Traffic Sign Recognition Based on Human Visual Perception

A Doctoral Thesis
submitted in partial fulfilment of the
requirement for the award of
Doctor of Philosophy
from Middlesex University

Author:

Kunbin HONG

Middlesex University

August 2005

Abstract

This thesis presents a new approach, based on human visual perception, for detecting and recognising traffic signs under different viewing conditions.

Traffic sign recognition is an important issue within any driver support system as it is fundamental to traffic safety and increases the drivers' awareness of situations and possible decisions that are ahead. All traffic signs possess similar visual characteristics, they are often the same size, shape and colour. However shapes may be distorted when viewed from different viewing angles and colours are affected by overall luminosity and the presence of shadows. Human vision can identify traffic signs correctly by ignoring this variance of colours and shapes. Consequently traffic sign recognition based on human visual perception has been researched during this project. In this approach two human vision models are adopted to solve the problems above: Colour Appearance Model (CIECAM97s) and Behavioural Model of Vision (BMV). Colour Appearance Model (CIECAM97s) is used to segment potential traffic signs from the image background under different weather conditions. Behavioural Model of Vision (BMV) is used to recognize the potential traffic signs.

Results show that segmentation based on CIECAM97s performs better than, or comparable to, other perceptual colour spaces in terms of accuracy. In addition, results illustrate that recognition based on BMV can be used in this project effectively to detect a certain range of shape transformations. Furthermore, a fast method of distinguishing and recognizing the different weather conditions within images has been developed. The results show that 84% recognition rate can be achieved under three weather and different viewing conditions.

Acknowledgements

I am grateful to my supervisor Dr. Xiaohong Gao for her work guidance, support and encouragement in this project. I am also grateful for all the constructive feedback she gave upon reading the final draft and this thesis.

I would also like to express my deepest appreciation to Dr. Peter Passmore, my second supervisor. His extensive knowledge and earnest attitude inspired me tremendously. I really enjoyed the inspiring discussions with him and his invaluable suggestions.

I also thank the opportunity to the collaborative institute, A.B.Kogan Research Institute for Neurocybernetics of Rostov State University in Russia. Special thanks go to Dr. Natalia A. Shevtsova. Her superb guidance on the research makes me continuous endeavours on the work to produce fruitful results.

Many thanks go to other former and current research students for the discussions and support. Great thanks go to Dr. Casper F. Nielsen for his valuable suggestions. I am also indebted to my friends Dr. Kok Fong Tan, Dr. Vooi Yap, and Dr. Stephen Batty, researchers at our school for their kindness in sharing their research experience with me.

I am greatly indebted to Prof. Ronnie Luo in Leeds University for his guidance and useful advice and Mr. Song Tao in Derby University (now in Loughborough University) for his useful information.

I offer my gratitude and sincere thanks to my parents for their encouragement love and support. All my love to my wife, Mrs. Qian Yu, who shares my life and work throughout the whole PhD study. Thanks also go to my baby son, Henry, who made so many efforts to be his "quiet".

This work is financially supported by the School of Computing Science, Middlesex University. Their support is gratefully acknowledged.

Contents

Captions

List of Figures

List of Tables

Chapter 1

Introduction

The development of traffic sign recognition systems for assisting drivers has been a growing interest over the past several years [1]. Traffic signs are represented by particular colours and geometric shapes and provide information to assist drivers to handle their cars and to regulate traffic. However, under certain weather conditions, traffic signs are harder to recognize quickly and correctly. Those conditions could lead to traffic accidents and even threaten drivers and passengers' safety. Therefore it is desirable to develop a traffic sign detection and recognition system to alert the driver to the presence of signs.

An automatic real time system requires identification of traffic signs correctly at the right time and at the right place. According to the features of traffic signs, there are two major methods applied for traffic sign recognition based on characteristics of signs. They are *colour-based* and *shape-based* recognition. Colour is a dominant visual feature and undoubtedly represents a key piece of information for drivers to handle. For example, red colour means stop, yellow means danger *etc.* [1]. Therefore it is widely used in the systems designed for traffic sign recognition. Shape is the other important feature of signs. The majority of signs possess four main geometric shapes: octagon, triangle, circle, and rectangle. For example, in the UK, a triangular shape with red colour usually represents a warning sign. Hence, shape also plays an important role in traffic sign recognition.

However, those methods meet problems when processed under different viewing conditions and in a real environment. *Colour-based* techniques in computer vision run into problems if the light source varies not only in intensity but also in chromaticity as well. For example, (1) weather conditions may affect colour appearance and decrease the visibility of traffic signs. (2) outdoor lighting conditions vary from day to night and may affect the apparent colours of traffic signs [2, 3]. No method has been widely accepted yet. Recognition based on shape also meets problems. Causes of errors include traffic signs being distorted when viewing angle changes, scale changes when the car is in motion and the presence of similarly shaped objects within the field of view.

This research aims to develop a new approach based on human visual perception to recognize traffic signs accurately and quickly under different viewing conditions. The theory is based upon the colour perception and behavioural attributes of human eyes. Based on colour contents, a standard colour appearance model CIECAM97s recommended by CIE (International Committee on Illumination/Commission Internationale de l'Eclairage) [4] is used to segment traffic signs from the rest of scene. After segmentation, traffic signs are identified by the other human vision model, called Behavioural Model of Vision (BMV) [5].

The structure of this thesis is organised as follows. Literature review is given in Chapters 2 and 3. Chapter 2 details the background, knowledge, and the current progress in this field of image processing, and specifically in relation to traffic sign recognition. In Chapter 3, a review of two human vision models: Colour Appearance Model of CIECAM97s and Behavioural Model of Vision (BMV), are given. Chapter 4 presents the research based on human vision models. Experimental results are given in Chapter 5. In Chapter 6, system evaluation is described. Finally, Chapter 7 summarises and concludes the work presented in the previous chapters and scope for further work.

Chapter 2
Background and Related Work

This chapter outlines the background and related work of traffic sign recognition. It starts with the importance and requirements of traffic sign recognition, as well as, introduction of features of traffic signs. Then, the relative image processing methods used for traffic sign recognition are presented. Finally, current progress on the development of traffic sign recognition system is overviewed with the indication of advantages and disadvantages. A corresponding solution is proposed.

2.1 Background

2.1.1 Importance of developing traffic sign recognition system

A sign is a device that provides a visual message by virtue of its situation, shape, colour or pattern, and sometimes by the use of symbols or alpha-numeric characters [6]. The function of traffic signs for road traffic is to improve the quality of service of the road network. This quality can be expressed in road safety, travel speed and road capacity, and also driver comfort and convenience.

However, sometimes, these signs may not be noticed until too late. For example, in some weather conditions, it is harder to recognize traffic signs fast and correctly. At night-time, drivers are easily distracted or blinded by headlights of the oncoming vehicles. These situations all make threats to driving safety and sometimes cause traffic accidents. A system of traffic sign detection and recognition system is thus desirable to alert the driver to the presence of signs. It requires the correct identification of traffic signs at the optimal time and place under different viewing conditions.

2.1.2 Requirements of developing traffic sign recognition system

In order to help drivers to recognise a sign accurately, it is very important that the computerised system recognise these signs correctly at the right time under different conditions. This brings two evaluation criteria of driving system: one is accuracy and the other is speed.

Accuracy is an important factor for recognition. As described, a sign recognition system should be used as an assistant for drivers, to warn about the presence of a specific sign such as "go ahead" or "stop", or some risky situation such as driving at a speed higher than the speed limit. Therefore, to recognize traffic sign correctly is a fundamental criterion.

Also, there are two kinds of speed to consider. One is image-processing speed to recognize the right sign at the right time. Another is camera's shutter speed while taking the pictures. Hence, the speed of processing needs to be considered to ensure that traffic signs are recognized within the correct time frame.

To achieve criteria mentioned above, traffic sign recognition uses image processing and computer vision techniques that are suited to the features of traffic signs.

2.1.3 Characteristics of traffic signs

Traffic signs are specially shaped and exhibit coloured patterns used to guide car/train drivers or pedestrians while driving or walking along the road. They are also designed to control traffic at road junctions. Usually, shapes include octagon, triangle, circle, and rectangle for traffic signs. Colour is another character of traffic signs. Usually, there are eight colours red, orange, yellow, green, blue, violet, brown and achromatic [7]. For example, in traffic signs red means stop or warning, blue always means traveller services information and yellow is normally used as background colour for signs, etc. [8, 9]. Table 2-1 illustrates some examples of traffic signs from British road.

Sign Examples	Meaning	Colour	Geometric Shape
	Turn left ahead	Blue	Circle
	One-way traffic	Blue	Rectangle
	No entry for vehicular traffic	Red	Circle
	Stop and give way	Red	Octagon
	Give way	Red	Triangle

Table 2-1: Examples of traffic signs in the UK

From table above, it is clear that a circular or rectangular blue directs the way for drivers and a red coloured sign, of any of the three shapes triangle, circle, or octagon, gives warning information to drivers. The table shows that colour and shape are the main features of traffic signs.

To detect traffic signs automatically, traffic signs are firstly taken by camera or video. Afterwards, these digital pictures are analysed using methods of image processing. The following parts, therefore, briefly introduce knowledge of image processing and colour knowledge relative to traffic sign recognition.

2.2 Image processing in traffic sign recognition

Traffic sign recognition is a field of applied image processing research. It is concerned with the automatic detection of traffic signs in traffic scene images [10]. Normally, traffic scene images taken from a camera/video camera are stored in computer numerically which are called digital images. A digital image (*I*) is described in a 2D discrete array and is divided into M rows and N columns. The intersection of a row and a column is termed a *pixel*. Therefore, a whole image is represented by a rectangular array of *pixels* (*picture elements*) [11]. The pixel value assigned to the integer coordinates [m, n] with {m = 0, 1, 2, ..., M-1} and {n = 0, 1, 2,..., N-1} can be described as I [m, n] and usually is digitized to [0, 255] for monochromatic images. An example of image is shown in the Figure 2-1 below [12].

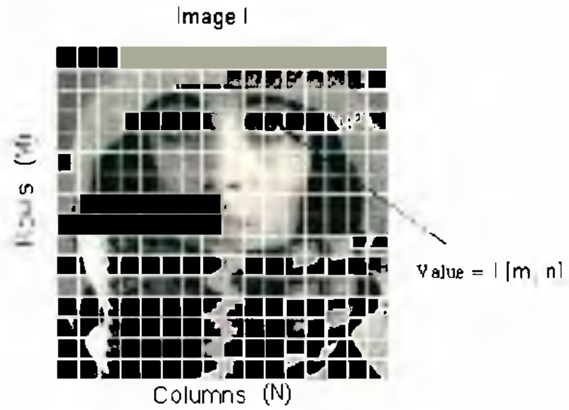


Figure 2-1: Example of an image [12]

Figure above illustrates a grey (monochromatic) image. Value of pixel I [m, n] is between [0, 255]. The whole image is an array of pixels values.

However, traffic signs images are colour images and therefore have a different description when compared to grey images. In general, colour images are described by utilizing colour spaces [13].

2.2.1 Colour representations — Colour spaces

In [7], the CIE points out that traffic signs are coloured to ensure the proper guidance and control of the various forms of transport, so as to increase safety and to facilitate rapid movement. In other words, colour is an important characteristic of traffic signs. Before describing any solution for sign detection based on colour, it is necessary to give a brief introduction to colour and its expression in image processing.

2.2.1.1 Basic theory of colour

In 1666, Newton has found that the white light consists of visible spectrum which includes all visible colours ranging from red, orange, yellow, green and blue to violet [14]. The experimentation is shown, simply, in Figure 2-2 [14].

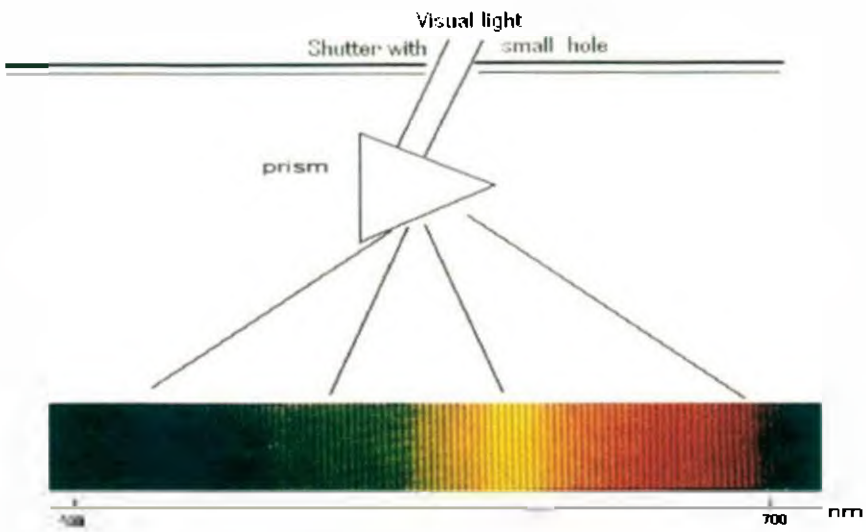


Figure 2-2: The experiment of Newton [14]

Based on the experimental facts of Newton's famous experiments, the road for colour investigation was open for progress. Now, scientists have found that colour is part of electromagnetic spectrum with energy in the range of 380nm to 780nm wavelength [3].

However, the human eye is incapable of analyzing colour into its spectral components by wavelength. Colour is the perceptual result of light, object and eyes [3, 15-17]. The light source generates light illuminating an object. Some part of the spectrum of the light is reflected from the object and is subsequently measured by an observer such as our light-sensitive eyes or by a colour camera. The measured light is then sent to our brain where the colour of the light is observed. The description above shows that an observed colour contains three essential elements: light, object and observer shown as Figure 2-3 [18].

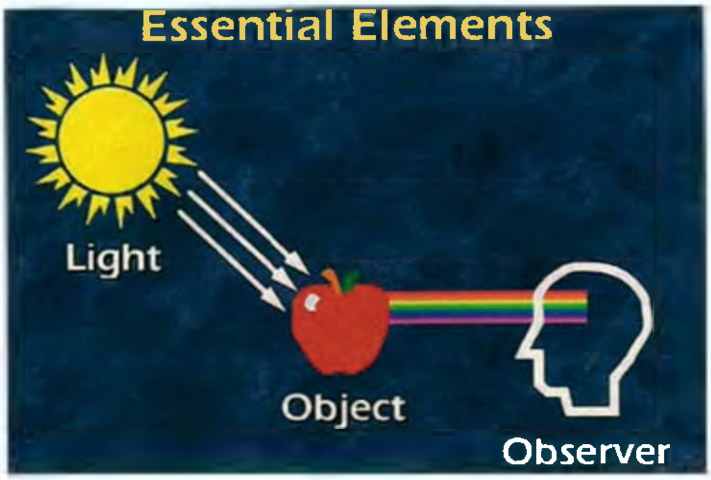


Figure 2-3: Example of colour perceptual by 3 components [18]

Light source

The main light source is the sun. Further, some of artificial light sources exist such as fluorescent lamps, or by heating up material. There are two ways to characterize light sources. One is light spectral power distribution (SPD). SPD is the amount of radiant power at each wavelength of the visible spectrum and denoted by $P(\lambda)$. The other common term to characterize light sources is colour temperature. Colour temperature corresponds to the temperature of a heated black body radiator. The colour of the black body radiator changes with temperature. For example, the radiator changes from black at 0 K (Kelvin), to red at about 1000 K, white at 4500 K to bluish white at about 6500 K. The colour temperature of the sun may vary during the time of the day (e.g. reddish at sunrise and bluish at noon) and the weather conditions(e.g., sky with/without clouds) [3]. The Commission Internationale de l'Eclairage (CIE) recommended that the average daylight has the colour temperature of 6500K and is denoted by D65.

The Object

Coloured materials are called objects or samples. The colour of an object is defined by the reflectance, a function of wavelength. Reflectance is the ratio of the light reflected from a sample to that reflected from a perfect reflecting diffuser identically irradiated, and is denoted by $R(\lambda)$ [17].

