classifier is a many-to-one mapping, which evaluates the validity of the colour of a pixel to a particular colour surface (see eqn. 2.27). The descriptive features used are intensity, saturation and hue. The algorithmic feature is the lookup-table-based classifier, $F(P_j)$ which is described in eqn.2.28. They evaluate the validity of a pixel to a surface colour, for those pixel colours that intersect with the colour set **A**. The beauty of this approach is that its simplicity and effectiveness can achieve real-time processing of colour information between the image and the database.

$p_j = (i_j, s_j, h_j)$, is the colour of a pixel p_j.

where i is the intensity,

   s is the saturation,

   h is the hue.

$A = \{c_1, c_2, c_3 \ldots\ldots, c_n\}$ is a set of sample colours

$$F(p_j) = \begin{cases} 1 & \text{for } p_j \in A \\ 0 & \text{for } p_j \notin A \end{cases} \quad \cdots\cdots(2.27)$$

Hung, Ellis and Chamberlain [HEC91] implemented a vision system for brick inspection. Different inspection strategies are employed dependent on the types of brick. The brick types are categorised into textured, multi-coloured and single coloured. For the inspection of multi-coloured types, there is a requirement for the segmentation algorithm to divide the image of a single brick into different colour regions, using a segmentation algorithm based on clustering. For the inspection of single coloured types, the colours of reference bricks are represented by their mean and variance. The colour of a sample brick will be tested against reference bricks based on the Mahalanobis distance criteria. Since the ISH colour system is used, as the representation of the brick's colours and the data is not correlated, the covariance matrix is a diagonal and the calculation of Mahalanobis distance is simplified. All equations are shown below. The descriptive features are the ISH colour attributes. The algorithmic features consist of the means and variances of the colour surface for calculating the Mahalanobis distance. The implementation can be done in a similar way to that of Massen and Böttcher [MB89]. Instead of calculating the distance for each brick's colour against the colour of the reference brick, a lookup table is used with the colour of sample bricks as the index, and the corresponding distances are the outputs.

$$r_{j,k} = \left(i_{j,k}, s_{j,k}, h_{j,k} \right)$$

r_j is pixel j of reference brick colour k.

The average $\tilde{r}_k$ of reference brick type k,

$$\tilde{r}_k = \sum_{j=1}^{n} r_{j,k} = \left(\sum_{j=1}^{n} i_{j,k}, \sum_{j=1}^{n} s_{j,k}, \sum_{j=1}^{n} h_{j,k} \right) \quad \cdots\cdots(2.28)$$

The variance σ_k of the reference brick colour k is:

$$\sigma_k = \frac{\sum_{j=1}^{n}(r_{j,k}-\widetilde{r}_k)^2}{n} \quad \cdots\cdots(2.29)$$

$$t_{j,l} = \left(i_{j,l}, s_{j,l}, h_{j,l}\right),$$

$t_{j,l}$ is pixel j of the test brick l.

The average $\widetilde{r}_k$ of test brick type l,

$$\widetilde{t}_l = \sum_{j=1}^{n}t_{j,k} = \left(\sum_{j=1}^{n}i_{j,l}, \sum_{j=1}^{n}s_{j,l}, \sum_{j=1}^{n}h_{j,l}\right) \quad \cdots\cdots(2.30)$$

The matching criteria is based on Mahalanobis Distance, M_k for k type of brick.

$$M_k = \sqrt{\frac{(\widetilde{r}_k-\widetilde{t}_l)^2}{\sigma_k}} \quad \cdots\cdots(2.31)$$

2.3.3 Localisation of Objects using Colour

Matas, Marik and Kittler [MMK92] argued that the segmentation of colour images and the localisation of colour surfaces can be achieved by generating and verifying the hypothesis of the relationship between a pixel to the target surfaces that have been selected from the database. The hypotheses are based on the probability functions of pixel colour and the colour of the target surface in the database. Both representations of colours in the database and the calculation of the probability function are shown below. The probability function is an *a priori* probability model based on Bayes formula. The definition of a colour surface is based on the dichromatic model [Sha85] and it is represented as a line in the three-dimensional colour (RGB) space. The database's colours are modelled with images of objects. The calculation is more complicated than the previous examples. The descriptive part is the line description in RGB colour space and the algorithmic part is the probability function.

L_{bc}^i is the body colour of object i in the database,

RGB is the pixel value of a corresponding surface patches.

$$P(RGB|M_{bc}^i) = f_n(\|L_{bc}^i, RGB\|) \cdots\cdots (2.32)$$

where $\|\cdot\|$ denotes Euclidean distance of the RGB triplet

of a pixel from line L_{bc}^i and f_n represents properties of noise.

$$p(M_{bc}^i|RGB) = \frac{p(RGB|M_{bc}^i)P(M_{bc}^i)}{\sum_j p(RGB|M_{bc}^j)P(M_{bc}^j) + P(RGB|bg)*P(bg)} \cdots\cdots (2.33)$$

where $P(bg)$ refers to the aprior probability that a pixel does not

correspond to any model and $P(RGB|bg)$ denotes the colour distribution

of non - model (background) pixels.

Swain [Swa90] proposed a technique called colour indexing for locating and identifying multi-coloured objects. The colour histogram of a multi-coloured object is used as an index to the database. The identification of objects is based on the intersection between object histograms in the database and the image. The localisation of the target object in the images is based on a technique called histogram backprojection. Brock-Gunn and Ellis [BE92] exploit this method in tracking coloured objects. Objects are segmented by movement and this is identified on a frame-to-frame basis using a normalised histogram correlation. Funt and Finlayson [FF95] modified the indexes and overcame the problem of changing illumination. The descriptive features are different in the three researches: Swain has used the opponent and RGB colour system, whilst Brock-Gunn and Ellis were using opponent colour attributes. Funt and Finlayson used a new illuminant invariant colour index. The algorithmic features are the intersection function between the histogram of the image colours and the colour histogram of each object (colour template). The equations are shown below, which describe the intersection of histograms.

$H(M_i)$ is the histogram of model object i.

$H(O_j)$ is the histogram of object j in the image.

$I(M_i,o_j)$ is the function to determine the fittness between two histograms. It is intersection between the two histograms.

$$I(M_i,o_j) = M_i \cap o_j \cdots\cdots(2.34)$$

2.4 Memory colour

In this section, I introduce the term "memory colour" which will be used in other chapters. Memory colour is defined as colour information that consists of the two essential and generic features: descriptive and algorithmic features.

The first essential feature is the *descriptive* feature: the colour attributes of the colour (surface). The second essential feature is the *algorithmic* feature that is used to evaluate the similarity between colour from the image and in the database. From the previous sections (2.3.1-3), the examples have been described.

Figure 2.8 A colour database.

The colour database (see figure 2.8) can be considered as an independent entity in a vision system. The two databases (colour and geometrical) are required to have links between each other. Any colour extracted from the image can be categorised as a particular memory colour, rather than categorised to any particular object. It can avoid the duplication of information stored in the database (with geometrical and colour information) for each object. There are situations where the separation of the information (colour and geometric) is appropriate for the overall vision system: if colour was used as the first cue for object localisation, the only information required is the memory colours. Likewise, only geometric information of the objects is required if the cue is based on geometric information; if the processes of recognising the two cues are running simultaneously in a parallel processing machine, the two pieces of information can be stored in different processors; since colour is one of the generic features of the object, it is not necessary for it to be associated with the geometric feature of the object. For example, if object 1 consists of colours A, B, C and E and object 2 consists of colours B, C, F and G, then colours A, B, C, E, F and G are considered as the generic features and are stored as separated entities which are not associated with any object. Similarly, if several objects have the same colours but different geometric structure, only one set of colour values is kept in the database. In both cases, the duplication of data can be avoided. For any new object with colours that are not present in the colour database, such colours can be inserted to the database directly. The next section describes a data structure that is used to accommodate all the necessary memory colour information.

2.5 Colour Information Lookup Table (CILUT)

The Colour Information Lookup Table (CILUT) is a data structure that accommodates all the necessary information of memory colour and provides direct access to the necessary data. The index to the CILUT entry is the value of the pixel or region colour that can be represented in any colour system, for example RGB, ISH, XYZ, YIQ, etc. The object and surface in the database (the surface s is the name of the memory colour and object o the name of the object, both are included in the CILUT and can be

seen in figure 2.9) that corresponds to the description of the image colour can be found in the CILUT. The value d is the fitness measure between the image colour and its corresponding memory colour. The graphical illustration of a CILUT is shown in figure 2.9. The CILUT can be pre-computed off-line and the information within the CILUT can be changed according to the environment in which objects that the vision system will likely to perceive. Furthermore, because the CILUT can be implemented in EPROM[1], it can provide fast access to the required information.
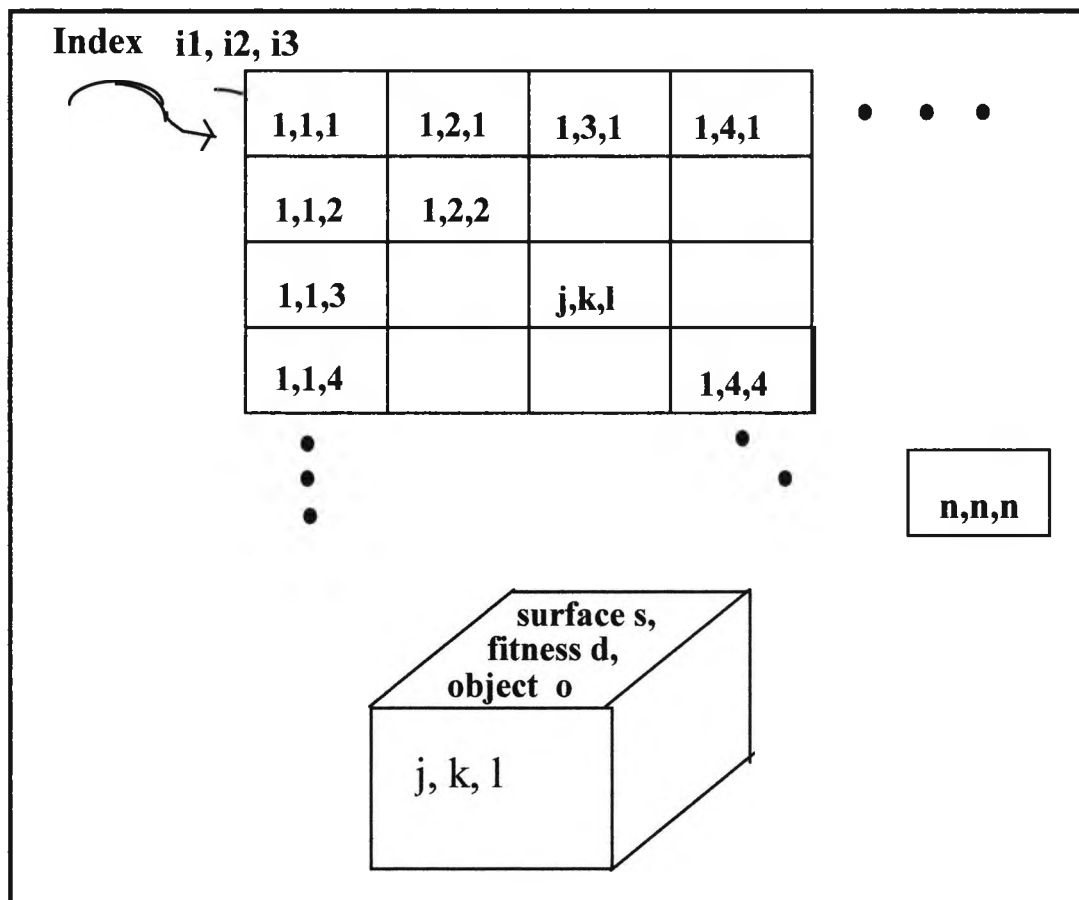
Figure 2.9 Colour Information Lookup Table (CILUT)

---

[1] Erasable Programmable Read Only Memories.

A colour histogram can be used as a lookup table after the information is accumulated. The difference between the two is that the output of a CILUT consists of a few values and the histogram only contains (by definition) the probability density. Also, the use of colour histograms [Swa90, BE92] as a template for matching can be replaced by a CILUT, which can accommodate several sets of object colours in the same table.

2.6 Summary

Colour information exploited at both low and high level in image processing is described in this chapter. The early work on the use of colour in low level vision concentrated on finding a set of colour features that are suitable for segmentation. With growing research into the use of colour object recognition, it is necessary to address the issue of colour information and its processing. From the review of previous work on colour object recognition, a generic but sufficient colour information (memory colour) is proposed as an element in a colour database and its corresponding data structure (CILUT) has been implemented for the work described in chapter 4 and 5.

Chapter 3
Use of Colour in Edge Detection

Abstract:

This chapter describes an investigation into the use of colour to improve the quality of edge detection. It is divided into two parts: firstly, to determine a quantitative and objective evaluation of the difference in detection performance between colour and grey level edge detectors. No significant difference is found, from experimental results, in the gradient magnitudes between the colour and grey level images. However it is observed, at genuine edges, that the three colour planes are correlated but only weakly so at noise edges. In the second part a corroboration function, based on consistency of the gradient orientations in the three colour planes, is introduced for selecting edgels and is derived from the physical properties of surfaces. The use of the corroboration function in conjunction with a colour Canny edge detector, forms an algorithm quite robust in operation in rejecting unwanted (noise) edges.

3.1 Introduction

Edge information had been used in many aspects of machine vision, such as object recognition, tracking, stereo correspondence, and medical image analysis. Edge detection in grey level images has been widely investigated and discussed [Dav75, MH80, Can86, TP86] in the last few decades. The use of colour in edge detection is considered as a natural extension, and the investigation of colour edge detection has normally followed in an evolutionary fashion from intensity edge detectors. For example: Nevatia [Nev77] has extended the Hueckel operator to colour images; Robinson [Rob77] has extended her compass gradient edge detection to colour images; Novak [NS87] has extended the Canny edge detector and Cumani [Cum91] has investigated the application of second-order differential operators to colour images.

These studies focused on whether the additional colour information improves the performance of the detector and also on how to combine the gradient magnitudes from the three colour planes into a single value. Nevatia and Robinson also transformed the RGB signal into other colour co-ordinate systems in which the chromatic and luminance components are separated. Both found that chromatic and luminance edges are correlated at the pixel locations that correspond to the desired boundaries, though with the limited set of images they used, they have concluded both visually and quantitatively that most of the edges can be extracted from the luminance component. Novak, without using any colour transformation, has found that the combined RGB edges appear better than edges from the equivalent intensity images under subjective evaluation. However, objective and statistical measures of performances were not reported.

The remaining problem, after computing the gradient magnitude on each image location (grey level or colour), is how to select the edgels which result from genuine contrast changes and to reject those caused by noise or other physical effects. The common approach is to set a threshold or thresholds based on the gradient magnitude; however the choice of threshold often requires human intervention. Furthermore, when the boundaries are within low contrast regions, determination of an appropriate

threshold can become difficult. Observation of our experimental results found that the existence of edgels on different colour planes, at the same location or close by, are highly correlated at the desired boundaries; this phenomenon can be explained by the physical properties of the surfaces' reflection spectra. However this correlation is much weaker for edges generated by noise. In addition, the gradient orientation can be used to identify genuine edges and help distinguish them from noise.

The aim of this chapter is two fold: one is to provide quantitative evidence on the difference in gradient magnitude between colour and grey level edge detection and to examine whether these differences can be used to improve the robustness of detecting edges; the other is to use the colour information to support the selection of edges, as an alternative to the conventional approach of determining an appropriate threshold of the gradient magnitudes. The selection functions are derived from experimental observation and the physical properties of the colour surfaces; they are compared with the conventional thresholding method based on the signal-to-noise ratio between detected images.

Section 3.2 reviews some of the work in edge detection and the sources of noise causing the degradation of the detection performance. Section 3.3 shows the absolute and relative difference in gradient magnitude between colour and grey level images. The measurement procedure and statistical results are shown in section 3.3. Section 3.4 shows the correlation of edges in three colour planes. Section 3.5 describes the selection algorithm for suppressing unwanted edges and the relation to some physical properties of the surface and is the conclusion of this chapter.
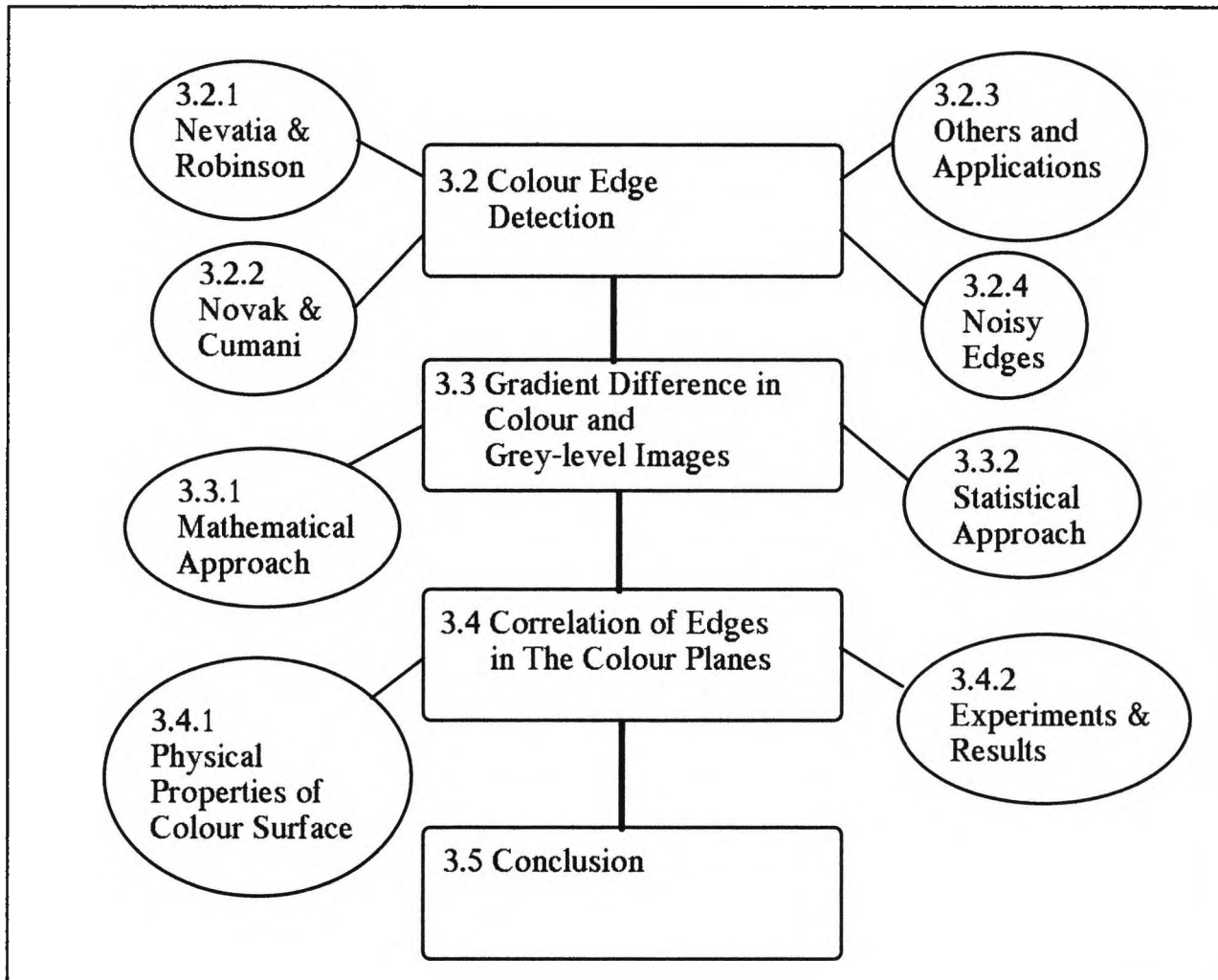
Figure 3.0 The structure for chapter 3.

3.2 Colour Edge Detection

Low level image processing is an important part of the overall vision problem. Higher level algorithms dealing with relational and cognitive aspects rely strongly on the efficiency, robustness and precision of low level operators such as edge detection [MH80, Can86]. Canny [Can86] summarised several criteria for designing an edge detection system: firstly a high signal-to-noise ratio, achieving low probabilities of false alarms and failures to mark real edges; secondly good localisation - which means the position of the marked edgels should be as close as possible to the real boundaries; thirdly a single response is required to indicate a step edge. Most of the current edge detection systems are divided into three steps to

satisfy these criteria. The first step is to regularize [TP86] the image by convolving it with an appropriate regularizing function; for example the Gaussian function which reduces the noise level and ensures that the image is differentiable. The second step computes the derivatives of the image in order to find the local intensity gradient at each pixel location. The final step is to detect a single response to a step edge by computing the zero-crossings in the second derivative [MH80, Cum91] or maxima detection in the first derivative [Can86].

Colour edge detection methods [NS87, Cum91] have typically performed the edge detection described above on each colour plane separately and then combined the separate edge images for each colour channel following computation of the derivatives. However this tends to destroy the edge characterising information, which causes ambiguous colour edges to be identified.

3.2.1 Nevatia & Robinson

Nevatia [Nev77] & Robinson [Rob77] investigated the use of colour. There are several similarities between them: both extend the existing edge detectors to colour; both use different colour systems and investigate which colour attributes are useful for detecting visually significant edges; both propose some definition of colour edges; finally they conclude that edges detected in different colour components are highly correlated.

Nevatia [Nev77] extended the Hueckel edge operator [Dav75] to colour. The detector was applied to luminance and chromatic channels of the images. The result indicates the edges detected in both channels are highly correlated. The paper proposed three different methods for finding a colour edge:-

1. Define a metric distance in the chosen colour space and use discontinuities in this distance to compute the presence of colour edges. The metric could be based on experiments with human visual perception.

2. Compute the edges in three chosen colour components to be largely separate. The edges in the three components may then be merged by some specified procedures.

3. Allow the edges in the three chosen colour components to be largely independent, but impose some uniformity constraints for the convenience of utilising all three edges concurrently

The first definition has a colour metric [HB88] considering a "colour distance" function that is derived as an estimate of the physical distance between normalised spectral power distributions. The following section describes two approaches based on the second definition. The first two definitions of colour edges do not retain the characteristics of the three edgels detected in different colour components. The final definition is the one chosen for the experiments in this chapter.

Robinson [Rob77] developed a compass gradient edge detection method and extended it to incorporate the colour information. Several colour features were evaluated for their suitability for edge detection. The colour co-ordinates under evaluation were (Y, I, Q) for NTSC, CIE (L, a, b), etc. Her view was "a better approach for colour edge detection would be to generalise the definition of an edge in the three-dimensional colour space and to develop colour edge extraction methods accordingly". The edge activity index (EAI) is used to measure the occurrences of edges in different colour planes. The comparison of the suitability of different colour components is based on EAI. She concluded that brightness seems to the best attribute to be employed in edge detection. Some results indicate that there is inherent cross-correlation of the edges from different colour components.

Those two observations reinforce the method for selecting edges based on the correlation of edgel detected in three colour planes, which is introduced in section 3.4.

3.2.2 Novak & Cumani

Consider a colour image as a function $F(x,y) = (f_r(x,y), f_g(x,y), f_b(x,y))$. There are two theoretical approaches: Novak [NS87] extends the Canny edge detection operator; first he uses the Jacobian matrix to describe the derivatives and then applies Canny's method [Can86] for selecting edgels. Cumani's [Cum91] method is based on the zero-crossings of a second-order differential operator. This formulation indicates that the extensions of a second-order differential operator to colour images will suffer ambiguity in the sign of the gradient and hence the closure of zero-crossing lines is not guaranteed. However both the computations of the magnitude of edges are the same and are shown in this section. Also it provides the mathematical background for section 3.3.

Novak [NS87] described the x and y derivatives in the three colour planes by a Jacobian matrix J and defined the gradient magnitude as the square root of the eigenvalue of $J^T J$.

$F(x,y)$ is the function of a colour image and

$$F(x,y) = [f_r(x,y), f_g(x,y), f_b(x,y)] \cdots\cdots (3.1)$$

Since F is a function (x,y) of the image,

and can be described by its variation at any point by

$$\Delta C = J\Delta(x,y)$$

where J is the Jacobian matrix containing the dervatives of each colour plane.

$$J = \begin{bmatrix} R_x & R_y \\ G_x & G_y \\ B_x & B_y \end{bmatrix} = \begin{bmatrix} C_x & C_y \end{bmatrix} \cdots\cdots (3.2)$$

where $C_x = \begin{bmatrix} R_x \\ G_x \\ B_x \end{bmatrix}, C_x = \begin{bmatrix} R_x \\ G_x \\ B_x \end{bmatrix},$

$R_x = \dfrac{\partial f_r}{\partial x}, G_x = \dfrac{\partial f_g}{\partial x}, B_x = \dfrac{\partial f_b}{\partial x}, R_y = \dfrac{\partial f_r}{\partial y}, G_y = \dfrac{\partial f_g}{\partial y} \text{ and } B_y = \dfrac{\partial f_b}{\partial y},$

$$J^T J = \begin{bmatrix} C_x C_x & C_x C_y \\ C_y C_x & C_y C_y \end{bmatrix} \cdots\cdots (3.3)$$

Also, three pragmatic approaches on computing the magnitude of colour edges were used: the first one is simply combining the derivatives in a Euclidean sense; the second one is the sum of absolute derivatives and the final one is to choose the absolute maximum magnitude among the derivatives. They reported that the final measure seems to be the best, but the detail of the comparison is not shown.

Cumani [Cum91] defined the gradient magnitude as the directional maximum of a local contrast function S. The function was derived by applying the variational method.

The function S(P,**n**) is defined as the squared local contrast at point P in direction **n**, where **n** is defined as a unit vector and indicates the gradient direction in the image plane (x_1, x_2) and $\mathbf{n} = (n_1, n_2) = (\cos\theta, \sin\theta)$.

$$S(P,\mathbf{n}) = \sum_{i=1}^{2}\sum_{k=1}^{2} \gamma_{i,k} n_i n_k = E n_1^2 + 2F n_1 n_2 + G n_2^2 \cdots\cdots (3.4)$$

where $\gamma_{i,k} = \sum_{j=1}^{m} \dfrac{\partial f_j}{\partial x_i} \dfrac{\partial f_j}{\partial x_k}$, $E = \gamma_{1,1}, F = \gamma_{1,2}$ and $G = \gamma_{2,2}$

f_j is the jth band of the image function.

m is number of band, $m = 3$ for RGB images.

Eqn. (3.4) is in a quadratic form, from which by varying **n**, a minimum and maximum can be found. The gradient magnitude is defined as the directional maximum of the contrast function. Since Novak inserts the derivatives into a Jacobian matrix J, and Cumani describes the derivatives in a quadratic form, both maxima are the same and can be calculated by finding the eigenvalue of $J^T J$ or from eqn. (3.3), and the direction of the maximum is the corresponding eigenvector. The eigenvalue λ_c is the square of the gradient magnitude and is given by:

$$\lambda_c = \frac{((E+G) + \sqrt{(E-G)^2 + 4F^2})}{2} \cdots (3.5)$$

Eqn. (3.5) can be used to compute the edge magnitude in monochromatic images by setting *m=1* [Cum91]. The expression is identical to the definition used in Marr & Hildreth [MH80] to define the grey level gradient. λ_l is the square of the gradient magnitude:

$$\lambda_l = \left(\frac{\partial f_i}{\partial x}\right)^2 + \left(\frac{\partial f_i}{\partial y}\right)^2 \cdots (3.6)$$

where f_i is a function of the greylevel image

3.2.3 Others and Applications

On the implementation and performance evaluation of colour edge detectors, Koschan [Kos95] had examined several detectors: Sobel operator, Laplace operator, the Mexican hat operator, Cumani and Alshatti-Lambert operator. The conclusion is that Cumani detectors yield the output with the best quality. Malowany [MM89] had considered the VLSI implementation of two algorithms and compared their performance on synthetic colour images with noise added and the equivalent grey level images. They found that, in image regions of low contrast and high noise content, the use of colour in both detectors gave a better performance. However, there were no details of any objective or quantitative measurements that had been used for the comparisons.

On the applications side there are examples [FF95, JB91, CGG91, GGS98] that exploit the additional characteristics of colour edges. Finlayson and Funt [FF95] use three gradient magnitudes of a colour edge as the indexes for matching colour objects. The advantage of such an approach is that the index is invariant to the changes of illumination. Jordan and Bovik [JB91] exploit the chromatic information to resolve the ambiguity in edge-based stereo-correspondence. Cumani, Grattoni and Guiducci [CGG91] use the colour edge as one of the components in encoded colour images, because

the characteristic of colour edge provides a more complete description of the images.

3.2.4 Noisy Edges

The remaining problem in edge detection is to select the correct edgels and reject those caused by spurious noise or other physical effects. Signal noise may be attributed to three principal sources: the sensor, digitisation and scene characteristics. Limitations of colour and monochrome CCD-cameras [KSK90] can cause errors in edge detection (in general, for all segmentation techniques) as a result of clipping, blooming and chromatic aberration. Quantisation errors that occur in the analogue to digital conversion reduce the precision of a pixel value. Scene characteristics related to texture and roughness of an object's surface, the mixture of body and surface reflection and inconsistent illumination can also produce inhomogeneities in surface colour. These are the major degrading factors which can produce spurious (or incorrect) edgels.

3.3. Gradient Difference in Colour and Grey Level Images

In colour images, each image point is typically represented by a vector consisting of three primary colour planes, namely red, green and blue. There are two principal attributes of an edgel: its (gradient) magnitude and orientation. The gradient magnitude indicates the significance of the local intensity changes in the image and the gradient orientation indicates the direction of these changes. If a gradient operator is applied to each component $F(x,y) = (f_r(x,y), f_g(x,y), f_b(x,y))$, there will be three pairs of gradient magnitudes and orientations for each pixel in the image, whilst the grey level image generates only a single pair per pixel. In order to examine the difference in gradient magnitude between colour and grey level images, a method is required to combine the three pairs of gradient magnitudes and orientation. This combination is based on the computation of the eigenvalue of eqn. (3.5).

3.3.1 Mathematical Approach

The difference between gradient magnitude in colour images and the equivalent grey level images can be expressed in eqn. (3.9). Also, I have examined the sign of the difference in two situations where the orientations of the gradient of the edgels are different and equal in all colour planes.

The magnitude of gradient in grey level image is expressed below. Consider the grey level version of image as shown in eqn. (3.7) where $f_1(x,y), f_2(x,y)$ and $f_3(x,y)$ correspond to the image components from red, green and blue respectively. The grey level image function is simply the summation of the three colour components.

$$F_l(x,y) = f_1(x,y) + f_2(x,y) + f_3(x,y) \cdots (3.7)$$

Differentiate $F_l(x,y)$ respect to x and y,

$$\frac{\partial F_l(x,y)}{\partial x} = \frac{\partial f_1(x,y)}{\partial x} + \frac{\partial f_2(x,y)}{\partial x} + \frac{\partial f_3(x,y)}{\partial x}$$

$$\frac{\partial F_l(x,y)}{\partial y} = \frac{\partial f_1(x,y)}{\partial y} + \frac{\partial f_2(x,y)}{\partial y} + \frac{\partial f_3(x,y)}{\partial y}$$

Using the eqn (3.6) for calculating the magnitude of the edge in a grey level image.

$$\therefore \lambda_1 = \left[\frac{\partial F_l(x,y)}{\partial x}\right]^2 + \left[\frac{\partial F_l(x,y)}{\partial y}\right]^2 \cdots (3.8)$$

$$= \sum_{i=1}^{3}\left[\frac{\partial f_i(x,y)}{\partial x}\right]^2 + \sum_{i=1}^{3}\left[\frac{\partial f_i(x,y)}{\partial y}\right]^2$$

$$2 \bullet \begin{cases} \left(\frac{\partial f_1(x,y)}{\partial y} \bullet \frac{\partial f_2(x,y)}{\partial y} + \frac{\partial f_2(x,y)}{\partial y} \bullet \frac{\partial f_3(x,y)}{\partial y} + \frac{\partial f_3(x,y)}{\partial y} \bullet \frac{\partial f_1(x,y)}{\partial y}\right) + \\ \left(\frac{\partial f_1(x,y)}{\partial y} \bullet \frac{\partial f_2(x,y)}{\partial y} + \frac{\partial f_2(x,y)}{\partial y} \bullet \frac{\partial f_3(x,y)}{\partial y} + \frac{\partial f_3(x,y)}{\partial y} \bullet \frac{\partial f_1(x,y)}{\partial y}\right) \end{cases}$$

$$= E + G + 2H \cdots (3.9)$$

$$\text{where } E = \sum_{i=1}^{3}\left[\frac{\partial f_i(x,y)}{\partial x}\right]^2, G = \sum_{i=1}^{3}\left[\frac{\partial f_i(x,y)}{\partial y}\right]^2$$

$$\text{and } H = \sum_{i=1}^{m}\frac{\partial f_i}{\partial x}\frac{\partial f_{(i+1)\bmod 3}}{\partial x} + \sum_{i=1}^{m}\frac{\partial f_i}{\partial y}\frac{\partial f_{(i+1)\bmod 3}}{\partial y}$$

Recall eqn. (3.5) below. The expression for E and G in both eqn. (3.5) and (3.9) are the same.

$$\lambda_c = ((E+G)+\sqrt{(E-G)^2+4F^2})/2\cdots\cdots(3.5)$$

$$\text{where } E = \sum_{i=1}^{3}\left(\frac{\partial f_i}{\partial x}\right)^2, F = \sum_{i=1}^{3}\left(\frac{\partial f_i}{\partial x}\frac{\partial f_i}{\partial y}\right) \text{ and } G = \sum_{i=1}^{3}\left(\frac{\partial f_i}{\partial y}\right)^2$$

The difference between the two squared gradient magnitudes is shown below. If eqn. (3.10) shows a negative difference, this implies that the gradient magnitude of the grey level edge is higher than that of the equivalent colour edge.

$$\therefore \lambda_c - \lambda_1 = (\sqrt{(E-G)^2+4F^2} - (E+G))/2 - 2H\cdots\cdots(3.10)$$

Using polar forms, we express the derivatives in terms of the orientation of gradient on each individual plane.

$$\frac{\partial f_1}{\partial x} = p\cos(\alpha)\cdots\cdots(3.10), \quad \frac{\partial f_1}{\partial y} = p\sin(\alpha)\cdots\cdots(3.11)$$

$$\frac{\partial f_2}{\partial x} = q\cos(\gamma)\cdots\cdots(3.12), \quad \frac{\partial f_2}{\partial y} = q\sin(\gamma)\cdots\cdots(3.13)$$

$$\frac{\partial f_3}{\partial x} = r\cos(\theta)\cdots\cdots(3.14), \quad \frac{\partial f_3}{\partial y} = r\sin(\theta)\cdots\cdots(3.15)$$

where α, γ and θ are the gradient orientations in r, g and b planes respectively,
p, q and r are the gradient magnitudes in r, g and b planes respectively

$$\delta = \lambda_c - \lambda_l \cdots\cdots(3.16)$$

Substitute polar form derivatives into eqn. (3.10).

$$\delta = \cfrac{\sqrt{\begin{array}{l}(p^2 + q^2 + r^2) + 8(p^2q^2\cos(\alpha)\sin(\gamma)\sin(\alpha - \gamma) + \\ q^2r^2\cos(\gamma)\sin(\theta)\sin(\gamma - \theta) + r^2p^2\cos(\theta)\sin(\gamma)\sin(\theta - \alpha)\end{array}} - (p^2 + q^2 + r^2)}{2}$$

$$\underbrace{\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad}_{\text{Part 1}}$$

$$- 2(pq\cos(\alpha - \gamma) + qr\cos(\gamma - \theta) + rp\cos(\theta - \alpha)\cdots(3.17)$$

If α, γ and θ are the same,

then $\sin(\alpha - \gamma), \sin(\gamma - \theta)$ and $\sin(\theta - \alpha)$ will all be equal to zero.

Since $\cos(\alpha - \gamma), \cos(\gamma - \theta)$ and $\cos(\theta - \alpha)$ are equal to 1,

$$\delta = -2(pq + qr + rp)\cdots\cdots(3.18)$$

If one of the angles is opposite to the others,

consider $\alpha = \gamma \pm 180$ *and* $\gamma = \theta$,

then $\sin(\gamma - \theta), \sin(\gamma - \theta)$ and $\sin(\theta - \alpha)$ will be all equal to zero.

In general there are two situations associated with where the edgel lies. In the first, edge point x is on the boundary between two regions (illustrated in figure 3.1a-d). In the second, edge point y is at the junction of three or more regions (illustrated in figure 3.2a-d).

The first situation can be sub-divided into two cases:

Case 1: if all gradient orientations from colour planes are in exactly the same direction, then the first part of eqn. (3.17) will be equal to zero (shown above) and H will be positive. Since *p, q* and *r* have the same sign, *pq, qr* and *rp* are all positive (see eqn. (3.18)). Hence the difference δ, must be less than zero. This implies that the computed gradient magnitude of the grey level edge is greater than the colour one.

Case 2: if one of the gradient orientations is opposite in direction to the others, then part 1 of eqn. (3.17) will be equal to zero, but the sign of H is not definite and depends on the difference between the two regions. This is because one of the values *p, q* and *r* has different sign (see eqn. (3.18)).
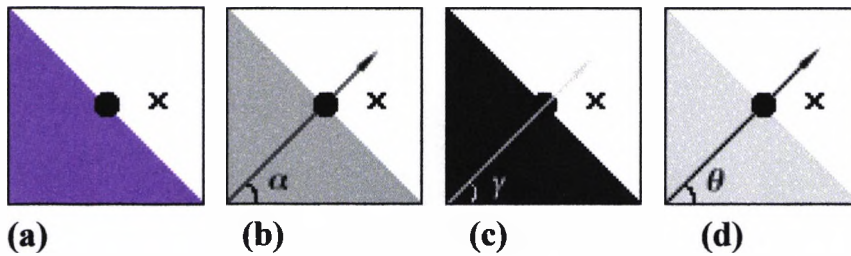
(a) **(b)** **(c)** **(d)**

Figure 3.1 a,b,c and d are showing the boundary of two regions in all colours, red, green, and blue plane respectively.
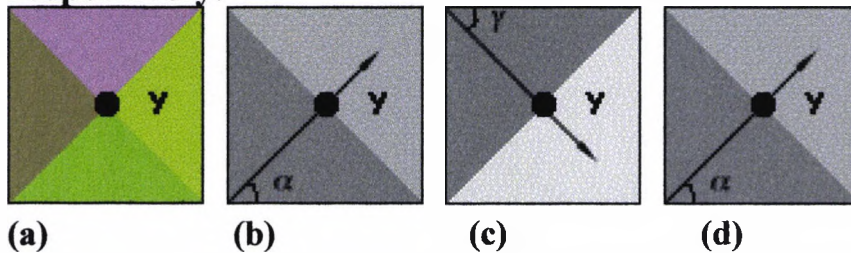

(a) **(b)** **(c)** **(d)**

Figure 3.2 a,b,c and d are showing the boundary of several regions in all colour, red, green, and blue planes respectively

In the second situation the problem can be divided into many cases, since all cases depend on the difference between their gradient orientations and the sign of δ in each case depends on the difference between all regions.

If colour edges are perceived to be visually better than grey level edges, as advocated by the previous work [Nev77, NS87], then one possible explanation is that the majority of edgels which support the geometry of the objects have positive or zero values in δ and unwanted edgels caused by noise or other physical effect have negative values in δ. Therefore, the strength of the colour edges is higher. The next section describes the investigation of this proposed explanation over a range of real images.

3.3.2 Statistical Approach

Using the mathematical approach provides the conclusion that if all the gradient orientations of an edge in three colour planes are the same, the magnitude of the grey level edge is greater than the colour one. There is no definite conclusion on whether the magnitude of the grey level edge is

higher or lower in the other situation where the gradient orientations of an edge in three colour planes are not the same. An alternative approach is to find out the average difference between the colour and grey level edge. The following measures are used to compute the statistics of the gradient magnitudes over a set of real images:

$$\delta_{r,i} = \frac{\sqrt{\lambda_{c,i}} - \sqrt{\lambda_{l,i}}}{\sqrt{\lambda_{c,i}}} \quad \cdots\cdots (3.19)$$

$$\widetilde{\delta}_r = \frac{\sum_{i=1}^{n} \delta_{r,i}}{n} \quad \cdots\cdots (3.20)$$

$$\sigma^2 = \frac{\sum_{i=1}^{n} (\widetilde{\delta}_r - \delta_{r,i})^2}{n} \quad \cdots\cdots (3.21)$$

$\widetilde{\delta}_r$ is the average of δ_r on a set of edgels.

σ^2 is the variance.

$\widetilde{\delta}_r$ is the average of δ_r and σ^2 the variance of δ_r. δ_r is used to reflect the Relative Difference between the gradient magnitude of colour and grey level edgel, since the value of λ_c varies for each edgel over the image.

These two measurements($\widetilde{\delta}_r$ & σ^2,) are able to provide an objective comparison between gradient magnitudes resulting from edge detection with a colour and grey level image. The measurement of the Relative Difference at each edge location indicates whether grey level images have sufficient information to reflect changes of surfaces and other physical effects such as shading and specularity. The range of test images, the resulting images and some statistical results are shown in next section.

3.3.3 Experiments & Results

Firstly, a colour version of the Canny edge detector based on [NS87] is applied to the colour images. A manual segmentation based on the edge magnitude is used to identify the desired and noisy edgels. Next, the

Relative Differences in the gradient magnitudes of the detected edge points are determined over the entire image. Finally, the statistics of the noisy edgel and desired edgel set are computed separately. Five sets of statistics extracted from the images in figure3.2a-f are shown in table 3.1.

Image	$\widetilde{\delta}_r$-real edges	$\widetilde{\delta}_r$-noise edges	σ^2-real edges	σ^2-noise edges
3.3a	-25.8%	-22.9%	1.58%	8.1%
3.3b	-19.0%	-19.0%	15.0%	10.0%
3.3c	-23.7%	-17.8%	7.0%	13.2%
3.3d	-18.8%	-16.9%	15.6%	14.8%
3.3e	-22.5%	-17.6%	10.0%	13.5%
3.3f	-24.2%	-20.8%	9.0%	12.4%

Table 3.1 Statistics of the gradient difference between colour and grey level images on five different images.
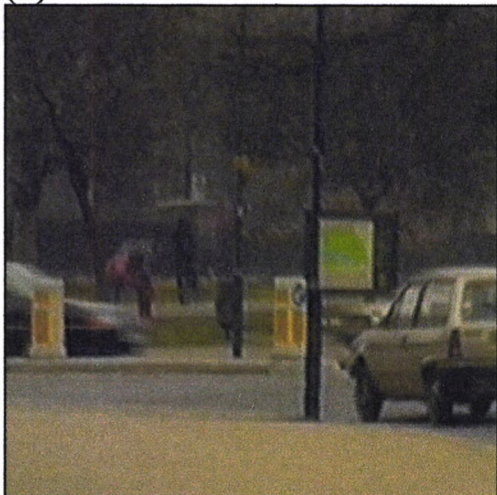
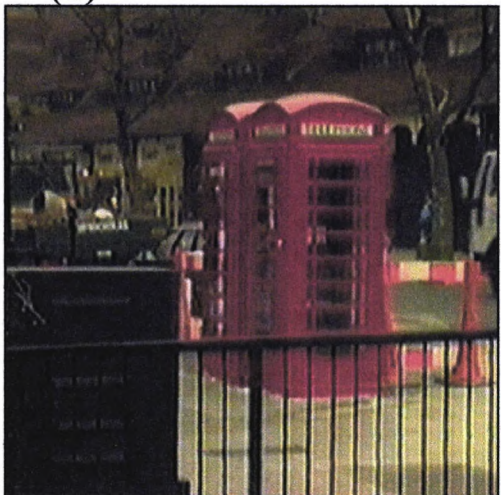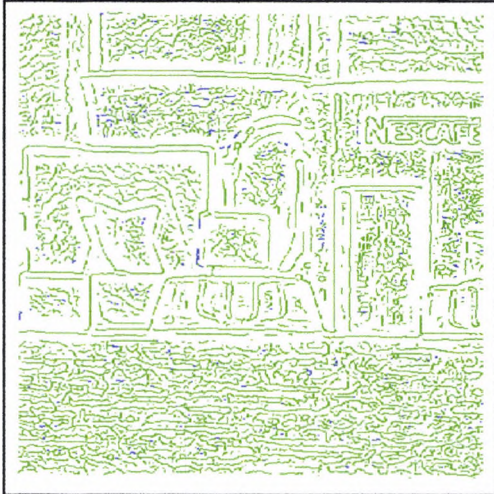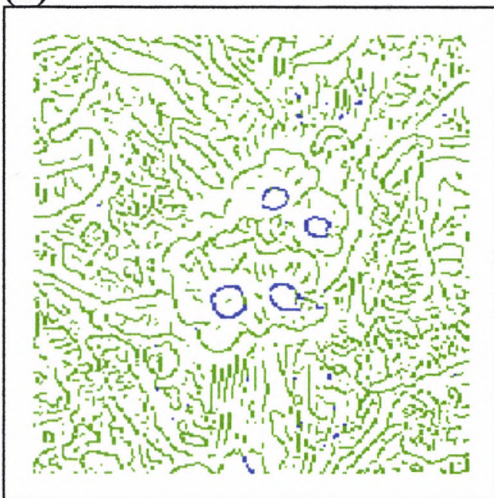**Figure 3.3a-f. The set of original test images.**

(a)

(b)

(c)

(d)

(e)

(f)

Figure 3.4a-f show the sign of the difference δ_r in gradient magnitude between colour and level images. Positive, and negative are in blue and green respectively.
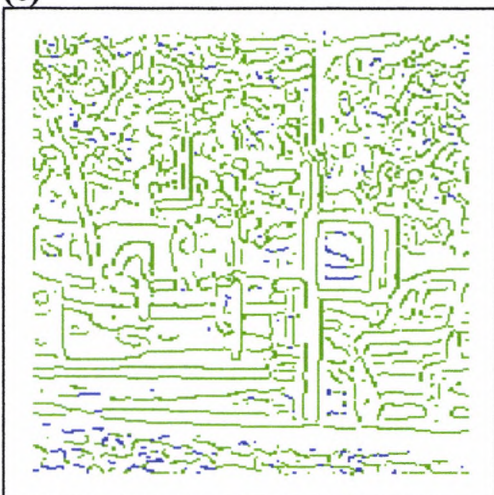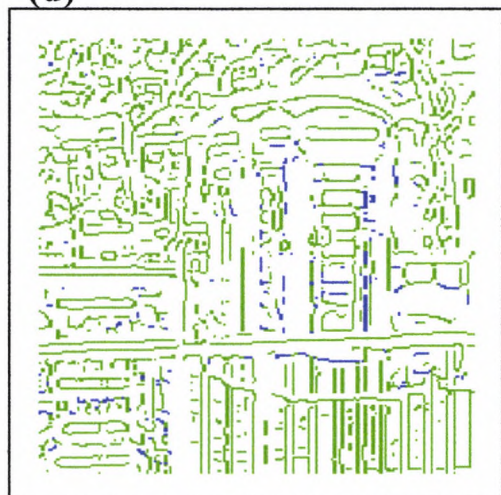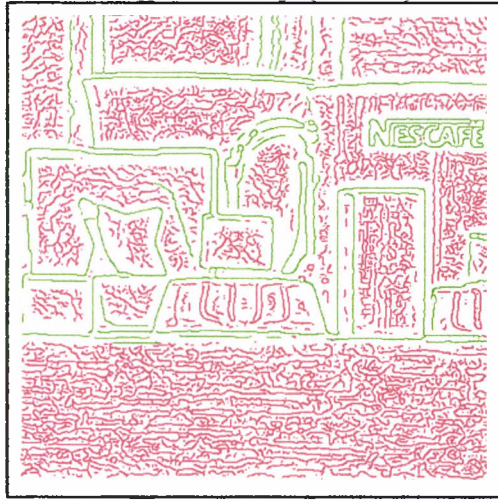

(a)

(b)
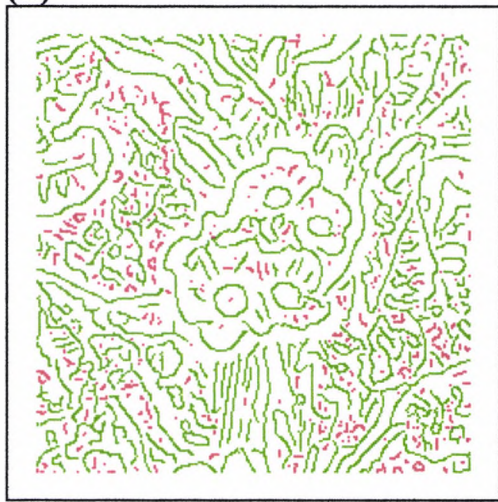
(c)

(d)

(b)

(f)

Figure 3.5(a-f) output of edge detector with the manual segmentation indicating the noisy (in red) and desired (in green) edgels.

(a)

(b)

(c)

(d)

(a)

(f)

The average of the Relative Difference on the desired edges over all the images is approximately -22%, indicating that on average, the gradient magnitude of the grey level is comparatively higher than its colour equivalent. There are some edges which have a higher gradient magnitude in the colour image, though the boundary on which they originate is not spatially consistent over the three colour planes or in regions with low contrast. Overall the boundary edgels are equally well detected by both the colour and grey level edge detectors.

3.4 Correlation of Edges in Three Colour Planes

Nevatia & Robinson [Nev77, Rob77] and my experimental results indicate that genuine edges are correlated in the three colour planes. This is a very useful parameter for distinguishing genuine from noisy edges formulated by using a simple logical operator and local correlation of the gradient orientation. Experimental results are compared, based on the relative signal to noise ratio between colour edge detection with corroboration and a grey level edge detector.

3.4.1 Physical Properties of Colour Surfaces

If two different surface reflection functions have significant differences in the red, green and blue planes, then the changes found in one colour plane are likely to be found in the other planes. The two different neighbouring surfaces should also share the same boundary and they should have the same orientation in each colour plane. My experimental results have demonstrated that genuine edgels are consistent in the three colour planes, but the edgels caused by the specularity, roughness of colour surfaces and other noise effects are not.

We apply two constraints in selecting edges. First, the presence of an edgel at a particular location in one plane can be corroborated by a consistent registration at the same or close-by location in the other planes.

Consequently they should have the same gradient orientation (or in the opposite direction), unless the edges are at the junction of three or more surfaces.

We have applied the algorithms to a synthetic image (figure 3.6a) and a noisy version (figure 3.6b), to which Gaussian noise (standard deviation of 20 grey levels) has been added. Firstly, a colour Canny edge detector [NS87] with standard deviation of 2.0 is applied to each colour plane and then the output edgels are tested against the two constraints. Figure 3.6c and d show the results of applying the colour Canny detector, indicating all the edges detected (threshold = 1). As can be seen, particularly in figure 3.6c, weak edges in the image resulting from the coarse surface shading in the synthetic image are mistakenly detected. The results shown in figure 3.6e and f (also using threshold = 1) were produced following the application of the corroboration function, and indicates that the constraints reject the weak edges. Quantitatively, taking figure 3.6e as our 'gold standard', the signal-to-noise ratios for figures 3.6c, d and f are 0.66, 0.097 and 0.19 respectively (see Table 3.2).

Table 3.2 The relative signal-to-noise of result images with and without corroboration.

	With corroboration Relative SNR.	Colour Canny Relative SNR.
Original synthetic image	1.0	0.66
Synthetic image with noise	0.19	0.097

Figure 3.6 (a) Synthetic CAD-generated image (b) with added Gaussian noise (s.d. = 20 intensity levels) per colour plane. Images (e) and (f) show the edges detected using edge corroboration, whilst (c) and (d) show those detected using the colour Canny edge detection.
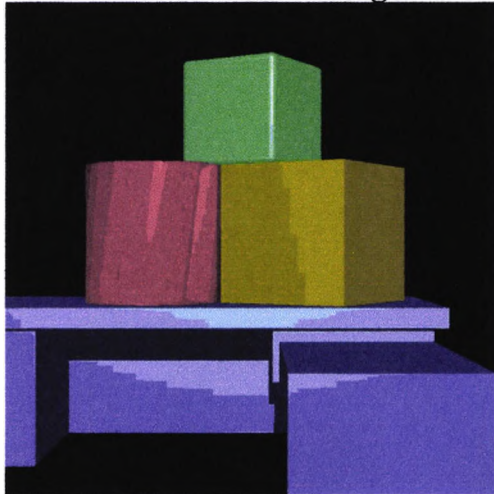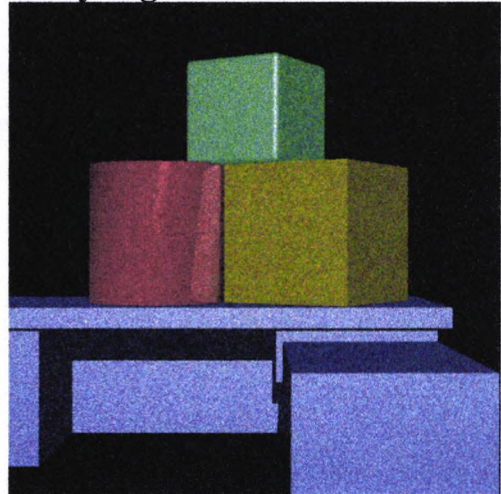
3.4.2 Experiments and Results

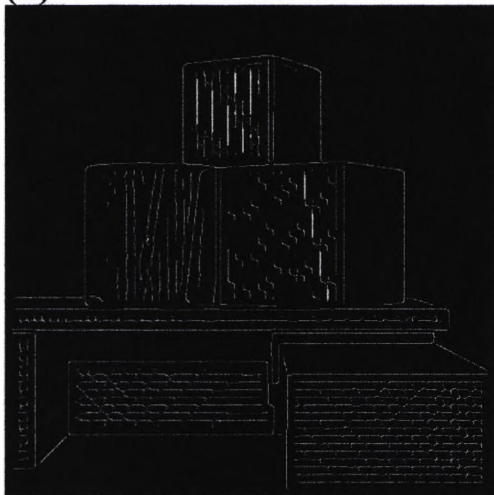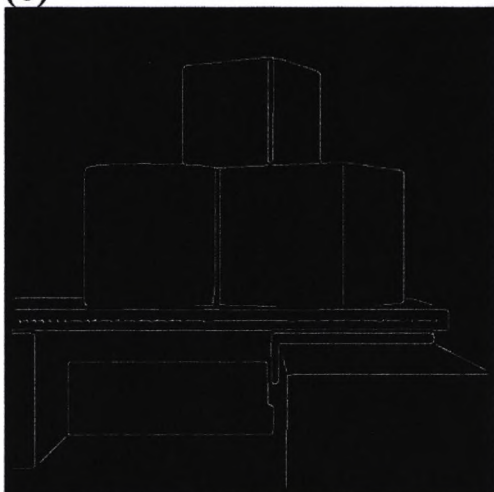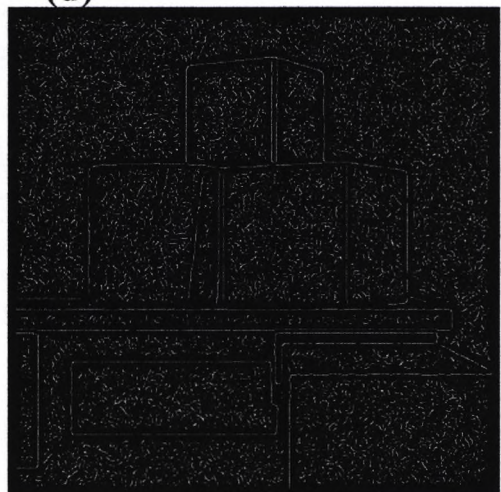In this section, there is an example of the corroboration function applied to a real image. The Canny edge detector is applied to the individual colour planes. Figure 3.7a is the original image and figures 3.7b, c and d show the result of the colour Canny detection on the red, green and blue planes respectively. Figure 3.7e is the combined image of the three output planes, generated by applying a logical 'OR' operation among the result of three colour planes (figure 3.7b, c & d), e.g. black = r +g +b, edges are found in all three colour planes. Figure 3.7f is the combined image of the three output planes, generated by applying a logical 'AND' operation among the edgel's orientation of the three colour planes. The edges correlated in three colour planes at the same or nearby (one pixel displacement) pixel locations are shown. This method is applied to the other test images (figure 3.3) and comparison measurements are shown in table 3.3.

A manual segmentation is applied to separate the desired and unwanted edgels. The noise edgels were identified as short edge segments of low gradient magnitude. Then the signal to noise ratio is computed by counting the number of detected edge pixels which correlate with the desired (signal) and unwanted (noise) edgels separately. Although this method proved satisfactory for the results on synthetic images shown in figure 3.5, identification of real and false edges proved more complex.

As can be seen from the results in figure 3.7 and table 3.3, the corroboration function produces a much less dramatic reduction in the number of noisy edgels detected. Nevertheless, it can be seen that the average reduction is by approximately a factor of 2. Two reasons can be identified to justify the poorer performance of the corroboration function in this case. Firstly the colour images were digitised using a conventional single CCD camcorder, for which the spatial localisation of the red, green and blue pixels are non-coincident. Secondly the edge detection can produce a slightly different localisation for each colour band, and hence the corroboration function must search its local neighbourhood for appropriate edges before the constraints can be applied to the gradient orientations.
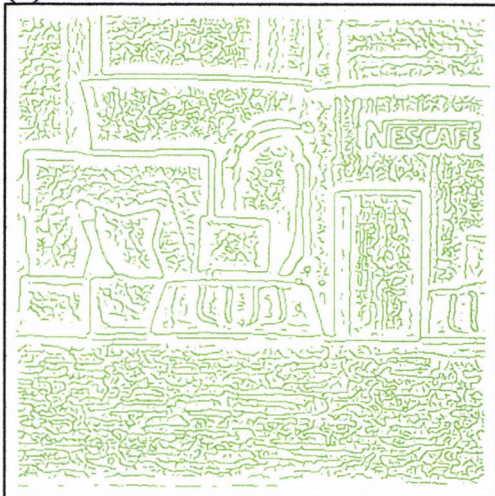
Figure 3.7 (a) Original image, (b), (c) & (d) are the result of Canny operator applied on red, green and blue plane respectively, (e) & (f) show the result of two combining methods.
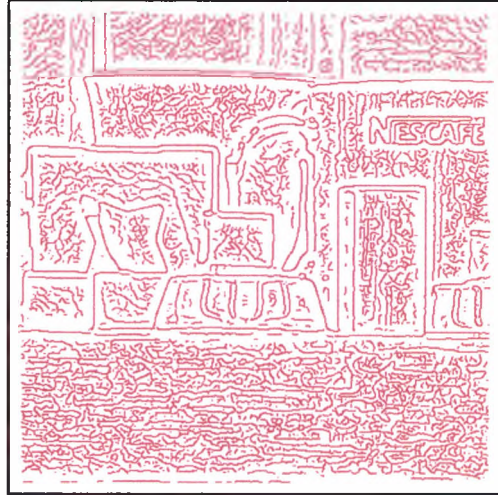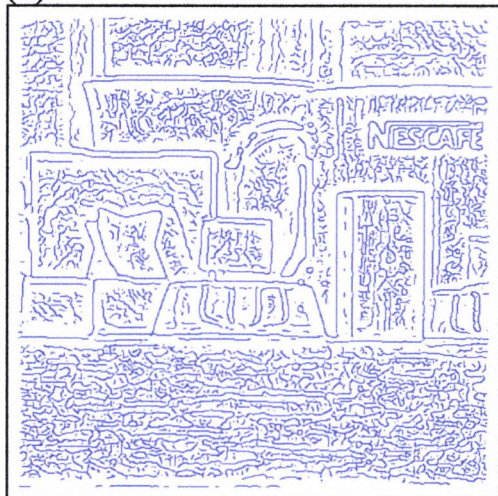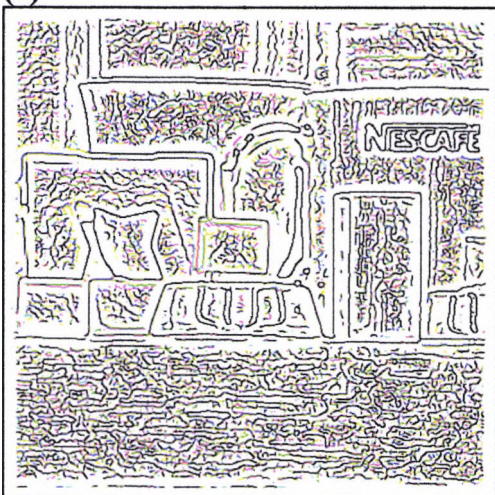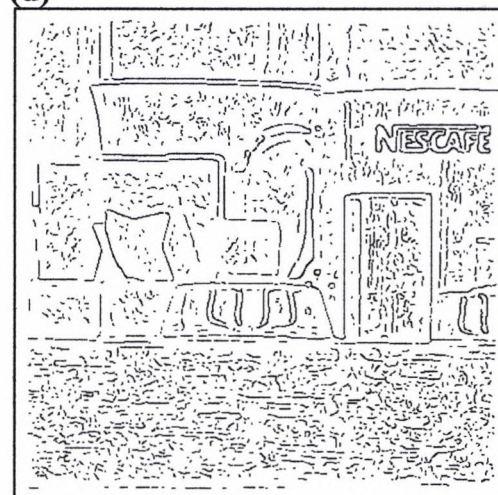
Table 3.3. Improvement on signal-to-noise ratio (relative to grey level image) for each of the 5 images (figure 3.3a-e) based on just the colour Canny edge detector, and then with the added gradient orientation corroboration.

Image	Grey level	Colour edges SNR ratio	Corroboration SNR ratio
1	1.0	1.02	2.62
2	1.0	1.37	1.63
3	1.0	1.55	1.55
4	1.0	1.01	1.43
5	1.0	1.63	2.07

3.6 Conclusions

This chapter has presented a mathematical and statistical comparison of colour and grey level edge detectors and has found that edges are detected equally well in colour and grey level images. The problem is how one can select an appropriate threshold or thresholds to include all the desired edgels and ignore the incorrect ones. The proposed solution is based on using the correlation of gradient orientations in the three colour planes as a parameter to distinguish the genuine and incorrect edgels. The corroboration function, which is described as a simple logical operation on the three colour planes, has been applied to both synthetic and real images. The initial experiment on synthetic images has shown a significant reduction in the number of false edges caused by slow changes of colour across a surface; also, the results on real images indicate an increase in the signal-to-noise level in selecting edgels, though the initial results on real images are much less dramatic than for the synthetic images. The problem of this approach is that it requires good localisation of the edge and consistency in the measurement of the gradient orientations in the three colour planes. Extracting a reliable measure of gradient orientation requires a high signal-to-noise ratio, and hence a Gaussian smoothing function with a large standard deviation is used; however, there is a trade-off between the amount of smoothing and good localisation of the edge [Can86]. A possible solution for improving the robustness of the corroboration function is to

choose other different possible parameters which can be applied to resolve the ambiguities in edge selection. An extension based on the hysteresis thresholding originally used by Canny can be adapted to preserve edge continuity based on preserving those edges with minimum deviation of gradient orientations.