Usually, the colour reflected from an object is the product of SPD of the illuminant ($P(\lambda)$) and the spectral reflectance of the object ($R(\lambda)$) and is computed by $PR(\lambda) = P(\lambda)R(\lambda)$ [3, 17].

The observer

The observer measures light coming directly from a light source $P(\lambda)$ or light which has been reflected (or transmitted) from objects in the scene $R(\lambda)$. The observer can be a colour camera or human eyes. For the human eye, the retina contains two different types of light-sensitive receptors, called rods and cones. Rods are more sensitive to light and are responsible for vision in twilight. The cones are responsible for colour vision and consist of three types of receptors sensitive to long (red), middle (green) and short (blue) wavelengths. The response of these three cones with wavelength is drawn in the Figure 2-4 below [19].

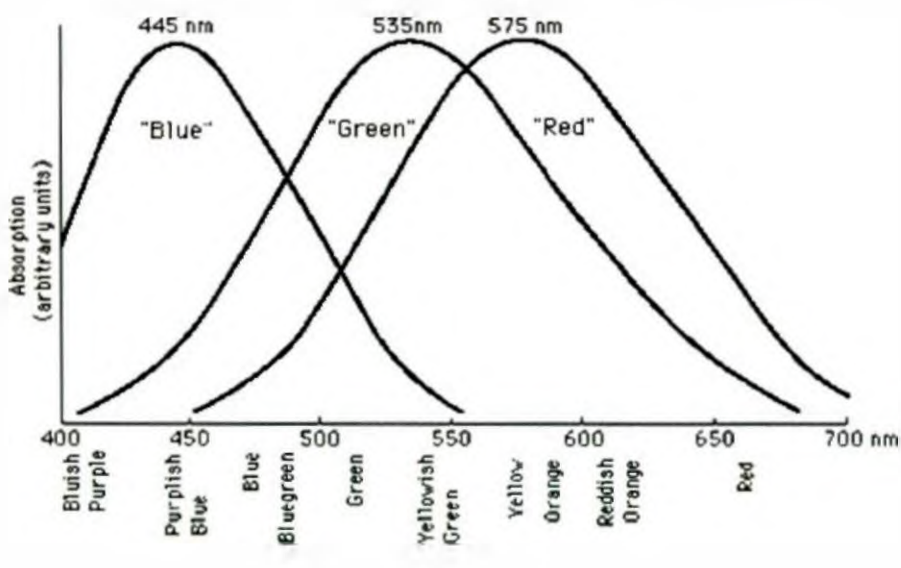


Figure 2-4: Response of three cones with wavelength [19]

However, the sensation of a human observer can not be measured by an objective instrument. Therefore, experiments have been conducted on human observers to measure the spectral sensitivities of the human eye. The observers were asked to match a test light, consisting of only one wavelength, by adjusting the energy level of three separate primary lights which are Red (700nm), Green (546.1nm), and

Blue (435.8nm) recommended by CIE. At each wavelength the amount of energy was recorded for the three primary colours. The results of this matching are called colour matching functions, usually denoted as $\bar{x}(\lambda)$, $\bar{y}(\lambda)$ and $\bar{z}(\lambda)$. Also these colour matching functions also can be thought as the colour response of the eye [2, 14]. It is displayed in the Figure 2-5 below [14].

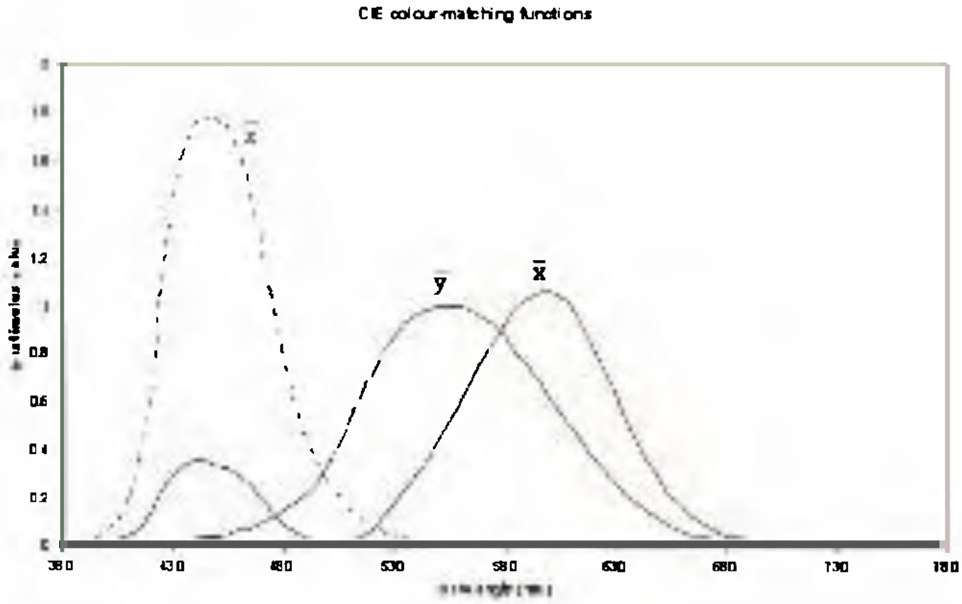


Figure 2-5: Colour matching function for human [14]

In conclusion, the colour can be measured as a vector of three measurement $\rho = [\rho_1, \rho_2, \rho_3]$ given by

$$\begin{aligned}
\rho_1 &= \int_\lambda P(\lambda)R(\lambda)x(\lambda)d\lambda \\
\rho_2 &= \int_\lambda P(\lambda)R(\lambda)\bar{y}(\lambda)d\lambda \\
\rho_3 &= \int_\lambda P(\lambda)R(\lambda)z(\lambda)d\lambda
\end{aligned} \tag{2-1}$$

where λ denotes wavelength, $P(\lambda)$ is the spectral power distribution of illuminant. $R(\lambda)$ is the spectral reflectance or transmittance factor of the object depending if the object is reflective or transmissive medium, and $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, $z(\lambda)$ are colour matching function of eyes [2, 14].

The description of colour above explains how human perceive and measure colour. Similarly to the human eye, which has three kinds of receptors (cones) for sensing specific parts of the spectrum [2], CCD cameras detect colour with three sensors, each one for a "primary" colour: (red, green and blue). Therefore, an object seen by a camera is represented by a collection of three-coordinate (R, G, B) pixels. This collection forms a data space containing the pixels. It can also be called colour space or called colour model [20].

2.2.1.2 Colour spaces used in Image processing

The previous section introduced the basic colour knowledge about physical colour properties, three basic elements of observed colours and colour matching functions. However, it is difficult to cope with those physical colour properties when dealing with digital colour images. Therefore, sciences used colour spaces/colour models to express colour so that a digital colour image can be easily handled.

The colour space is a kind of mathematical representation, which is three-dimensional orthogonal coordinate system [21] . Apart from the *RGB* colour space, there are many ways to represent colour depending on the application. In different colour spaces, the three axes represent different meanings. The digital image can be treated within different colour spaces, which facilitates the operations [21, 22]. The colour spaces that are utilized in conjunction with traffic sign recognition are described in following sections.

2.2.1.2.1 RGB colour space

The *RGB* colour space is the most used colour space for image processing. This space is the basic one because colour cameras, scanners and displays are most often provided with direct R (Red), G (Green), B (Blue) signal input and output.

To represent the *RGB* colour space, a cube can be defined on the R, G, and B axes shown in Figure 2-6 below. Each colour being described by its components (R, G, B), is represented by a point and can be found either on the surface or inside the cube. All grey colours are placed on the main diagonal of the cube from black ($R=G=B=0$) to white ($R=G=B=$max).

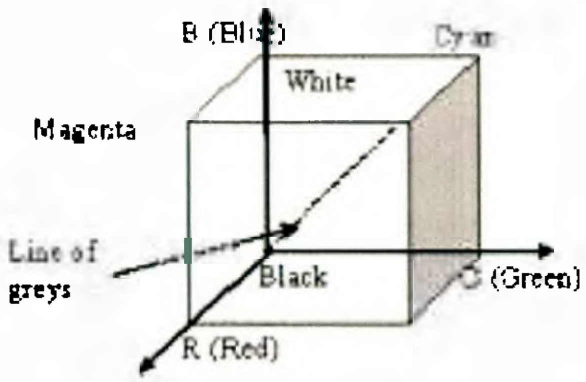


Figure 2-6: *RGB* colour space [21]

However, it is hard to visualize a colour based on R, G, and B components. Also, the three coordinates are highly correlated [21]. As a consequence of this strong correlation, variations in ambient light intensity have a disastrous effect in RGB by shifting the clusters of colour pixels toward the white RGB = [255, 255, 255] or the black corner RGB = [0, 0, 0] of the cubic space. From a colour point of view, an object can thus be unrecognizable if it is observed under different intensities of illumination.

RGB colour space is not directly related to the intuitive notions of hue, saturation and brightness. For this reason, other colour spaces have been developed which can be more intuitive, in manipulating colour and were designed to approximate the way human's perceive and interpret colour. They are the *HSI*, *HSV* and *HSL* colour spaces.

2.2.1.2.2 HSI, HSL and HSV colour spaces

HSI, HSL and HSV are perceptual colour spaces. In the perception process, a human can easily recognise attributes of colour: intensity, hue and saturation. Hue represents the actual colour or tint information. Saturation indicates how deep or pure the colour is. Intensity is simply the amount of light. *HSI* colour space can be easily transformed from *RGB* colour space.

According to [13, 23], the conversion from *RGB* to *HSI* is:

$$
\begin{aligned}
&H = \arccos\left|\frac{[(R-G)+(R-B)]/2}{\sqrt{(R-G)^2+(R-B)(G-B)}}\right|, \text{ where } H = 360^{\circ} - H \\
&\text{if } (B/I) > (G/I) \\
&S = 1 - \frac{3\min(R,G,B)}{R+G+B} \text{ or } \max(R,G,B) - \min(R,G,B) \\
&I = (R+G+B)/3
\end{aligned}
\qquad \textbf{(2-2)}
$$

The *HSI* colour space can be described as the following Figure 2-7 [24].

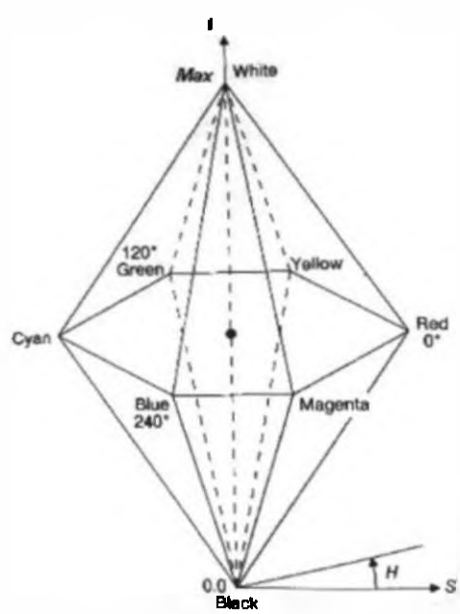


Figure 2-7: *HSI* colour space [24]

This picture shows that intensity *I* is changing from 0 to max (usually is 255), saturation *S* is changing from centre (0) increasing to max (1) or from 0 to 255, and hue *H* is changing from red as a circle ranging from 0 to 360. *HSL* and *HSV* colour space are similar to *HSI* colour space. Although these *three* colour spaces give us a more intuitive description of colour, they do not establish uniform colour spaces. Therefore, CIE introduced a colour space which is CIELUV colour space [22, 25].

2.2.1.2.3 CIE-based colour spaces --- CIELUV

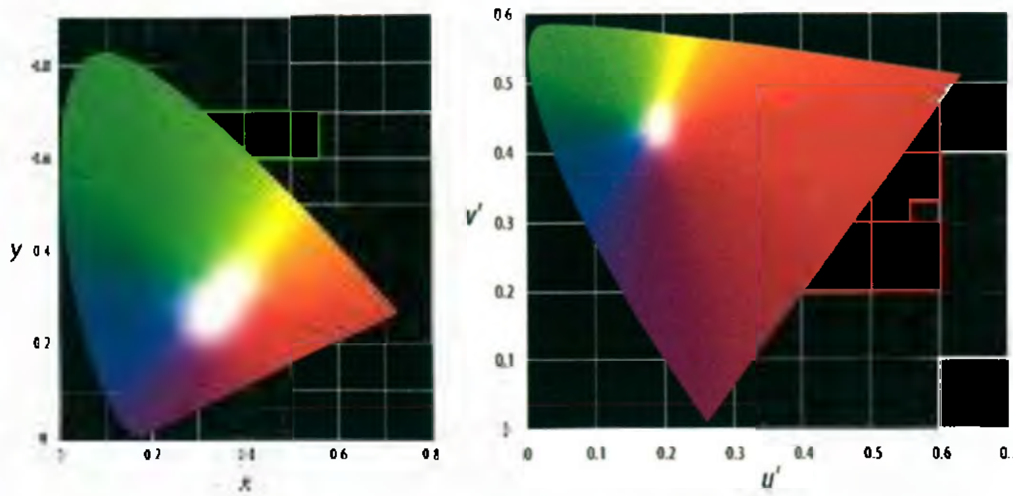
In 1976, the CIE defined a new colour space CIELUV to enable us to get more uniform and accurate models. Sometimes, it is also called *universal colour space* [26]. This colour representation results from work carried out in 1931 by the Commission Internationale d'Eclairage (CIE). The CIE LUV colour space is a perpetually uniform derivation of a standard CIE*XYZ* space [2, 27]. Hence, it is essential to briefly introduce the CIE*XYZ* colour space and chromaticity colour space before giving more detail of CIELUV colour space.

- **CIE*XYZ* and chromaticity colour space *xy (u'v')***

As described before, colour has commonly been measured by viewing combinations of the three standard elements. It can be expressed by equation (2-1). Based on the formula and measurement, CIE in 1931 has defined the CIE*XYZ* colour space which is relative to the standard observer [2] and also is called the tristimulus colour space [28, 29]. Therefore, each colour can be represented by *XYZ* colour space. However, the tristimulus values of them are not correlating [14]. Hence, CIE has defined a colour space normalized from *XYZ* colour space. It is defined as following formula [14].

$$
\begin{aligned}
x &= \frac{X}{X+Y+Z} \\
y &= \frac{Y}{X+Y+Z} \\
z &= \frac{Z}{X+Y+Z}
\end{aligned}
\tag{2-3}
$$

Because $x+y+z=1$, x and y can be used to describe the colour, which is called xy chromaticity co-ordinates colour space. The example of x, y chromaticity diagram is shown in the following Figure 2-8 (a) [18].



(a) (b)

Figure 2-8: Chromaticity colour space *xy* and *u'v'* [18]

However, the distribution of the colours on *xy* chromaticity coordinate is not uniform. Then CIE has recommend a new colour chromaticity which is *u'v'* in 1976 [14, 18]. The example of *u'v'* colour space is shown above in Figure 2-8 (b) and the values are obtained by the following equation (2-4).

$$u' = \frac{4X}{X + 15Y + 3Z} \qquad v' = \frac{9Y}{X + 15Y + 3Z} \tag{2-4}$$

- **CIELUV colour space**

Chromaticity diagrams show only proportions of tristimulus values, and not their actual magnitudes and they are only strictly applicable to colours all having the same luminance. Colours however, differ in both chromaticity and luminance, and some methods of combining these variables are required. In 1976, the CIE used the CIELUV colour space as the perceptually uniform colour spaces whose expressions are defined below [14, 25].

$$L^* = 116 f(\frac{Y}{Y_0}) - 16 \text{ , if Y/Y0 >0.008856, else } L^* = 903.3 \cdot (\frac{Y}{Y_0})$$

$$u^* = 13 \cdot L^* \cdot (u' - u'_0)$$

$$v^* = 13 \cdot L^* \cdot (v' - v'_0) \qquad \textbf{(2-5)}$$

$$H = \arctan gent(v^* / u^*)$$

$$C = \sqrt{(u^*)^2 + (v^*)^2}$$

where u'_0, v'_0 are the values of u', v' for the appropriately chosen reference white. The L component has the range [0,100], the U component has the range [-134, 220], and the V component has the range [-140, 122]. *H* is the angle of Hue which express by angle ranging between 0 and 360. *C* is the value of Chroma which ranging from 0 to 260. So, CIELUV can also produce a colour space is *HCL*. Also [30] gives the example figure of CIELUV which is shown in Figure 2-9 below.