Chapter 4
Segmentation of Colour Space

Abstract:

The aim of this chapter is to study the segmentation of colour images through the segmentation of colour space. The study divided into parts: first is the segmentation method based on merging colours into clusters; the second is a method based on classifying colours into different classes.

First part: *An initial processing is performed to reduce the subset of colours in the colour space. The reduction method employs the Peano scan to transform the three-dimensional colour histogram into one dimension and then conventional histogram equalisation is applied. Hence the merging algorithm is performed on the encoded images. The merging criteria are constructed based on the available information of each colour in the histogram and exploit the assumption that the distribution of each colour cluster in the colour histogram is unimodal. Each cluster is represented in a tree structure for easy manipulation. The result shows that the algorithm is able to adapt to different shapes of clusters with no a priori knowledge about the clusters.*

Second part: *This describes a segmentation algorithm based on classification of colours. Assumptions are made on the images that some of the colours appearing on the image are known to the database. With such assumptions, the problem of segmentation becomes an exercise of classification. A Colour Information Look Up Table(CILUT) is used in the process.*

4.1 Introduction

Image descriptions nearly always refer to parts (regions, features or objects) of the given images, thus the initial step in the image understanding process is segmentation, in which the image is divided into parts. Also the divided parts will be further processed to obtain their features, such as their shape, size, colour, position etc. Thus image segmentation and its performance are important to the overall image understanding system.

Most of the research work on image segmentation methods are mainly applied to the grey level image [HS85]. Their limitation can be easily understood, when two adjacent surfaces are the same in brightness but different in colour. Colour provides provide a larger discriminant space and the use of colour becomes an important issue on image segmentation. Some may consider that the segmentation method for colour images is an extension of grey level image method. This chapter reviews some of the early and current work in colour image segmentation based on partition the colour space.

Colour images may be segmented through the segmentation of the colour histogram. The colours of the pixels from a homogenous region will be close in the colour space. Also, the colours of the pixels from regions with different colour values will be located at a different position in the colour histogram.

Two approaches to segmenting the colour space are studied in this chapter: first the image data is considered as a set of different distributions in the colour space and assumes that each is unimodal, using a mode (peak) finding algorithm to merge colour groups. A colour compression process is applied to reduce the resolution of the colour histogram. If the cell size of the histogram is small, the data will be sparsely distributed and the segmentation results in many small clusters. If the cell size is increased, several adjacent distributions may be merged. The overall algorithm has been applied to a variety of images and the result shows that different shapes of clusters are formed in the histogram.

The second approach is to classify image colours into different classes. Such an approach is required to model the characteristic of each surface and store it in a database (e.g. the colour database described in chapter 2). The algorithm has been applied to real images and the results show that this technique is effective.
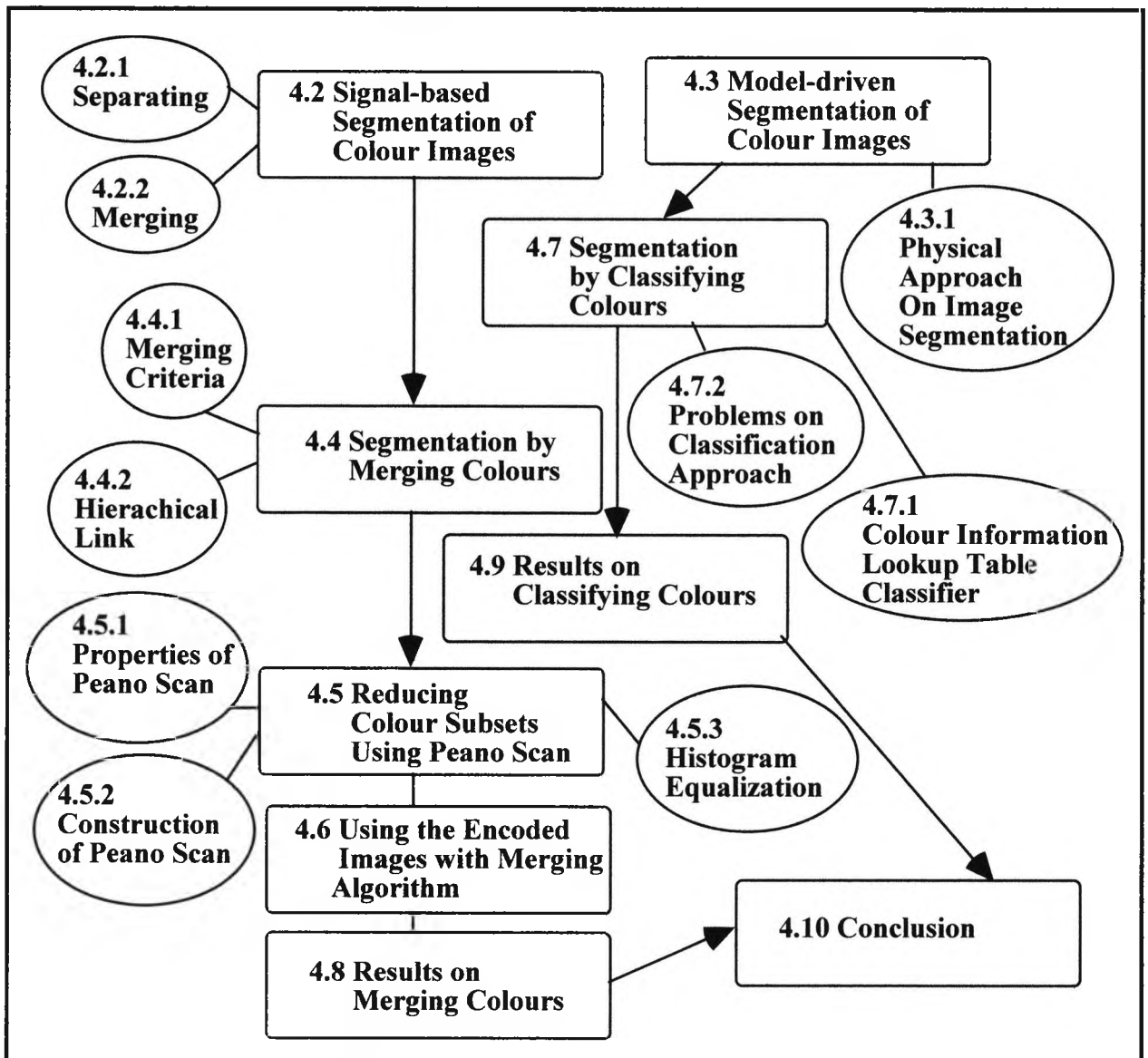

Figure 4.0 The structure of chapter 4.

4.2 Signal-based Segmentation of Colour Images

In signal-based segmentation, the image data is considered as a set of random signals, and minimum *a priori* knowledge about the surface is required. The approaches are commonly statistical. Most of techniques exploit the characteristic of the distribution in the colour histogram or detect the discontinuity of the colour in the image. Generally the techniques are divided into four main categories: the first and second methods principally operate in the spatial domain of the image (edge detection and region growing); the other two mainly operate in the spectral domain (histogram thresholding and clustering). One may consider the studies of these techniques are an extension of grey level image segmentation. Edge detection on colour images will simply detect the discontinuities in three colour planes rather than the one plane for grey level image. More details on the edge detection on colour images can be found in chapter 3. Region growing methods first divide the image into basic regions, then either adjacent identical points or arbitrary small regions are merged into one region if they are similar, where the similarity criteria depends on the actual system.

Segmentation techniques based on analysis of the spectral distribution of the colour are more complicated. The two common techniques are histogram thresholding and clustering: histogram thresholding is based on separating colours; clustering is based on merging colours. More details are described in section 4.2.1 and 4.2.2. Although the algorithm for segmenting the colour histogram is more complicated, the use of three-dimensional colour space provides a larger discriminative space. Research on segmenting the colour space have shown good results on the segmented images [KNF76, DBE77, CF83, Oht85, KB90].

4.2.1 Separating Colours

The segmentation algorithm based on separating colours in the histogram is commonly applied to three 1-D histograms rather than a single 3-D histogram. Each histogram summarises the distribution of values in one of

the colour attributes, by searching for the appropriate thresholds that act as the separators between different distributions of data in the 1-D histogram, different regions can be found. Ohlander et al [OPR78] showed an example of using histogram thresholding techniques on their region splitting method. Their segmentation method takes a region of the image, and using histograms of the features' values in this region determines a threshold for one feature to be used to split the region into sub-regions. These sub-regions are then further segmented if needed. Determination of the thresholds of the histogram has been considered the problematic part of their segmentation techniques. Ohta [Oht85] attempted to ease the problem by finding a suitable set of colour features, such that each feature provides greater discrimination between the different distributions of data.
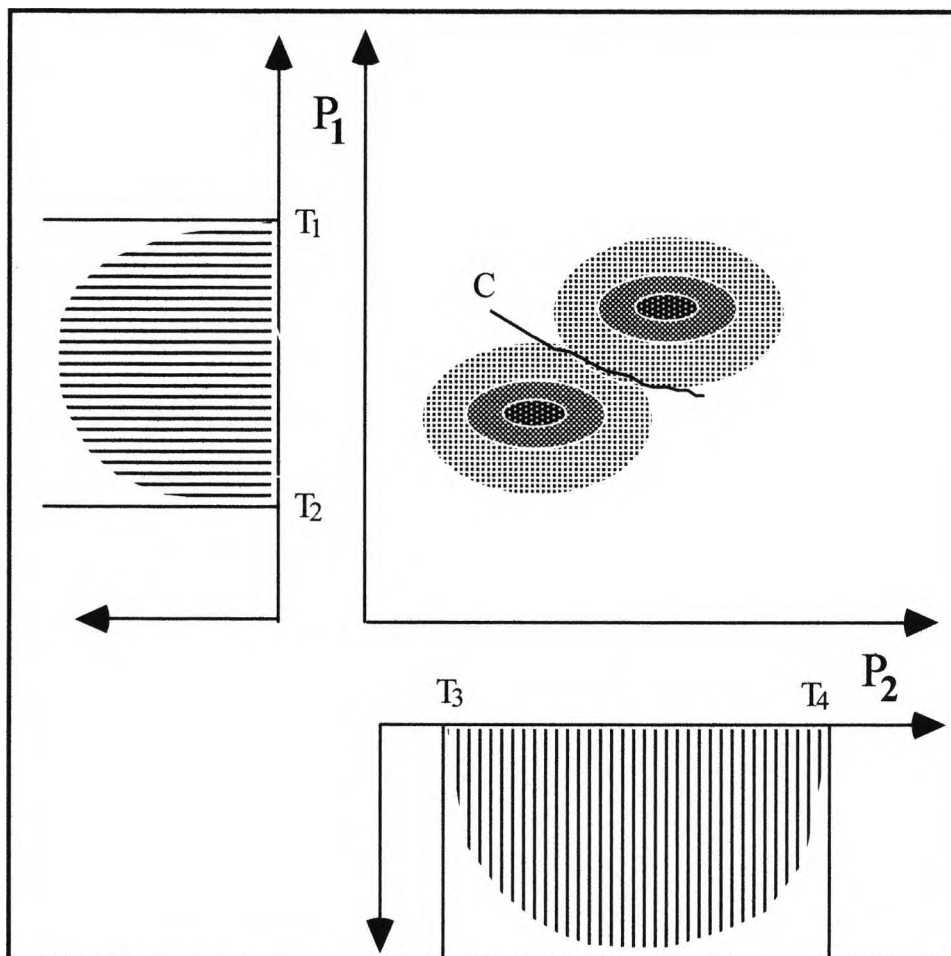

Figure 4.1 A 2-dimensional histogram of a colour images and its corresponding two 1-dimensional histograms.

Tominaga [Tom86] used the same algorithm with a set of human perceptual colour attributes. There are some limitations with their algorithm, because distributions of colour in the 3-D colour histogram are not always mirrored in three 1-D histograms Figure 4.1 shows an example of a 2-D histogram and its corresponding 1-D histograms on different features. The first pair of thresholds T_1 & T_2 of the histogram with feature P_1 are not able to separate the two distributions, and neither is the pair of thresholds T_3 & T_4 on the histogram with feature P_2. There was an attempt at separating colours in a two-dimensional histogram by Sarabi & Aggarawal [SA81]; the algorithm used *a priori* knowledge for the determination of thresholds. The threshold is a curve in 2-D space (seen in figure 4.1, where C is the separator between the two distributions).

4.2.2 Merging Colours

The segmentation algorithms based on merging colours are considered as a process of finding different clusters in the colour space, since the colours situated within a cluster are assumed to be from a homogenous region. Also merging colours to form groups requires the use of some knowledge or assumptions about the distribution of the clusters.

The early algorithms only operate on the vector representation of colour data. Most of these clustering algorithms require *a priori* knowledge about the images (maximum number of clusters or the location of the cluster centres), fine tuning on the parameters (neighbourhood criteria), and enormous processing time is required because of the iterative nature of the algorithms, for example, k-means and ISODATA algorithms.

To date the most plausible clustering techniques operate on the features' (texture or colour) histogram and assume the cluster distributions are unimodal [KNF76, Kit76], and results of the analysis have shown that such an assumption is appropriate. Because of the unimodal assumption, the clustering algorithms [KNF76, DBE77, CF83, KB90] become a process of mode (peak) finding and linking relevant colours to their corresponding

peak. Also the advantage of the algorithm is that no parameter is required and the process is not iterative [KB90].

4.3 Model-driven Segmentation of Colour Images

Model-driven segmentation can be considered as a process of classification with some predefined knowledge about the images. In comparison with signal-based segmentation, model-driven segmentation puts more restrictions on the images. The images must consist of surfaces that are similar to those in the model. Also the system is required to model the surfaces that are likely to appear in the image, and the process of modelling and classification should be performed under the same conditions. This technique is more applicable in an active vision system, because only the required surfaces are segmented from the entire image. Unnecessary processing on segmenting unrelated regions can be ignored. Furthermore it provides a shortcut for colour recognition, because the regions which are being segmented are those in the model.

A typical example can be found in Gonzalez-Rodriquez et al [GBS90] and Matas et al [MMK93] where they proposed a system for segmentation and recognition of coloured objects. The system consists of learning and recognition steps. The learning includes a clustering process that is used to discriminate different colour surfaces. The segmentation is achieved by recognition; the process is based on the colour distance between the image colour and the cluster colour representation. Other examples of segmentation by recognition are described in chapter 2.

4.3.1 Physical Approach on Image Segmentation

Specular highlights formed on the surface of a dielectric material will result in two or three distinctive colour regions in the image. These colour regions will be segmented into different regions by most of the signal-based segmentation algorithms, despite the fact that regions of different colours originated from the same surface. The specular highlight is due to the

inhomogenous physical properties of the surface, the viewing orientation and the position of the illuminant. Shafer [Sha85] proposed a model that explains the relationship between these physical effects. Later Klinker et al [KSK90] analysed the relationship between the colour of the highlight and the object in the colour histogram. Using these relationships, it is possible to identify whether the highlight originated from the same object. Therefore the over-segmented regions resulting from the signal-based segmentation can be minimised.

The other common physical effect resulting from different colour regions on a homogenous surface is shading. The effect is caused by the difference in viewing and illuminant orientation at each location of the surface. Healey[Hea89] used the normalised colour system (described in chapter 2), which allows the algorithm to process a representation of a colour image in which the contribution to pixel values that are only due to the illumination and the materials' properties are preserved, but in which factors that are due to scene geometry have been removed. Hence the classification of pixel values are more stable.

4.4 Segmentation by Merging Colours

The colour image is segmented by separating the colour in its corresponding colour histogram. The algorithm exploits the distribution of colours in the histogram. It assumes that the pixel colours from a homogenous region are similar and are close in the colour space. Also the cluster they form in the colour histogram is a unimodal distribution [KB90, Kit76, MMK95, Oht85]. Therefore the problem on segmenting the colour histogram becomes an exercise in finding all the unimodal distributed clusters. Such assumptions are widely used in colour image segmentation and they are important to the cluster finding strategies. The algorithm is based on merging similar colours and mode (peak) finding within each cluster; such methods can be found in some early works in the machine vision community [KNF76, DBE77, CF83]. The advantage of the method is that it requires no *a priori* knowledge about the clusters. Also it is able to adapt to different shapes of cluster. The detail is described in section 4.4.1.

Each cluster is represented by a hierarchical structure and the merging of two colours is based on a link established between them. The hierarchical structure for the clusters is described in section 4.4.2.

On studying the cluster finding algorithm, there is a problem that is associated with the cell size of the histogram [KB90] or in the other words, the quantisation step of each feature. If the cell size is small, the distribution extends over a large space and is fairly even within the cluster. On the other hand, if the cell size is large, several clusters may be merged because of the lack of resolution. In section 4.5, an encoding method is employed to adapt to the distributions of the histogram. The algorithm divides the colour histogram into a number of bins that have the same number of pixel counts. The bin size is determined by the size of the image and the required spectral resolution. The end product of the process is a reduced subset of colours which forms an equalised histogram, while it preserves characteristics of the distributions in the histogram.

4.4.1 Merging Criteria

In order to merge the colours within each cluster in the colour histogram, the algorithm exploits the characteristics of the clusters and the available information from the colour histogram.

The merging criteria are based on assumptions about the clusters: i) the colours from the homogenous region are close to each other in the colour histogram; ii) similar colours are close together and they form a cluster whose distribution is unimodal. The available information in the histogram is the location and the occurrences of the colour:

Each colour C_i is represented as a vector (red, green and blue) in the 3-dimensional colour space.

$$C_i = [r, g, b] \quad \cdots\cdots (4.1)$$

Each colour cell $H(C_i)$ in the histogram contains the number of pixels that are represented by the corresponding colour vector. $H(\ )$ is the histogram.

Conditions for merging:

I. Closeness (similarity).
This is based on a distance measure between two colours. A Euclidean metric is used in this algorithm.

$$\text{Distance} = \left\| C_i - C_j \right\| \quad \cdots\cdots (4.2)$$

where C_i is the colour i, C_j is the neighbourhood colour.
If the neighbourhood colour C_j is the closest colour C_i, then merge.
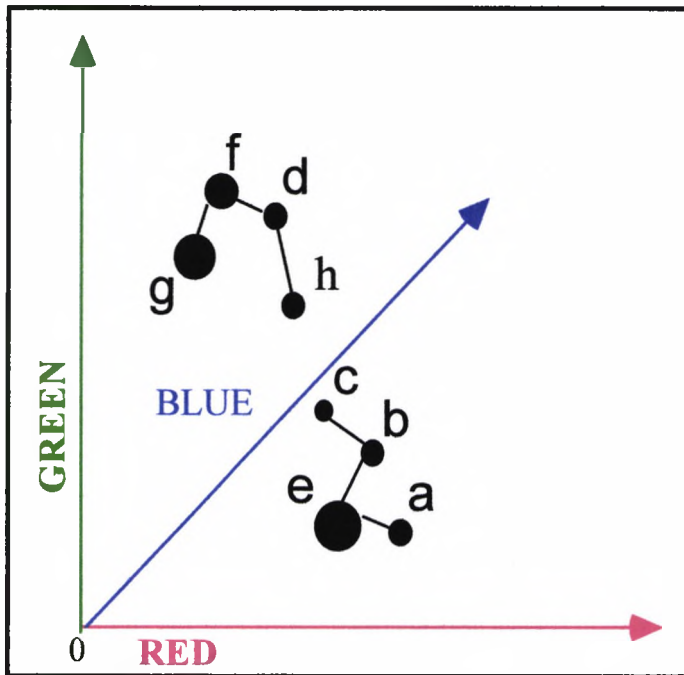
II. Pixel counts

$H(C_i)$ is the pixel count of the colour i.
If the neighbourhood colour C_j has the highest pixel count among those which have the same similarity (or distance) and is higher than or equal to colour C_i, then merge.

When these two conditions are met, the two colours are merged. The two conditions are complementary. The first condition can resolve the ambiguity when several neighbourhood colours have an identical number of pixel counts. The second condition can resolve the ambiguity when two neighbourhood colours are equally close to the colour. This condition also provides part of the mechanism for the peak climbing.

An example of the colour merging is shown in figure 4.2. Consider colour C_a, the closest colour to C_a is C_e and the pixel count of C_e is higher than C_a, therefore C_a and C_e are merged.

The non-uniformity of colours within each homogeneous region is assumed to be caused by Gaussian noise, generated during the image capturing or as a result of other physical effects such as the reflection functions of the surfaces which are a mixture of body and surface reflection. Therefore, the position of the peak within each cluster will be equal to the mean colour of each cluster.

Figure 4.2 A Colour histogram with two connected clusters .
(the larger the spot, the higher number of pixel counts)

4.4.2 Hierarchical Link

In the cluster analysis, two similar colours are merged if the merging criteria are met and a link is established between the two colours, for example, C_a and C_e in figure 4.2. When all the links are established in the graph, a set of data colours is formed for each cluster. Since the distribution of the cluster is assumed to be unimodal, each cluster only contains one peak (mode). Each set of data colours can be easily mapped to a tree structure, in which the root of each cluster tree is the peak colour of the cluster. Therefore the peak searching become the process of finding the root in each cluster tree.

An example of the merging procedure applied to a set of colour data as shown in figure 4.3. The corresponding connected cluster structure is shown in figure 4.4.

An example is shown in figure 4.5. In order to perform the peak-climbing process, it simply includes the direction in each link, where the direction is

from a low to a high pixel count. The reasons for using low-to-high direction is that all colours within a cluster are pointing to another colour, except the peak which is not pointing to any colour. Therefore the peak of each cluster can be easily found. The highest order of tree is the root, the lowest order of the tree is the leaf. Therefore when the cluster is mapped to the tree structure, the peak is the root and the colours around the boundary of the each cluster are the leaves. The tree structure of the cluster in figure 4.5 is shown figure 4.6.
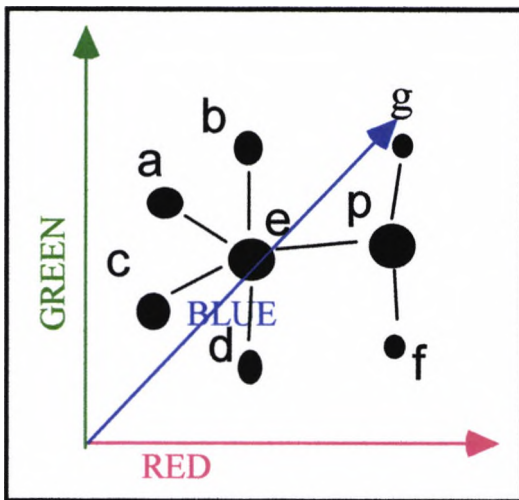
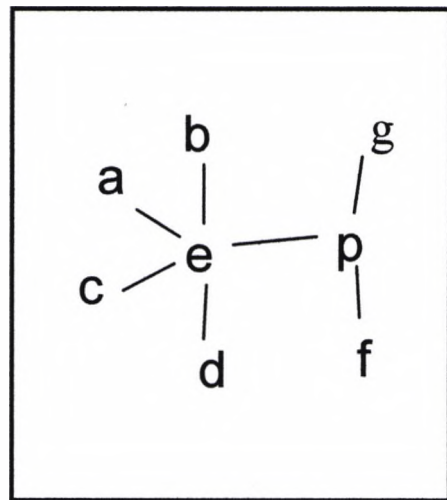**Figure 4.3 A set of merged colours in the histogram.**
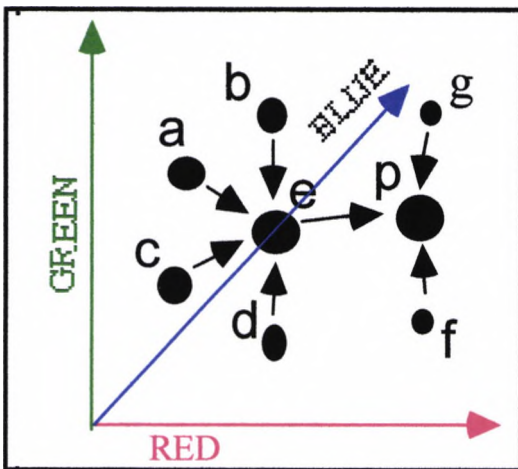
Figure 4.4 A set of connected colours.

Figure 4.5 A linked cluster in the colour histogram.

Most of the links are unidirectional, the direction is from low to high because it sufficient for the labelling and mapping. For example, if the root *p* of *cluster 1* (see figure 4.6) is labelled, since all colours within the cluster are pointing toward their highest neighbourhood colours, there is a simple path from any colour to the peak colour. However if the direction is high to low, the labelling and mapping becomes more complicated. It is because in this structure, the colours around the boundary of the clusters have to be labelled instead of the peak (see figure 4.7). The difficulty is to find all the leaves within the same cluster and label them appropriately. For example, colour *e* only points to one colour in the low-to-high scheme, but in high-to-low, it points to many.
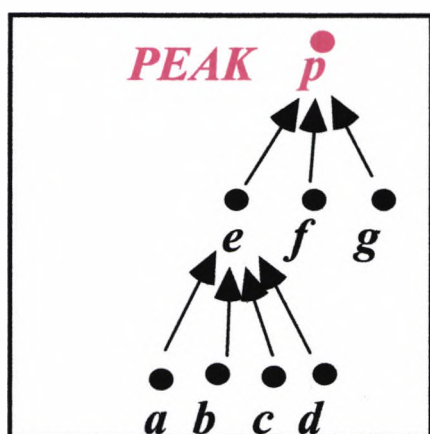
Figure 4.6 The structure of tree which represents the cluster shown in figure 4.5
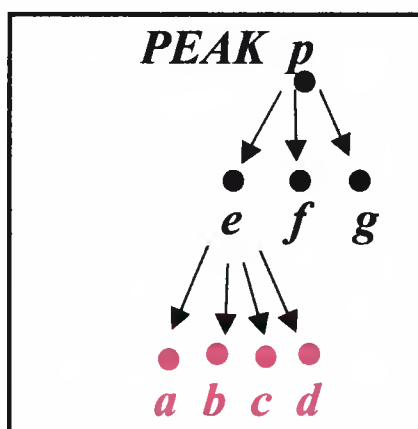
Figure 4.7 A tree structure of a cluster which has high-to-low directional link.

4.5 Reducing Colour Subsets Using Peano Scan

Colour compression is performed on the spectral domain of the image, for example, 24 bits per pixel is compressed into 6-8 bits per pixel. Colour compression is performed before the merging algorithm. This simplifies the merging by reducing the number of links needed to form the cluster tree (see section 4.4). In addition, the compressed image is more easily manipulated in the merging algorithm and it allows the images to be displayed with a frame store with a limited number of bytes for each pixel.

Colour compression is achieved by merging a certain number of colours to form a small subset. It is necessary to ensure that the elements within the subsets are similar to each other. A Peano scan acts as a transformation between n-dimensional space and a 1-D space retaining some of the spatially associative properties of the space. More details about the properties and construction of Peano scan are described in section 4.5.1 and 4.5.2. As a result, analytic techniques to process the data, which are essentially single channel operations, can be applied, for example, histogram equalisation. This histogram equalisation operation results in a reduced subset of colours. The dominant colours (with high pixel count) are emphasised whilst the less noticeable colours (with low pixel count) are merged with their (dominant) neighbourhood colours and the pixel counts are accumulated. Although each subset has a similar number of pixel counts, the density of the set of less noticeable colours is lower. The histogram equalisation operation used in this chapter is described in section 4.5.3.

4.5.1 Properties of the Peano Scan

In 1890 mathematician Peano [Sim86] showed that a curve can be defined which has the properties of space filling on an n-dimensional digital space. Later Hilbert provided the geometrical definition for such a curve which can be obtained with a recursive process. The scan is able to transform any dimensional space to a one-dimensional curve without losing the spatial compactness of the data.

The relevant properties of the Peano scan can be summarised as follows:

1) The unbroken curve passes once through every element in the digital space and without intersecting itself at any point.

2) Points close on the curve are close in space.

3) Points close in space are likely to be close on the Peano curve.

4) The curve acts as a transform to and from itself and n-dimensional space, preserving some of spatial properties on the scan.

A Peano-Hilbert scan is a mapping

$$H_p^n : U(n, p) \rightarrow U(1, np) \cdots (4.3)$$

$U(n, p)$ for the non-ordered set of hypercubes in n dimensions (2^{np}).

$U(1, np)$ for the corresponding segments in one dimension (2^{np}).

The mapping establishes a one-to-one correspondence between 2^{np} (n is the number of dimensions and p the precision) hypercubes in n dimensions and 2^{np} intervals in one dimension, i.e. on a line. As a Peano-Hilbert scan orders the 2^{np} hypercubes, there is a corresponding order of segments in the one-dimensional space.

An example of the Peano scan (H_5^2 and H_2^3) in 2-dimensional and 3-dimensional space is shown in figure 4.8 and 4.9 respectively.
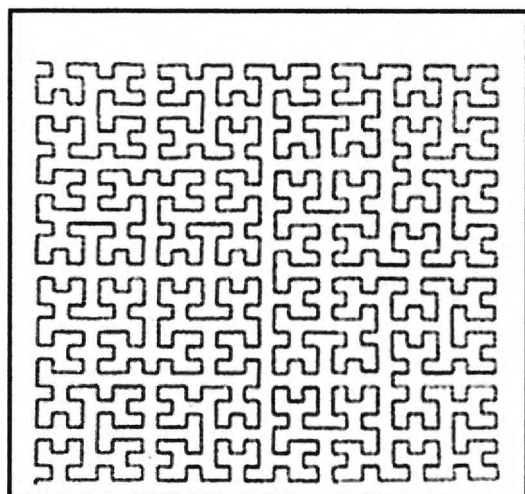

Figure 4.8 H_5^2 scan, 2-dimensional Peano Scan with the resolution 32.
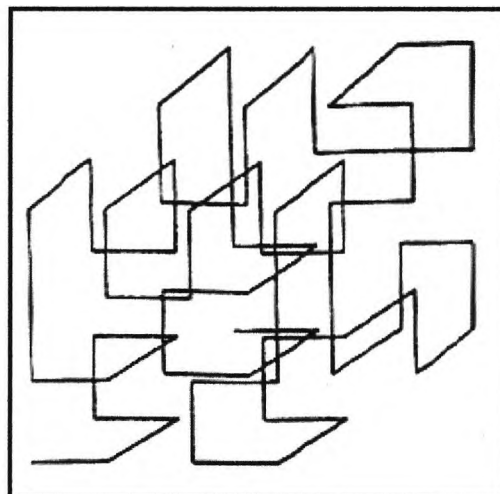
Figure 4.9 H_2^3 scan, 3-dimensional Peano Scan with the resolution 4.

Applications of the Peano scan in image processing are found in image segmentation and classification. Quinqueton's clustering method [Qui78]

employed the Peano scan to transform an n-D histogram to a 1-D histogram and a minimum spanning tree was used to group the data. Lehar & Steven [LS84] used a Peano Scan encoding method [SLP83] (as described in this chapter) to generate a subset of colours and then applied these compressed colour sets for fast image enhancement. Lambert et al [LCB91] employed a Peano scan on both spectral and spatial domain of the images: firstly, an encoding method [SLP83] is used to reduce the colours to a maximum of 256; secondly, a split and merge algorithm was used on the images which was spatially transformed using a two-dimensional Peano scan rather than the quad-tree representation. Therefore, splitting of the quad-trees is replaced with a process of splitting the Peano curve.

4.5.2 Construction of Peano Curve

The scan is an operation that maps the hypercubes in the n-dimensional space into points (or intervals) on a 1-D line. The mapping can be implemented in the form of a lookup table in which the index is the location of the hypercube in n-D space and the output is the position along the curve. Such a construction requires a large amount of memory if the dimension and precision of the digital space are high. For example, for a 3-D colour space with 256 levels for each dimension, the total number of locations is over 1.6 million. Also the process of entering the data (over 1.6 million values) into the lookup table is time consuming if the mapping is generated manually. Alternatively the scan can be generated dynamically using a recursive process. The scan can be represented in a hierarchical structure. The scan can be decomposed into a sequence of seed scans. Using the definition used in section 4.5.1, the scan in the three-dimensional colour space with precision of 2^8 in each axis is represented by H_8^3. The 3-dimensional seed scan is H_1^3. There are 24 possible 3-dimensional seed scans [Hun88] with different starting and ending directions. For example, for the colour space with a red, green and blue axis, the starting direction of the scan is in the red axis and the ending in the green axis. The scan **A** is shown in figure 4.10. r+ and r- move one step forward and backward on the red axis; g+ and g- move one step forward and backward on the green axis, and b+ and b- move one step forward and backward on the blue axis.
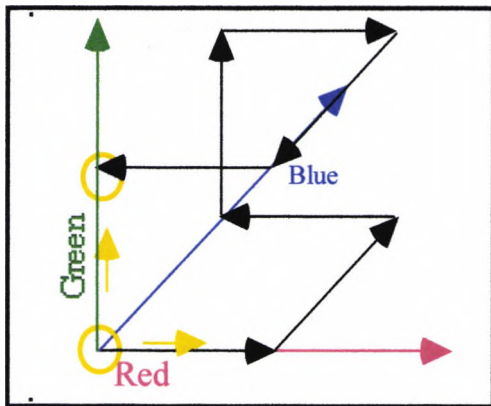
Figure 4.10 A 3-dimensional seed scan H_1^3.
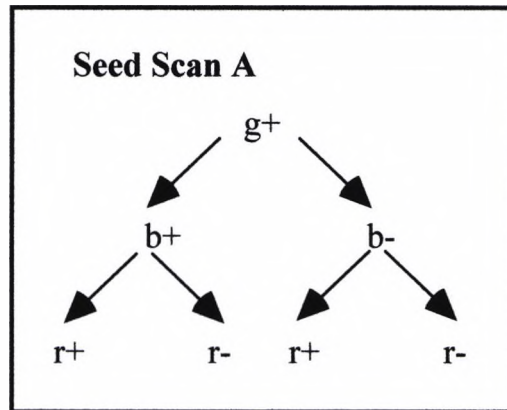
Figure 4.11 The seed scan A is represented in a tree structure.

The scan **A** can be represented in a tree structure and is shown in figure 4.11. The pseudo code of scan **A** is shown below: The complete set of seed scans and the corresponding procedures for filling the 3-dimensional space can be found in Hung's [Hun88] report. There is also an example of a Peano scan H_3^3 shown in figure 4.12.

Example of seed scan procedure:

```
procedure SeedscanA(i)
        if (i > 0)
                begin
                call procedure SeedscanB(i-1);
                r+;
                call procedure SeedscanD(i-1);
                g+;
                call procedure SeedscanC(i-1);
                r-;
                call procedure SeedscanF(i-1);
                b+;
                call procedure SeedscanB(i-1);
                r+;
                call procedure SeedscanE(i-1);
                g-;
                call procedure SeedscanC(i-1);
                r-;
```

```
            end;
      end;
Where 2ⁱ is the resolution of the scan.
```

Where 2^i is the resolution of the scan.
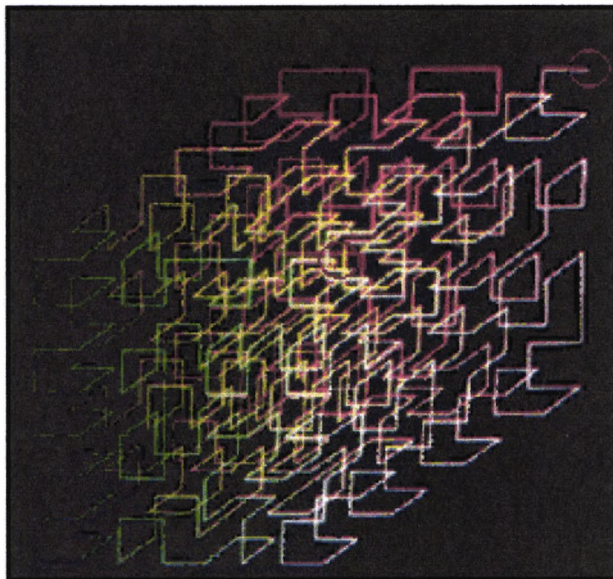
Figure 4.12 Peano scan H_3^3 in 3-dimensional colour space, the starting and end point is red circled.

4.5.3 Histogram Equalization

By tracking through RGB colour space using the Peano scan and recording the pixel counts at each location in the 3-dimensional histogram, a 1-D histogram is formed. Because of the ordering along the Peano line, approximations can be made without severe colour errors, as close points are close in colour space. The problem of choosing the colours (LUT values) now becomes analogous to the technique of histogram equalization, where each bin in the 1-D histogram has similar number of pixel counts.

An initial "bin size" is the total number of pixels divided by the maximum number of required subset of colours.

$$b = \frac{p}{n} \quad \cdots \cdots (4.4)$$

where b is the bin size.

    p is the total number of pixels in the image.

    n is the maximum number of possible bins.

For example, for a 512*512 image.

Number of pixels = 512 * 512 = 262144.

The maximum number of possible bins is 256.

$\therefore$ bin size = 512 * 512 / 256 =1024.

An alternative way to calculate the "bin size" is to utilise all the available entries in the LUT. Therefore a larger subset of colours can be formed without exceeding the predefined maximum. The calculation is similar to the one described above, except an extra scan of the histogram is performed before the calculation. This process will pick up the colours that have a number of pixel counts greater than the initial "bin size" and put them into the LUT. The total number of pixels will be deducted by each number of pixel counts which correspond to those selected colours. The maximum number of LUT entries will be deducted from the number of selected colours. Then the new "bin size" will be calculated by the total number of remaining pixels divided by the remaining number of LUT entries. The selection process and the calculation of the new bin size can be repeated until the optimal bin size is found. The equations are shown below.

$$b' = \frac{p - p'}{n - n'} \quad \text{......(4.5)}$$

where b' is the new bin size.

    p' is the sum of pixel counts which correspond to the picked colours.

    n' is the number of selected colours.

There are different ways to choose the colour in each bin as the representative of the colour in the LUT. Three different colour representatives of subset are used

1) the peak colour.

2) the average colour

3) the end point.

The average colour is found to be very similar to the peak colour in each interval and is visually better than the end point value.

4.6 Using the Encoded Images with Merging Algorithm

When the images have been encoded with histogram equalisation, the distribution of the LUT colours is quite even. The number of pixel counts in each bin is similar, but the size of each bin is different. Therefore the density of bins is used for the merging criteria, and less dense bins are linked to the denser bins. The peak of each cluster is the densest bin. Using the density of the bin is identical to using the number of pixel counts of the colour location. Both are measuring the density of each colour subspace in the colour histogram. The number of pixel counts of each colour measures the density in one colour location, but the density of each bin measures the average density of each bin.

4.7 Segmentation by Classifying Colours

The colours of different pixels from a homogenous region should have similar colour features; therefore by classifying the similar colours into the same class, it segments the image into regions. Two studies on classifying colours of the pixels are described in chapter 2: Maseen & Böttcher implemented a system for classifying the colour of the pixels into different object classes, and the class is modelled with real images; Matas et al [MMK93] argued that segmentation can be performed by classifying the colours of the pixels. Their system require one process for learning different classes of colours and another process to classify the pixel colours into different object classes. The learning process models the characteristics of different coloured objects. In studying this process, it is required to consider what necessary information must be extracted from the model object colours. The classification process is to separate the pixels into different classes. It is required to define a similarity criteria between the pixels and the different colour classes.

In this chapter, an attempt is made to use memory colours as the representation for different classes in the database and the classification process is implemented based on a lookup table classifier. The memory colours and the lookup table are identical to those described in chapter 2. Each class of colour is modelled from real images. Recall from section 2.4, memory colour consists of two components: a descriptive part and an algorithmic part. The descriptive part uses the mean colour of each class, whilst, the algorithmic part is based on a Mahalanobis distance measure. The index to the lookup table is the colour value of pixel and the output is the class number to which the pixel colour should belong.

Mahalanobis distance $MD_{i,j}$ between the pixel colour i and the mean colour of class j.

$$MD_{i,j} = \left[\underline{c}_i - \underline{c}_j \right]^T Cov_j \left[\underline{c}_i - \underline{c}_j \right] \quad \cdots\cdots(4.6)$$

where Cov_j is covariance matrix of class j

4.7.1 Colour Information Lookup Table (CILUT) Classifier

The major part of the system is the lookup table and the implementation is identical to the Colour Information Lookup Table (CILUT) which is described in chapter 2. Since the distance measure and corresponding class number can be processed off-line, the time spent on the classification process is only the time taken to search the information in the required lookup table location.

4.7.2 Problems On Classification Approach

There are some problems with this approach. The conditions for modelling the colour of each class are essentially the same as the one used for the test image. Changes of lighting will result in changes of the pixel colours (more details are described in chapter 5). However if colour constancy is to be added before the classification, the changes of pixel colour would be

minimised. Also some of the colours are not classified in the image, because the corresponding colour surfaces are not modelled.

4.8 Results on Merging Colours

The merging algorithm with a compression process was applied to a range of images. The resulting colour histograms showed that different clusters with different shapes are formed. Also the resulting colour histogram is used as the mapping between the original image data and the segmented image. The following equations show the steps to map the original data to the resulting image.

$I(x,y)$ is the original image.

$$I(x,y) = [r,g,b] \cdots\cdots (4.6)$$

where r, g and b is red, green and blue value respectively.

After the merging has applied, the resulting histogram $H(r,g,b)$

$$H(r,g,b) = l \cdots\cdots (4.7)$$

where $l = 1...n$, n is number of clusters found.

$$H(r,g,b):I(x,y) \rightarrow R(x,y) \cdots\cdots (4.8)$$

$R(x,y)$ is the resulting image.

On the result, it shows the image is segmented into a set of regions and can be seen in figures 4.14 and 4.15. Also it shows the algorithm is able to adapt to different shapes of clusters, and that can be seen in figures 4.19-28.

The following figures 4.13-4.16 show the original image and the output images. Figure 4.14 shows the output image (using a 'rainbow' colour lookup table to enhance the merged region) of the merging method, it can be seen that the similar colours from a homogenous region are merged. Figure 4.15 shows the output image using its 'natural' lookup table, in which each entry of the table is the average colour of cluster. With the use of this lookup table, any region of a particular range of colour can be easily selected. For example, in figure 4.16, the region of the brick wall that is

light brown in colour is selected. (The selection process has been done manually).

Figure 4.13 Original house image.

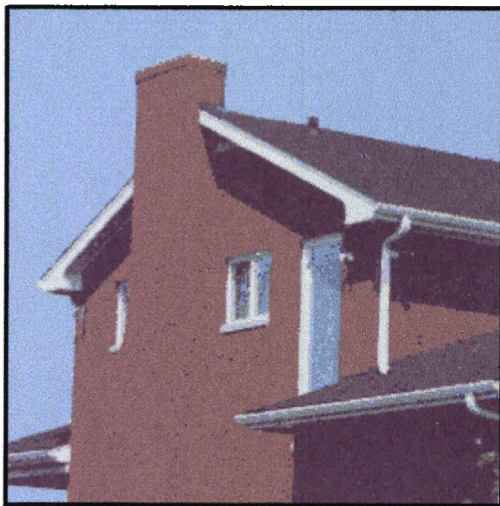**Figure 4.14 The resulting image.**

Figure 4.15 The output image with its 'natural' lookup table

Figure 4.16 The output image with the manually selected region

The colour histogram of the house image is shown in figure 4.17. The complete colour distribution of the house image is segmented. The segmented colour histogram is shown in figure 4.18 and each cluster is shown in figures 4.19-4.28. The results have shown that the algorithm is able to adapt to different shaped clusters which are not easily separated by a hyperplane.
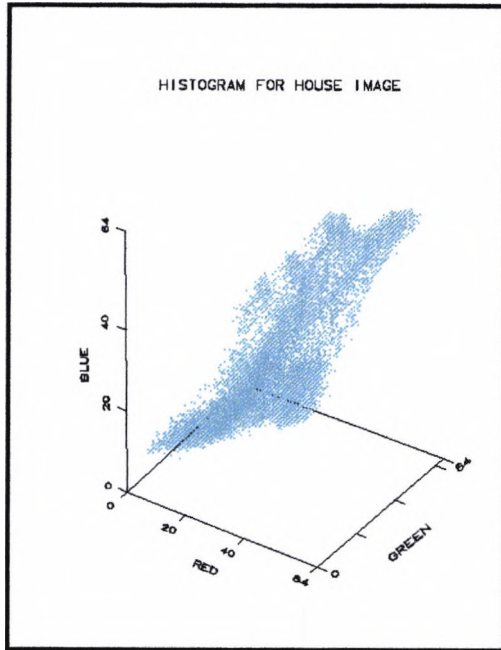
Figure 4.17 The overall cluster
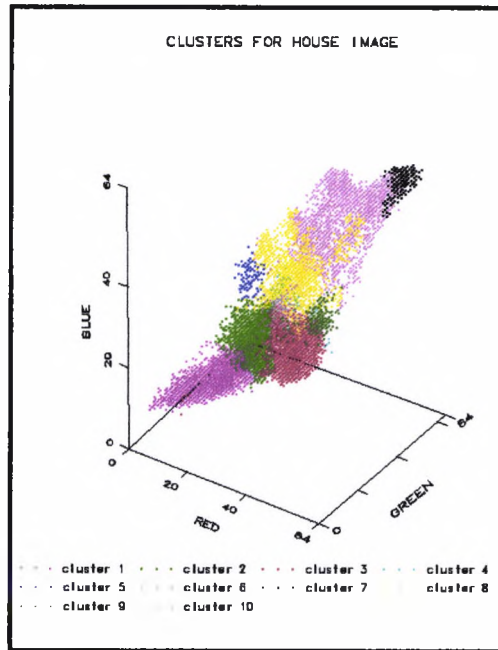
Figure 4.18 The segmented colour histogram, it consists of ten major clusters.
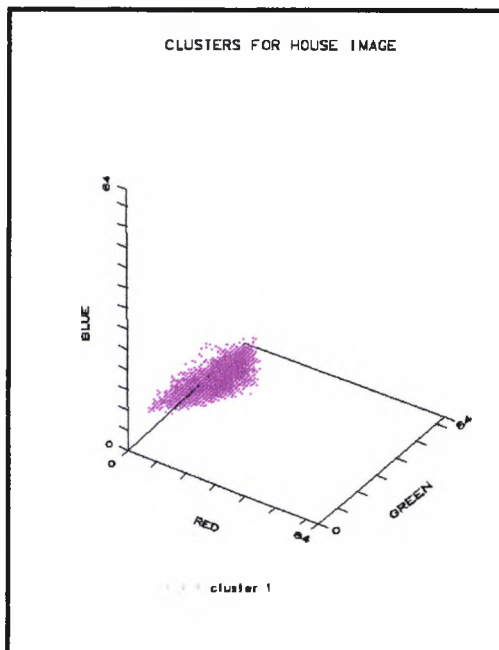
Figure 4.19 Cluster 1
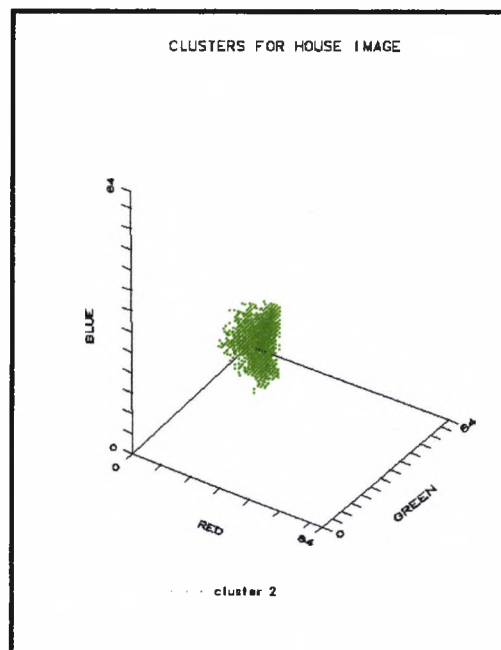
Figure 4.20 Cluster 2
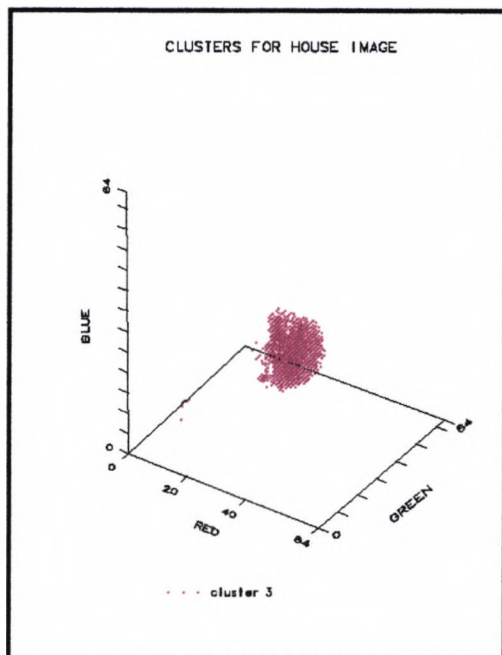
Figure 4.21 Cluster 3
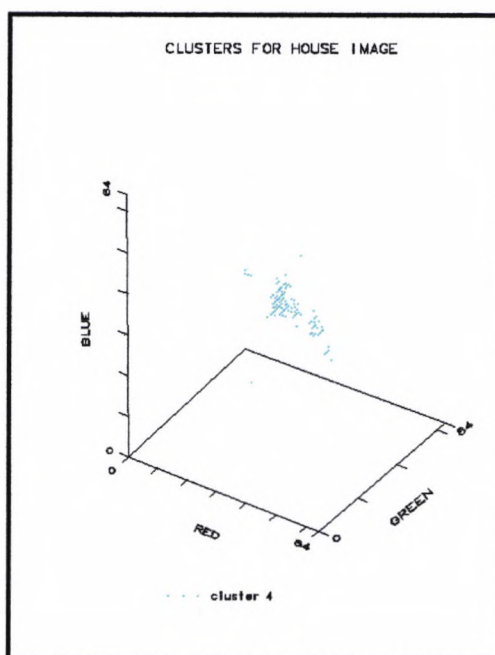
Figure 4.22 Cluster 4
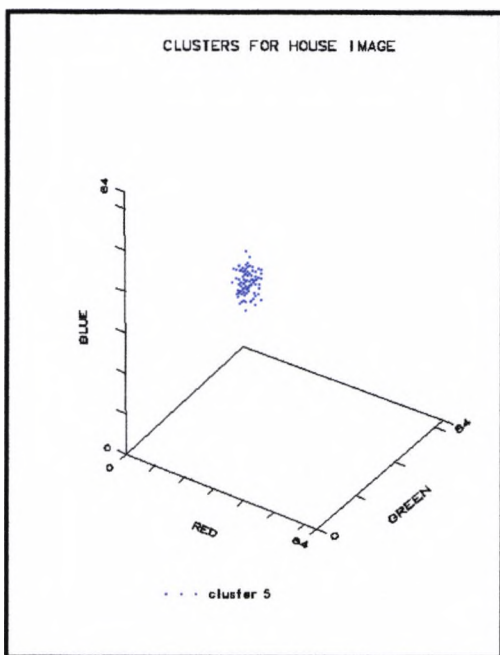
Figure 4.23 Cluster 5
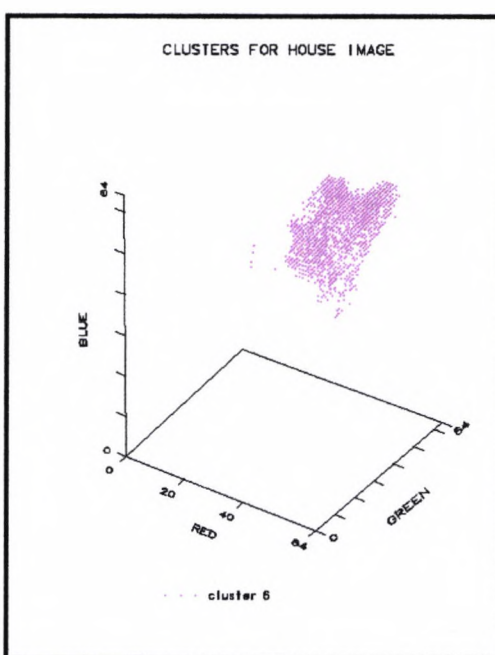
Figure 4.24 Cluster 6
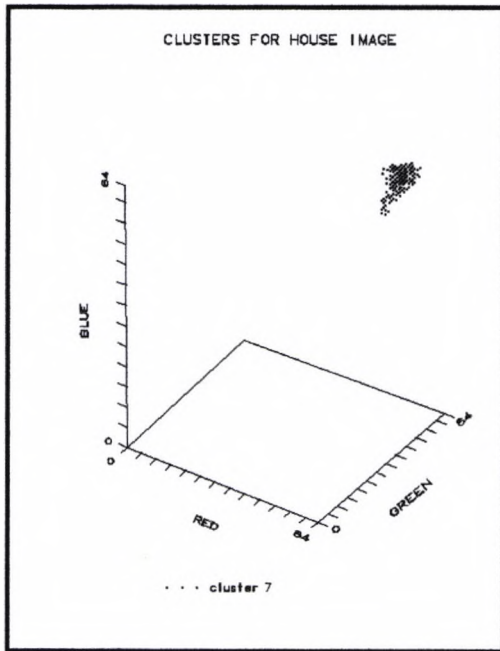
Figure 4.25 Cluster 7
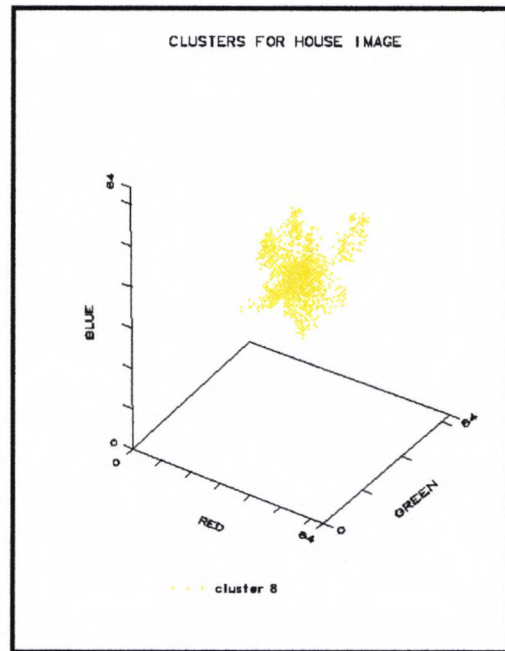
Figure 4.26 Cluster 8
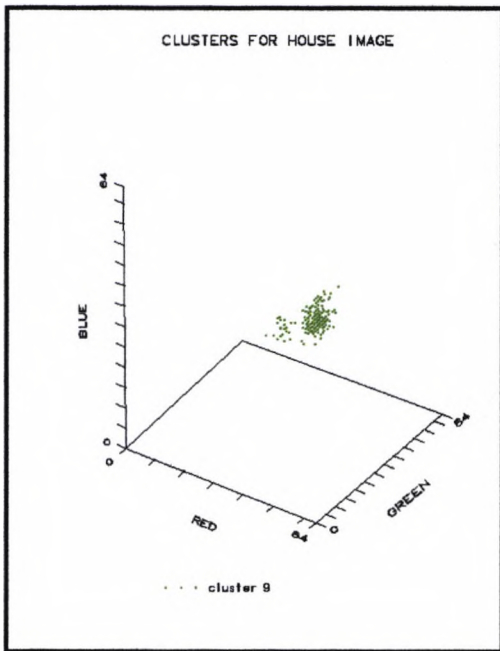
Figure 4.27 Cluster 9
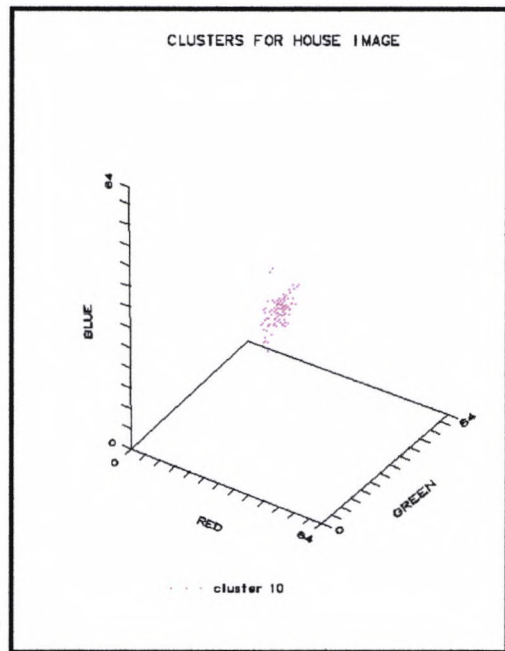
Figure 4.28 Cluster 10

Another test image is the jelly bean image (see figure 4.29) in which four different colours of jelly beans are used. The purpose of using the jelly bean image is to test the algorithm initially and then compare the result with the colour classification algorithm (the algorithm is described in section 4.7).

The image consists of five main colours including the background and is segmented from the colour histogram using the merging method. Figures 4.30-4.33 show the jelly beans with a homogeneous colour which are selected with the use of its 'natural' lookup table.

Figure 4.29 The original image of jelly beans (size is 256 x 256)
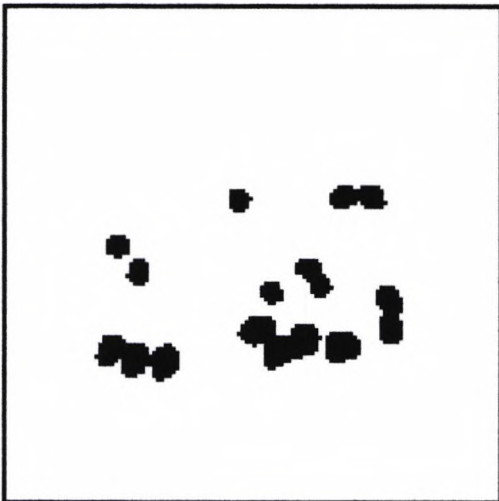
Figure 4.30   Green regions

Figure 4.31 Black regions
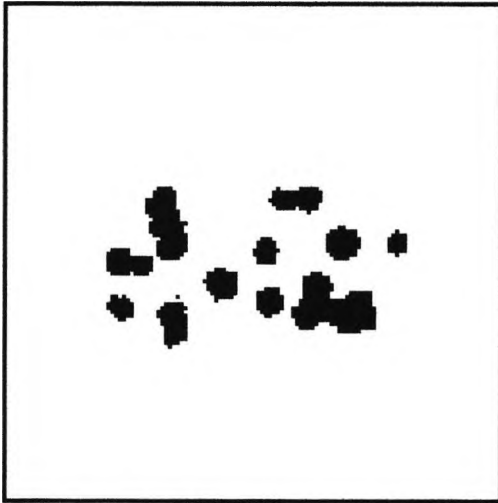
Figure 4.32   Red regions   **Figure 4.33 Yellow regions**
(merging method and colour is selected with the use of lookup table)
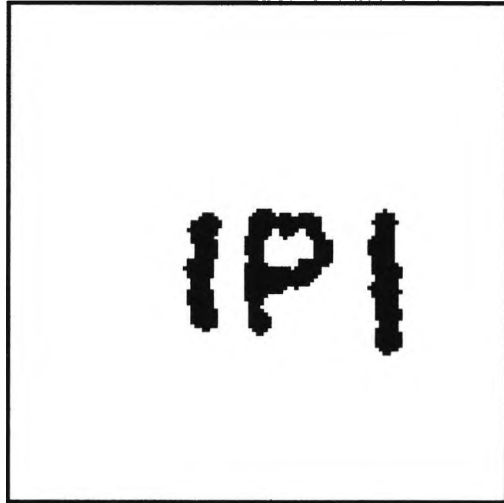
4.9 Results on Classifying Colours

This section consists of two sets of results. The first set is that of the jelly bean image applied to the classifier. The resulting images shown in figures 4.34-4.37 are very similar to the resulting images from the merging algorithm. This simple image (with only five dominant colours) provides a test image for the classifier and the results show the image can be segmented correctly. Since the classifier is simply a Lookup table, the processing time is proportional to the number of pixels. The second set is an image (see figure 4.38) which consists of several different coloured objects. The classified image is applied to the spatial clustering process (pixels with the same class are merged) and the average colour of the region is calculated. Figure 4.39 shows the resulting image. Also the results show that the different coloured objects are reliably segmented (see figures 4.40-4.43). However the region of specular highlight on the object surface is also segmented, because the colours are different from each other and the colour of the specular highlight is closer to white. Although there are specular highlights on the red jelly bean in figure 4.29, their colours are less close to white when compared with the distance to the red.
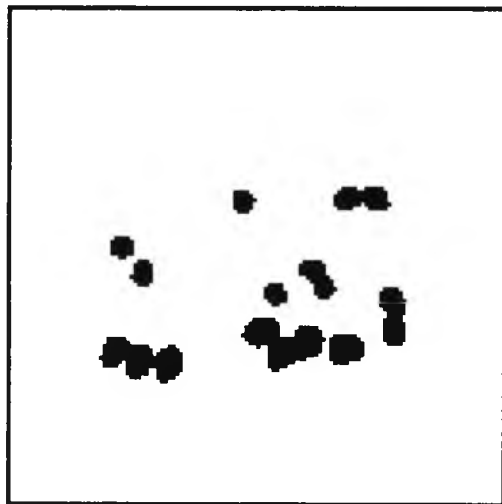
Figure 4.34   Green Regions
(Classifier method)
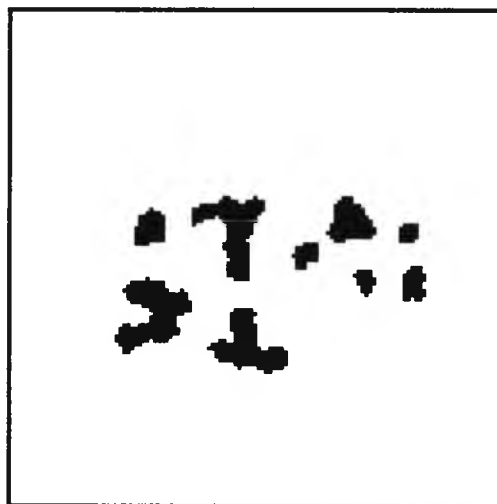
Figure 4.35 Black Regions
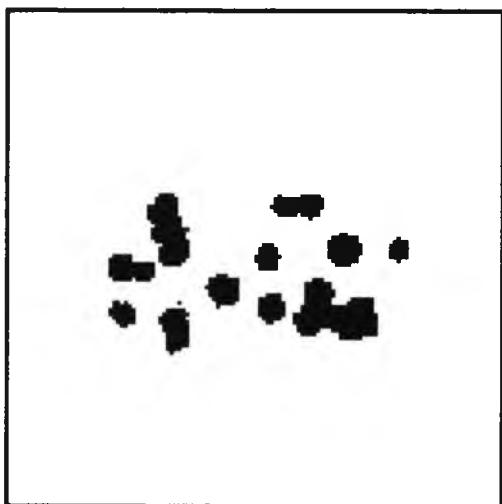
Figure 4.36   Red Regions
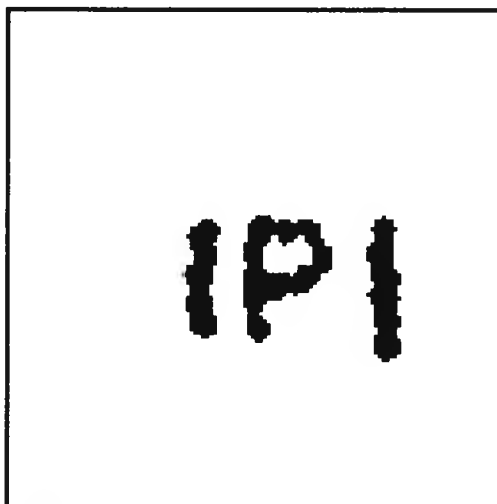
Figure 4.37 Yellow Regions

Figure 4.38 The original image, it consists of different colour objects.
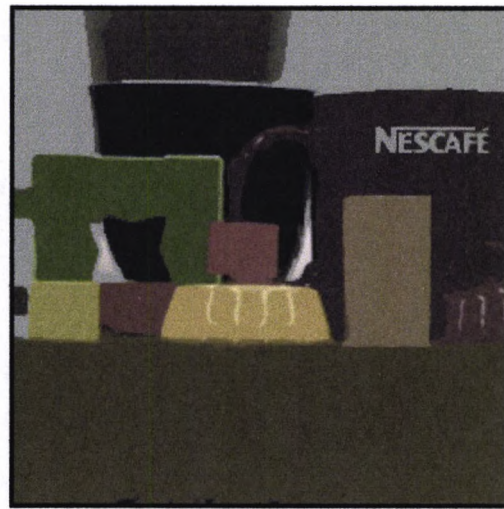
Figure 4.39 The resulting image from the classifier.
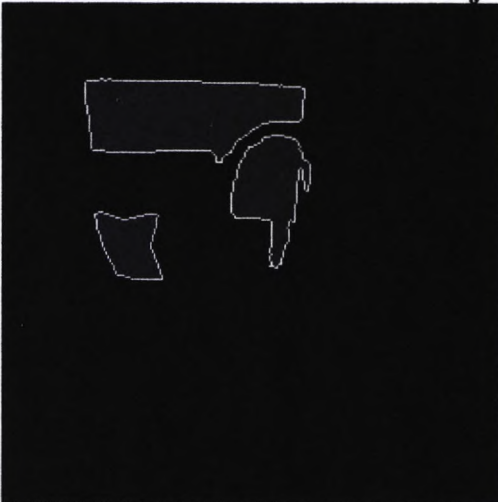
Figure 4.40 Blue Regions (Using the classifier approach)

Figure 4.41 Green Region
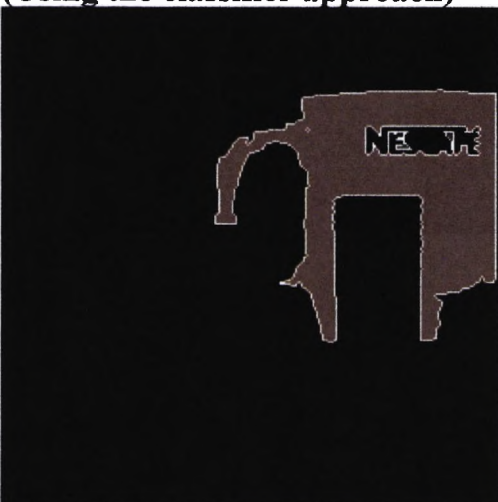
Figure 4.42 Red Region (Using the classifier approach)
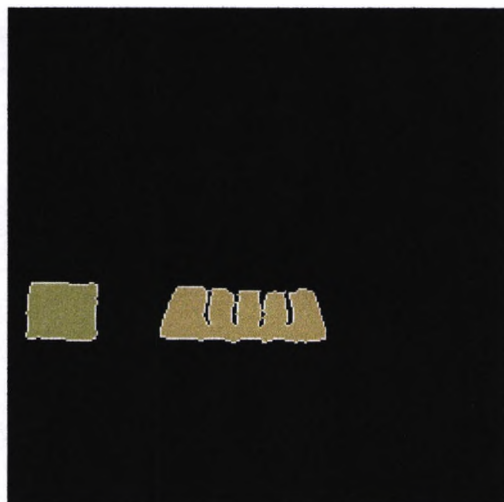
Figure 4.43 Yellow Regions

4.10 Conclusion

Two approaches for segmenting colour space have been studied.