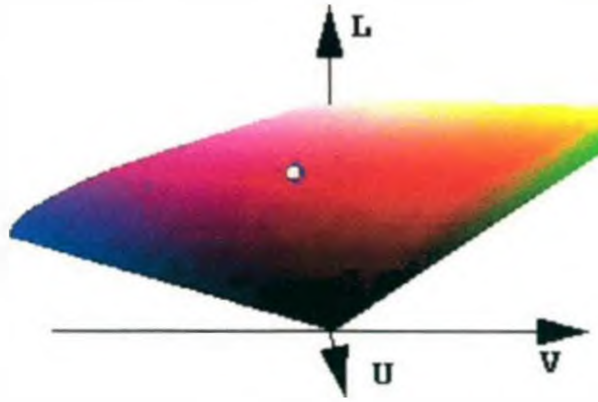


Figure 2-9: CIELUV colour space [30]

2.2.1.2.4 Video colour space --- YIQ

RGB colour space is widely used in storage. However, it is not suitable for television transmission. Therefore, the colour space of *YIQ* colour space is created for colour television transmission, in which Y represents luminance and the other two components carry colour information [26, 31]. It can be easily transferred from *RGB* colour space.

$$
\begin{aligned}
Y &= 0.299R + 0.587G + 0.114B \\
I &= 0.596R - 0.274G - 0.322B \\
Q &= 0.212R - 0.523G + 0.312B
\end{aligned}
\qquad \textbf{(2-6)}
$$

The Y parameter has the range [0, 255], the I parameter has the range [-152, 152], and the Q parameter has the range [-133.36, 133.62].

The colour spaces described here are major colour models used in image processing. According to different colour spaces, a colour image can be described by three components such as R, G, and B; or H, S and I, or Y, I and Q. The examples are shown in the Table 2-2 below. The *RGB* value of example colour patch of traffic signs below were taken from Paint Shop 7.0 [32].

Signs	Colour pixel	*RGB* Colour space	*HSI* Colour space	*YIQ* Colour space
		R: 0, G: 119, B: 192	H: 202°, S: 1, I: 103.67	Y: 91.74, I: -94.43 Q:-2.33
		R: 237, G: 23, B: 31,	H: 359.43°, S: 0.92, I: 97	Y:87.90 I:124.96 Q: 47.89

Table 2-2: Examples of colour spaces expression for colour blue and red of traffic signs

Table 2-2 illustrates that one colour can be expressed by different colour spaces. Hence, a whole digital colour image can be expressed by different colour spaces according to the application of image processing.

2.2.2 Shape representations and matching

Shape is another important property of traffic signs, which can be easily understood. For example, in the UK, triangle means warning; circle with colour red can mean 'No' [33]. Hence, shape based image processing also plays an important role [9] in traffic sign recognition. Usually, shape-based recognition considers three kinds of methods: shape finding by Hough transform, shape description by chain code and template matching.

2.2.2.1 Shape finding by Hough transform

Borders are usually used to describe the shape of objects, which is called contour-based shape representation and description [34]. In order to find the contour of shape, Hough transform is normally used.

Hough transform has excellent shape description abilities [34] and is a widely used method to find features of a particular shape such as line and circle within an image. It enables the recognition of curves within an image (Figure 2-10 a) by recognizing points in a transformed parameter space (Figure 2-10 b) [34]. The simple example of *Hough* transform for finding a straight line is stated below.

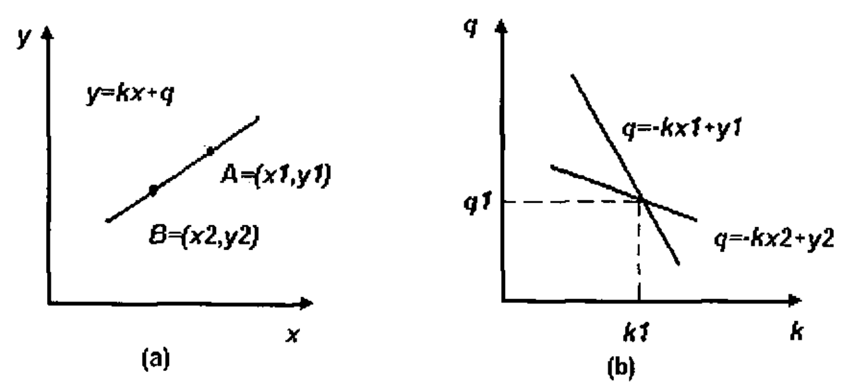
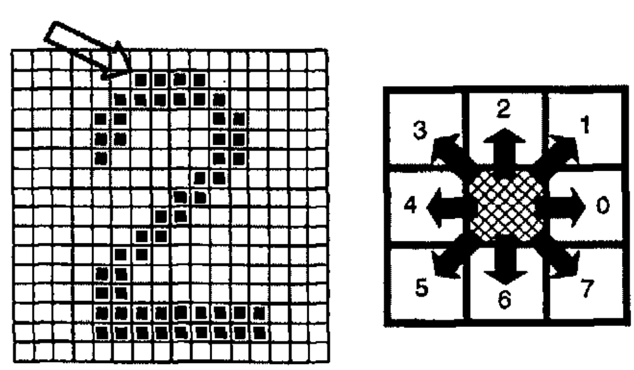


Figure 2-10: Example of Hough transform for finding a straight line [34]

A straight line in an image can be defined by two points A (x1, y1) and B (x2, y2) shown in Figure 2-10 (a). All straight lines through the point A and B are given by the expression $y1 = k * x1 + q$ for some values of *k* and *q* (*k* defines the slope and *q* defines intercept). This means that the equation of $y1 = k * x1 + q$ can be interpreted as an equation $q = -k * x1 + y1$ in the parameter space k, q shown in Figure 2-10 (b). The only common point ($k1$, $q1$) of both straight lines in the k, q parameter space is the point, which represents the only existing straight line connecting points A and B with k and q (slope and intercept) in the original image. By using the same technique, a circle in an image can be obtained by recognizing the point in its corresponding parameter space [34, 35].

2.2.2.2 Shape description by chain code

Hough transform provides a method to find shape contours. *Chain code is* used for the description of object borders [34, 36, 37]. The border is defined by the co-ordinate of its reference pixel and the sequence of symbols corresponding to the line of the unit length in several pre-defined orientations [38]. Figure 2-11 (b) shows that one pixel with its 8 pre-defined orientations. These 8 directions hint this pixel's 8 neighbours. It starts from horizontal (0) and rotates anti-clockwise to direction (7) in figure 2.7 (b). One example of chain code for letter "2" is described below Figure 2-11 (a) [34, 38].



(a) (b)

Figure 2-11: Example of chain code description [34]

Chain code description of letter "2" begins from the starting referent point marked by an arrow following clockwise direction. If the following pixel exists, the orientation of previous pixel pointed to the following pixel is recorded. Finally, chain code of this pattern “2” is:

0007766555555660000000644444442221111112234445652211.

2.2.2.3 Template matching

Template matching [34, 39] is another widely used method for identifying objects in image processing. Templates of objects are first created; then the result of matching will be obtained by comparing the template to the image through Euclidean distancc and correlation [40].

Euclidean distance can be described using the following formula:

$$Eu\_dis = \sqrt{\sum_{i,j}(f(i,j) - g(i,j))^2} \qquad \textbf{(2-7)}$$

where $f(i, j)$ is the image, $g(i, j)$ is the template. i, j is the coordinate of corresponding positions of f and g. Therefore, the original matching is 0, and the smaller the Euclidean distance is, the better matching.

Correlation is another matching criterion, which describes a match between f and g located at position (u, v) [34]. It is expressed by the formula below.

$$C(u,v) = \frac{1}{\max_{(i,j)\in V}\left|f(i+u, j+v) - g(i,j)\right| + 1} \qquad \textbf{(2-8)}$$

where V means the set of all image pixels in the processed images, f is the image to be processed and g is the template. Therefore, perfect matching is 1 and the bigger the value of C, the better the matching is.

This section introduced the techniques of image processing that are widely used in the field of traffic sign recognition. It includes both colour based and shape based techniques. The next sections will overview current progress in the development of traffic sign recognition systems, based on the above techniques.

2.3 Overview of the current progress on development of traffic sign recognition systems

Progress on Road Sign Recognition (RSR) or Traffic Sign Recognition (TSR) research started in Japan in 1984 [10]. There were several laboratories working on the area of developing an intelligent recognition system. Piccioli [41] in Italy developed a method to detect the triangular and circular traffic signs. In USA, Douville [42] focused on classification traffic signs into different groups based on traffic signs meaning (warning, stop, cross, speed). Shneier [43] developed an algorithm to detect the signs in a video stream. The research focused on the daytime environment. Belmann [44] focused on using neural network to detect traffic signs taking by a video camera under daytime conditions. In Spain, Escalera [45, 46] leaded a research group to develop a driver support systems. The research on traffic signs recognition used deformable template to detect the road signs taking by a video camera. In Germany, Büker [47] and Gavrila [48, 49] leaded two rescarch groups to detect traffic signs separately. In Taiwan, Fang [50] built a neural network system to detect the 'speed' traffic signs and 'warning' signs taking by a video camera.

To recognize the traffic signs, there are two steps used in research groups. First, colour characteristic is normally used to extract traffic signs from rest of scene of an image. This extraction is called *segmentation* [34]. After segmentation, the shapes are used to recognize traffic signs. The procedure can be described by the following Figure 2-12 below.

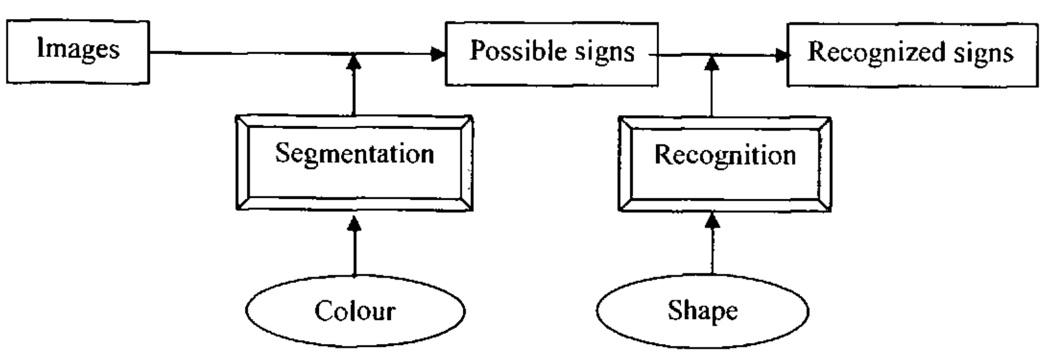


Figure 2-12 Traffic sign recognition procedures

The figure above apparently describes the two procedures of traffic sign recognition: segmentation and recognition. According to the features of traffic signs, these systems applied the two major methods mentioned, colour-based and shape-based methods.

2.3.1 Traffic sign recognition based on colour

Colour provides very significant information for drivers. Most traffic sign recognition systems found in the literatures are based on colour and acknowledge its importance. Normally, colour is used to segment traffic signs from the images. To segment traffic signs from the rest of image, the normally used method is first to get the ranges of three components based on colour spaces as thresholds. Then, by comparing these thresholds, the potential regions of traffic signs can be extracted. This segmentation method is called *thresholding* [34]. The following systems used in sign detection provide evidence that colour is one of key features. Most research uses *RGB*, *HSI* colour spaces to segment traffic sign.

Kehtarnavaz *et al.* [51] processed 'stop' signs, which are mainly in red colour, using colour to segment signs. The colour images were transferred from *RGB* to hue, saturation and intensity (*HSI*) colour space. Kehtarnavaz defined the thresholds according to the statistical study on 'stop' signs. The recognition was based on the shapes. Only specific sign, "Stop", was detailed in the paper.

Kellmeyer [52] has created a system to detect 'warning' signs by using two colour spaces *HSI* and *RGB*. Firstly, Kellmeyer converted each pixel's *RGB* values to *HSI* values and then used the natural log function to limit the saturation of pixels to reduce the number of non-colourful pixels. After reducing non-colourful pixels, the segmentation was based on *RGB* colour space, which was obtained by transferring the *HSI* space back to *RGB*. The system was able to detect 55% of the 'warning' signs within the 55 images. Kellmeyer recognised that this test only represents a single day's viewing conditions and that these results should not be extrapolated for all conditions.

Nicchiotti *et al.* [53] transferred the colour information of the 'danger' and 'prohibition' signs from *RGB* to hue, saturation and lightness (*HSL*) colour space. And the segmentation was performed on *HSL* colour space. The template matching was carried out for recognition following the segmentation. However, there was no recognition result and processing time mentioned in the project.

Paclík *et al.* [54] tried to classify traffic signs into different colour groups. By this method, a pixel was classified into six basic colours (white, black, red, blue, green and yellow). The hue, saturation, and brightness (*HSV*) colour space was used by transformation from *RGB* space. White and Black were detected by the thresholds of V and S values because they do not have chromatic information. Red, blue, green and yellow were detected by the thresholds of H, V and S values. In this project, Paclík concentrated on classifying signs into different colour groups. However, the results were obtained from traffic signs boards directly, which means the image scene only contain traffic signs. Also, the images were taken under general illumination conditions.

Zadeh, Kasvand, and Suen [55] created sub-spaces in *RGB* space which enclosed most of the variations for each colour in the image of each traffic signs. The sub-spaces in *RGB* space was formed by training cluster of signs. The sub-spaces were the ranges of colour traffic signs and used to segment the traffic signs. There are no detailed results expressed in the paper.

An approach was made by Priese and Rehmann [56-59] applying a parallel segmentation method based on *HSV* colour space. The subsets of pixels in the *HSV* space were grouped so that every object can be represented. The colour classes have been set up according to the types of signs. The project concentrated on the classification of traffic signs such as 'prohibition' signs. The approach only focused on which group the traffic sign was.

Colour plays an important cue for traffic sign recognition. From those work cited here, many research groups pay attention to use this characteristic to segment and recognize traffic signs. On the other hand, shape also plays an important role for traffic signs as much research has also concentrated on identifying traffic signs based on shape.

2.3.2 Traffic sign recognition based on shape

After extracting potential traffic signs from images, the following step is to recognize traffic signs. To recognize traffic signs, shape, another feature of traffic signs, is widely used. As described in section 2.1.3, traffic signs are typically in the shapes of triangle, rectangle, circle or octagon. Therefore, any shape-based traffic sign detection algorithm will need a substantial amount of research and development to correctly identify traffic signs reliably. This section introduces the various research groups that have taken this shape based approach.

Kehtarnavaz *et al.* [51] tried to identify the 'stop' sign. Kehtarnavaz performed edge detection and applied the *Hough* transform to describe the shape of signs, whose boundary contour was represented by eight straight lines. In this paper, only 'stop' signs have been testified and few of samples have been mentioned.

Piccioli *et al.* [41] concentrated on geometrical reasoning for sign detection. In their research, the authors assumed that the set of interesting road signs only contain triangular and circular signs. Hence those possible signs similar to triangle can be segmented using the horizontal or having a slope of the ranges [60-ε, 60+ε], [-60-ε, -60+ε] degrees, where ε is the deviation from 60 calculated from samples. They also used *Hough* Transform to detect the circles. However, they just focused on the triangular and circular signs recognition and also the recognition results were not given.

Prince [60] also detected traffic signs based on geometrical analysis of the edges. The method was to describe the road signs by a minimum set of interest points of geometrical characteristics of signs. Furthermore, signs were recognized by comparing those points between model and signs. However, he only tried to identify triangular-shaped and rectangular-shaped road signs.

A Hierarchical Structure Code (HSC) was presented by Büker *et al.* [61] to use for traffic sign recognition including triangle, rectangle and circle. The researchers created the sign database using HSC which consisted of an encoding and a linking procedure. By using this method, a transition was formed from the signal space of an image into the space of its symbolic representation. When recognising the signs, the symbolic representation was used to compare to the representation of signs in a database. However, no more experimental results have been explained in the paper.

A method was proposed by Gavrila *et al.* [48, 62]. They used a template-based correlation to identify the potential traffic signs in the image. By matching the similarity between a segmented template *T (feature template)* and a segmented image *I (feature image)*, the sign was recognised. The research aimed to recognise only circular and triangular (up/down) traffic signs as seen on highways and secondary roads.

Puntavungkour *et al.* [63] also used template matching by the normalized Euclidean Distance. The research focused on the grey-level image and few tests were demonstrated.

Douville [42] has developed a system to detect the 4 traffic signs (Warning, Stop, Speed, Cross) based on their shapes. Warning signs have triangular shapes, stop signs have octagonal shapes, speed limit signs have rectangular shapes, the cross signs have diamond shapes. It was assumed that the traffic signs have well-defined shapes and edges. Those geometrical features of traffic signs were built into a database. One neural network was created to classify these 4 traffic signs. The features of the signs were inputs of the network system. The system has the four outcomes which correspond to with the 4 signs. Modified template matching method was used as criteria for classification.

Many research groups concentrated on identification algorithms based on geometrical shapes of traffic signs. Those methods show that shape-based traffic sign recognition can identify some specific traffic signs.

The systems reviewed above utilized the two properties of traffic signs as their indices. One is colour and the other is shape. The following table, shown in Table 2-3, summarizes those systems.

Research Groups	**Colour**			**Shape**		**Traffic Signs**
	RGB	HSI	CIELUV	Representation	Template Matching	
Kehtarnavaz		•		•		STOP
Kellmeyer	•	•				Warning
Nicchiotti		•			•	Danger and Prohibition
Paclik		•				*
Zadech	•					Speed
Priese		•				*
Piccioli				•		Triangle and Circle
Prince				•		Triangle and rectangle
Gavrila					•	Circle and Triangle
Puntavungkour					•	Few signs
Douville					•	*

Table 2-3 Summary of traffic sign recognition mentioned above

•: which method is used

*: Classify traffic signs into different groups

Table 2-3 illustrates each system and the methods it utilizes for recognizing. From table above, it can be easily found that the systems above are built for special sets. For example, Kehtarnavaz concentrated on 'Stop' sign. Prince concentrated on triangular and rectangular signs. Some research groups concentrated on classifying traffic signs to different group based on colour or shape features. Nevertheless, because limited computer power was available at that time, trade-off between accuracy and speed of image processing existed for traffic sign recognition. Furthermore, although those methods can identify some types of traffic signs, it has some limitations when running in real environment.

2.4 Limitations of existing methods and solution

Colour-based and shaped-based traffic sign recognition are two main methods applied in traffic sign recognition. Most studies solve the problem of identifying specific traffic signs, such as 'stop' sign. The recognition rate analysed by such studies is built on special sets such as 'stop', or 'warning' signs.

However, in real environment, these two methods will meet problems. For example, colour appearance changes under different weather conditions; shapes are distorted when viewing from different viewing angle. Most of the studies mentioned above do not consider the change affected by the viewing conditions.

- In the real environment, colour appearance changes in different viewing conditions. The observed colour may vary with a change in the intensity and colour of light sources. For example, (1) weather conditions may affect colour appearance and decrease the visibility of traffic signs, e.g., sky with/without clouds; [2, 3, 64]. (2) outdoor lighting conditions vary from day to night and may affect the apparent colours of traffic signs [6, 65]. This results in the changing of colour appearance. One example of colour appearance changing under different weather conditions is shown in the Figure 2-13. Two pictures in Figure 2-13 are taken in the same place on a sunny day and rainy day respectively. Top picture (a) taken on a sunny day looks more colourful than the bottom picture (b) taken on a rainy day. From the work cited above, most colour-based techniques in computer vision run into problems if the illumination source varies not only in intensity but also in colour as well.

(a) Picture taken on a sunny day

(b) Picture taken on a rainy day

Figure 2-13: Examples of colour appearance change under different weather conditions

- In the real envirnnment, shape has been affected by many factors. The distorted signs affected by different viewing angles also bring forward the difficulty of recognition. The recognition changes with even small differences in rotation, scale and noise. Some signs do not always have a perfect shape (corners may be torn). Moreover, there are signs with different meanings with same shape. For instance, '20' (MPH) limited speed is the same 'circle' shape with '40' (MPH) limited speed. Some examples of traffic signs in real environment are shown in the Table 2-4. In this table, picture *(a)* shows a distortion example viewed from different angle. Picture *(b)* gives an example of rotation traffic signs. In picture *(c)*, noise disturbs on the sign by leaves and shadows. Picture *(d)* shows an example of similar shape with different meaning signs. Besides, there are also many similar shapes present within a natural urban environment that are not traffic signs. This also brings forward the difficulties of recognition. Hence, the recognition based on shape is not reliable.

			
Distortion	Rotation	Noise	Same shape (circle) different meaning
(a)	*(b)*	*(c)*	*(d)*

Table 2-4: Some examples of traffic signs under real environments

Due to the limitations of previous work, new methods to recognise traffic signs in real environments accurately and quickly are proposed. The solution may be tuned to the human visual system.

This chapter reviewed the significance of traffic sign recognition and the characteristics of traffic signs. The knowledge of image processing based on sign features of shape and colour are introduced. Following the introduction of image processing, previous researches on traffic sign recognition are reviewed based on colour-based and shape-based method. Finally, the limitations of current research on traffic sign recognition in different viewing conditions are presented. To overcome the limitations, the method simulating human eye's working theory is adapted to recognize traffic signs under different viewing conditions. The next chapter will introduce two human vision models utilizing in this project for traffic sign recognition.

Chapter 3

Study of two Human Vision Models

This chapter will examine two human vision models based on human colour appearance and human eyes' behavioural processes respectively. Colour and shape are the most informative features for traffic signs. In order to take full advantage of these characteristics in this study, they are again used for traffic sign recognition. Colour is applied for segmentation and shape is for recognition. To overcome the shortcomings of existing methods, two human vision models are used for segmentation and recognition respectively. One of vision model is CIE Colour Appearance Model (CIECAM97s) [66]. This standard model was recommended by CIE (Commission Internationale de l'Eclairage) in 1997 to measure colours under a wide range of viewing conditions. It models how the human visual system perceives the colour of an object under different lighting conditions and with different backgrounds. In other words, this model can estimate a colour appearance as accurately as an average observer can. The other vision model is developed by Russian scientists I.A.Rybak *et al.*[5, 67] and is called the Behavioural Model of Vision (BMV). BMV has the capability of recognizing complex images (e.g. faces) invariantly with respect to shift, rotation, scale and some degree of distortion.

3.1 Colour appearance model

As described in chapter two, colour is the perceptual result of looking at objects illuminated by light through human eyes [3, 15-17] and colour appearance changes with viewing conditions. For example, the red tomato seems to be more colourful outdoor in bright sunlight than indoor. Taken another example, a piece of paper is seen indoor and outdoor, in which a white patch and a grey patch are drawn side by side. When we take the paper into room and put it outside under bright shinny day, the patches of grey and white look less bright in the room than outside. Also, when we see same objects under different illumination, the colour appearance of objects is different. The figures below are taken in the laboratory, which illustrates the difference under different light sources. When taking photos, the laboratory is dark and the lighting comes from the box.

(a) Picture taken under light source D65

(b) Picture taken under light source F

Figure 3-1: Different colour appearance for colour chart taken under different light source

The picture in Figure 3-1 (a) is taken under light source D65 and the picture in Figure 3-1 (b) is taken under light source F. Figure 3-1 shows that colour appearance is affected by the light source.

This phenomenon can cause severe problems in colour control and recognition. For instance, in the surface industries the colourist needs to know the degree of colour change across a wide range of illumination conditions so that the industry can produce the correct surface colour, which can be predicted same by human under different illumination conditions. To overcome those problems, human observers are used to distinguish colours. The reason is that human observers are best in distinguishing objects regardless of those variances by human vision system [68-70]. For instance, considering the example of a piece of paper where a white patch and a grey patch are drawn, the white still looks white and grey still looks grey although they look different indoors and outside of the room [14].

Hence, it is helpful to model the human vision system, colour control system and recognition systems under different viewing conditions [4, 5]. To interpret colour appearance model, the basic theory of colour appearance need to be understood first. Then, the standard colour appearance model of CIECAM97s is introduced.

3.1.1 Perceptual attributes of colour appearance

Colour appearance can be described not just by the physical methods in chapter 2. It can be described by perceptual attributes (hue, saturation, lightness etc.) too [2, 14]. Based on [14], it includes two types of perceptual attributes: one is basic perceptual attributes including hue, brightness and colourfulness. The other is relative perceptual attributes which includes lightness, chroma and saturation.

3.1.1.1 Basic perceptual attributes

Hue

Hue is the attribute of a visual sensation according to which an area appears to be similar to one, to proportions of two, of the perceived colours red, yellow, green and blue[14]. That is to say, one can mix paints of adjacent colours in this series and obtain a continuous variation from one colour to the other. Normally, hue can be described as a circle [16, 68]. One example of hue circle is described in the Figure 3-2 below. Usually hues can be measured by hue angle which ranges ranged from 0° to 360° in Figure 3-3 (a), and hue quadrature which ranges from 0, 100, 200, 300 and 400 in Figure 3-3 (b) corresponding to psychological hues of red, green, blue and back to red.

Figure 3-2: Example of Hue circle [68]

Hue angle
(a)

Hue quadrature
(b)

Figure 3-3: Expression of Hue [14]

Brightness

Brightness is another visual sensation which describes an area exhibits more or less light. Francisco [68] gives a concise idea that brightness is the quantity of light apparently coming from an object.

Colourfulness

Colourfulness is a measure of the intensity of the hue in a given colour. It is an attribute of human vision feeling according to the area that appears to exhibit more or less of its hue [14, 25].

3.1.1.2 Relative perceptual attributes

In order to explain these perceptual attributes, Hunt [14] gives an example of a white and a grey patch seen side by side on a piece of paper. If the patches are observed in bright sunlight they will look very bright. and if the paper is taken into the shade, or indoors, the patches will look less bright. But the white will still look white, and the grey will look grey. The colour attributes of grey are judged relative to the white colour under the same illuminant. This white is called reference white. This is such an important phenomenon that certain relative perceptual attributes of colours are given separate names, called lightness, chroma and saturation.

Lightness

Lightness is the brightness of an area judged relative to the brightness of a similarly illuminated reference white. The relationship can be written as a mathematical formula [68]:

$$Lightness = \frac{Brightness}{Brightness(White)} \tag{3-1}$$

The lightness scale runs from black (0) to white (100).

Chroma

Chroma is the colourfulness of an area judged as a proportion of the brightness of a similarly illuminated reference white. The mathematical relationship can be written below:

$$Chroma = \frac{Colourfulness}{Brightness(White)} \tag{3-2}$$

Saturation

Saturation is the colourfulness of an area judged in proportion to its brightness. It is a relative measure of colour purity. The mathematical formula is seen as:

$$Saturation = \frac{Colourfulness}{Brightness} \tag{3-3}$$

Three attributes above provide the relative perceptual attributes of colours. Chroma and lightness are defined relative to reference white. Whilst, saturation judgment does not require the concept of a similarly illuminated white; it is necessary to judge colourfulness relative to the brightness of the same area.

A colour can be expressed through last six terms (Hue, brightness, colourfulness, lightness, chroma and saturation). These are used to describe human visual perception when seeing a colour under different conditions or environments. Hence, to model how human eyes work, the above colour perceptual attributes must be taken into consideration. It has a significant usage for object recognition and industrial colour control under real environments so that many research scientists want to develop a human vision model which imitates the human eye's working theory to predict colours.

3.1.2 Colour appearance model of CIECAM97s

A human vision model used to predict the colour appearance under a wide range of viewing conditions is called colour appearance model [17]. Colour appearance models are based on colour vision theories. The theory of colour vision for colour appearance model is simply introduced below.

3.1.2.1 Brief theory of colour appearance model

Colour is the perceptual result of the human eye looking at an illuminated object. Colour perceptual depends on not only spectral distribution of the colour stimulus (object) and light but also structure of the eye and surroundings of the stimulus. Hence, a simple structure of human eyes introduced by Hunt [14] is introduced below and the relative stimulus area will also be discussed.

It is well known that the human eye perceives colour by the receptors of cones and monochromatic vision under low levels of illumination by the receptors of rods. Then colour is further processed cognitive and biologically by following procedures [14, 71]. When light is captured by receptors, the molecules of the photosensitive pigment are excited and a change in electrical potential is produced. This change is transmitted through nerve fibres to the brain, as a result the colour is perceived. Hunt [14] used the following Figure 3-4 to explain the basic working theory of cones and rods. In Figure 3-4, ρ, γ and β express three cones of human eyes. The rod and cone receptors are connected to neurons (nerve cells).

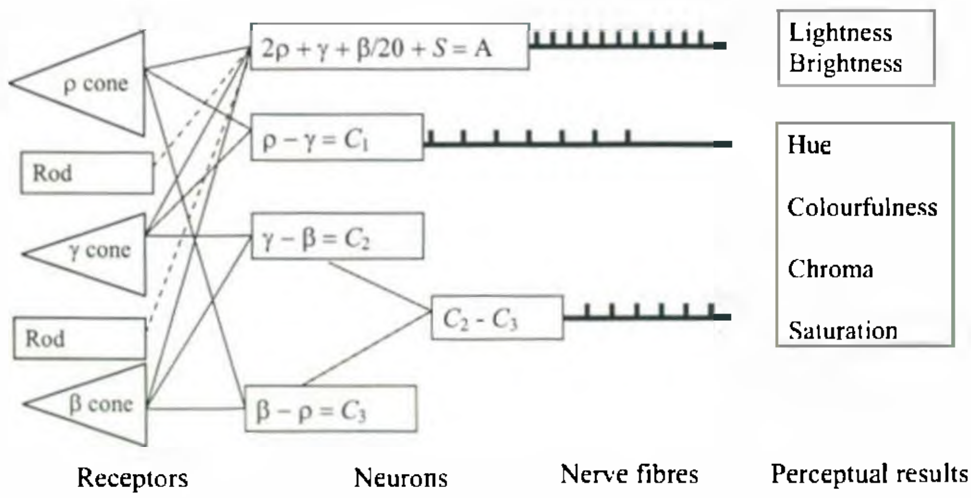
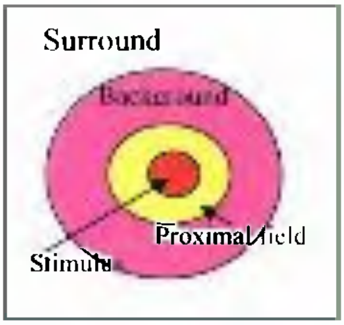


Figure 3-4: Hypothetical Diagrammatic representation of possible types of connections between retinal receptors and some nerve fibres [14]

Figure 3-4 above is the basic frame work for colour appearance model construction [2, 72]. It considers working theory of how human eyes perceive colour. Receptors of human eyes (cones and rods) receive signal of colour and send them to brain through neurons and nerve fibres. The result is perceptual attributes of colour.

Perceptual colour changes with the environment. In order to describe the effects of surround, colour scientists give five different visual fields definition in the model [4, 14, 73]. They are colour stimulus, proximal field, background, surround, and adapting field.

Francisco [68] utilizes the schematic diagram of the visual fields in Figure 3-5 based on the definition above.