The merging algorithm based on mode (peak) finding has been used by a number of researchers and good results have been shown in their experiments. The algorithms require no parameters, such as the number of clusters or number of iterations. It is also able to adapt to different shapes of clusters. However the remaining problem is the determination of appropriate cell size of the histogram. With the use of Peano scan encoding as a pre-processing step, it is possible to adjust the cell size with respect to the distributions of histogram, and the result shows that most of the significant clusters are found in the histogram. Moreover, the technique can be applied to other feature spaces, for example, texture, but it would require an appropriate representation of texture of a region.

The segmentation based on classifying pixel colours into different classes is more effective because no pre-processing is required. It is simply a direct mapping between the pixel values and the classes through the Colour Information Lookup Table (CILUT, see section 2.6 in chapter 2). The objects should be viewed under the condition that does not cause any shading or specular highlights. It is necessary to capture the test images and model the object surfaces under the same lighting conditions. This is because the image colour is dependent on the light impinging on the surface and its reflectance. More details on how to achieve constant image colour under different lighting conditions are described in chapter 5. Despite the constraints of this approach, it is appropriate for object inspection in a manufacturing plant where the viewing and lighting conditions can be controlled. The beauty of the algorithm is its efficiency. The classifier is only applied to the pixel and such a process can be run in parallel. Also, the classifier and Colour Lookup Table are simple to construct.

Chapter 5
Colour Constancy

"All objects that are known to us from experience, or that we regard as familiar by their colour, we see through the spectacles of memory colour."
Hering 1878[1]

Abstract:

In this chapter, it is shown that the use of memory colour can be found in various colour constancy algorithms and the mechanics of these algorithms are to drive the observed colours close to the memory colour(s) in the colour space. A new colour constancy algorithm is proposed based on fitting the observed colours in the image to a set of memory colours in the colour database. The fitness criterion employs a least mean square error measure. A by-product of the match is a diagonal matrix for spectral transformation which enables the system to adapt to varying illumination conditions. Also, it presents the details of this simplified colour constancy algorithm, and provides results of its application to a number of test images

---

[1] Original sentence is in Germen, translated to English by L.M. Hurvich & D. Jameson [HJ64]

5.1 Introduction

It is common practice in current vision systems to perform object recognition by matching geometric model features with appropriate image primitives (e.g. corners, lines, surface detail). However, when the geometric objects' features are poorly detected in the image, due to lack of resolution (e.g. distant objects) or occlusion, such methods are unreliable. Although colour is an important attribute of objects, there are particular problems associated with employing colour analysis to aid object recognition strategies. Because an object's sensed colour results from an interaction between the spectral reflection properties of the object surface and the spectral properties of the illuminant, the colour of the object is not constant under different illuminants. For instance, an object's colour observed under a tungsten filament lamp will be quite different from that if neon or natural lighting was employed. Comparing the difference between the two colour distributions in the colour space which are shown in figure 5.0 and 5.1, one sees that there is a significant difference between the colour of the same set of objects when viewed under different lighting conditions. The problem is further complicated by the effects of mutual illumination from proximal surfaces, which will contain only a modified subset of the original illuminant, and by the presence of specular highlights.
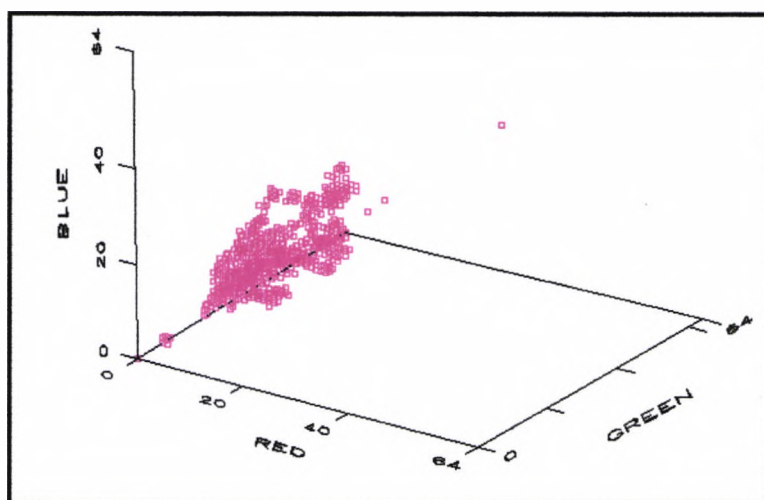
Figure 5.0 The distribution of a set of objects' colours observed under a tungsten filament lamp.
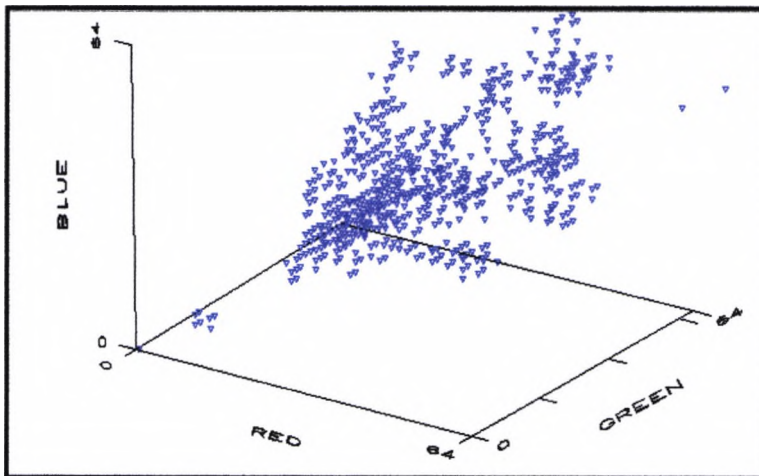
Figure 5.1 The distribution of same set of objects' colours observed (using the same settings of the camera) under bright sunlight.

Colour constancy (or spectral adaptation) is an important mechanism in human colour vision and has been widely investigated. This process enables humans to perceive constant colour of a surface over a wide range of lighting conditions. There have been several studies on constructing an algorithm which can mimic such a useful mechanism and they are based on the physical properties of surfaces, illumination and sensors, using mathematical tools and heuristic assumptions for investigation. Such studies can be split into two major issues: First, is what form of a transformation should be used in the algorithm to discount the effect of the illuminant; second, how are the parameters of the desired transformation computed. In this chapter, the work is focused on the second issue.

To date the most powerful colour constancy algorithm was proposed by Forsyth [For90] and was named Crule. Finlayson [Fin95] later reduced the complexity of the Crule algorithm by reducing the dimensionality of the colour space. A review of their algorithms and others are described in section 5.2. Section 5.3 illustrates the similarities between various algorithms including Crule. It shows that they were using memory colour(s) in their algorithms and the mechanics of the algorithms are to drive the observed colours close to the memory colour(s) in the colour space.

In this chapter, I propose a new algorithm based on the matching between memory colour and the observed colour from the image. It exploits some of the existing methods. A brief description of the algorithm is given below.

The process of matching geometric features typically relies on extracting primitives from the image, followed by hypothesis generation and verification with geometric information stored in the form of a CAD database. This method can also be applied to matching colour information. Two questions have to be considered: what parameters should be used for the hypothesis; and what should be the criterion for the verification of the hypothesis. The parameters of the hypothesis used in this chapter are the elements of a diagonal matrix transformation which had proven sufficient [Fin95] for discounting the effect of changing illuminants. The verification of the hypothesis is based on the fitness between the two sets of colours (observed colours and memory colours in the database). The fitness is quantified by employing the least mean square error measure. A colour database is employed, which must contain a number of colours that are similar to those which appear in the scene (The database colours may have been determined under different lighting conditions). The philosophy of this approach and the others that are described in section 5.3 are found to be cognate with the belief of Hering, *"All objects that are known to us from experience, or that we regard as familiar by their colour, we see through the spectacles of memory colour".*

Mondriaan images are used to generate the colour database, and have also been used initially to test the operation and reliability of the algorithm, which has proved to be robust in estimating the colour of coloured surfaces imaged under varying illumination conditions. Results of applying the method to a number of more realistic images are also presented. The results in section 5.7 show the algorithm performing effectively to aid the localisation and identification of objects. The major components of this chapter are displayed in figure 5.2
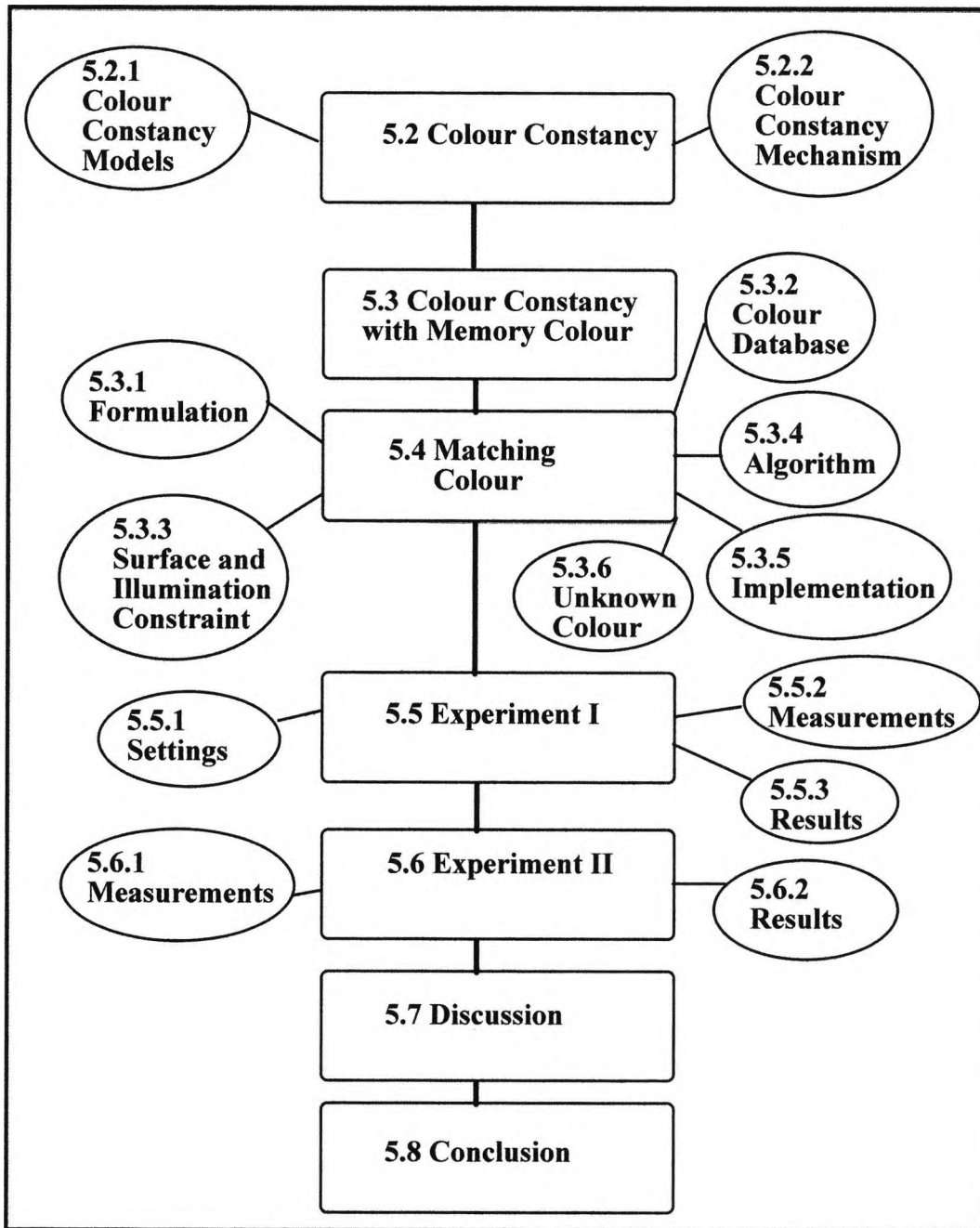
Figure 5.2 The structure of chapter 5.

5.2 Colour Constancy

The colour of a surface is an important feature of an object that can be used for its localisation and identification. However the observed colour of a surface is dependent on the illumination, because light reflected from the surface is the product of illumination and surface reflectance (see eqn. 2.1,

2.2 and 2.3 in chapter 2). Therefore when the spectral content of the illumination changes, the observed colour of a surface will change. Colour constancy is the ability to recover surface descriptors from observed colours that do not change as the illumination changes. This ability is embedded in the human visual system, allowing us to approximate the constant colour of the surface over a range of illumination conditions.

In this chapter, I split the reviews on colour constancy into two major parts: first is the model of colour constancy; second is the mechanism of the algorithms. The first part describes the several colour constancy theories and different approaches. The second part concentrates on the mechanism of various algorithms.

Most of the models are based on a number of assumptions which are related to the physical properties of colour patches and the spatial and spectral distribution of the observed colour. The assumptions can be divided into two sets: first is the Mondriaan world assumptions and dielectric material assumptions; second is the heuristic assumptions about spectral distribution of the observed colours.

The Mondriaan world is commonly used for generating test images for the constancy algorithms or stimuli for human colour vision research. Some assumptions of the Mondriaan world which are commonly employed in those algorithms are listed below:

• A set of squared colour patches.
• A reasonable number of different coloured patches.
• The surfaces properties of the colour patches are Lambertian.
• A single source of illumination.
• An even illumination over the scene.

This set of assumptions is commonly considered as the standard for the research on colour constancy algorithms.

5.2.1 Colour Constancy Models

The earliest and simplest coefficient adaptation proposed by Von Kriss has been proven [Fin95] to be the most general model for colour constancy. This model changes the coefficient of each colour channel or the coefficient of the basis functions which are used to describe the colour response of the surfaces. Other models [LM71, MW86, FD88, For90] are considered in the same class of coefficient adaptation models with appropriate sensor transformations [Fin95]. Since the sensor transformation is fixed and dependant on the spectral characteristics of the sensors rather than the illuminant, it is not considered as part of the adaptation process. The following paragraphs describe several colour constancy models.

Land's Retinex theory [Lan85] attempted to explain the mechanism of human colour perception. The colour descriptors of the observed surfaces are determined by a triplet of numbers, and each computed on a single waveband. The three wavebands are low, medium and high which are referring to red, green and blue respectively. The algorithm computes a relative colour descriptor for each surface, and the value of the relative colour descriptor is dependent on the spatial arrangement of surfaces. This is because the colour descriptor of each patch is computed with respect to its neighbourhood patches' colour (See eqn. 5.9 & 5.10).

A finite dimensional model is popular for approximating the surface reflectance functions and illuminants. Several researchers [JMW64, WS67 Dix78, PHJ89] have shown that low dimensional models (e.g. using three to five basis functions) are adequate for approximating the surface reflectance and illuminant function. Such a model allows the analysis and manipulation of these functions with linear algebra. Therefore the colour descriptor of a surface can be represented by a set of coefficients of the basis functions. The colour constancy models described in the next few paragraphs employ the finite dimensional model for a more accurate approximation of the spectral characteristic of the reflectance and illumination.

Maloney & Wandell's [MW86] world of colour surface reflectance and illuminant functions are different from other researchers. The surface reflectance functions are modelled with two basis functions and the spectral density function of the illumination is modelled by three basis functions. The number of sensors must be greater than the number of basis functions for describing surface reflectance functions. The number of colour surfaces that appeared on the images must be greater than the number of basis functions of the illuminations. With all the required conditions, the algorithm is able to uniquely compute a matrix transformation that discounts the effect of the illumination, by solving a set of simultaneous equations. The equations are shown below. By computing the inverse of Λ_ε and multiplied by equation 5.4, the colour descriptor of each surface will be independent of the illuminant (see eqn. 5.5).

$$\rho_k^x = \int E(\lambda)S^x(\lambda)R_k(\lambda)d\lambda \cdots\cdots (5.1)$$

ρ_k^x, the sensor respone at location x,
and k class of sensors.

$$S^x(\lambda) = \sum_{j=1}^{n} \sigma_j^x S_j(\lambda) \cdots\cdots (5.2)$$

$S^x(\lambda)$ is the surface reflectance function at location x,

$$E(\lambda) = \sum_{j=1}^{n} \varepsilon_j E_j(\lambda) \cdots\cdots (5.3)$$

$E(\lambda)$ is function for ambient illuminant.
substituting eqn. 5.2 & 5.3 into eqn. 5.1

$$\rho^x = \Lambda_\varepsilon \sigma^x \cdots\cdots (5.4)$$

where Λ_ε is representing the illuminant matrix.

$$\Lambda_\varepsilon^{-1}\rho^x = \sigma^x = \Lambda_\varepsilon^{-1}\Lambda_\varepsilon\sigma^x \cdots\cdots (5.5)$$

Forsyth [For90] had developed a body of mathematics to explain the transformation that discounts the effect of changing illumination and can be sufficiently represented by a diagonal matrix, if the photoreceptor spectral sensitivities are nearly disjoint. His algorithm is different from others: the approach is based on the surface reflectance constraints that will limit the feasible set of parameters for matrix $M_{i,c}$. $M_{i,c}$ is described in eqn. 5.6.

$\underline{c}_c$ is the colour descriptor of a surface observed under a canonical illuminant.

$\underline{c}_i$ is the colour descriptor of a surface observed under illuminant i.

$\underline{c}_{m,i}$ is the estimated colour descriptor.

Use a diagonal matrix $M_{i,c}$ to compensate for the difference between the canonical illuminant and illuminant i.

$$\underline{c}_c \approx \underline{c}_{m,i} = M_{i,c}\,\underline{c}_i \quad \cdots\cdots (5.6)$$

Ho, Funt and Drew [HFD90] attempted to separate the colour signal into two components, illumination and reflection. The model required the measurement of the complete spectral response of an observed colour surface, which is not simple to obtain from a conventional colour CCD camera. However, they mentioned that such a measurement can be obtained from the image based on the effect of chromatic aberration [FH89]. In order to estimate the coefficients of the basis functions for the illumination and surface reflection functions, a few constraints are applied to limit the characteristic of the basis functions. Then, the coefficients of the estimated function will be approximated, based on the error between the measured spectral response of colour surface and the estimated functions (eqn. 5.7). The experiment was made on the simulated surface reflection and illumination data, no real images were used.

Combine eqn 5.2 and 5.3

$$I(\lambda) \approx \sum_{i=1}^{m} \sum_{j=1}^{n} \varepsilon_i \sigma_j E_i(\lambda) S_j(\lambda) \cdots\cdots (5.7)$$

Approximate models with limited number of basis functions.

$M(\lambda)$ is the spectral response of a colour surface.

$$\text{error} = \sum_{\lambda} \left[\sum_{i=1}^{m} \sum_{j=1}^{n} \varepsilon_i \sigma_j E_i(\lambda) S_j(\lambda) - M(\lambda) \right]^2 \cdots\cdots (5.8)$$

Finding the coefficients of σ_j, ε_i for which error is minimum.

5.2.2 Colour Constancy Algorithms

For a machine vision system to perform colour object recognition with minimum restrictions on the condition of the illuminant, it is desirable to have a reliable colour constancy algorithm [For90, Swa90, MMK93, Fin95]. Also, some have assumed that the colour constancy processing should be performed before the model matching.

In this section, several algorithms are described which are generally limited to cope with Mondriaan-type images. All are based on a set of assumptions about the spatial and spectral distribution of the colours within the scene and these assumptions play an important role in the way of constructing the algorithms.

Some algorithms [Buc80, GJT87a,b] are based on the assumption that the average of the spectral response of all surfaces in the scene is constant for a wide range of situations. Hence, their surface colour descriptors can be computed by rescaling the response of all surfaces according to the ratio between average response of all surfaces and the predefined average. Such assumptions for processing a single image are quite easy to confound, since the combination of colour surfaces can be biased toward one particular colour. However, if a wide range of images is included in the process, the biased arrangement of colour surfaces may be reduced.

The final part of the Retinex theory [Lan85] is an algorithm to compute the illuminant invariant colour descriptors of surfaces. The colour descriptor of any surface i is represented by the average of all the logarithmic ratios between surface i and another (see eqn. 5.9-5.11). The assumptions of the algorithm are quite restrictive, because only one set of images is shown in the experimental results. Brainard & Wandell [BW86a,b] have made an analysis of the algorithm, showing that it is a form of normalisation process with a reference patch and it is not a plausible model for human colour vision. Later, Forsyth [For90] compared his algorithm with the Retinex algorithm, concluding that the different spatial arrangements of a set of colour surfaces will result in different colour descriptors for each surface.

The relationship of surface *i* and *j*:

$$R^{\Lambda}(i,j) = \sum_k \delta \log \frac{I_{k+1}}{I_k} \cdots\cdots (5.9)$$

Λ is representing one of the principle wavebands (red, green and blue)

$$\delta \log \frac{I_{k+1}}{I_k} = \begin{cases} \log \dfrac{I_{k+1}}{I_k} & if \left| \log \dfrac{I_{k+1}}{I_k} \right| > threshold, \\ 0 & if \left| \log \dfrac{I_{k+1}}{I_k} \right| < threshold. \end{cases} \cdots\cdots (5.10)$$

Average of relationships at surface i:

$$\overline{R}^{\Lambda}(i) = \frac{\sum\limits_{j=1}^{N} R^{\Lambda}(i,j)}{N} \cdots\cdots (5.11)$$

Another set of assumptions [DL86, Lee86] is that the material is an inhomogeneous dielectric and the spectral response of its specularity highlight is identical to the spectral density function of the illumination. Therefore, the surface descriptors can be computed by the multiplication of the inverse of the illumination matrix and the observed colour values. The algorithms rely on detecting the specularities of the observed materials. Their techniques for identifying the specularity are based on the dichromatic reflection model [Sha85], described in eqn. 5.12. The model represents two parts of light reflected from a surface. The interface and body parts are represented by L_f and L_b respectively. This assumption is different from that for Mondriaan images, where all surfaces are Lambertian and the algorithm will fail for this case. From eqn. 5.11, one can see the proportion of body and interface is also dependent on the angle of light impinging on the surfaces, so the illuminant is not homogeneous over the image. Again, this is different from the common assumption for Mondriaan images. Consequently the algorithms can only compute the chromatic feature of the illumination, since the intensity is not constant over the image.

$$L(\lambda,i,e,g) = L_f(\lambda,i,e,g) + L_b(\lambda,i,e,g)$$
$$= m_f(i,e,g)c_f(\lambda) + m_i(i,e,g)c_b(\lambda) \cdots\cdots (5.11)$$

where i is the angle of incidence of light,

e is the angle of exitance (reflected light), g is the phase angle.

A practical approach to colour constancy was suggested by Tsukada & Ohta [TO90] by which the geometrical feature of the objects is used for initial identification. When an object is identified, the spectral adaptation begins by computing the difference between the colour of an observed surface and its colour in the object database. It should be considered that this approach is a constrained algorithm since it required certain objects to be present in the scene.

Forsyth's novel algorithm called Crule [For90] exploits the surface reflectance constraints to limit the mappings ($M_{i,c}$, see eqn. 5.6 in section 5.2.1) between the colour gamut of an observed scene under an unknown illuminant and the colour gamut of the same scene under a canonical illuminant. Furthermore, Finlayson [Fin95] limited the mappings by imposing an illumination constraint and reducing the parameters of the mapping. It is easier to understand through the graphical representation of the colour gamut and the feasible mappings. Figure 5.3 shows the graphical representation of an observed image gamut in a two-dimensional colour space. Each set of feasible mappings is constructed by finding the mappings between one of the hull points of the observed colour gamut under an unknown illuminant to all the hull points of the colour gamut of possible surface reflectances. The colour gamut of possible surface reflectances is assumed to include all the surface colours in the world (ideally) and can be seen in figure 5.4 and the convex hull of the data is considered as the surface reflectance constraints. The intersection of all sets of feasible mappings can be seen in figure 5.5, showing the mappings for $M_{i,c}$ with two parameters, m_1 and m_2.

The key of this approach is to explore the physical properties of the surface reflectance and illumination, then apply the constraints to limit the set of mappings (see figure 5.5). The blue patterned area in figure 5.5 is the remaining set of mappings which satisfy all the physical conditions. The last part of both algorithms is to compute the trace[1] of each remaining

---

[1] The sum of the diagonal matrix elements.

mapping. The assumption in the algorithm is that the best mapping is the one with the maximum trace. In other words, the mechanism of the algorithm is to expand the volume of the observed colour gamut, such that the boundary of the observed colour gamut moves close to and within the boundary of colour gamut of all the possible surface reflectances. From the experimental results [For90], they reported that the increase in the number of distinctive colour surfaces will increase the constancy of the colour descriptor of the observed surface. It can be explained that more distinctive colours appearing in the images will increase the probability of the occurrence of these colours, which correspond to the hull points of the colour gamut of the possible surface reflectance, and more observed colours correspond to the hull points of the possible surface gamut will reduce the set of remaining feasible mappings.
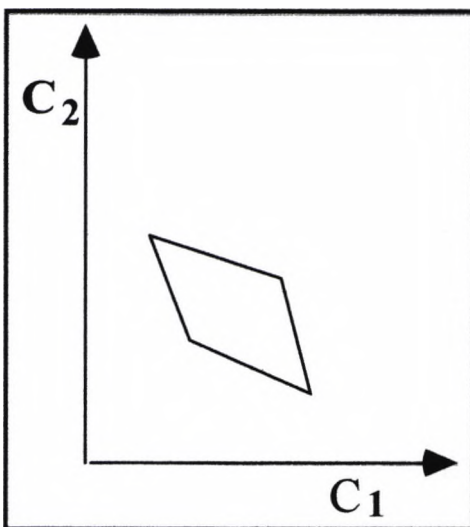
Figure 5.3 The colour gamut of an observed image under an unknown illuminant.
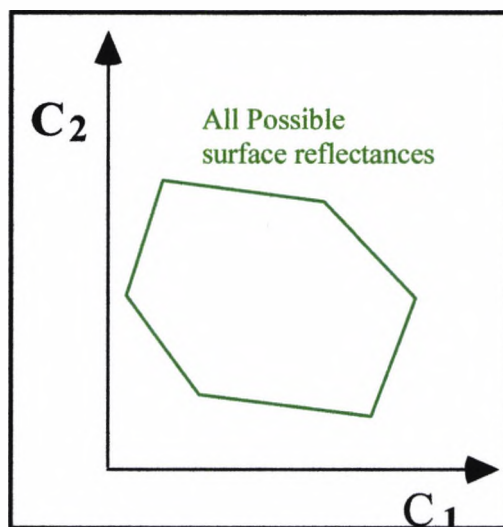
Figure 5.4 The colour gamut of all possible colour surfaces observed under a canonical illuminant, which forms the basis of surface reflectance constraints.
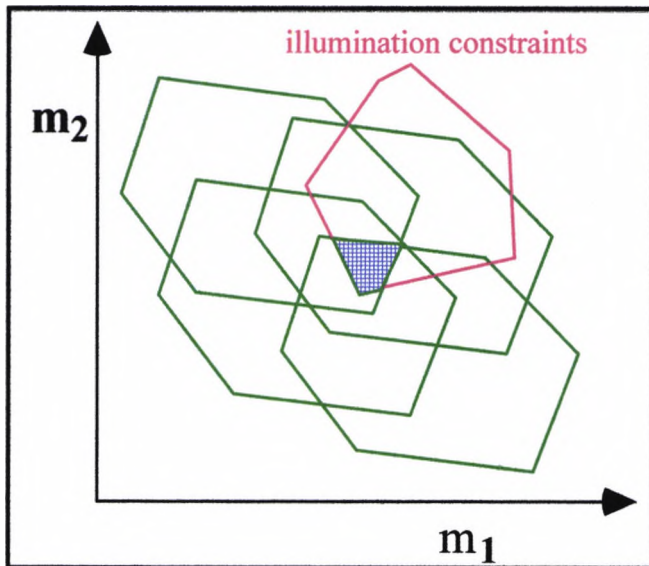
Figure 5.5 The set of feasible mappings and illumination constraints.

5.3 Colour Constancy with Memory Colour

There are several examples [Bri78, Lan85, For90, Fin95, LM71, Buc80, GJT87, NS94, TO90] in which memory colours or reference colours are used in algorithms. The mechanics of the algorithms are to drive the observed colours close to the memory colours or reference colours. The number of memory or reference colours that are used varies from a single colour to a set of colours. All of them expect that the reference or memory colours is obtained by observing the image colours under the canonical lighting condition (ideally it is white light[1]).

The earliest algorithm proposed by Von Kriss [WB86] was based on a local averaging process, which assumes the average of the observed colours corresponds to a pre-defined average. Figure 5.6 shows a graphical visualisation of the algorithm: the gamut of the observed colours is in pink and the transformed colours are in blue. The mechanism is to transform the observed colour close to the predefined average colour. Later Buchsbaum [Buc80] and Gershon et al [GJT88] followed this heuristic assumption with different ways to compute the average of the observed colours.

---

[1] Evenly distributed power density function (pdf) of illuminant.
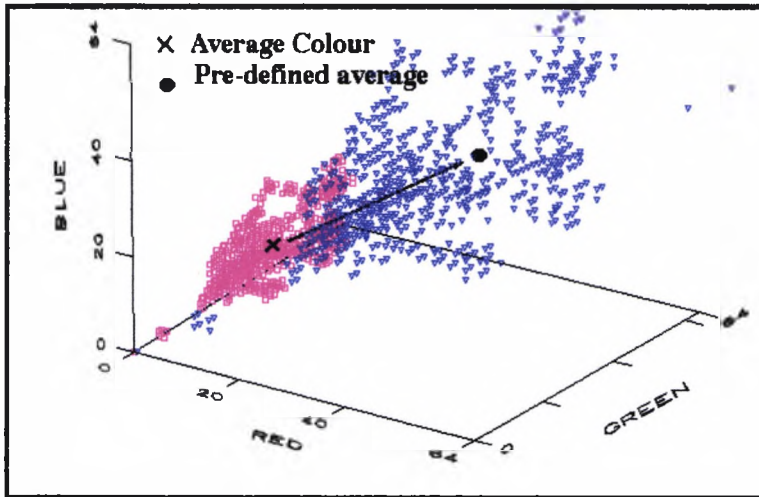
Figure 5.6. The changes of observed colours in the colour space when the averaging colour constancy algorithm is applied.
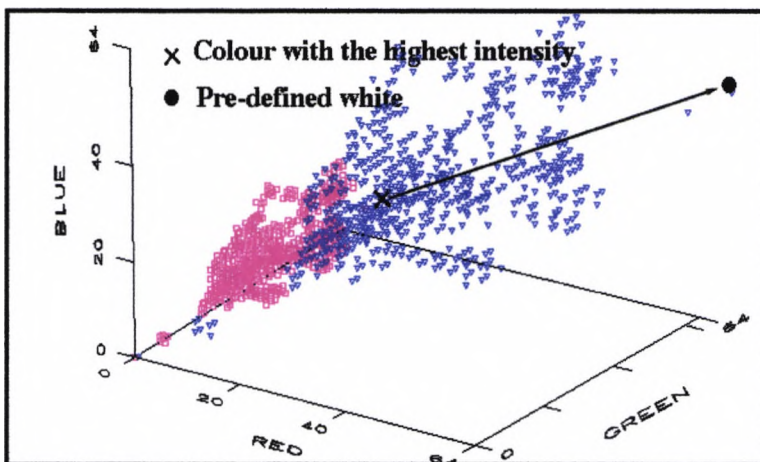
Figure 5.7 The changes of observed colours in the colour space when the colour constancy algorithm with a predefined white is applied (Under a transformation, the colour at location X will become the colour at location ●).

Another version of the algorithm [Lan85] assumes that the brightest patch in the scene is a white patch and it corresponds to a predefined white. Therefore computing the mapping between the colour descriptor of the brightest patch and the predefined white will provide the mapping that transforms the observed colour under an unknown illuminant to the canonical illuminant. Figure 5.7 shows the effect of the algorithm, where the gamut of the observed colours is in pink and the transformed colours are

in blue. Alternatively, others [Bri78, NS94, NS91, TV91] suggested that a white patch or several distinctive colour patches are placed at a known location(s) in the scene, so that the algorithm becomes a calibration process.
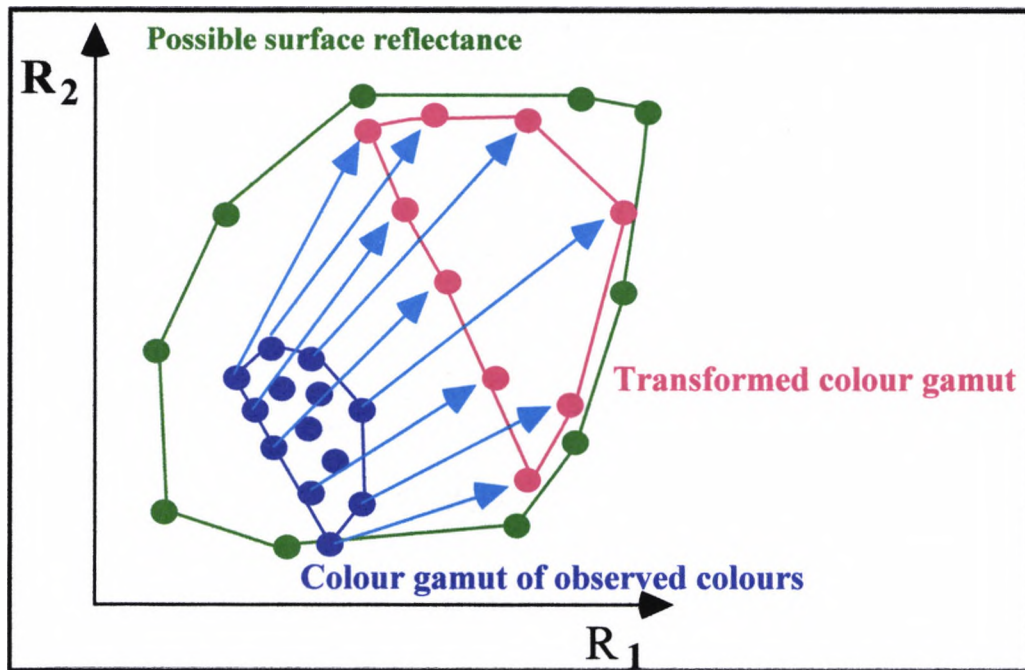
Figure 5.8 The changes of observed colours in the colour space when the Crule algorithm is applied.

Forsyth [For90] has shown that Crule computes a more consistent colour descriptor for each colour surface than other previously proposed algorithms when applied to Mondriaan images. Figure 5.8 shows the effect of the algorithm; the gamut of observed colours, the transformed colours and all the possible surface reflectances are bounded by the blue, red and green lines respectively. The Crule algorithm is applied to a three-dimensional colour space, though figure 5.8 only illustrates the principle of the algorithm in a two-dimensional colour space. In this algorithm, the boundary colour of all possible surface reflectances can be considered as the memory colours and the mechanics of the algorithm is to transform the colour close to but without moving beyond, the boundary. Also, one will realise that if the observed colours do not correspond to any of the boundary colours of all possible surface reflectances, the constancy of the

colour descriptors of observed colour patches will be reduced. Forsyth had concluded in his thesis that more distinctive colour patches appearing in the scene will increase the constancy of the colour descriptors. It can be easily explained since more distinctive colours will increase the chance of the occurrence of the boundary colours in the scene.

All these algorithms are easily confounded by a situation where the majority of the observed colours are biased toward a small range of similar colours. Also changes of the predefined average, white or the boundary colours will affect the values of the recovered colour descriptors. For example, given the set of observed colours, but using a different predefined average colour, the resulting colour descriptors of the patches will be different. Therefore Hering's belief, "All objects that are known to us from experience, or that we regard as familiar by their colour, we see through the spectacles of memory colour", can be applied in these situations.

The current thinking in solving such a problem is to constrain the mappings between the observed colour and memory or reference colours. Surface reflectance constraints [For90] and illumination constraints [Fin95] are the typical examples which are able to limit the mappings. In this chapter, the algorithm proposed has exploited those constraints, and further constrains the possible mappings by introducing a fitness measure between the observed colours and memory colours.

5.4 Colour Matching

In this research, the problem is defined differently from other research on colour constancy algorithms for machine vision, since the colour database is incorporated into the overall system. It computes a relative colour descriptor of each patch with respect to the colour in the colour database. If a set of colour patches is observed under two different lighting condition, the computed colour descriptors with respect to the same colour database will be the same. The process of computing the relative colour descriptor is based on transforming the observed colours (which are captured under a different lighting condition) to a 'normal' colour space where the colour

database is formed, and find the optimal match between the observed colours and the colours (memory colours) in the database. The colour constancy is assumed to be the by-product of this matching.

Therefore, for a set of observed colours under an unknown illumination, the question is how to compute the parameters of a transformation that will transform the observed colours close to colours in the colour database. The colour gamut of a set of observed colour surfaces under unknown illuminant is shown in figure 5.9. The set of memory colours (colour database) as modelled from the real images is shown in figure 5.10. All the observed colours correspond to the colours in the database and the changes of a colour's position in colour space are due to changes of the illumination. Therefore if an appropriate and consistent mapping can be applied to the observed colours to reduce the discrepancies between observed colours and the colours in the database, then we can achieve colour constancy. Figure 5.11 shows the result of such a transformation.
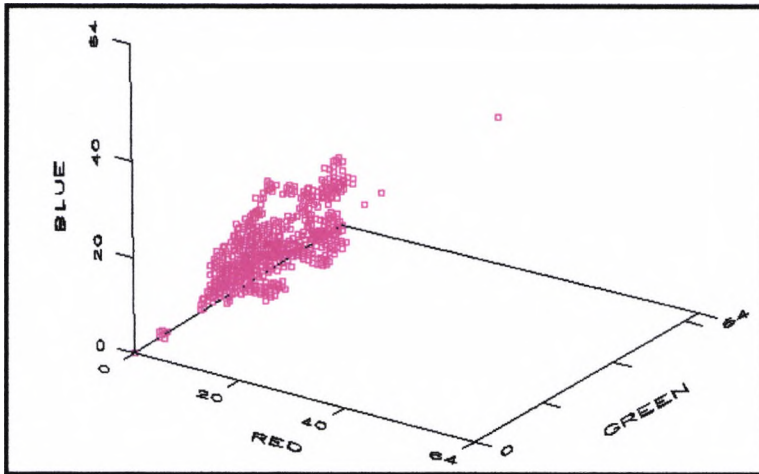
Figure 5.9 The colour gamut of observed colour surfaces in the colour space.
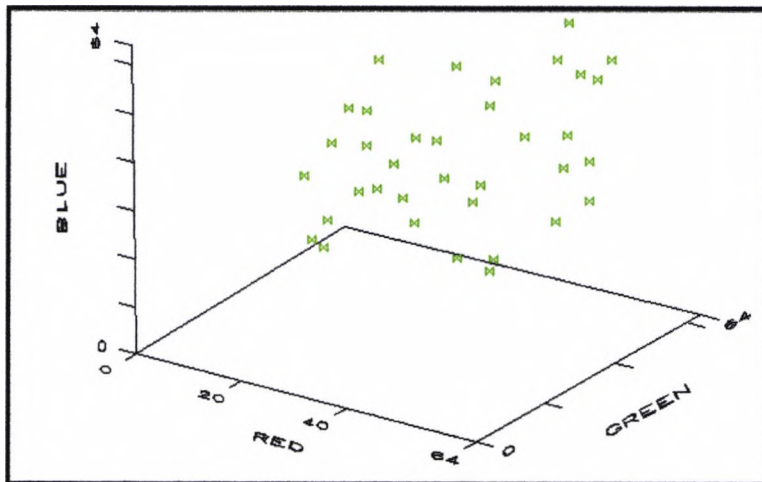
Figure 5.10 The plot of memory colours in colour space.
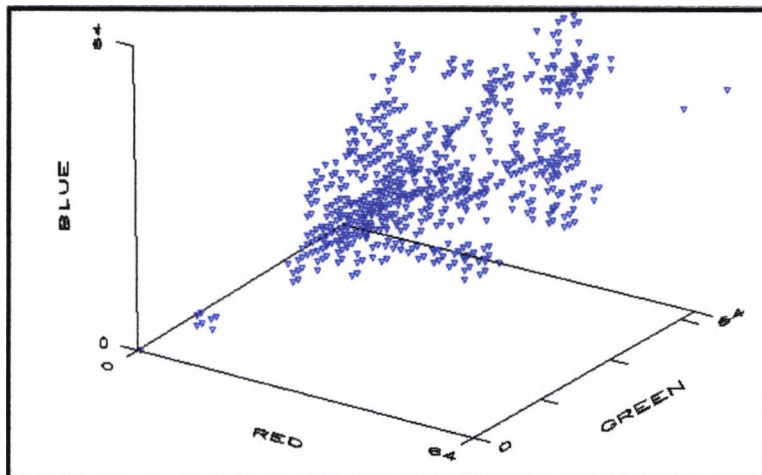
Figure 5.11 The transformed colour gamut in the colour space.

Now the problem has been defined, the following questions are raised:

- what is the form of the transformation which will be sufficient to discount the unknown illuminant?
- what physical properties of illumination and surface can be exploited by the system?
- what are the settings for the experiments?
- how to model the colour surfaces?
- how to quantify the measure of fitness between the set of observed colours and the memory colours in the colour database?

Some of these questions have been addressed by other researchers. The diagonal matrix transformation has been shown [Fin95] to give a simple solution for colour constancy and provides an appropriate transformation when applied with respect to the response of the filters [Fin95, For90] that are used in the image acquisition process. Also, a heuristic approach [WB86] based on the analysis of the Von Kriss adaptation has shown that the choice of filters is important for the accuracy of the colour constancy system.

The set of all possible surface reflectances has been proven to be convex [For90]. This is useful for limiting the set of feasible mappings between the observed colours under an unknown illuminant and the colours under a canonical illuminant. In model-based vision, most of object surfaces can be modelled and stored in a form of database. This database can be used to construct the surface reflectance constraints. Therefore, the colour response of significant or dominant colour surfaces found in the image should be within the colour gamut of all possible surface reflectances. Another useful constraint proposed by Finlayson [Fin95] is the illumination constraint that further limits the possible illumination mappings. The construction of an illumination constraint can be considered as another modelling routine in model-based vision where certain elements of the environment are controllable and within a limited set. For example, in a laboratory environment where the possible sources of illuminations are daylight, fluorescent light or a mixture of both, the illumination constraint can be easily modelled. In the Mondriaan world, these constraints are applicable, because other physical effects, such as mutual illumination, three dimensional objects and dielectric materials are assumed to be absent from the scene. Also Finlayson [Fin95, FH97] constructed an algorithm, 'colour in perspective', to reduce the complexity of the Crule approach by decreasing the dimensionality of the colour space.

The system was initially tested with a Mondriaan image, which provides a common test ground for different algorithms; their simplicity provide a good starting point for any system to evolve to cope with more complicated and realistic assumptions about the world.

5.4.1 Formulation

The fitness measure between the observed colours from the image and the colours in colour database is based on a least mean square error measure. The hypothesised transformation is first generated and then verified based on this fitness measure. It assumes that the observed colours correspond to the existing colours in the colour database and that the changes of value of the colour descriptor in the colour space are due to changes of illumination. The following equations define the error between the two sets of colours.

Let the colours modelled using illumination i in the database be represented by $\underline{C}_{i,1}, \underline{C}_{i,2}, \underline{C}_{i,3}, \cdots \underline{C}_{i,j}, \cdots \underline{C}_{i,n}$. n is number of colours in the colour database. The observed colours from an image under an unknown illuminant are represented by $\underline{C}_{u,1}, \underline{C}_{u,2}, \underline{C}_{u,3}, \cdots \underline{C}_{u,k}, \cdots \underline{C}_{u,m}$. m is number of observed colours in the image. The colour descriptor of a modelled surface j observed under illuminant i is represented by $\underline{C}_{i,j}$. The colour descriptor of surface k observed under illuminant u is represented by $\underline{C}_{u,k}$

$$\underline{C}_{i,j} = M_i \underline{S}_j \cdots\cdots (5.12)$$

where M_i is the illumination matrix i

and $\underline{S}_j$ is the vector for surface reflectance j.

$$\underline{C}_{u,k} = M_u \underline{S}_k \cdots\cdots (5.13)$$

where M_u is the illumination matrix u

and $\underline{S}_k$ is the vector for surface reflectance k.

Suppose the surfaces observed in the image correspond to a colour in the database. It is required to compute a transformation $M_{i,u}$ to account for the difference between illuminant u and i.

$$\underline{C}_{i,j} \approx M_{i,u} \underline{C}_{u,k} \cdots\cdots (5.14)$$

This transforms the observed colours to their corresponding colours in the colour database as close as possible. $err_{j,k}$ is the error between the transformed colour k from the image and the colour j in the database. $M_{i,u}$ is the hypothesis transformation between illuminant i and illuminant u. $M_{i,u}\underline{c}_{u,k}$ and $\underline{c}_{i,j}$ are vectors; the error between them is measured by calculating the Mahalanobis distance. The variability of the distribution of surface j is represented by covariance matrix $Cov_{i,j}$.

Mahalanobis distance is $err_{j,k}$,

$$err_{j,k} = \left[M_{i,u}\underline{c}_{u,k} - \underline{c}_{i,j} \right]^\mathsf{T} Cov_{i,j} \left[M_{i,u}\underline{c}_{u,k} - \underline{c}_{i,j} \right] \cdots\cdots (5.15)$$

For each hypothesis p and $p = 1...n$, $M_{i,u,p}$, the overall error measure associated with the accumulation of distances between the transformed colours and their corresponding colours in the colour database is computed.

The overall error $Oerr_p$ for the hypothesis transformation matrix $M_{i,u,p}$, is :

$$Oerr_p = \sum_{k=1}^{m} \left[M_{i,u}\underline{c}_{u,k} - \underline{c}_{i,j} \right]^\mathsf{T} Cov_{i,j} \left[M_{i,u}\underline{c}_{u,k} - \underline{c}_{i,j} \right] \cdots\cdots (5.16)$$

All the hypothesis transformations are tested against the constraints and the corresponding $Oerr_p$ will be compared.

5.4.2 Colour Database

The use of a colour database is a particular feature of the overall system, because the observed colours under an unknown illumination are transformed on the basis of the fit between the set of observed (image) colours and the memory colour in the colour database. Memory colours in the colour database can be modified based on knowledge about the local environment (see chapter 2, section 2.4).

The colour database was generated from a Mondriaan image of colour patches which are Lambertian. (It is also possible to model some colour

patches with a dielectric characteristic if an appropriate highlight detection process is used for pre-processing. However, this type of material is not considered in this thesis.) All colour patches are imaged under the same illumination. This is the learning process for the system. The system records the average colour descriptor of each colour patch and the variability of the distribution in the colour descriptors. The variability originates from the slight inconsistency of the surfaces' reflectance. These distributions are assumed to be Gaussian. The common approach for representing a Gaussian distribution in multi-dimensional space is to compute the mean vector and covariance matrix. The equations are shown below.

Colour surface j.

Mean colour descriptor of surface j under illumination i

$$\widetilde{\underline{c}}_{i,j} = \frac{\sum_{q=1}^{n} \underline{c}_{i,j,q}}{n} \cdots\cdots (5.17)$$

where n is the number of example pixels.

Covariance matrix of colour surface i.

$$Cov_{i,j} = \frac{\sum_{q=1}^{n} \left(\widetilde{\underline{c}}_{i,j} - \underline{c}_{i,j,q} \right)^2}{n} \cdots\cdots (5.18)$$

where n is the number of example pixels.

5.4.3 Surface and Illumination Constraint

The set of all possible surface reflectances has been shown to be convex, but this set has to be modelled by real data since no mathematical model is found for constructing the set. In this system, all available colour patches are imaged under the same illumination (the ideal illuminant is white

light[1]). The set is formed by finding the convex hull of the colour gamut. This set provides the part of the constraint to confine any hypothesis transformation. The other part of the constraint is dependent on the colours found in the image (see section 5.2.2, Crule algorithm). The greater the variety of colours appearing in the image, the greater the constraint. The hypothesis transformation is applied to all observed colours and the transformed colours are tested against the set of all possible surface reflectances. However, Forsyth [For90] proved that the process can be simplified by finding the intersection set of all mappings that map the hull points of the colour gamut of the image to the set of all possible surface reflectances (see the diagram in section 5.2.2). The intersection set is the surface reflectance constraints which is used to confine the set of hypothesis transformations.

The illumination constraint is the set of the inverse of illumination matrix multiplied by the illumination matrix that is used to model the memory colour. The hypothesis transformation aims to discount the changes of illuminants and the illumination constraint is used to confine them.

$\underline{c}_{u,j}$ is the sensed colour under an unknown illuminant u

$$\underline{c}_{u,j} = M_u \underline{s}_j$$

where M_u is the matrix for an unknown illuminant.

Since the colours in the database is modelled under different illuminant i,

$\underline{c}_{i,j}$ is the memory colour j modelled under an unknown illuminant i

$$\underline{c}_{i,j} = M_i \underline{s}_j$$

where M_i is the matrix for illuminant i.

To transform the observed colour j to memory colour j, requires two matrix multiplications.

---

[1] Evenly distributed power density function (pdf) of illuminant.

$$\underline{c}_{i,j} = M_i M_u^{-1} M_u \underline{s}_j \cdots\cdots (5.19)$$

Therefore for each possible illuminant m, it is required compute $M_i M_m^{-1}$ and the convex hull of this set will be the illumination constraint. The illumination constraint only uses the chromatic part of the illumination because the intensity of the illumination will not be confined.

5.4.4 Algorithm

The system is divided into two parts. The first part is an off-line process that is used to model the variability of colour surfaces. The representation of a colour surface consists of its mean colour and covariance matrix. The other part is the on-line process which begins by generating hypothesis transformations to map observed colours to each colour in the database. Parameters for a transformation are generated by the match, and a verification procedure is used to determine the consistency of the transformation in mapping the other observed colours into the database. The first verification procedure is that the set of hypothesis transformations will be tested against the surface and illumination constraints. For the next step, an error measure $err_{j,k}$ (see eqn. 5.15) of each pair of matched colours is computed and the average of the overall error $Oerr_p$ is used as the criterion for estimating the degree of fitness between the observed colours and the colour database. Computing the error measure $Oerr_p$ in eqn. 5.16 requires: firstly finding the closest colour in the colour database to each transformed colour; secondly computing the $err_{j,k}$ between these two colours. The final step of the algorithm provides a set of transformations which result in a good fit between the observed colours and the memory colours.

The colour matching algorithm is summarised:-

1. Generate a transformation by matching every observed (image) colour and the memory colour in the database.

2. Test the hypothesis transformation against the surface and illumination constraints

3. Apply the hypothesis transformation to the observed colours and compute the total error measure.

4. Choose the transformation which resulted with the minimum total error measure and apply this to the observed colours (for further verification of assumptions with other object feature).

There are some implementation considerations which can simplify and speed up the performance of this process. For step 1 of the algorithm, an accumulator space can be used where the parameters of the hypothesis transformations are used as the index. The accumulation space will reflect the consistency of the transformations and is shown in figure 5.12.
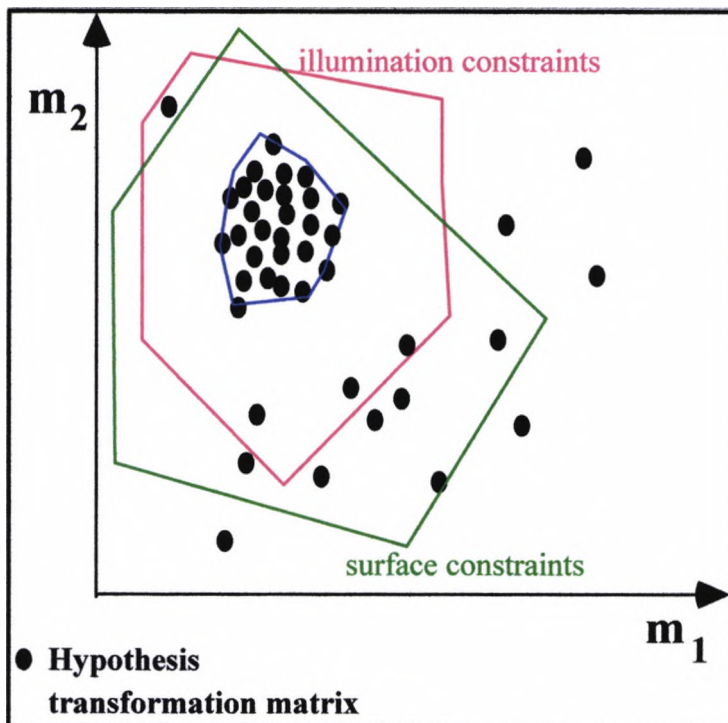

Figure 5.12 Accumulation space of hypothesis transformations.

The hypothesis transformations within the blue boundary form a dense cluster. It indicates that one of them will provide a consistent

transformation which transforms the observed colours close to the colours in the database. Therefore, instead of verifying all the hypothesis transformations, it is only required to consider those within the cluster constraints. Figure 5.12 is an imaginary example which shows that the colour vector has two bases which are laid on the axis of m_1 and m_2.

Step 3 of the algorithm requires an exhaustive search and computes the distance between a transformed colour and colours in the database. Therefore it will be computationally expensive. One alternative is to use a lookup table where the index is the transformed colour descriptor and the output is the error measure associated with the index and represents the nearest colour from the database. More detail on the Colour Information Look Up Table (CILUT) can be found in chapter 2, section 2.5.

5.4.5 Unknown Colours

A observed colour is defined as an unknown colour when it does not correspond to any of the colours in the colour database. Therefore in the process of accumulating the error of each match, if the error, $err_{j,k}$, exceeds a defined value, the observed colour will be considered as an unknown colour and excluded. If the number of unknown colours is greater than the known colours, the hypothesis transformation is rejected. Therefore there are two parameters that have to be determined: first, the allowable error $err_{j,k}$ of a match; second, the ratio between the number of known and unknown colours. The allowable error is defined according to the distributions of memory colours in the colour space. If the memory colours are close to each other in colour space, the allowable error is set to a small value, because the observed colour may be close to its corresponding memory colour. However if the memory colours are sparsely distributed in the colour space, the corresponding memory colour and observed colour may be far apart. The allowable error is determined by the minimum of the distances between a pair of memory colours. The distances between a pair of memory colours is shown in eqn 5.19. The ratio is the number of pixels with known colour to the unknown colour. The equation is shown in eqn 5.21. Since the algorithm assumes that majority of the

observed colours will have a colour corresponding to a memory colour, therefore the ratio will be smaller than one.

c_i is memory colour i and c_j is memory colour j

Mahalanobis distance between memory colour i to colour j is $err_{i,j}$,

$$d_{i,j} = [c_i - c_j] \, Cov_i [c_i - c_j] \quad \cdots\cdots (5.20)$$

where $i \neq j$

θ is the ratio

kc is number of pixels are represented by known colour,

uc is number of pixels are represented by unknown colour.

q is the constant to ensure the denominator is not zero.

$$\theta = \frac{uc}{kc + q} \quad \cdots\cdots (5.21)$$

These values are not used in the experiment I which is described in section 5.5, since no unknown colours are included in these images. However, they are used in the experiment II which is described in section 5.6.

5.5 Experiment I

This experiment is designed to test the operation of the algorithm with Mondriaan images. It demonstrates the qualitative and quantitative differences of the images observed under different lighting conditions. The observed colours of these images are input to the algorithm, and the resulting colours are observed and measured. There are two quantitative comparsions between the memory colours in the colour database, the input colours and the output colours of the Mondriaan images . The performance measure indicates how well the algorithm performs in computing the constant colour descriptors of the patches over a range of illuminants. The two measurements used in the experiment are described in section 5.6.2 and the resulting images and data are shown in section 5.6.3.

Firstly a Mondriaan image is used to build the colour database. Secondly, the same image is observed under several unknown illuminants and used as the input to the algorithm. All images are captured under the settings decribed in section 5.5.1. The input images to the algorithm contain no unknown colours. Therefore no parameter is required to be set for the rejection of colours. The selection of colours for building the colour database is subjective, based on two factors: the colours of the patches are well distributed in the colour space; the colours are saturated and bright. The chosen colour patches have the colours that are similar to those that appear in the images in experiment II (see section 5.6).

5.5.1 Experimental Settings

The experimental settings described below are used for building the memory colours in the database. The test images used in this section and section 5.7 are captured under the same settings.
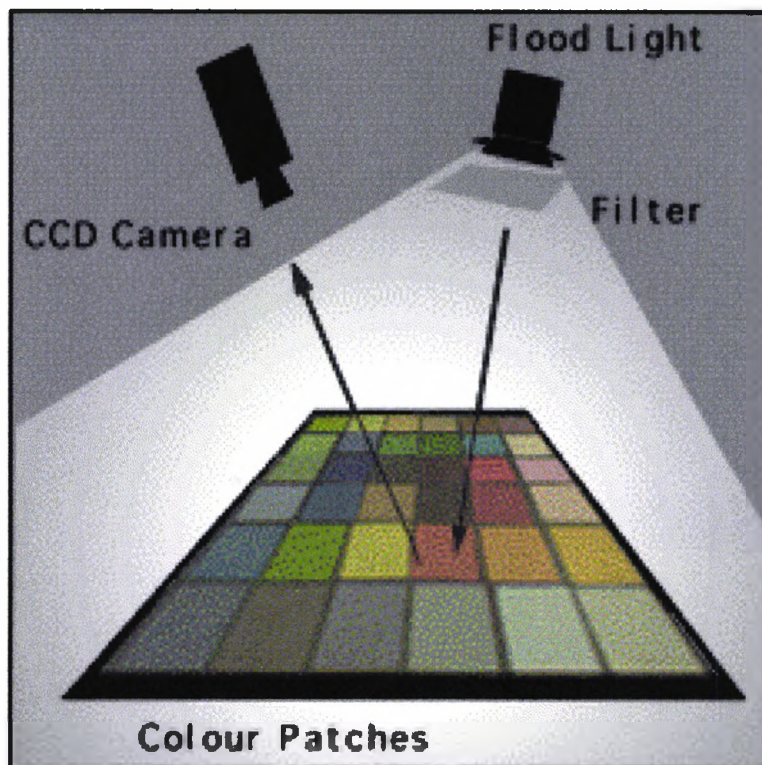
Figure 5.13 The settings for capturing images.

Initially the Mondriaan picture observed under canonical illuminant is used to build the representation of colour in the database. Canonical illuminant is assumed to be the reference, the one chosen in these experiments is one similar to the sunlight and comparatively brighter than other available sources. Also the colour observed under any other illuminant will be transformed to the value when the colours are observed under the canonical illuminant. The spectral characteristics of the canonical illuminant (curve shown in yellow) are shown in figure 5.14.