Different Visual fields

Figure 3-5: A simple explanation of vision fields [68]

The colour stimulus area:

Typically a uniform patch of colour pigment.

The proximal field:

The immediate environment of the colour element considered, extending typically for about two degree from the edge of the colour element considered in all or most directions.

The background:

The environment of the colour element considered, extending typically for about 10 degree from the edge of the proximal field in all, or most directions.

The surround:

The field outside the background.

The adapting field:

The total environment of the colour element considered, including the proximal field, the background, and the surround, and extending to the limit of vision in all directions.

According to this scheme presented in Figure 3-5, the human vision system receives information from different elements of the visual field. According to the inputs described earlier, the system reaches an internal state corresponding to the perception of colour. Those are described through six observation variables: Hue, brightness, colourfulness, lightness, saturation and chroma. This regulation that allows us to predict the observation values from the input signals defines the colour appearance models of human vision system.

In 1997, the CIE technical Committee TC1-34 [4, 74] introduced a colour appearance model of CIECAM97s. This model can predict the accurate colour appearance under a wide range of viewing conditions by imitating human eye's working theory [74].

3.1.2.2 CIECAM97s colour appearance model

The colour appearance model of CIECAM97s tries to model how the human visual system perceives the colour of an object under different lighting conditions and with different backgrounds. A schematic of CIECAM97s is shown in Figure 3-6. The figure is modified from original figure in [68].

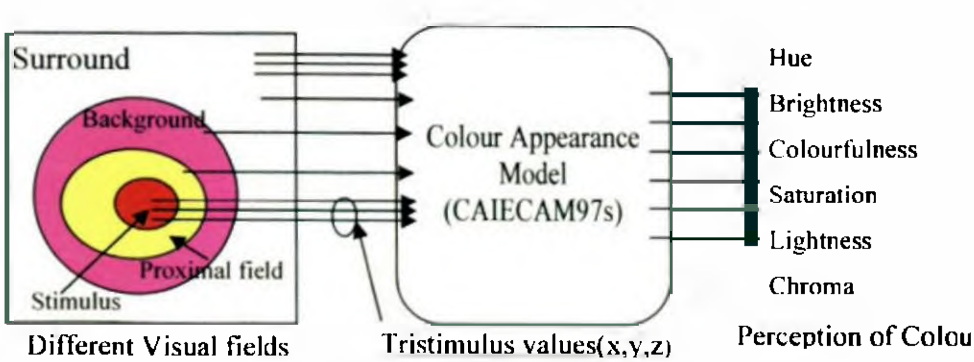


Figure 3-6: A schematic of colour appearance model

From the figure above, it can be seen that colour appearance model considers visual fields such as colour stimulus and its background, surround and proximal field and outputs colour perception of human.

3.1.2.2.1 Structure of CIECAM97s model

CIECAM97s has input data of vision fields expressed by tristimulus value and output the colour appearance correlation to 6 perceptual colour attributes, this is done through imitating human eyes working theory of chromatic adaptation [4]. A more detail schematic diagram is drawn in the Figure 3-7 to show the structure of CIECAM97s. The figure is modified from original figure in [17].

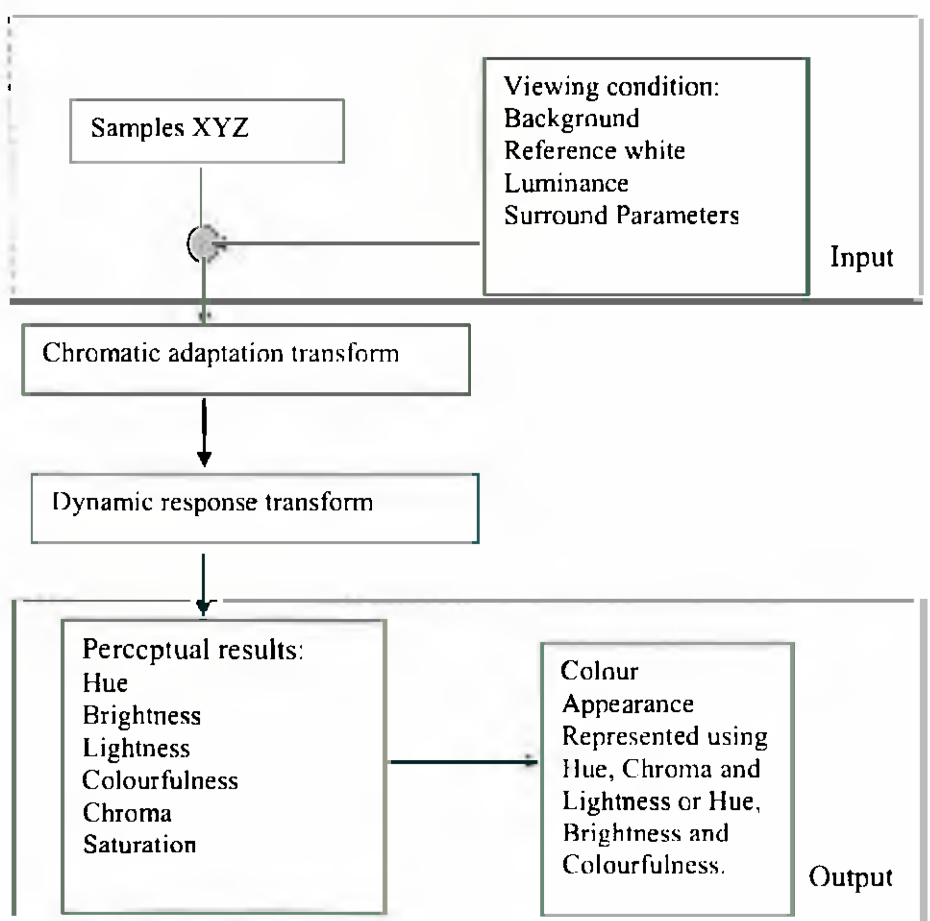


Figure 3-7: A structure of CIECAM97s colour appearance model

It can be seen from Figure 3-7 that CIECAM97s consider the vision field inputs and perceptual colour outputs, as attributes through a chromatic adaptation transform and a dynamic response function. Those outputs can be used to represent colour appearance [17, 66].

A chromatic adaptation transform is capable of predicting the corresponding colour in terms of a colorimetric specification (such as tristimulus values) from one set of illumination conditions to another. A pair of corresponding colours would look the same when viewed under two illuminants, for example illuminant A and D65.

The dynamic response functions are used to predict the extent of changes of responses of stimuli of different luminance factors across a wide range of luminance levels, i.e. from very dark scotopic to very light photopic vision.

The above two steps are the representation of simulating human eyes chromatic working theory. The detail mathematics calculation is interpreted in [4, 66]. Figure 3-7 gives the input of CIECAM97s. It includes colour sample which expressed by CIE chromaticity space of XYZ, and viewing conditions which includes background, reference white, surround parameters and luminance. It outputs six perceptual colour attributes. The next section will detail the input and output of CIECAM97s model.

3.1.2.2.2 Inputs and outputs of CIECAM97s model

Input of CIECAM97s

Input of CIECAM97s should have tristimulus value of colour, reference white and viewing conditions. The detail of inputs are the relative tristimulus values of the samples in the source condition, XYZ, the relative tristimulus values of source white in the source conditions, $X_W Y_W Z_W$, the adapting field luminance in cd/m^2 La, and the relative luminance of the source background in the source conditions, Y_b. Additionally, the constants c for the impact of surround, N_c for a chromatic induction factors, F_{LL} for a lightness contrast factor, and F for a factor for degree of adaptation. Nc, F_{LL}, F, and c consists of viewing conditions parameters [4].

Those inputs are listed below.

- *XYZ*: Relative tristimulus values of colour stimulus
- $X_W Y_W Z_W$: Relative tristimulus values of white
- La: Luminance of the adapting field (cd/m*m)
- Y_b: Relative luminance of the background
- Surround parameters: c, Nc, F_{LL}, F

The main feature of CIECAM97s model is to predict colour appearance under different viewing conditions. There are four defined surrounds: average, dim, dark and cut-sheet. The surround is categorical and is defined based on the relationship between the relative luminance of the surround and the luminance of the scene [4]. The surround can also be defined by comparing the surround luminance to the average luminance of the viewing field [66]. An average surround has a relative surround luminance of greater than 20% of the luminance of the scene white. A dark surround has a relative surround luminance that is 0% of the luminance of the scene white. A dim surround has a relative surround luminance between 0% and 20% of the luminance of the scene white. The cut sheet surround is a specific surround for viewing of cut sheet transparencies. According to different surrounding, the surround parameters have different values. The examples are taken from Table 3-1 [4, 66, 72].

Viewing condition	c	Nc	F_{LL}	F	Examples
Average Surround, samples subtending >4°	0.69	1.0	0.0	1.0	Viewing surface colours
Average Surround	0.69	1.0	1.0	1.0	
Dim Surround	0.59	1.1	1.0	0.9	Viewing television
Dark Surround	0.525	0.8	1.0	0.9	Viewing film projected in a dark room
Cut-sheet Transparencies (on a viewing box)	0.41	0.8	1.0	0.9	Viewing cut-sheet films in light boxes

Table 3-1: Surround parameters setting [4, 66, 72]

Output of CIECAM97s

Output of CIECAM97s should give perceptual attributes of colour, which are hue including hue composition (H) and hue angle (h), brightness (Q), lightness (J), saturation(S), chroma (C) and colourfulness (M).

CIECAM97s provide mathematical scales to correlate with perceptual appearance attributes. Lightness, chroma and hue correlates can be used to construct a colour space by considering J, C and h (H) as cylindrical coordinates. Alternatively, a brightness-colourfulness space could be constructed using CIECAM97s Q, M, and h as cylindrical coordinates [4]. Also one can be constructed using the normal means for cylindrical-to-rectangular coordinate transformations. Therefore, J, Ccos(h) and Csin(h) or Q, Mcos(h), and Msin(h) could be used as rectangular coordinates [75]. The Table 3-2 gives an example of those colour spaces based on CIECAM97s.

Colour space by CIECAM97s	Lightness-Chroma-Hue	Brightness-Colourfulness-Hue
Cylindrical Description	J, C, h (H)	Q, M, h (H)
Rectangular Description	J, Ccos(h) and Csin(h)	Q, Mcos(h), and Msin(h)

Table 3-2: Colour space by CIECAM97s [75]

CIECAM97s also includes an inverse step. If the colour perceptual values are known, tristimulus values of colour stimulus can be obtained according to different reference white and viewing conditions.

In summary, CIECAM97s affords an effective tool for predicting colour appearance in different viewing conditions. So far, it has been used in some colour production in laboratories, and different industries and colour management system [76]. The relative calculation formulas are listed in Appendix 1.

3.1.2.2.3 Application of CIECAM97s

CIECAM97s model is applied to use in cross-media image reproduction, colour rendering of light and colour difference. Janne [77] used this model to control the appearance of an image showed on a single monitor display under varying lighting conditions. Janne performed three experiments when reproducing colour images showing on the monitor. The reference image is initially displayed on a monitor under a reference lighting conditions and observed by different subjects. By changing different parameters of CIECAM97s, the images are reproduced and displayed in monitor under different lighting conditions. The subjects were asked which reproduced image can represent the reference image under such lighting conditions. The results prove that CIECAM97s can effectively predict the colour changes when lighting conditions changing.

Jin-Seo Kim [76] used this model in colour management system to reproduce colour in an monitor. In Jin-Seo Kim work, the CIECAM97s is used to reproduce colour image scanned by scanner on a monitor. Two viewing conditions were adopted. One is average conditions for forward CIECAM97s and dim condition for reverse CIECAM97s. The results show the reproduced image by CIECAM97s provides better matching to original images than other methods.

From the application described above, CIECAM97s generally use in the colour reproduction, colour rendering etc. The reason is that the output of this colour appearance model express how human look at the colour in different environment. It helps industries to reproduce colour in different environments. It gives us the clue that we can use it in image processing. From the definition and structure of CIECAM97s, the pair of corresponding colours looks the same in different environments and the lightness and chroma do not change a lot. We can use this model to segment images to overcome the shortage mentioned in section 2.4.

3.2 Human behaviour model

Human can recognize objects and patterns independently of changes in shifts in the object or changes in orientation and scale [78-80]. Gibson [81] hypothesis that the human visual system is strongly tied to the ability to recognize invariants. This is supported by psychophysics [82-84] and neurophysiology [85]. It has been known that when looking at an object, people only achieve high resolution for that portion of the object whose image falls directly on the fovea. Most objects we examine are much larger than this. To build a high-resolution impression of them, we must move our eyes so that all portions of their image fall on the fovea within a short time. These movements are carried in a series of jerky *saccades*, interspersed with stationary *fixations* [86]. By perception and analysing those fixations, an object is recognized.

3.2.1 Brief behavioural theory of human eyes movement

In 1967, Yarbus [87] demonstrated that human eyes move and successively fixate at the most informative parts of the image during visual perception and recognition. One example of Yarbus experiments is shown in Figure 3-8 [87]. The left one is the picture used to test human eyes movements. The right one is the tracking result of human eyes movements. Those points in right picture are the fixation points of human interesting. From experiments, Yarbus also pointed out that the eyes actively perform points' selection according to problem definition. And that information of point is processed under the control of visual attention. As a result, visual perception and recognition may be considered as behavioural processes [67].

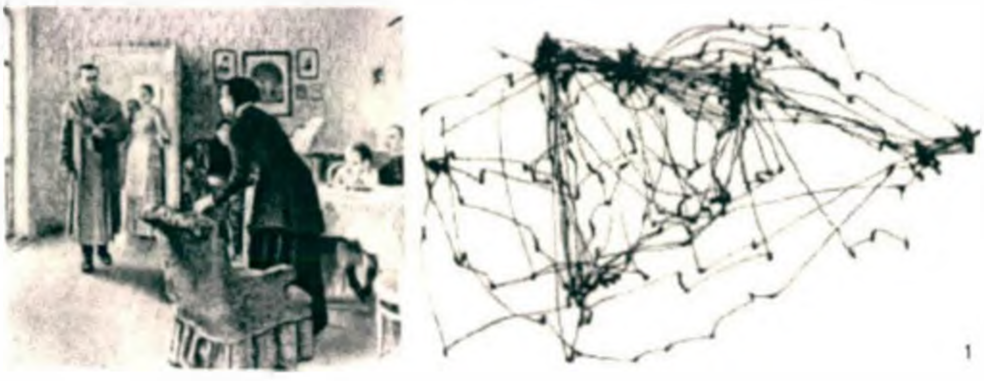


Figure 3-8: One example of eye movements [87]

After Yabus's work, the behavioural theory of eye's movements was thoroughly investigated. Theories behind the behavioural processes of human vision (eye movements) have been formed through studies of work about behavioural, visual perception and also recognition of psychological [87, 88] and of neuro-anatomical and psychological [89-92].

From the behavioural point of view [93], the ***behavioural paradigm*** can be described as two parts: internal representation and an external object understanding. An *internal representation* (model) of new circumstances is created in the brain during conscious observation and active examination. The active examination is aimed toward finding and memorizing of functional relationship between the applied actions and the resulting changes in sensory information. An *external object* becomes "known" and may be recognized when the system is able to subconsciously manipulate the object and to predict the object's reactions to the applied actions. Therefore, the internal object representation contains chains of alternating traces in "motor" and "sensory" memories. Each of these chains reflects an alternating sequence of elementary motor actions and sensory (proprioceptive and external) signals which are expected to arrive in response to each action. The brain uses these chains as "behavioural programs" in subconscious "behavioural recognition" when the object is (or is assumed) known. This "behavioural recognition" has two basic stages:

(i) conscious selection of the appropriate behavioural program (when the system accepts a hypothesis about the object),

(ii) subconscious execution of the program. Matching the expected (predicted) sensory signals to the actual sensory signals, arriving after each motor action, is an essential operation in the program execution.