Figure 5.13 shows the setting for capturing a colour image. Photoflood light is used with a diffuser to provide constant illumination over the colour patches. Illumination with different spectral characteristics was created by using different colour filters (Lee filters) placed in front of the light source.
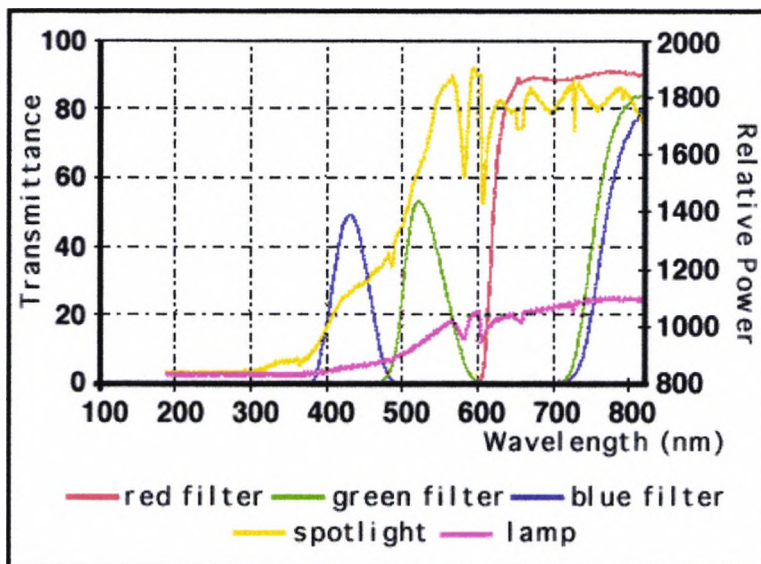
Figure 5.14 The response of three different filters and the spectral density function of two different illuminants.

A Mondriaan style image, made up of a set of colour patches (which have matte surfaces and are considered Lambertian) is imaged normal to the camera axis, with the light source position arranged to minimise shading or highlighting. A monochrome CCD camera with gain control disabled is

used with three primary colour filters (Kodak Wratten filter nos. 29, 47B and 58) for colour separation. The spectral transmittance characteristics of the three filters are shown on figure 5.14. Also, because the sensors' spectral response is beyond the visible range (400nm-700nm), an infra-red cut-off filter is used. The images are quantised to 8 bits per colour channel.

The Mondriaan image is captured under a canonical illuminant. The mean colour and covariance matrix of each patch is calculated. Figure 5.15 shows the 36 mean colours which are used in the database. The colour information lookup table is built based on these real data and used in the experiments described in this section and section 5.6.
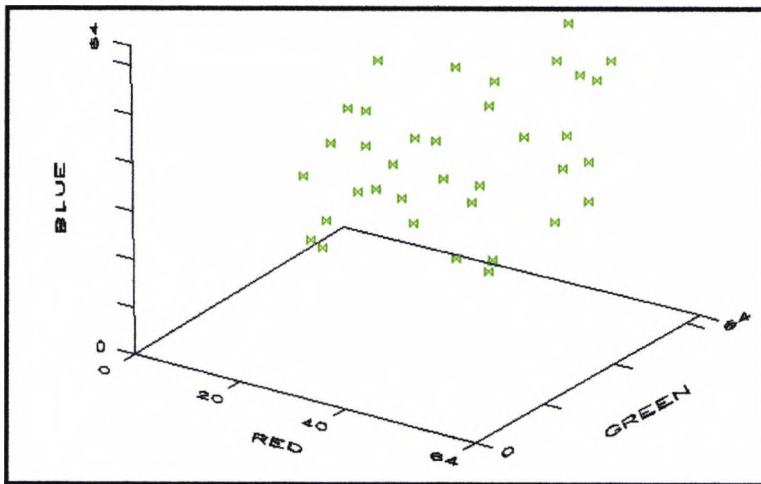
Figure 5.15 The memory colours of the colour database in the colour space.

5.5.2 Measurements

Two measurements are used for the comparison of the input and output colours to the memory colour in the colour database: the average of the absolute (see eqn. 5.23) and the relative (see eqn. 5.24) differences between the two colour sets. The input and processed images are interpreted in terms of these measurements and the data are summarised in table 5.2 & 5.3.

$D_{r,i}$, $D_{g,i}$ and $D_{b,i}$ are the absolute differences between the average colour of an observed patch i and its corresponding colour j in the database.

$$\left.\begin{array}{l} D_{r,i} = \overline{c_{r,i} - c_{r,j}} \\ D_{g,i} = \overline{c_{g,i} - c_{g,j}} \\ D_{b,i} = \overline{c_{b,i} - c_{b,j}} \end{array}\right\} \cdots \cdots (5.22)$$

Subscript r, g and b correspond to red, green and blue planes respectively.

$\widetilde{D}_{r,i}$, $\widetilde{D}_{g,i}$ and $\widetilde{D}_{b,i}$ are the average of $D_{r,i}$, $D_{g,i}$ and $D_{b,i}$ respectively.

$$\left.\begin{array}{l} \widetilde{D}_r = \dfrac{\sum\limits_{i=1}^{n} D_{r,i}}{n} \\[20pt] \widetilde{D}_g = \dfrac{\sum\limits_{i=1}^{n} D_{g,i}}{n} \\[20pt] \widetilde{D}_b = \dfrac{\sum\limits_{i=1}^{n} D_{b,i}}{n} \end{array}\right\} \cdots \cdots (5.23)$$

The average of relative differences in percentage, $\widetilde{p}_r$, $\widetilde{p}_g$ and $\widetilde{p}_b$ are :

$$\left.\begin{array}{l} \widetilde{p}_r = \dfrac{\sum\limits_{i=1}^{n}\left(\dfrac{D_{r,i}}{c_{r,j}} * 100\%\right)}{n} \\[24pt] \widetilde{p}_g = \dfrac{\sum\limits_{i=1}^{n}\left(\dfrac{D_{g,i}}{c_{g,j}} * 100\%\right)}{n} \\[24pt] \widetilde{p}_b = \dfrac{\sum\limits_{i=1}^{n}\left(\dfrac{D_{b,i}}{c_{b,j}} * 100\%\right)}{n} \end{array}\right\} \cdots (5.24)$$

5.5.3 Results

Mean colour values for each colour surface are shown in table 5.1 (column 2), corresponding to each colour surface in raster order (each colour is marked with reference to table 5.1). Figure 5.17 shows the set of colours imaged under the first unknown illuminant and the average colour values are also listed in table 5.1 (column 3). In order to demonstrate the operation of the matching algorithm on a reduced set of colours, an arbitrary subset of the image from figure 5.17 have been selected. The output of the system corresponding to figures 5.18 & 5.19 and the average colour values are shown in table 5.1 (columns 4 and 5). Although the resulting colours in figures 5.18, 5.20 & 5.22 look identical to the colours in figure 5.16 (the colours in the database), there is still a small difference between them and these values are indicated in table 5.2 and 5.3.

Figure 5.16 Mondriaan image under canonical illumination.

The experiment with selected regions of figure 5.17 indicates that even with fewer colours in the image (figure 5.19), the system still able to adapt to the illumination. Also, the average of absolute differences with respect to the reference is approximately only 7-10 grey-levels for each channel. These results indicate that the diagonal matrix transformation is sufficient to generate a well matched set of colours. Results in table 5.3 indicate the significance of changes when the algorithm is applied to the input images

and also shows that the algorithm is accurately approximating the colours of the patches.
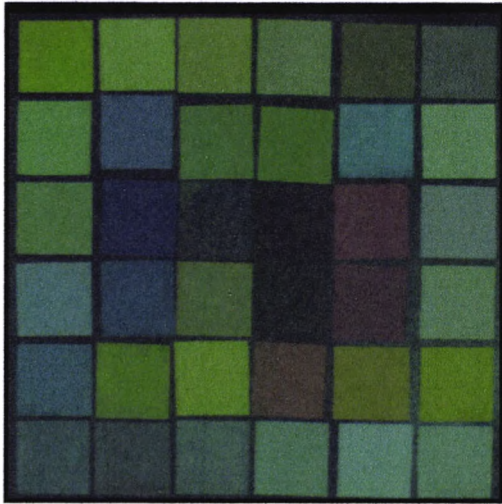
Figure 5.17 Input image (under bias illumination)
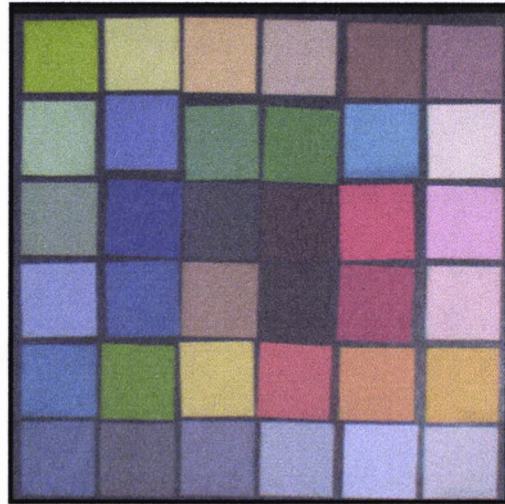
Figure 5.18 Output image of the algorithm with the input image shown Figure 5.17
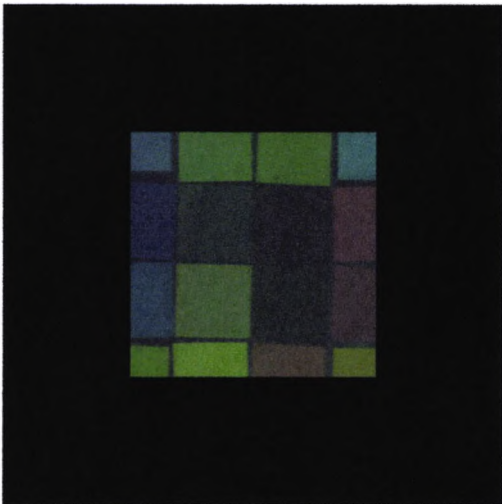
Figure 5.19 Selected regions from image shown in Figure 5.17.
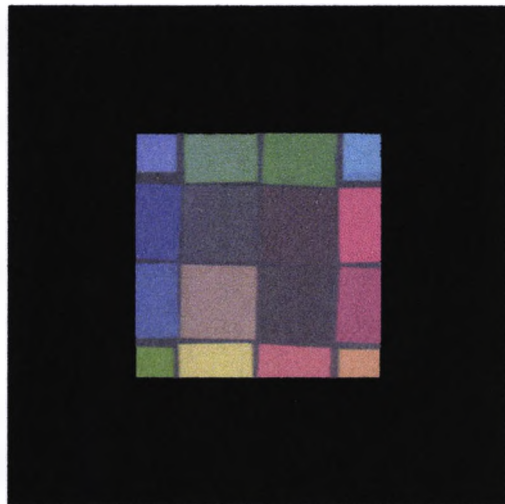
Figure 5.20 Output image of the algorithm with the input image shown in Figure 5.19.
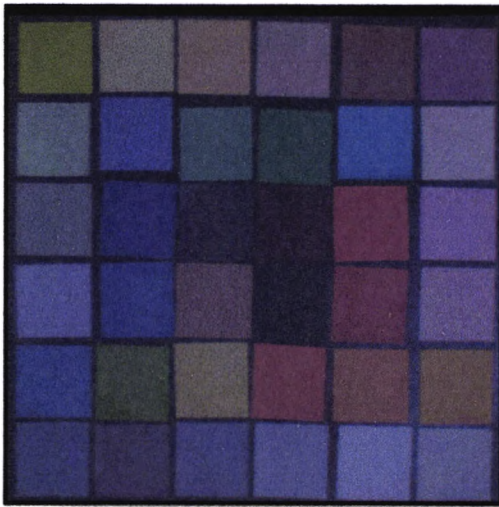
Figure 5.21 Input image (under bias illumination).
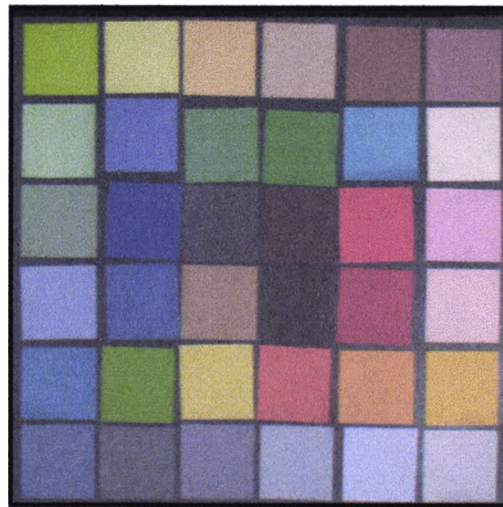
Figure 5.22 Output image of the algorithm with the input image shown in Figure 5.21

Table 5.1 Average colour values r,g,b, for the patches under different illuminants, also shown in figure 5.16, 5.17, 5.18 & 5.20.

	figure5.16 under canonical illumination r,g,b	figure 5.17 under bias illumination r,g,b	figure 5.18 all colours input to the system, and its output r,g,b	figure 5.20 selected region input to the system, and its output r,g,b
Patch no.				
1	144,176,92	51,127,57	144,184,82	
2	196,200,148	68,139,100	192,204,147	
3	208,176,136	72,125,95	206,184,138	
4	188,164,160	65,119,113	188,168,165	
5	144,108,104	52,78,72	144,112,105	
6	160,116,140	57,85,99	160,121,144	
7	148,188,164	54,133,109	152,192,160	
8	104,116,180	42,85,120	116,121,176	108,125,183
9	104,148,124	41,109,89	116,156,128	108,162,137
10	88,132,104	38,98,75	109,140,108	102,146,117
11	92,148,204	39,110,140	108,160,212	104,162,220
12	224,204,212	79,141,143	224,204,211	
13	132,152,148	50,113,100	139,164,148	
14	80,80,136	36,52,97	96,72,140	93,74,145
15	92,88,100	39,60,68	111,84,101	104,86,108
16	96,76,84	40,47,57	103,67,84	104,66,88
17	216,80,96	74,51,66	208,72,96	201,75,100

	figure5.16 under canonical illumination r,g,b	figure 5.17 under bias illumination r,g,b	figure 5.18 all colours input to the system, and its output r,g,b	figure 5.20 selected region input to the system, and its output r,g,b
Patch no.				
18	224,168,216	78,123,145	219,182,212	
19	144,156,208	52,115,132	156,170,196	
20	88,100,156	38,72,111	100,102,162	96,104,176
21	173,141,129	60,102,92	169,148,134	183,146,135
22	90,80,89	38,49,58	104,68,85	96,70,81
23	193,84,102	67,53,71	189,75,103	204,73,108
24	228,190,206	74,132,140	215,190,210	
25	92,124,177	38,91,115	108,130,170	
26	112,155,100	43,113,65	124,164,92	108,160,94
27	213,201,133	72,138,94	204,201,138	223,208,143
28	211,98,101	73,66,69	208,96,100	224,96,100
29	222,150,110	79,109,78	224,156,112	234,157,112
30	227,182,115	78,129,82	223,184,124	
31	108,118,154	42,86,101	118,121,148	
32	118,112,128	45,79,86	128,115,127	
33	136,131,161	50,96,113	142,139,168	
34	165.172,197	59,124,134	167,178,198	
35	199,201,240	71,139,155	196,200,228	
36	198,191,214	70,135,143	196,196,212	

Table 5.2 Average difference of the colours in the input or output image to reference colours, for Mondriaan image in figure 5.17-22.

	$\widetilde{D}_{r,i}$	$\widetilde{D}_{g,i}$	$\widetilde{D}_{b,i}$	$\widetilde{D}_{r,o}$	$\widetilde{D}_{g,o}$	$\widetilde{D}_{b,o}$
set 1	98.2	41.2	46.7	7.42	6.22	4.17
set 2	84.7	35.2	37.4	10.0	7.69	8.06
set 3	56.9	50.5	87.6	8.03	7.88	5.56

set1 = image under biased illumination 1 (36 colours), figure 5.17 and output image, figure 5.18.

set2 = image under biased illumination 1 (16 colours), figure 5.19, and output image, figure 5.20.

set3 = image under biased illumination 2 (36 colours), figure 5.21, and output image, figure 5.22.

$\widetilde{D}_{r,i}$ is the average difference, where r, g and b denote red, green and blue respectively; where i and o denote input and output respectively.

Table 5.3 Average difference in percentage, for Mondriaan image in figure 5.17-22.

	$\widetilde{p}_{r,i}$	$\widetilde{p}_{g,i}$	$\widetilde{p}_{b,i}$	$\widetilde{p}_{r,o}$	$\widetilde{p}_{g,o}$	$\widetilde{p}_{b,o}$
set 1	61.9	29.3	31.5	6.61	5.07	2.90
set 2	60.9	30.9	30.8	8.29	7.14	6.54
set 3	36.9	35.6	59.9	7.56	7.88	3.73

set1 = image under biased illumination 1 (36 colours), figure 5.17 and output image, figure 5.18.

set2 = image under biased illumination 1 (16 colours), figure 5.19, and output image, figure 5.20.

set3 = image under biased illumination 2 (36 colours), figure 5.21, and output image, figure 5.22.

$\widetilde{p}_{r,i}$ is the average difference in percentage, where r, g and b denote red, green and blue respectively; where i and o denote input and output respectively.

5.6 Experiment II

The experiment II is designed to test how well does the algorithm transform the observed colour data into the domain of the 'normal' colour space. If a set of colour patches is observed under different illuminants, their observed values are different. Since the aim of algorithm is to recover their colour values when the patches are observed under canonical illuminant, and it is considered the values in the domain of the 'normal' colour space. The first part demonstrates that the difference of objects' colours observed under different illuminants. The second part is to examine the resulting images of the object colours after the application of the algorithm; all colours observed under another illuminants are mapped to the 'normal' colour space, which consists of all the memory colours.

In the results of experiment I, there is a difference in the colour values between the original Mondriaan image under the canonical illuminant and the transformed image which was originally observed under an unknown illuminant. Since both images are captured under real lighting conditions, it is not possible to determine whether the error is caused by the diagonal matrix transformation or the recovery algorithm. In experiment II, the images are captured using a flat-bed scanner (the illuminant is a built-in fluorescent lamp; the sensor is charge coupled device which is similar to those used video cameras; 8 bits per channel) and the different lighting

conditions are generated by changing the scale of each channel of the images. Therefore it is able to determine the reliability of the algorithm which is independent of the diagonal matrix transformation. As it has been mentioned in the section 5.5.1 and 5.5.2, the transformation and recovery algorithm are related but separated issues.

Seven different cereal boxes are used for the test images. Most of the dominant colours (i.e. with large area) are presumed to be similar to the memory colours in the database. The algorithm uses two parameters for excluding unknown colours (see section 5.4.5). An unknown colour is defined as the colour that is not similar to any memory colour in the database. The complete or partial surface of the cereal boxes imaged under different illuminants are input to the algorithm and compared. The comparison procedure is based on two measures described in section 5.6.1. The difference between eqn. 5.23 and eqn 5.25 is that the later is calculating the difference between each corresponding pixel. Finally seven sets of input and output images and the corresponding measurements are shown in section 5.7.2.

5.6.1 Measurements

Two measurements are used to measure the difference between the two resulting images. The first one is the average difference between the two images. The second one is the standard deviation of their differences. These measurements are used to interpret the quantitative differences between the output images (figures 5.23-9(b) against 5.23-9(d) respectively) of the process and the data are shown in table 5.4.

$\widetilde{D}_r$, $\widetilde{D}_g$ and $\widetilde{D}_b$ are the average differences between the two output images. i indicates the corresponding pixel in both output images.
Subscript r, g and b represent red, green and blue plane respectively.

$$\begin{rcases} \tilde{D}_r = \dfrac{\sum\limits_{i=1}^{n} D_{r,i}}{n} \\[2em] \tilde{D}_g = \dfrac{\sum\limits_{i=1}^{n} D_{g,i}}{n} \\[2em] \tilde{D}_b = \dfrac{\sum\limits_{i=1}^{n} D_{b,i}}{n} \end{rcases} \cdots\cdots(5.25)$$

S_r, S_g and S_b are the standard deviations (s.d) of differences between the two output images.

$$\begin{rcases} S_r = \sqrt{\dfrac{\sum\limits_{i=1}^{n}\left(\tilde{D}_r - D_{r,i}\right)^2}{n}} \\[2em] S_g = \sqrt{\dfrac{\sum\limits_{i=1}^{n}\left(\tilde{D}_g - D_{g,i}\right)^2}{n}} \\[2em] S_b = \sqrt{\dfrac{\sum\limits_{i=1}^{n}\left(\tilde{D}_b - D_{b,i}\right)^2}{n}} \end{rcases} \cdots\cdots(5.26)$$

5.6.2 Results

The first part of the procedure captures the images of the cereal boxes under an unknown illuminant (by manually changing the coefficient of each channel) and these are shown in figures 5.23-29(a). Figures 5.23-29(c) show selected regions of the cereal boxes imaged under another unknown illuminant (generated by manually changing the coefficient of each channel). It has been shown in experiment I of section 5.5 that if the same set of colour patches are captured under two different illuminants and the colour constancy algorithm is applied to the images, the computed colour descriptors are very similar. The experiment also investigates the operation of the algorithm if only part of the surfaces is visible. In this experiment,

occlusion is simply achieved by truncating half the image and at the same time trying to preserve the appearance of dominant colours.

Figures 5.23-29(b) and 5.23-29(d) show the resulting images after the constancy algorithm has been applied to images shown in figures 5.23-29(a) and 5.23-29(c) respectively. Results shown:

1. The ratios of known colour θ (see eqn.5.21 in section 5.4.5) for figure 5.23-29(a) are less than one. It indicates that the number of known colours is greater than unknown colours. The results confirm those of the synthetic images, with the proviso that the colour content of the scene must exhibit some similarities to those found in the colour database (see section 5.6).

2. Each pair of the resulting images is visually similar to the selected region (white boundary seen in figure5.23-29(c)). Table 5.4 shows the quantatitive differences between each pair of images. In comparing these pairs of images, the overall average difference in red, green and blue channels are about 5 (sd = 5), 7 (sd = 6) and 6 (sd = 7) grey-levels respectively. This experiment demonstrates the difference in the resulting images is independent of the use of diagonal matrix (because the change of illuminant is achieved by changing of the coefficient of each colour channel, see section 5.6). It provides a quantatitive difference in colour values when two different images of the same cereal box are mapped to the 'normal' colour space.

Figure 5.23-29 are Coco-pops, Corn-flakes, Corn-pops, Frosties, Rice-krispies, Ricicles and Crunchy-nut respectively.

In each set, (a) The image observed under the first unknown illuminant, and (c) is the resulting image. (b) The selected part of the image observed under the second illuminant and (d) is the result image.

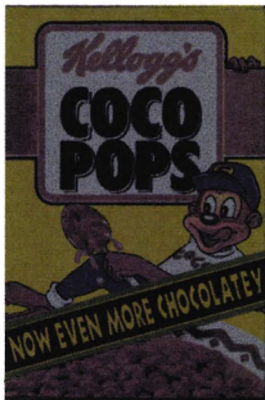Figure 5.23a   Figure 5.23b   Figure 5.23c   Figure 5.23d

Figure 5.24a   Figure 5.24b   Figure 5.24c   Figure 5.24d
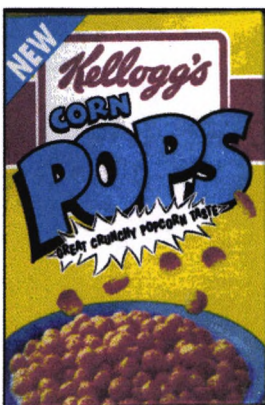
Figure 5.25a   Figure 5.25b   Figure 5.25c   Figure 5.25d
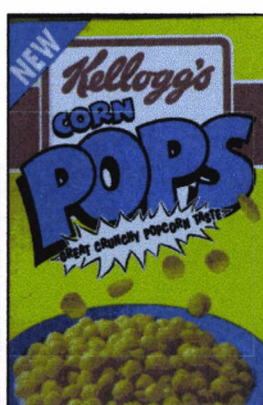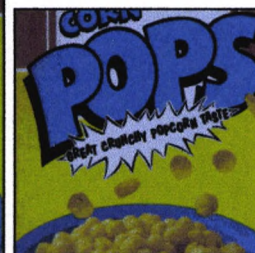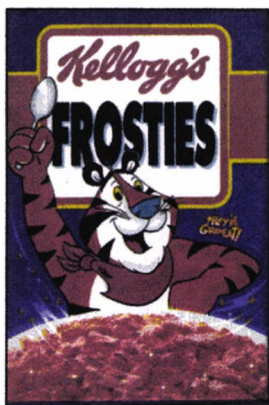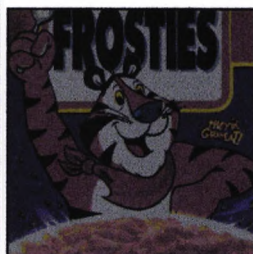
Figure 5.26a   **Figure 5.26b**   **Figure 5.26c**   **Figure 5.26d**
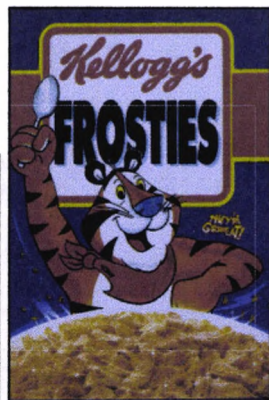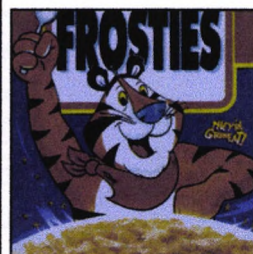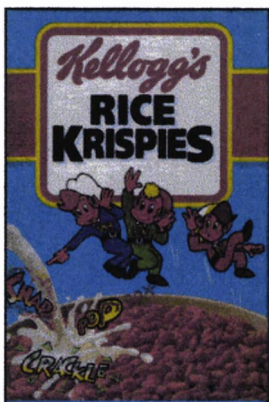
Figure 5.27a   **Figure 5.27b**   **Figure 5.27c**   **Figure 5.27d**
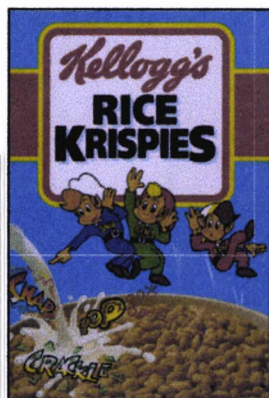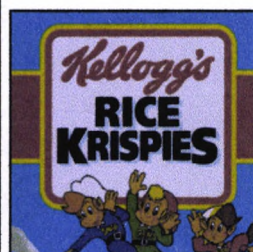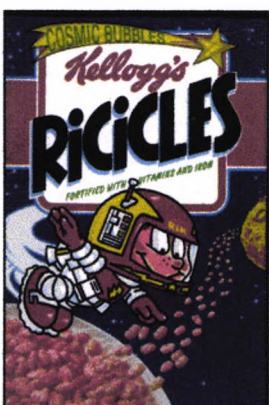
Figure 5.28a   **Figure 5.28b**   **Figure 5.28c**   **Figure 5.28d**
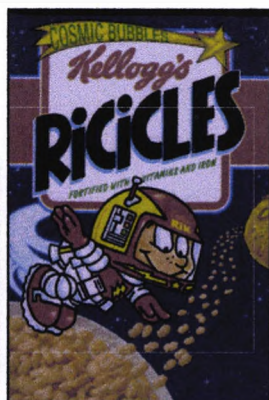
Figure 5.29a   Figure 5.29b   Figure 5.29c   Figure 5.29d
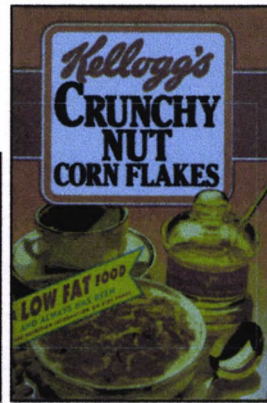
Table 5.4 Mean and standard deviation of grey-level difference on six sets of test images, see figure 5.23-29.

set	$\widetilde{D}_r$	$\widetilde{D}_g$	$\widetilde{D}_b$	S_r	S_g	S_b
1	7.12	6.63	3.51	6.11	5.87	6.63
2	3.94	5.35	6.05	4.07	3.86	7.87
3	2.52	13.02	7.08	3.94	7.44	8.20
4	7.42	9.12	6.84	6.89	6.73	6.12
5	4.03	3.90	4.07	4.52	4.47	4.29
6	5.27	3.97	9.07	5.81	5.44	9.12
7	3.61	7.62	4.72	3.96	8.60	7.48

1=Coco-pops,   2=Corn-flakes,   3=Corn-pop,   4=Frosties,   5=Rice-Krispies,   6=Ricicles, 7=Crunchy-nut.

5.7 Discussion

The use of a colour constancy algorithm for object localisation and identification has been emphasised recently. Swain [Swa90] has mentioned that if the objects are illuminated under different lighting conditions, a colour constancy operation should be performed before histogram intersection to improve the overall performance of identification process. Matas, Marik and Kittler [MMK95] assumed that their input images are processed by a colour constancy algorithm before the building of the object representation and matching against the models. The colour constancy

algorithm proposed in this chapter has been demonstrated to perform well in computing the constant colour descriptor for surface. Although it is limited to performing under the assumptions of the Mondriaan world, where every object is flat and its surface property is Lambertian, it provides the starting point for the algorithm to be evolved for real open world images. With the use of the proposed colour constancy algorithm [HE95] as a pre-processing step for the colour histogram correlation [WE95], it has shown that a significant improvement on performance of matching observed objects' colour colours and models' colours can be achieved. The possible explanation for the improvement is that the objects are viewed from a far distance, the specularity on the object surface and other physical effects are reduced (low resolution and the average effect). Therefore the scenes are close to the Mondriaan world assumptions and the algorithm becomes applicable.

5.8 Conclusion

It has been shown that a form of memory colour is used in some existing algorithms and the new algorithm presented here forms an important part in computing constant colour descriptors.

A simplified algorithm for colour constancy is proposed based on a diagonal matrix transformation. The transformation is determined by comparing the colours in the unknown image with those in a pre-determined database. A minimum error criterion is used to select the 'best' mapping of the unknown colour set into the database colour set, and provides a mechanism for transforming the unknown colour set into a 'normalised' colour set, according to the database. The results indicate that the algorithm seems to give a low error in the matching process.

On the testing the reliability of the colour constancy algorithm, it is necessary to use artificially generated lighting in order to evaluate the inaccuracies of the result, which is independent of the approximation of the diagonal matrix transformation.

Other issues have been found in the implementation of the overall colour constancy process. There are several problems associated with generating reliable transformations related to the dynamic range of the colour in different parts of the image, and to the quality of the colour information under low illumination lighting conditions (which may result a quantisation error).

The results shown in this chapter reinforce the belief of Hering, *"All objects that are known to us from experience, or that we regard as familiar by their colour, we see through the spectacles of memory colour."*

Chapter 6
Conclusions & Further Work

6.1 Conclusion

The work reported in this thesis is concerned with the use of colour in different modules of a machine vision system. The modules outlined in chapter 1 were colour edge detection, colour segmentation, colour constancy and the colour database. Since there is no general architecture for a vision system and the general vision task is not defined, these modules can be used to provide some of the building blocks for a goal specific vision system. Extracting geometrical features of objects from a colour image using colour edge detection and segmentation of colour space have been studied. Extracting surface features of an object from a colour image by segmentation of colour space and the process of estimating the colour description of object's surface, have also been studied. The issues addressed include the use of colour in edge detection, signal-based and model-based segmentation of colour space, colour constancy and colour information representation. From these studies, the relationships between colour information and colour constancy are established. Thus a new representation of colour information is proposed for surface or object recognition. This can be employed for both colour space segmentation and colour constancy. The summary for each issue and their relations are presented in following sections 6.1.1-4.

6.1.1 Colour Information Representation

Colour information exploited at both low and high levels of image processing is described in chapter 2. The early work on the use of colour in low level vision has been concentrated on finding a set of colour features that are suitable for segmentation. The suitability set is based on the one which is the least correlated and has the lowest number of attributes.

In the review of previous research, it is found that there are two essential and common features in colour information representation

for surface colour in object or surface recognition systems: the algorithmic and descriptive features.

A colour database for surface colours is proposed with a structure based on a lookup table. It is termed the Colour Information LookUp Table (CILUT). This lookup table (CILUT) accommodates memory colours, which represents surface colour information. Memory colour consists of two features: one is the algorithmic or operational, and is used for the evaluation of similarities between colour from the image and model; the other is descriptive, and forms the colour attributes.