The above behavioural paradigm was formulated and developed in the context of visual perception and recognition. The process of recognition was supposed to consist of an alternating sequence of eye movements (recalled from the motor memory and directed by attention) and verifications of the expected image fragments (recalled from the sensory memory). This work has been researched extensively [88, 94] using Yarbus' approach, which compares the individual scanpaths of human eye movements in two phases: during image memorizing, and during the subsequent recognition of the same image. They found these scanpaths to be topologically similar and suggested that each object is memorized and stored in memory as an alternating sequence of object features and eye movements required to reach the next feature. The results prompted the consideration of eye movement scanpaths as behavioural programs for recognition.

The neuro-anatomical and psychological data complementary to the above behavioural concept are presented by Ungerleider and Mishkin [89], Mishkin, Ungerleider and Macko [90], Van Essen [91], and Kosslyn *et al.* [92]. It was found that the higher levels of the visual system contain two major pathways for visual processing called "where" and "what" pathways. The "where" pathway leads dorsally to the parietal cortex and is involved in processing and representing spatial information (spatial locations and relationships). The "what" pathway leads ventrally to the inferior temporal cortex and deals with processing and representing object features.

The behavioural theory assumes that visual perception and subsequent recognition of an object are the result of behavioural processes. “Behavioural Processes” has two basic stages:

(1) conscious selection of the appropriate behavioural program (*a plan of how the eye will move*)
(2) subconscious execution of the chosen behavioural program (*actual eye movement*).

In program execution, the expected features (sensory signals) are matched with the actual features. Moreover, there are two major neural pathways that process the visual information, the “where” and “what”, which separately, lead dorsally to the parietal cortex and ventrally to the inferior temporal cortex. The “where” pathway performs and represents spatial information, while the “what” pathway deals with all information relating to semantic features.

Based on the theory above, a behaviour model of vision (BMV) was developed by *A.B.Kogan Research Institute for Neurocybernetics* in Russia [5].

3.2.2 Behavioural Model of Vision (BMV)

The Behavioural Model of Vision (BMV) has been shown to reliably recognize that it has ability to recognize complex images invariantly with respect to shift, rotation and scale. The principle theory of BMV model is described as following.

3.2.2.1 Model description

BMV consists of three levels, (a) low level, (b) intermediate-level processing and (c) high level subsystem. The schematic of this model is described in the Figure 3-9 [5].

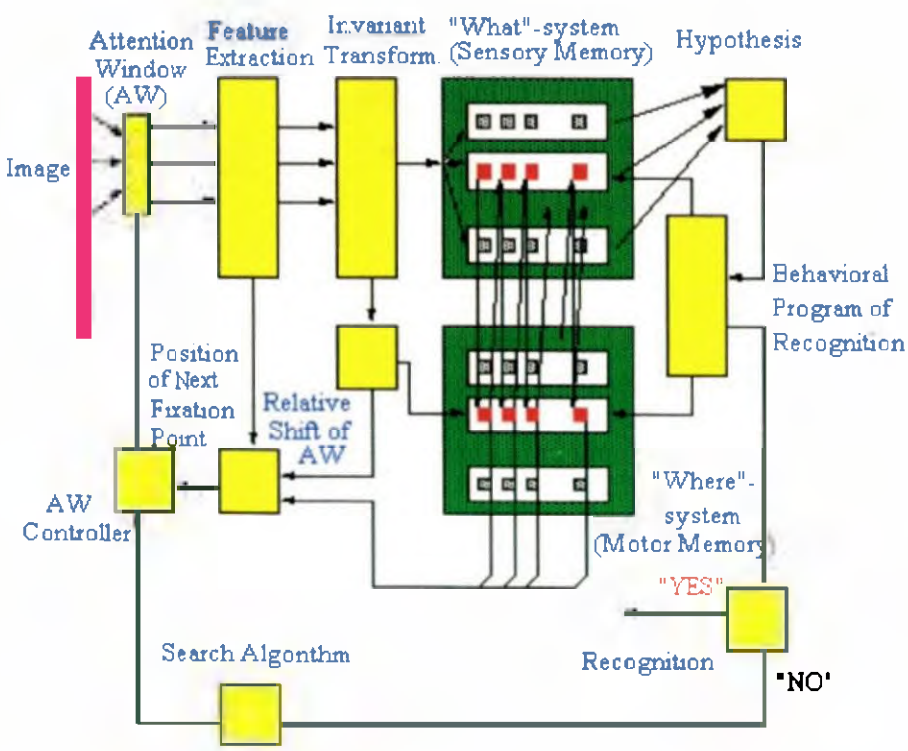


Figure 3-9: Schematic of Behavioural Model of Vision [5]

In the first level called low-level, primary features of image are detected. The Attention Window (AW) performs a primary transformation of the image into a 'retinal image' at one point. The retinal image is characterised by a set of edges, which is extracted by primary feature detection unit. This set of edges is described by one 'basic' edge in the AW centre and several 'context edges'. Figure 3-10 is the schematic of the attention window. It consists of the centre with 3 concentric circles divided by 16 lines with the discrete angle step of 22.5°. It forms 49 context points that is 3*16 plus one central point (49 = 3 x 16 + 1). The context points are located at the intersections of sixteen radiating lines and three concentric circles. At every context point, the edge feature is detected by 5*5 windows using Gaussian function. The size of AW is selected according to the complexity of images. The size of the AW determines the detail of the representation, the smaller the window the more representative it is of visual features. However, if the size of AW is too small, it increases the processing time.

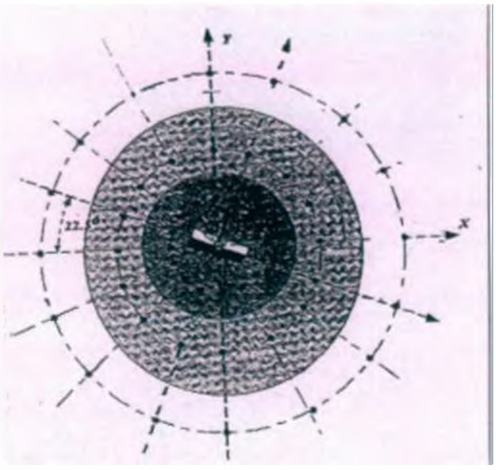


Figure 3-10: Schematic of the AW [5]

The next stage is the intermediate processing level called "Invariant Transformation". Within this level, primary features are described by using mathematic methods. It transforms the set of primary features into invariant second-order features using a co-ordinate system attached to the basic edge in the centre of the AW and oriented along the brightness gradient of the basic edge. The relative orientations and angular locations of the context edges, with respect to the basic edge, are considered as invariant second-order features.

The last one is the high level subsystem. It contains three modes:

1. In the *memorising* mode, the image is processed at sequentially selected points. At each point, the set of edges is extracted from the AW, transformed into the invariant second-order features and stored in the sensory memory ('what' structure). At the same time, the shift of the AW ('eye movement') is controlled by a special module. The relative of the shift from one point to the next one is stored in the motor memory ('where'-structure).

2. In the *search* mode, the image is scanned by the AW under the control of a search algorithm. At each fixation, the current retinal image from the AW is compared to all retinal images of all objects stored in the sensory memory. The comparison will last until an input retinal image similar to one of the stored retinal images at some fixation point is found. When such a retinal image is found, a hypothesis about the image is formed.

3. In the *recognition* mode, the behavioural program is executed by way of consecutive shifts of the AW and consecutive verification of the expected retinal images recalled from the sensory memory.

Simulation carried out shows that the features of an image can be expressed by using this model. The features describe an image with some fixation points and the corresponding mathematics expression, which includes the invariant second-order feature description and edge description etc. The application examples of recognition based on BMV model is stated below.

3.2.2.2 Application of BMV model

BMV model has been used in face recognition and simple scene objects recognition. In 1999, G.V.Golovan *et al.* [95] used this model to extract the face features and test recognition in a certain range of facial image transformations. In the work of G.V.Golovan *et al.* a face is tested via different view angle of face (left, frontal and right) and scale changing of image. The results prove that face recognition based on this model can reach a high recognition rate in a certain range of facial image transformations. Besides facial image recognition, this group also test the simple scene objects recognition (box, tripod etc.) in a range of shift, rotation and scale transformations [67]. The results showed that it can also detect the simple objects from the scene.

Although BMV model was tested that it can recognize facial images and simple scene objects, traffic signs are the special objects with different contents. It brings a challenge that BMV model can be utilized for traffic sign recognition. The relative calculation formulas are listed in Appendix 1

3.3 Hypothesis of traffic sign recognition based on two human vision models

As described in chapter 2, current traffic sign recognition has some limitations when meeting environments and viewing angle changing. According to the review in this chapter, CIECAM97s provides a method that peoples how to look at the objects under different viewing conditions and BMV model provides a method that peoples how to recognize an object invariance of scale, shift, rotating changes. Therefore, the approach based on these two models is investigated in this research.

When weather changes, colour appearance of traffic signs changes. Figure 2-13 gives an example of colour appearance of blue and red traffic signs changes when weather changing from sunny day to rainy day. However, people can still recognize colour blue and red regardless of weather changes. CIECAM97s can predict this change by simulating human eyes working theory. It considers the environment changing and produces 6 colour perceptual attributes. According to the description above, the environment changing is expressed by the parameters and reference white. Based on different reference white and parameters, the output of basic colour perceptual attributes of CIECAM97s changes. However, the relative colour perceptual attributes keep same by comparing to reference white. This is the reason why humans perceive blue as blue, red as red even under weather changes [14]. Equations 3-1 to 3-3 also shows this property. A colour can be represented by three colour perceptual properties: hue, chroma and lightness. Those give us a clue that these three properties can be used to create an accurate range of colour properties. By comparing those ranges, the traffic signs can be segmented from the background under different viewing conditions.

When the car is moving, the viewing angle is changing. The size of traffic signs change from small to big. Also the image distorts (the perceptual angle changes). Sometimes, the signs have been rotated manually by accident. Those situations can cause the recognition difficulty. Table 2-4 gives some examples in real environment. Based on the review of the BMV model, it could solve the problems met above. By simulating how humans recognize objects, the BMV model tries to recognize objects by detecting those informative points and path of eye movements. That is to say, the objects are recognized whatever the objects distortion, scaling or shift, if the information (informative points and path of eye movements) are matched to database. Therefore, this model is used in the research to recognize traffic signs.

In this chapter, two human vision models: colour appearance model of CIECAM97s and behavioural model of vision (BMV) are interpreted. The CIECAM97s model takes viewing conditions into account and predicts colours as accurately as an average observer. The BMV model recognises complex objects accurately. It is invariant to the shift, rotation, or scaling and a degree of distortion. These two human vision models have the properties required for the development of traffic sign recognition system. Therefore, in order to recognize traffic signs quickly and accurately under different viewing conditions and in real environments, a new approach based on these two models is put forward. The CIECAM97s model will be used to segment the region of interest (ROI) from the images and the BMV model will be used to recognize traffic signs after segmentation.

Chapter 4
Methodology

In this chapter, a new approach to detect the traffic signs under different viewing conditions, in real environment and based on human visual perception is presented. The approach is based on two human vision models. One is colour appearance model of CIECAM97s and the other is behavioural model of vision of BMV. The overall procedure is schemed in the following Figure 4-1.

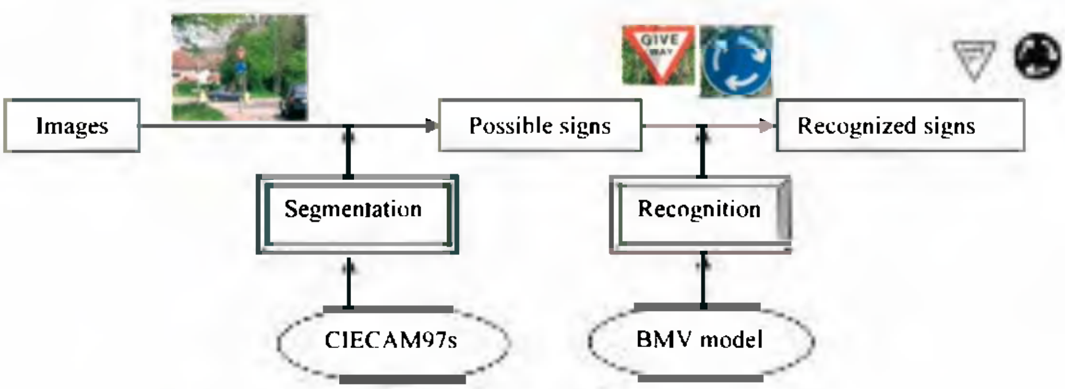


Figure 4-1: The procedure of traffic sign recognition based on human vision models

The figure above shows there are 2 major stages in this approach. Firstly, the CIECAM97s is used to segment images to get the regions of interest (ROI). Secondly, the BMV model is then applied to recognize those ROI to identify traffic signs

To implement this approach, the following procedures are carried out (1-4).

1. Image collection and analysis
2. Classifying images into different weather conditions
3. Segmentation of images taken under different weather conditions based on CIECAM97s colour appearance model

- CIECAM97s parameter setting under different weather conditions in image processing
- Segmentation according to thresholds of colour attributes based on CIECAM97s under different weather groups

4. New application of BMV model to the field of traffic sign recognition.

A schema of these procedures is also shown below in Figure 4-2.

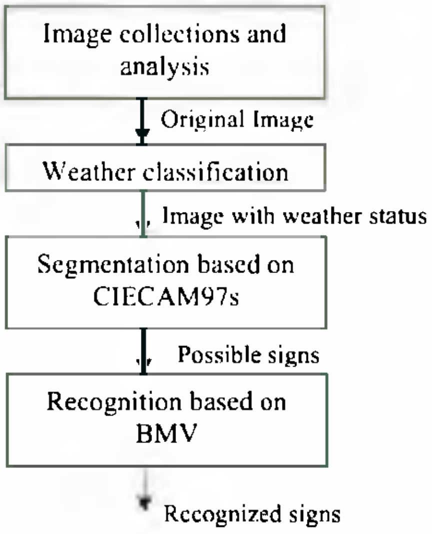


Figure 4-2: Schema of traffic sign recognition based on CIECAM97s and BMV

The processes above are described in section 4.2 to section 4.5. They have been designed and developed with the aim of obtaining a high recognition rate and a fast identification. All of the equipments that have been used throughout this project are firstly introduced in section 4.1.

4.1 Equipment and software

4.1.1 Camera

A high quality Olympus Digital Camera C-3030 Zoom is used to capture pictures in the real viewing conditions. The specifications of this camera are:

- High resolution: It has 3.34-mega pixel Charged Couple Device (CCD).
- Flexible manual control: The camera has complete manual control which includes shutter, aperture or manual exposure modes. These help to obtain the required effect.
- The pictures can be saved as TIF format pictures, which are not compressed and are consequently a lossless image format.
- The pictures can be easily transferred from the camera to the computer using cable through computer Universal Serial Bus (USB) port.
- The camera has the LCD screen. It helps us to see the picture captured in time and to analysis the effect of capture.

4.1.2 Macbeth colour checker board

Macbeth colour checker board shown in Figure 4-3 is the industry standard colour checking chart for cinematographers and photographers alike. It provides the needed standard with which to compare, measure and analyze differences in colour reproduction in various processes.

This colour checker is a checkerboard array of 24 scientifically prepared colour squares in a wide range of colours. Many of these squares represent natural objects of special interest, such as human skin, foliage and blue sky. These colour patches are the same colour as their counterparts. Because of this unique feature, the colour patches will match the colours of natural objects under any illumination and with any colour reproduction process.

Figure 4-3: Macbeth Colour Checker board

In our research. this colour patches is used as the standard colour stimulus. They will be measured by colour measurement instrument such as luminance meters.

4.1.3 Colour measurement instrument

Colour measurement instruments are designed to measure colours in terms of reflectance. radiance. and the CIE colorimetric values such as tristimulus values (X, Y, Z) interpreted in Chapter 2. In our study. a Minolta Luminance meter LS-100 is used. Figure 4-4, shown below. details this instrument.

Figure 4-4: LS-100

The luminance meter LS-100 has advantages on measuring not only illuminant but also surface colours. The result of measuring is shown directly by the chromaticity value. It has the following features:

- It provides a measurement with an accuracy of 0.001 to 0.999 cd/m^2 depending on distance of measurement.
- It provides a bright field of view and the centre spot indicates exactly what is being measured.
- It has an easy-to-read viewing system called through-the-lens (TTL). A 4-digit LCD panel on the side of the meter shows measured values plus operation and error indications; in addition, an LCD inside the viewfinder allows readings to be taken while viewing the object. Also this meter is easy to calibrate and to handle.

Therefore, in our project, LS-100 is used to get the colour tristimulus value.

4.1.4 Imaging processing software

The image processing software for this project is Matlab 6.5. Matlab software is a high-performance language for technical computing and is provided by Mathworks Company (www.mathworks.com). The features of the software are as follows.

- It has a predefined image tool box which supports many fundamental image processing operations like edge extraction, transforms etc.
- It has a toolbox which uses to analysis the relationship among data and can produce the suitable curve to describe the relationship.
- It has an easy-to-use environment where problems and solutions are expressed in familiar mathematical notation [96].

For these reasons Matlab is widely used, and it is considered an excellent piece of software for technical computing. It enables fast and efficient data analysis and image processing.

This study considers process under the real environment. Therefore, picture collection and analysis is the first stage that is carried out.

4.2 Image collection and analysis

In this study, still images are considered. The images studied in this project are taken, from real environments in the UK so as to reflect future application of the system, by using a high quality Olympus Digital Camera C-3030 Zoom.

4.2.1 Image collection

The image collection needs to reflect different viewing conditions and also the variation in traffic sign size caused by differences in distance between traffic signs and driver (the position taking pictures). The viewing conditions contain two factors.

1. Weather conditions. The appearance of colour is different in different weather conditions. For example, the red in a sunny day appears more colourful than in a cloudy day. In order to recognise the traffic signs in different weather conditions, photos are needed to be taken in different weather. On the other hand, the appearance of colour looks different at different times in a single day. Here, sunny, cloudy and rainy weather conditions are considered.

2. Viewing angles. In order to identify the accurate traffic signs for cars from the opposite direction, the states of traffic signs need to be considered. In some junctions, there are complex traffic signs positions for the cars. On the other hand, whether the road is descending, ascending or flat need to be considered. Also, there are more than two signs at some junctions. Thus, we take photos of traffic signs positions from different viewing angles at junctions.

As mentioned, a parameter of image collection is distance between traffic signs and the viewing positions for taking images. The distance determines the size of traffic signs inside the images and is relevant to the recognition speed. According to The Highway Code [9] the stopping distance is more than 10 meters under 30MPH (miles per hour), which means the shortest distance should be 10 meters. On the other hand, if the distance of taking pictures is far from traffic signs, the images of signs will be very small to process. It takes about 4 seconds from the traffic signs to the site of taking pictures if the distance is 50 meters under 30MPH. It is long enough for a driver to cognitively process. We select 10, 20, 30, 40, and 50 meters as the distance for taking images.

Totally, there are 145 pictures being taken under different weather conditions. 52 pictures are taken in sunny day, 60 rainy day and 33 cloudy day images. These images are taken from different positions and distance to traffic signs. Every picture includes traffic signs and should represent the weather condition and viewing conditions at that time. To avoid the individual judgement, two people took the pictures together. After image collection, the images must then be analyzed manually.

4.2.2 Image analysis manually

The goal of image analysis in this project is to asses the quality of images. The quality of an image is whether the image can represent the real viewing conditions or not. Not all images can reflect the real viewing conditions because the quality of images is affected by shaking, individual judgement etc. Hence, image analysis is necessary.

There are two main factors which affected the quality of captured pictures. One is the model setting of camera and the other is aperture and exposure setting.

According to the property of camera, there are two main models of setting: auto-setting and manual control setting. If the auto-setting mode of the camera is used, the camera will adjust the aperture and exposure itself and the pictures will not reflect some weather situations. See examples below Figure 4-5 (a) and (b) which were taken in rainy day. The Figure 4-5 (a) was taken using auto-setting mode control and Figure 4-5 (b) was taken using manual control. After taking pictures, two people decide which picture will be taken by comparing images showing on the LCD screen of camera to the situation at that time. If the picture did not represent the situation, it will be marked and will be rejected. For example, the picture shown in Figure 4-5 (a) is considered brighter than the weather status at that time by two people. However, the picture shown in Figure 4-5 (b) is considered same as the weather status at that time. Therefore, we only use manual model to take pictures.

(a) The image taken in the rainy day using auto mode control

(b) The image taken in the rainy day using manual mode control

Figure 4-5: Two images taken at same position by different setting models of camera

Sometimes because of unsuitable setting like a small aperture or improper exposure time using manual control. the pictures cannot reflect the viewing conditions either. Analysing the following pictures Figure 4-6 (a) and (b). we can find the differences using different setting. Those two images Figure 4-6 (a) and (b) were taken with different aperture at the same place on a cloudy day. The first one was taken using the suitable aperture and the second one was taken by using a small aperture. Considering the images below. the bottom image seems darker than that in real environment because of a small aperture. Therefore Figure 4-6 (b) cannot be used as cloudy day.

(a) The picture taken in cloudy day with suitable aperture

(b) The picture taken in cloudy day with small aperture

Figure 4-6: Two pictures taken at same position by different aperture setting of camera

After collection and analysis, those pictures which have been marked not representing situation are rejected. Finally, 128 pictures containing traffic signs are obtained for further processing. There are 48 pictures of sunny day, 53 pictures of rainy day and 27 pictures of cloudy day respectively.

4.3 Images classification by weather conditions

Colour appearance model, CIECAM97s considers different viewing conditions such as weather changes or luminance changes and predicts colour appearance attributes such as lightness, hue and saturation *etc.* under different viewing conditions [17]. The inputs of this model require colour stimulus, reference white and viewing surround conditions parameters. As described in chapter 3, relative perceptual colour is judged by the reference white under such viewing conditions. The reference white changes with the viewing condition, so people can still perceive colour by judging with reference white that red still is red, grey still looks grey. This means that the reference white is an important input. Under different viewing conditions such as different weather conditions, the reference white is changed [2, 14, 97]. In other words, the weather status of an image is as important as the input of colour in the colour appearance model of CIECAM97s. It is relevant to the reference white, one of input parameter for CIECAM97s. To classify those collection images into different weather conditions is the next step taken.

4.3.1 Weather feature description study

To classify images into different weather groups, manual classification is applied. In this study, there are three weather conditions being considered. They are sunny, cloudy and rainy day. Dominant colour features, such as colourfulness, brightness, of each group are then extracted subjectively by 10 subjects. The 10 participants are all research students from the same department. The method is as follows:

Images from each group are displayed on a computer monitor (21 inch) by software *of Paintshop 7.0*. For each image, the participant was told that this picture was taken under which weather condition. Then the colour terms that the participants used to describe the weather conditions of image are recorded. The information which is recorded will be used as the evidence for classifying weather by computer.

4.3.2 Weather identification by computer

According to study, the weather status of an image can also be recognised through an algorithm based on local features of an image. A scheme in Figure 4-7 below illustrates the procedures of recognizing weather status from an image.

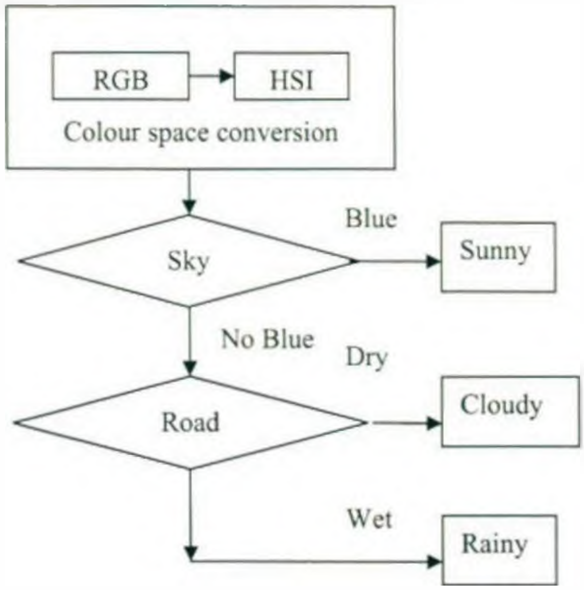


Figure 4-7: Procedure of identification of weather status from an image

Figure 4-7 above shows that the three steps that are adopted. Firstly, the whole image is converted into *HSI* colour space from *RGB* colour space. Secondly, according to the sky colour, a sunny day is recognized. If there is no blue sky, then the road texture is used to distinguish cloudy day and rainy day based on whether the road is wet or dry.

4.3.2.1 Colour space transformation

An image captured by digital camera Olympus C-3030 is stored as *RGB* colour space. Then the *RGB* space is transformed into *HSI* space by the following formula:

$$H = \arccos\left[\frac{[(R-G)+(R-B)]/2}{\sqrt{(R-G)^2+(R-B)(G-B)}}\right],$$
$$\text{where } H = 360^\circ - H \text{ if } (B/I) > (G/I) \qquad \textbf{(4-1)}$$
$$S = 1 - \frac{3\min(R,G,B)}{R+G+B}$$
$$I = (R+G+B)/3$$

Where *H* is the hue, *S* is the saturation and *I* is the intensity. To distinguish the weather status of an image these three perceptual colour features are used. According to the study, blue sky is the feature of sunny day. Hence *Hue* and *Saturation* of blue sky are used to identify sunny day. Based on the participants' result, the feature of road is expressed by texture of road. They are used to describe wet or dry. Therefore, the intensity of the road is used to detect wet or dry properties of road so that the rainy day can be identified. That is, how to use the *H, S, I* to distinguish blue sky and wet road is the solution for weather classification.

4.3.2.2 Sunny day identified by sky colour

In order to identify a sunny day or cloudy day quickly, a simple colour threshold method based on *HSI* colour space is used. Figure 4-8 below shows the procedures of sunny day image identification.

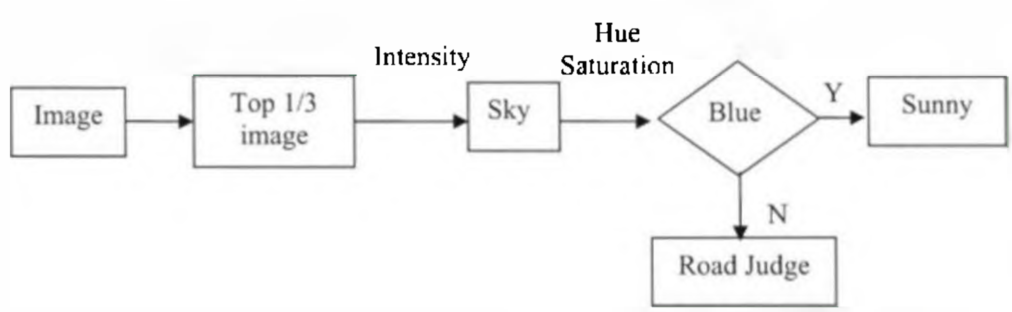


Figure 4-8: The procedures of sunny day classification

According to the weather feature description study, sky is the main feature which used to identify sunny day from an image by participants. Meanwhile, the sky parts are in the top 1/3 of an image. Therefore, to classify the sunny day, the procedure is finished by following the procedure which is shown in Figure 4-8. Firstly, the top 1/3 of an image is cropped and this part is converted to *HSI* colour space. The intensity of *HSI* of this 1/3 cropped image is used to segment the sky from the background by thresholding. Then, hue and saturation are used to judge if the sky is blue or not. If the sky is blue, then this image was taken in sunny day. Otherwise, it will be considered as rainy day or cloudy day. and will be judged by road features to detect rainy or cloudy day.

Examples of sky blocks of sunny day, rainy day and cloudy day cropped from images are listed in the following figures of Figure 4-9.

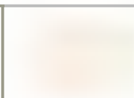
		
Two example blocks of sky in sunny day	Two example blocks of sky in cloudy day	Two example blocks of sky in rainy day

Figure 4-9: Examples of sky blocks in sunny day and cloudy day

The examples of sky in Figure 4-9 show that sunny day has more colourful blue sky than cloudy day and rainy day. Some cloudy days. however posses blue sky with clouds. Nevertheless. a cloudy day or rainy day has a less colourful blue sky or no blue sky features. Hence. this colour feature can be used to classify a sunny day or a cloudy/rainy day. In *HSI* colour spaces, hue represents the colour tone and saturation means extent of tone. Therefore. hue and saturation of sky are used to represent the colour feature of sky. 20 blocks of blue sky are selected from 10 samples of sunny day. 20 blocks of sky are selected from 10 pictures of cloudy day. Also. 20 blocks of sky are selected from 10 pictures of rainy day. Those pictures are calculated by using formula (4-1). The hue and saturation of those blue sky samples are obtained to calculate the thresholds for detecting sunny day or cloudy/rainy day. If the value of hue and saturation are bigger than thresholds. the weather of an image is classified into sunny day. Otherwise, it will be considered as cloudy/rainy day. Hue and saturation of each block and their thresholds are described in Chapter 5.

4.3.2.3 Rainy day and cloudy day identified by texture feature of road — Fast Fourier Transform (FFT)

To identify the rainy day and cloudy day, the FFT is used to describe the features of road and the average magnitude (AM) value of FFT is used to identify cloudy and rainy day. The procedures are listed in the following Figure 4-10.

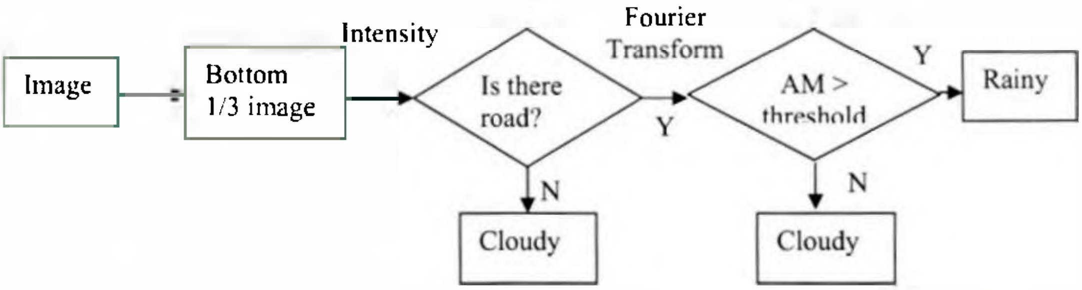


Figure 4-10: The procedures of rainy day and cloudy day classification

According to the weather feature description study, road is the main feature which used to identify rainy day from an image by participants. Meanwhile, the road parts are in the bottom 1/3 of an image. Therefore, to classify the rainy day, the procedure is finished by the following procedure which is shown in Figure 4-10. Firstly, the bottom 1/3 of an image is cropped and is transferred to *HSI* colour space. The road is then segmented from cropped bottom 1/3 of image via intensity of *HSI*. If there is no road in an image, this image is classified into cloudy day. If the road is segmented from the image, the road is further described by FFT. Finally, the average of Magnitude (AM) of FFT is used to judge cloudy day and rainy day. If the value of AM is bigger than threshold, the weather of this image is classified into rainy day. Otherwise, it will be identified as cloudy day.

Fast Fourier Transform (FFT) is used to describe as texture feature of road to identify rainy day and cloudy day. Usually, texture refers to visual properties like roughness, granulation and regularity [98]. In an image of a rainy day, the road looks more rough and dark. However, in cloudy day the road looks more smooth and bright. There are two examples of road representing rainy day and cloudy day respectively in the following Figure 4-11.