A colour database is employed in chapter 4 and 5. Since no comparison can be made with other databases of this type on efficiency, only some general comments can be drawn. Retrieving data using direct addressing (i.e. a lookup table) is very efficient. Adding or deleting memory colours may require a complete scan of the entire lookup table because the lookup table locations associated with the input memory colour can be widespread. The implementation of the lookup table requires a large memory, for example, a typical 24 bit addressable memory is needed for colour using three attributes of 8 bits each.

6.1.2 Use Of Colour In Edge Detection

Colour edges are perceived to be visually better than grey level edges, as advocated by some previous work [Nev77, Nov87]. One possible explanation is that the majority of edgels which support the geometry of the objects have positive or zero values in δ. Unwanted edgels caused by noise or other physical effects have negative values in δ, where δ is the difference between strength of colour and grey level edgel and if δ is positive, it means the colour edge is greater. A mathematical and statistical comparison of colour and grey-level edge detectors based on this definition found that edges are detected equally well with the use of colour and grey level images. On the mathematical approach, there are two situations where the edge can

be found, one is at the boundary between two regions and the other is the junction of three or more regions. In both situations, in general, δ is not always positive. In the statistical analysis, an average δ measure for an image is applied to a range of real images and recorded. It indicates that in general, the average δ of the detected edgels is not positive. Finally, a manual segmentation of edgels was applied, and two groups of edgels were formed; one is considered as the genuine edge (which supports the geometry of the object), the other is the noisy edge. A set of statistical measures was applied to both groups in a range of images, and was found that the average δ are very similar, except where the variances are different and the variances for the noisy edges are higher.

From the observation of the colour edge detector outputs, it was found that the edgels detected from three colour planes are highly correlated at a pixel or within a local neighbourhood, but the edgels caused by the specularity, roughness of the surfaces and the other noise effects are weakly correlated. A new method of combining edgels from three colour planes is proposed and named the corrobration function. It is based on exploiting the correlation of the colour edges. Two criteria are made for selecting edgels in the corroboration function. First, the presence of an edgel at a particular location in one plane can be corroborated by a consistent registration at the same or close-by location in the other planes. Second, they should have the same gradient orientation (or be in the opposite direction). The corroboration function was applied on synthetic and real images. There was an evaluation of the performance based on measuring the signal-to-noise ratio of the output edgels. The signal part is considered as the set of edgels which supports the geometry of the objects, and the noise part is the complement of the signal part. From the results shown in tables 3.2 and 3.3, the corroboration function applied to real images produces a much less dramatic reduction in the number of detected noisy edgels, in comparison with its performance on the synthetic images. Two reasons can be identified to justify the poorer performance of the corroboration in this case: first, the colour images were digitised using a conventional

single CCD camcorder, for which the spatial localisation of the red, green and blue pixels are non-coincident; second, the edge detection produces a slightly different localisation for each colour band. One possible solution is that the corroboration function must search its local neighourhood for appropriate edges before the constraints can be applied to the gradient orientations.

6.1.3 Segmentation Of Colour Space

Partitioning an image into two or more homogenous regions based on the segmentation of colour space was studied. Two approaches for segmenting colour space have been considered: one is signal-based, where the colours are considered as a set of random data, the other one is model-based, where the image is considered as a subset of colours from the modelled colours in the colour database.

A merging algorithm based on mode (peak) finding has been used by a number of researchers and good results have been shown in their experiments. The algorithm requires no parameters, such as the number of clusters, number of iterations etc. Also it is able to adapt to different shapes of cluster. However the remaining problem is the determination of an appropriate cell size of the histogram. With the use of the Peano scan encoding as a pre-processing step, it is possible to adjust the cell size with respect to the distribution of the histogram, and the result shows that most of the significant clusters are found in the histogram. Moreover, the technique can be applied to other feature spaces, for example, texture, but it would require an appropriate representation for the texture of a region.

Segmentation based on classifying pixel colours into different classes is more effective because no pre-processing is required. It is simply a direct mapping between the pixel values and the classes through the Colour Information Lookup Table (CILUT, see section 2.6 in chapter 2). The objects should be viewed under the same conditions the model is generated. It should not contain any shading

or specular highlights, because in general this changes the appearance of an object and will result in different colour values. It is necessary to capture the test images and model the object surfaces under the same lighting conditions (same colour temperature of illumination), because the image colour is dependent on the light impinging on the surface and its reflectance. More details on how to achieve constant image colour under different lighting conditions are described in chapter 5. Despite the constraints of this approach, it is appropriate for object inspection in a manufacturing plant where the viewing and lighting conditions can be controlled. The beauty of the algorithm is its efficiency. The classifier is only applied to the pixel and such a process can be run in parallel. Also, the classifier and Colour Lookup Table are simple to construct.

6.1.4 Colour Constancy

The philosophy of the approach in this thesis and the other attempts at colour constancy that are described in section 5.3, are found to be cognate with the belief of Hering, *"All objects that are known to us from experience, or that we regard as familiar by their colour, we see through the spectacles of memory colour"*. Chapter 5, illustrated the use of memory colour in different algorthms, and the mechanics of the algorthms are to drive the observed colours close to the memory colours (or reference colours). It showed that memory colour plays an important role in evaluating the colour description of surfaces, because the calculated values of colour descriptors are related to memory colours.

A simplified algorithm for colour constancy is proposed, based on a diagonal matrix transformation. The transformation is determined by comparing the colours in the unknown image with those in a pre-determined database. A minimum error criterion is used to select the 'best' mapping of the unknown colour set into the database colour set, and provides a mechanism for transforming the unknown colour set into a 'normalised' colour set, according to the database. The

results indicate that the algorithm seems to give a low error in the matching process. Although the lighting conditions were restricted in the experiment, this memory colour approach may provide a foundation for building a more sophisocated algorithm (see section 6.2). When the testing the reliability of the colour constancy algorithm, it was necessary to use artificially generated lighting in order to evaluate the inaccurcies of the result which is independent of the approximation of the diagonal matrix transformation.

Other issues have been found in the implementation of the overall colour constancy process. There are several problems associated with generating reliable transformations related to the dynamic range of the colours in different parts of the image, and to the quality of the colour information under low illumination lighting conditions (which may result in quantisation error).

The results shown in this chapter reinforce the belief of Hering, *"All objects that are known to us from experience, or that we regard as familiar by their colour, we see through the spectacles of memory colour."*

6.2 Further Work

Extracting a reliable measure of gradient orientation requires a high signal-to-noise ratio, and hence a Gaussian smoothing function with a large standard deviation is used; however, there is a trade-off between the amount of smoothing and good localisation of the edge [Can86]. A possible solution for improving the robustness of the corroboration function is to choose other parameters which can be applied to resolve the ambiguities in edge selection. An extension based on the hysteresis thresholding originally used by Canny can be adapted to preserve edge continuity based on preserving those edges with minimum deviation of gradient orientations.

On colour constancy, the memory colour approach may provide a foundation for building a more sophisticated algorithm which can perform robustly under more relaxed lighting conditions. The algorithm should cope with a set of less restricted lighting. One possible suggestion is to modify the contents of the colour database to deal with the relaxed conditions, such that specularity or shading of the objects is allowed. The modification of the colour database is required to include the specular highlights and the shaded parts of a surface in the memory colour representation.

Finally the colour constancy algorithm is found useful in an object recognition system, one of example can be seen in the work of Ellis & Walcott [WE95] on localisation of colour objects. Their method is based on using colour histogram correlation and my algorithm has been performed as a pre-processing step. Another suggestion is that the algorithm can be applied to other object recognition systems that require colour constancy, in order to cope with the varied lighting conditions.

Reference

[Ant82] Antonisse, H.J. "Image Segmentation in Pyramids", Computer Graphics and Image Processing, No.19, 1982, P.367-383.

[AR86] Arend, L. & Reeves, A. "Simultaneous Color Constancy", Journal of Optical Society America. A., Vol-3, No.10, October 1986, P.1743-1751.

[Bal91] Ballard, D.H. "Animate Vision" Artificial Intelligence 48, 1991, P.57-86.

[BB82] Ballard, D.H. & Brown, C.M. "Computer Vision", Prentice-Hall, 1982.

[BD97] Buluswar, S.D. & Draper, B.A. "Color Recognition by Learning: ATR in Color Images" British Machine Vision Conference 1998, CD-ROM.

[BE92] Brock_Gunn S. & Ellis T.J. "Using Colour Templates for Target Identification and Tracking" Proceedings of The British Machine Vision Conference, September 1992, P.207-217.

[Bec72] Beck, J. "Surface Colour Perception" Cornell University Press, 1972.

[Ber87] Berry, D.T. "Colour Recognition Using Spectral Signatures", Pattern Recognition Letters, No.6 1987, P.69-75.

[BH88] Binford, T.O. & Healey, G. "A Color Metric for Computer Vision", DARPA Vision Workshop, 1988, P.854-861.

[BLL90] Bajcsy, R, Lee S.W. & Leonardis, A. "Color image Segmentation with Detection of Highlight and Local Illumination Induced by Interreflections. " Proceedings of ICPR 1990, pp.785-790

[Bor91] Borges, C.F. "Trichromatic Approximation Method For Surface Illumination" Journal of Optical Society America. A, Vol. 8, No. 8, August 1991, P.1319-1323.

[Bri78] Brill, M.H. "A Device Performing Illuminatant-invariant Assessment of Chromatic Relations" Journal of Theory Biology, 71, 1978, P.473-478.

[Bri89] Brill, M.H. "Object-based Segmentation and Color Recognition in Multispectral Images", Proc. SPIE - International Society of Optical Engineer. (USA), Vol-1076, 1989, p.97-103.

[Bri90] Brill, M.H. "Image Segmentation by Object Color: A Unifying Framework and Connection to Color Constancy" Journal of Optical Society America. A, vol.7, No.10, October 1990, P.2041-2047.

[BT90] Borghesi, L. & Toscano, R. "A Colour Image Segmentation Process Suitable for Real Time Implementation", Proc. of the 5th Int. Conference on Image Analysis and Processing, Progress in Image Analysis and Processing, World Scientific, 1990, P.23-30.

[Buc80] Buchsbaum, G. "A spatial processor model for object colour perception" Journal of the Franklin Institute, No. 310, 1980, P.1-26.

[BW81] Brill, M.H. & West, G. "Contributions to the Theory of Invariance of Color Under the Condition of Varying Illumination" Journal of Mathematical Biology, No.11, 1981, P.337-350.

[BW86a] Brill, M.H. & West, G. "Chromatic Adaptation and Colour Constancy: A Possible Dichotomy", Color Research and Application, vol. 11, No. 3 Fall 1986, P.186-204.

[BW86b] Brainard, D.H. & Wandell, B.A. "Analysis of the Retinex Theory of Color Vision", Journal of Optical Society America. A., vol.3, No.10, October 1986, P.1651-1661

[BW90] Brainard, D.H. & Wandell, B.A. "Calibrated Processing of Image Color", Color Research and Application, Vol.15, No.5, October, 1990, P.266-271.

[BW91] Boult, T.E. & Wolff, L.B. "Physically-based Edge Labelling", Proc. 1991 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1991, P.656-662.

[Can86] Canny, J. "A Computational Approach to Edge Detection", IEEE Trans. on PAMI, Vol-PAMI-8, No.6, Nov. 1986, P.679-698.

[CB87] Chou, P.B. & Brown, C.M. "Probabilistic Information Fusion for Multi-modal Image Segmentation" Proceedings of the 10th International Joint Conference on Artificial Intelligence, 1987, P.779-782

[CF83] Connah, D.M & Fishbourne, C.A. "Segmentation by Colour" The Radio and Electronic Engineer, vol.53, No.4, April 1983, P.153-156.

[CGG91] Cumani, A., Grattoni, P. & Guiducci, A. "An Edgebased Description of Color Images", CVGIP: Graphical Models and Image Processing, Vol-53, No.4, July 1991, P.313-323.

[CLK89] Conolly, C., Littlewood, S. & King, E.S. "A system for the Segmentation of Colour Images", Third International Conference on Image Processing and its Applications, IEE, 1989, P.441-444.

[CMP90] Choo, A.P. & Maeder, A.J. & Pham, B. "Image Segmentation for Complex Natural Scenes", Image and Vision Computing, May 1990, Vol-8, No.2, P.155-163.

[Cum91] Cumani, A. "Edge Detection in Multispectral Images", CVGIP: Graphical Models and Image Processing, Vol-53, No.1, Jan. 1991, P.40-51.

[Dav75] Davis, L.S. "A Survey of Edge Detection Techniques", Computer Graphics and Image Processing, No.4, 1975, P.248-270.

[Dav91] Davidoff, J. "Cognition through Color" The MIT press 1991

[DBE77] Dahlqvist, B., Bengtsson, E., Eriksson, O., Jarkrans, T. Nordin, B. & Stenkvist, B. "Algorithm for Cluster Analysis" Proceedings of The Second Scandinavian Conference on Image Analysis, 1977, P.134-140.

[DG87] Dreschler-Fischer, L.S. & Gnutzmann, F. "Features Selection in Colour Images for Token Matching", Proc. of the 10th International Joint Conference on Artificial Intelligence, 1987, P.749-751.

[Dix78] Dixon, E.R "Spectral Distribution of Australian daylight", Journal of Optical Society America. A. No.68, P.487-450, 1978.

[DL86] D'Zmura, M. & Lennie, P. "Mechanisms of Color Constancy" Journal of Optical Society America. A., vol.3, No.10, October 1986, P.1662-1672.

[DL98] Duffy, N. & Lacey, G. "Colour Profiling Using Multiple Colour Spaces", British Machine Vision Conference 1998, P. 245-256.

[DZI94] D'Zmura M. & Iverson, G. "Color Constancy. III. General linear Recovery of Spectral Descriptions for Lights and Surfaces" Journal of Optical Society America. A, vol.11, No.9, October 1994,

P.2389-2400.

[DZm92] D'Zmura, M. "Color Constancy: Surface Color from Changing Illumination" Journal of Optical Society America. A, Vol. 9, No. 3, March 1992, P.491-493.

[EHO91] Ellis, T.J., Hung, T.W.R. & Omarouayache, S. "Structural Elements In Colour Images: A Parallel Approach" The third International Conference on Applications of Transputers, August 1991.

[FD88] Funt, B.V. & Drew, M.S. "Color Constancy Computation in Near-Mondriaan Scenes using a Finite Dimensional Linear Model", Proc. IEEE Computer Vision Pattern Recognition, June 1988, P.544-549.

[FDB92] Funt, B.V., Drew, M.S. & Brockington, M. "Recovering Shading from Color Images" European Conference on Computer Vision, 1992, P.124-132.

[FDH91] Funt, B.V., Drew, M.S. & Ho, J "Color Constancy from Mutual Reflection", International Journal of Computer Vision, 6:1, 1991, P.5-24.

[FF95] Funt, B.V. & Finlayson, G.D. "Color Constant Color Indexing", IEEE Trans. on PAMI, Vol.17, No.5, May 1995, P.522-529.

[FH89] Funt, B.V. & Ho, J. "Color from Black & White", International Journal of Computer Vision, Vol-3, No.2, June 1989, P.109-117.

[FH97] Finlayson, G. & Hordley, S. "Selection for Gamut Mapping Colour Constancy", British Machine Vision Conference 1998, CD-ROM.

[Fin92] Finlayson, G.D. "Colour Object Recognition" Master Thesis, Simon Fraser University, Canada, 1992.

[Fin95] Finlayson, G.D. "Coefficient Color Constancy", DPhil Thesis, Simon Fraser University, Canada, 1995.

[For90] Forsyth, D.A. "A Novel Algorithm for Colour Constancy" International Journal of Computer Vision, Vol. 5, No. 1, 1990 pp. 5-36.

[GBS90] Gonzalez-Rodriquez, M., Benitez-Diaz, D. & Suarez-Araujo, C.P. "Segmentation and Recognition in Visual Chromatic Spaces", Cybernetics and Systems: An International Journal (USA), Vol-21, No.23, March-June 1990, P.241-247.

[GGS98] Gevers, T., Ghebreab, S. & Smeulders, A. "Color Invariant Snakes" British Machine Vision Conference 1998, pp. 578-588.

[GJT87a] Gershon, R., Jepson, A.D. & Tsotsos, J.K. "The Use of Color in Highlight Identification", Proc. of the 10th International Joint Conference on Artificial Intelligence, 1987, P752-754.

[GJT87b] Gershon, R., Jepson, A.D. & Tsotsos, J.K. "From [R,G,B] to Surface Reflectance: Computing Color Constant Descriptors in Images" Proc. of the 10th International Joint Conference on Artificial Intelligence, 1987, P.755-758.

[GK89] Gupta, M.M. & Knopf, G.K. "Theory of Edge Perception for Computer Vision Feedback Control", Journal of Intelligent Robot System Theory Application (Netherlands) Vol-2, No.2-3, 1989, P.123-151.

[Gol78] Goldberg, M. "A Clustering Scheme for Multispectral Images", IEEE Trans. on SMC, vol-8, No.2, 1978, P.86-94.

[GP87] Geiger, D. & Poggio, T. "An Optimal Scale for Edge

Detection", Proc. of the 10th International Joint Conference on Artificial Intelligence, 1987, P.745-748.

[Gre90] Gregory, R.L. "Eye and Brain, The Psychology of Seeing" Fourth Edition, George Weidenfeld and Nicolson Ltd, 1990.

[Gre92] Gregson, P.H. "Angular Dispersion of Edgel Orientation: the Basis for Profile-insensitive Edge Detection", Proc. SPIE - International. Society of Optical Engineering, USA, Vol-1607, 1992, P.217-224.

[HB87a] Healey, G. & Binford, T.O. "Color Algorithms for General Vision System", Proceedings of the 10th International Joint Conference on Artificial Intelligence, 1987, P759-762.

[HB87b] Healey, G. & Binford, T.O. "The Role and Use of Colour in a General Vision System", Proc. DARPA Image Understanding Workshop, USC, CA, USA, 1987. P.599-613.

[HB88] Healey, G. & Binford, T.O. "A Colour Metric for Computer Vision", Proc. DARPA Vision Workshop, 1988. P.854-861.

[HD85] Huntersberger, T.L. & Descalzi, M.F. "Color Edge Detection" Pattern Recognition Letters, No.3, 1985, P.205-209.

[HE95] Hung T.W.R. & Ellis, T.J. "Spectral Adaptation With Uncertainty Using Matching" IEE International Conference on image processing and its application, Edinburgh, July 1995, pp. 786-790
[
Hea89] Healey, G. "Using Colour for Geometry-insensitive Segmentation", Journal of Optical Society America. A., Vol-6, No.6, June 1989, P.920-937.

[Hea91] Healey, G. "Estimating Spectral Reflectance Using Highlights" Image, Vision and Computing (UK), Vol-9, No.5, October 1991, P.333-337.

[Hea92] Healey, G. "Segmenting Images Using Normalized Color." Physics-Based Vision: Principles and Practice: COLOR, edited by Healey, G. E., Shafer, S. A. & Wolff, L.B. Jones and Barlett Publishers, 1992

[HEC91] Hung, T.W.R., Ellis, T.J. & Chamberlain, D.A. "A Colour Vision Facility" International Conference on Systems Engineering, Coventry, September 1991.

[Her64] Hering, E. "Outlines of A Theory of The light sense", Translated by L.M.Hurvich and D. Jameson, Harvard University Press 1964.

[HFD90] Ho, J., Funt, B.V. & Drew, M.S. "Separating a Color Signal into Illumination and Surface Reflectance Components: Theory & Application", IEEE Trans. on PAMI, Vol-12, No.10, October 1990, P.966-977.

[HNP82] Hong, T.H., Narayanan, K.A., Peleg, S., Rosenfeld, A. & Silberberg, T. "Image Smoothing and Segmentation by Multiresolution Pixel Linking: Further Experiments and Extensions", IEEE Trans. on SMC Vol-SMC-12, No.5, September 1982, P.611-622.

[Hor74] Horn, B.K.P. "Determine Lightness from an Image", Computer Graphics and Image Processing, No.3, 1974, P.277-299.

[HP88] Hurlbert, A.C. & Poggio, T.A. "Synthesizing a Colour Algorithm From Examples" Science, Vol.236, January 1988, P.482-485.

[HS85] Haralick, R.M. & Shapiro, L.G. "Survey: Image Segmentation Techniques", Computer Vision, Graphics and Image Processing, No.29, 1985, P.100-132.

[HS94] Healey, G. & Slater, D. "Global Color Constancy: Recognition of Objects by Use of Illumination-Invariant Properties of Color Distributions" Journal of Optical Society America. A., Vol.11, No.11, November 1994, P.3003-3009.

[Hun88] Hung, T.W.R. "Colour Image Processing" B.Sc. report City University, London, 1988.

[ID94] Iverson, G. & D'Zmura, M. "Criteria for Color Constancy in Trichromatic Bilinear Models", Journal of Optical Society America. A., Vol. 11, No. 7, July 1994, P.1970-1975.

[Jan90] Jang, Y. "Identification of Interreflection in Color Images Using a Physics-based Reflection Model", Master Thesis, University of South Carolina, 1990.

[JB91] JordanIII, J.R. & Bovik, A.C. "Using Chromatic Information in Edge-Based Stereo Correspondence", CVGIP: Image Understanding, Vol-54, No.1, July 1991, P.98-118.

[JMW64] Judd, D.B., MacAdam, D.L. & Wyszecki, G. "Spectral Distribution of Typical Daylight as a Function of Correlated Color Temperature", Journal of the Optical Society of America, Vol-54, No.8, August 1964, P.1031-1042.

[KB90] Khotanzad, A. & Bouarfa, A. "Image Segmentation by a Parallel, Non-parametric Histogram Based Clustering Algorithm" Pattern Recognition, 1990, P.961-973.

[Kit76] Kittler, J. "A Locally Sensitive Method for Cluster Analysis" Pattern Recognition, No.8 1976, P.23-33

[KNF76] Koontz, W.L.G., Narendra, P.M. & Fukunaga, K. "A graph-theoretic Approach to Nonparametric clustering" IEEE Trans. on Computers, Vol.25, 1976, P.936-944.

[Kos95] Koschan, A. "A Comparative Study on Color Edge Detection" Proc. on Second Conference on Computer Vision, December 1995, P.574-578.

[KSK90] Klinker, G.J., Shafer, S.A. & Kanade, T. "A Physical Approach to Color Image Understanding", International Journal of Computer Vision, No.4, 1990, P.7-38.

[Lan85] Land, E.H. "Recent Advance in Retinex Theory" In Ottoson, D. & Zeki, S (eds.), "Central and Peripheral Mechanisms of Colour Vision, Macmillan, New York 1985 or Vision Research, Vol.26, No.1, 1986, P.7-21.

[LB92] Lee, S.W. & Bajcsy, R. "Detection of Specularity Using Color and Multiple Views" Proceedings on European Conference Computer Vision, 1992, P.99-114.

[LBS90] Lee, H-C., Breneman, E.J. & Schulte, C.P " Modelling Light Reflection for Computer Color Vision", IEEE. Trans. on PAMI vol-12, No.4, April 1990, p.402-409.

[LCB91] Lambert, R.A., Chan, J.P. & Batchelor, B.G. "Segmentation of Colour Images Using Peano Curves on a Transputer Array" Proc. SPIE-Machine Vision Architectures, Integration, and Applications, Vol.1615, 1991, P.187-193.

[Lee86] Lee, H-C. "Method for Computing the Scene-illumination Chromaticity from Specular Highlights" Journal of Optical Society America. A., Vol-3, No.10, October 1986, P.1694-1699.

[LM71] Land, E.H. & McCann, J.J. "Lightness and Retinex Theory", Journal of Optical Society America. No.61, 1971, P.1-11.

[Mac81] MacAdam, D.L. "Color Measurement: Theme and Variations", Optical Sciences, Springer-Verlag 1981.

[Mar82] Marr, D. "Vision", W.H.Freeman and Company 1982

[MB89] Maseen, P. & Bottcher, P. "Colour and Shape: A Real Time Colour Vision System for Industrial Inspection", Proc. 9th Conference Automated Inspection & Product Control, May 1989, P.53-62.

[MH80] Marr, D. & Hildreth, E. "Theory of Edge Detection", Proc. Royal Society London, Vol.B-207, 1980, P.187.

[Mil88] Miller, J.W.V. "Image Processing Techniques for Illumination Correction", Proc. SPIE - International Society of Optical Engineering USA, Vol-956, 1988, P.18-25.

[MM89] Malowany, M.E. & Malowany, A.S. "Colour Edge Detectors for a VLSI Convolver", Proc. SPIE - International Society Optical Engineering (USA), Vol-1199, Part-2, 1989, P.1116-1126.

[MMK93] Matas, J., Marik, R. & Kittler, J. "Generation, Verification and Localization of Object Hypotheses based on Colour", Proc. On British Machine Vision Conference, September 1993, Vol.2, P.539-548.

[MMK95] Matas, J., Marik, R. & Kittler, J. "On Representation and Matching of Multi-coloured Objects" Proc. on International Conference on Computer Vision, 1995

[MW86] Maloney, L. & Wandell, B.A. "Color Constancy: A Method for Recovering Surface Spectral Reflectance", Journal of Optical Society America. A., January 1986, Vol-3, No.1, P.29-33.

[Nev77] Nevatia, R. "A Color Edge Detector and its Use in Scene Segmentation", IEEE Trans. on SMC., Vol-SMC-7, No.11, November 1977, P.820-826.

[NS87] Novak, C. & Shafer, S. "Color Edge Detection", Proc. of ARPA Image Understanding Workshop, USC, 1987, P.35-37.

[NS94] Novak, C.L. & Shafer, S.A. "Estimating Scene Parameters from Color Histograms" Journal of Optical Society America. A., Vol. 11, No.11, November 1994, p.3020-3036.

[ODE99] Olatunbosun, S., Dowling, G.R. & Ellis, T.J. "Is Colour an Invariant? " Paper printed in 'Proceedings of the IASTED International Conference, Signal and Image Processing (SIP'99), Nassau, Bahamas, Oct 1999, pp. 179-182, ISBN: 0-88986-267-2.

[Oht85] Y.Ohta, "Knowledge-based Interpretation of Outdoor Natural Color Scenes" Research Note in Artificial Intelligence, Pitman Advanced Publishing Program 1985.

[OKS80] Ohta, Y., Kanada, T. & Sakai, T. "Colour Information for Region Segmentation" Computer Graphics and Image Processing 13, 1980, p.222-241.

[OPR78] Ohlander, R., Price, K & Reddy, D.R. "Picture Segmentation Using a Recursive Region Splitting Method", Computer Graphics and Image Processing 8, 1978, p.3-13.

[OSA53] Committee on Colorimetry Optical Society of America, "The Science of Color" New York: Thomas Y. Crowell company 1953.

[OSA87] Optical Society of America, "Color Appearance" Technical Digest Series, Postconference Edition, Volume 15 1987

[PHJ89] Pakkinen, J.P.S., Hallikainen, J. & Jaaskelainen, T. "Characteristic Spectra of Munsell Colors", Journal of Optical Society America. A., No.6, 1989, P.318-322.

[PK94] Petrov, A.P. & Knotsevich, L.L. "Properties of Colour Images of Surfaces under Multiple Illuminants", Journal of Optical Society America. A., Vol.11, No.10, October 1994, P.2745-2749.

[Qui78] Quinqueton, J. "A New Clustering Method Using a Peano Scanning", International Joint Conference on Pattern Recognition, Kyoto, 1978, P.292-295.

[Rob77] Robinson, G.S. "Color Edge Detection", Optical Engineering, Vol-16, 1977, P.479-484.

[SA81] Sarabi, A. & Aggarawal, J.K. "Segmentation of Chromatic Images", Pattern Recognition vol-13, No.6, 1981, p.417

[SH89] Spann, M. & Horn, C. "Image Segmentation Using a Dynamic Thresholding Pyramid", Pattern Recognition, Vol-22, No.6, 1989, P.719-732.

[Sha85] Shafer, S.A. "Using Color to Separate Reflection Components", Color Research and Application, Vol-10, No.4, Winter 1985, P.210-218.

[SHS99] Seaborn, M., Hepplewhite, L. & Stonham, J. "Fuzzy Colour Category Map for Content Based Image Retrieval", British Machine Vision Conference 1999, P.103-112.

[SI94] Sato, Y. & Ikeuchi, K. "Temporal-Color Space Analysis of Reflection", Journal of Optical Society America. A., Vol.11, No.11, November 1994, P.2990-3002.

[Sim86] Simon, J-C. "Patterns and Operators: The Foundations of Data Representation", Translated by Howlett, J., North Oxford Academic, 1986.

[SLP83] Steven, R.J., Lehar, A.F. & Preston, F.H. "Manipulation and Representation of Multidimensional Image Data Using the Peano Scan", IEEE Trans. on PAMI, Vol-PAMI-5, No.5, September 1983, P.520-526.

[Sti78] Stiles, W.S. "Mechanisms of Colour Vision" Academic Press 1978.

[Swa90] Swain, M.J. "Color Indexing", DPhil Thesis, University of Rochester, New York, 1990.

[Sye92] Syeda-Mahmood, T.F. "Data and Model-driven Selection Using Color Regions", Proceeding on European Conference on Computer Vision 1992, P.115-123.

[Taj84] Tjahjadi, T. "Computer Vision Using Colour Shadows For Object Recognition", DPhil Thesis, The University of Manchester Institute of Science and Technology, United Kingdom, 1984.

[TO90] Tsukada, M. & Ohta, Y. "An Approach to Color Constancy using Multiple Images", Proc. on Third International Conference on Computer Vision, IEEE Computing Society Press, 1990, P.385-389.

[Tom86] Tominaga, S. "Color Image Segmentation Using Three Perceptual Attributes", IEEE Trans. on PAMI, Vol-PAMI-8, No.6, 1986, P.628-630.

[Tom87] Tominaga, S. "Expansion of Color Images Using Three Perceptual Attributes", Pattern Recognition Letters, No.6, 1987, P.77-85

[TP86] Torre, V. & Poggio, T. "On Edge Detection", IEEE Trans. on PAMI, Vol-PAMI-8, 1986, P.147-163.

[TW90] Tominaga, S. & Wandell, B.A. "Component Estimation of Surface Spectral Reflectance" Journal of Optical Society America. A., Vol.7, No.2, February 1990, P.312-317.

[WB81] West, G. & Brill, M.H. "Contributions to the Theory of Invariant of Color under the Condition of Varying Illumination", Journal of Mathematical Biology, Vol-11, 1981, P.337-350.

[WB86] Worthey, J.A. & Brill, H.B. "Heuristic Analysis of Von Kries Color Constancy", Journal of Optical Society America. A., Vol3, No.10, October 1986, P.1708-1712.

[WE95] Walcott, P. A. & Ellis, T. J. "A Colour Object Search Algorithm", Proceedings of 2nd ACCV Singapore 1995, P.243-247

[WE98] Walcott, P.A. & Ellis, T.J. "A Colour Object Search Algorithm", British Machine Vision Conference 1998, pp. 296-305.

[Wes79] West, G. "Color Perception and the Limits of Color Constancy" Journal of Mathematical Biology 8, 1979, P.47-53.

[Wri89] Wright, W.A. "A Markov Random Field Approach to Data Fusion and Colour Segmentation", Image and Vision Computing, Vol-7, No.2, May 1989, P.144-150.

[WS67] Wyszecki, G. & Stiles W.S. "Colour Science: Concepts and Methods, Quantitative Data and Formulas" John Wiley & Sons, Inc.1967

[Zah71] Zahn, C.T. "Graph-theoretical Methods for Detecting and Describing Gestalt Clusters", IEEE Trans. on Computers, Vol-20, No.1, January 1971, P.68-86.

[Zek92] Zeki, S. "The Visual Image in Mind and Brain", Scientific American, September 1992, P.43-50.