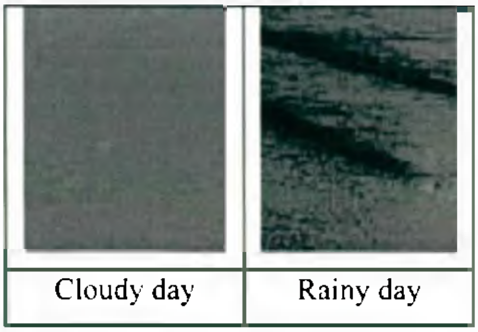


Figure 4-11: Examples of road blocks under cloudy and rainy day

The discrete *Fourier* transform can be defined as following formula [99]:

$$F(u,v)=\frac{1}{MN}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} f(m,n)\exp\left[-2\pi i\left(\frac{mu}{M}+\frac{nv}{N}\right)\right] \tag{4-2}$$

$$u = 0,1,\ldots,M-1 \qquad v = 0,1,\ldots,N-1$$

where $f(m, n)$ is the image, here it represents the intensity I of *HSI* colour space, n, m are the pixel co-ordinates, N, M is the image size, and u, v are frequency components [99].-

To measure the feature based on the power spectrum of *Fourier* transform, average magnitude (AM) is used to identify road feature in cloudy day and rainy day. It is defined as the formula [100] below:

$$AM = \sum_{j,k}|F(j,k)|/(N) \tag{4-3}$$

where $|F(j,k)|$ is the amplitudes of the spectrum and N is the number of frequency components, here is the image size. 15 blocks of road are selected from 10 pictures of cloudy day and 12 blocks of road are selected from 10 pictures of rainy day to calculate thresholds. AM value of each block and its thresholds will be described in Chapter 5.

4.4 Image segmentation based on CIECAM97s

After weather classification, and before the recognition of traffic signs can take place, potential areas of interest are segmented from an image. Colour appearance model, CIECAM97s, is used for this stage of segmentation. Firstly, based on CIECAM97s, the ranges of colour used in signs, mainly red and blue, will be decided as thresholds of segmentation. Secondly, those potential areas are segmented from an image by comparing pixel value of colour to the thresholds.

4.4.1 Obtaining colour ranges based on CIECAM97s

Threshold is an important method in image processing. In this study, the threshold is found by calculating the ranges of perceptual colour based on CIECAM97s. The Figure 4-12 below shows the procedure.

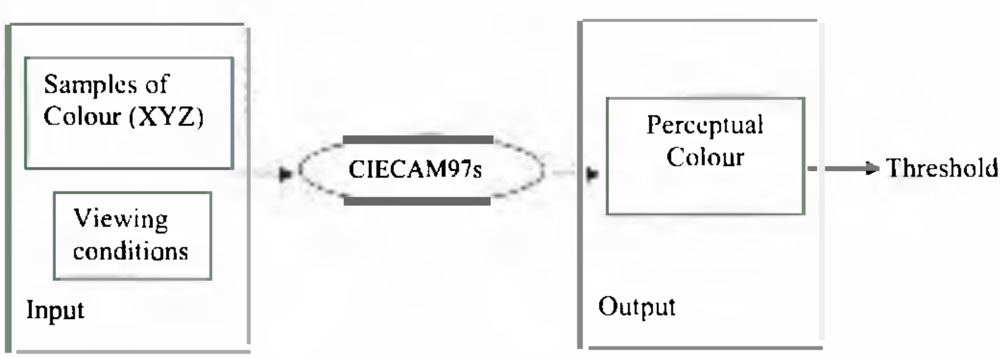


Figure 4-12: The procedure of thresholds obtaining

Figure 4-12 above shows the procedures of thresholds obtaining. The sample colours of traffic signs and viewing conditions are input to CIECAM97s model. After calculation, this model outputs the perceptual colours of samples. The thresholds are obtained from those output perceptual colours.

4.4.1.1 Input of CIECAM97s

Input of CIECAM97s includes sample colours expressed by tristimulus value of *XYZ*, and viewing conditions containing reference white, surround parameters and background value.

4.4.1.1.1 *XYZ* tristimulus value of colours

Tristimulus values of colour *XYZ*, defined by CIE, are an input of CIECAM97s. However, digital images taken by a digital camera are expressed using *RGB* colour space. Hence, it is necessary to convert *RGB* colour space to *XYZ* tristimulus value. Therefore, a transform between *RGB* colour spaces of digital camera output and *XYZ* colour spaces of samples need to be created. This step has been called as colour camera characterisation [101, 102].

Camera spectral sensitivity can be identical to, or a linear transform of, the human sensitivity of the CIE colour-matching functions [101]. That is to say, camera responses *RGB* can be mapped to CIE tristimulus values *XYZ* using a simple matrix equation:

$$T = M \cdot R \qquad \textbf{(4-4)}$$

Where T is a 3*1 column vector of tristimulus values, here is $[X, Y, Z]^T$, R is a 3*1 column vector of camera responses $[R, G, B]^T$, and M is a 3*3 transfer matrix that defines a linear transform.

The matrix M is characterization value of one camera for transforming response values to tristimulus values. For this characterization procedure to obtain matrix M two main processes are performed. Firstly, the camera system is used to ascertain sensor values for targets with known colour characteristics (CIE values). This includes taking photos to get *RGB* value and measure *XYZ* value from the standard colour chart. Secondly, these sensor values are transformed to match the target CIE co-ordinates by mathematics calculation [101].

Data collection

In the first step, 2 sets of data should be collected: *XYZ* values of standard colours and *RGB* values of those standard colours. In this study, standard colour checker is *Macbeth 24 colour checker chart*. The colour chart is measured by luminance meter LS100 to get *XYZ* value. The *RGB* data of camera sensor is obtained using Matlab after transferring images from digital camera to PC. A digital camera Olympus C3030 is used in this study. The Figure 4-13 below shows this procedure.

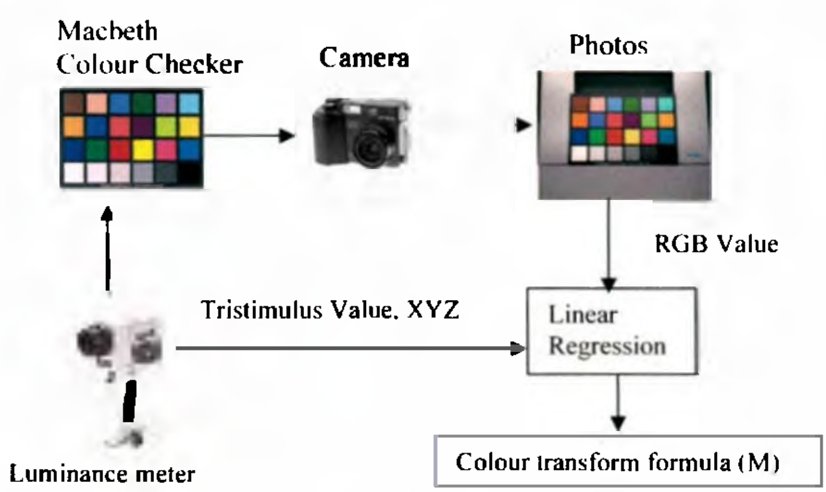


Figure 4-13: Method of obtaining colour transform matrix between *RGB* and *XYZ*

Figure 4-13 above shows that the colour chart including 24 colour pigments are taken by camera. *RGB* values of each 24 colour pigments can be obtained from images. At the same time, *XYZ* values of each 24 colour pigments are measured by luminance meter. In order to reduce the errors four pictures are taken, and colour chart is measured 4 times as well. Finally, in total 96 (24*4) data of *RGB* values and their corresponding 96 data of *XYZ* values of standard colours are determined.

After these two sets of data are obtained, the next step is to use them to calculate the matrix M expressing characterization of camera c3030 by linear regression.

Linear Regression--- Least squares estimation

Least squares estimation is used to find the prediction equation. The formula T=M*R above can be modified to

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{4-5}$$

The coefficients of matrix M, a_{11} to a_{33}, are obtained by calculating the estimated value of them, which can minimize the sum of the squared residuals: $\sum_{i=1}^{n}\left(X_i - \hat{X}_i\right)^2$,

here X_i is an observed response of measuring data, and $\hat{X}_i$ is a point on the prediction equation, $X - \hat{X}$ is the residual. n is the number of samples, here is 96.

The following procedures [103] are used to get M.

Step 1: Take the derivatives of $\sum_{i=1}^{n}\left(X_i - \hat{X}_i\right)^2$ with respect to the parameters a_{ij}, such as a_{1j} (j=1,2,3), the formula can be obtained below:

$$\begin{aligned} a_{11}\sum R_i^2 + a_{12}\sum R_i G_i + a_{13}\sum R_i B_i &= \sum X_i R_i \\ a_{11}\sum R_i G_i + a_{12}\sum G_i^2 + a_{13}\sum G_i B_i &= \sum X_i G_i \\ a_{11}\sum R_i B_i + a_{12}\sum G_i B_i + a_{13}\sum B_i^2 &= \sum X_i B_i \end{aligned} \tag{4-6}$$

Step 2: By calculating the above formula, a_{11} to a_{13} can be obtained. It is same to get a_{21} to a_{23} and a_{31} to a_{33}.

Finally, matrix M repressing camera characterisces is obtained. The *RGB* images can be transformed to *XYZ* colour space by the following formula.

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.2169 & 0.1068 & 0.048 \\ 0.1671 & 0.2068 & 0.0183 \\ 0.1319 & -0.0249 & 0.3209 \end{bmatrix} \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{4-7}$$

Therefore, in our project, tristimulus values of colour *XYZ* can be obtained using colour transform formula above.

The inputs of CIECAM97s include not only tristimulus of colour *XYZ*, but also viewing conditions. After tristilmulus value of colour, *XYZ* is obtained, the viewing conditions such as reference white, surround viewing conditions need to be set up/obtained.

4.4.1.1.2 Viewing conditions parameters selection

As described in chapter 2, viewing conditions include tristimulus values of reference white $X_W Y_W Z_W$, and surround viewing conditions parameters (c, FLL, F, Nc), also background Y_b.

Reference white ——$X_W Y_W Z_W$

Reference white (R_W) of CIECAM97s is relevant to the tristimulus values of source white in the source conditions [72, 73]. During the day time, the sun is the main source condition and weather plays an important role in the determination of colour temperature of the sun [3]. Hence, source white has different appearance under different weather conditions i.e. rainy day, cloudy day and sunny.

Normally, when using CIECAM97s, the reference white should be measured firstly. This can be done by measuring a white board in viewing conditions [73]. According to [6, 7] white colour usually forms one component of a traffic sign. By measuring and comparing values of white board and white part of traffic signs, white of signs can be used as R_W. An experiment of measuring reference white from white board and white colour of traffic signs at the same viewing condition described as below Figure 4-14 will explain this point of view.

White board measurement

White measurement from real signs

Figure 4-14: Reference white measurement from white board and white colour of traffic signs at the same view condition

Figure 4-14 above shows the method employed to measure reference white from white board and traffic sign at the same position under same weather condition. There are 5 separate white sections, on both traffic sign and board, which are measured. Usually, they are middle, top, bottom, right and left. Every section is itself composing of 3 patches that are measured by light meter-LS100. The measuring value is expressed by chromaticity value (x, y). Finally, the mean of 15 chromaticity value of white from signs is (0.3143, 0.316). The mean of 15 chromaticity value of white board is (0.313, 0.311).

These two mean values of white obtained from white board and traffic signs, are drawn in the xy chromaticity diagram of Figure 4-15.

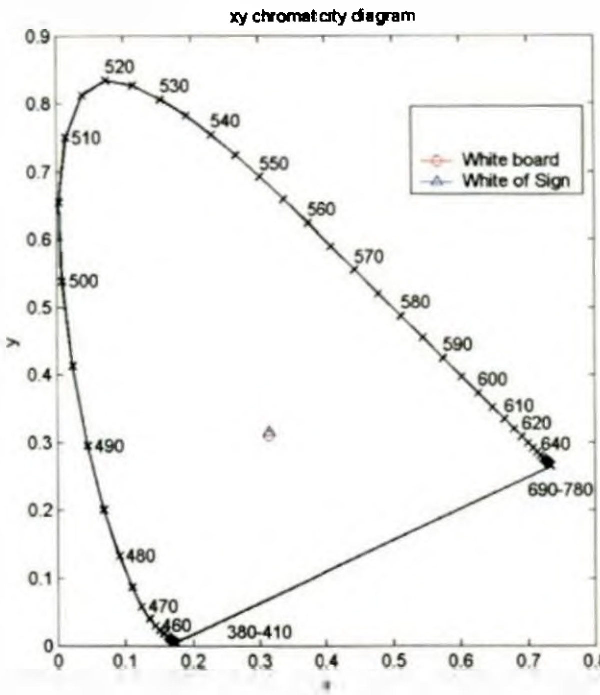


Figure 4-15: White point of white hoard and signs drawn in *xy* chromaticity diagram

Figure 4-15 above shows that these two white points are nearly at the same position. This example proves that white of traffic sign can be used as the reference white. In order to simplify calculation, a set of *RGB* values are selected from white part of traffic signs in pictures under different weather conditions respectively. Then, this set of *RGB* values are transferred to *XYZ* value. The mean value of *XYZ* value is used as *XYZ* value of reference white. For example, in sunny day, 10 pixels white are selected from white part of traffic signs in one picture. 10 pictures are selected from group of sunny day. Totally, 100 pixels of *RGB* values of white are used to determine reference white in sunny day. Those *RGB* values of white should be transferred to *XYZ* value by using formula (4-7) as tristimulus values of reference white ($X_W Y_W Z_W$). Finally, the mean value of this set is used as reference white in sunny day.

Surround parameters——(c, FLL, F, Nc)

Surround viewing conditions are other input parameters of CIECAM97s. Surround viewing conditions parameters have previously, in Chapter 3, been categorized into 4 groups. *Average* such as viewing surface colours in daytime, *dim* such as viewing

TV, *dark* such as viewing projected film in dark room and *cut-sheet* such as viewing transparency. As defined in [6, 7], the colours of traffic signs are surface colours. Therefore, viewing traffic signs in daytime is in the group of viewing surface colours. The surround viewing condition parameters are set to average surround in day light conditions as c=0.69, FLL=1.0, F=1.0, Nc=1.0 [66, 73].

Yb value

The Yb value is defined as the relative luminance of the source background in the source conditions. Normally it is fixed [14, 73], therefore, in our study it is set to typical value of 20.

4.4.1.2 Ranges of perceptual colour

The outputs of CIECAM97s are perceptual colour attributers. The ranges of perceptual colours are used as criteria for segmentation. In [14], one colour can be represented in an environment by hue, lightness, and chroma (*HCJ* colour space in chapter 3). In our study, red and blue colours of traffic signs are considered. The figure below shows the procedures of obtaining a range of perceptual colours.

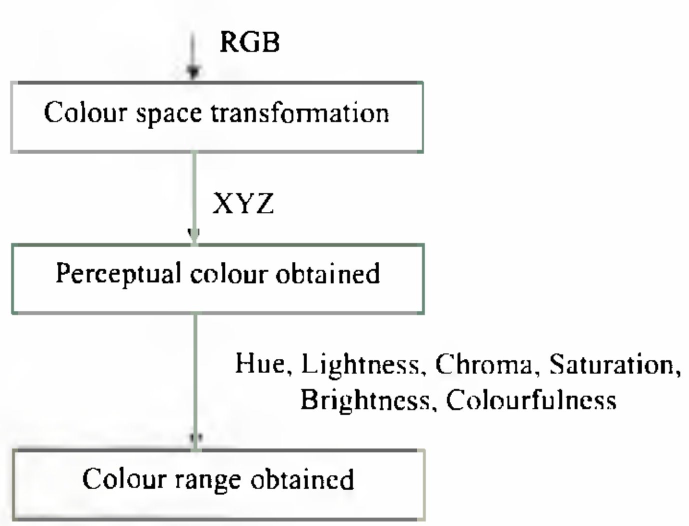


Figure 4-16: The procedure of obtaining perceptual colour

In the first stage, the samples of colour pigments are acquired from traffic signs in pictures under different weather conditions respectively. 10 pictures are taken from following weather conditions such as sunny, cloudy, rainy day respectively. In every picture, 15 points from sign are selected for ranges calculation. Finally 150 red pixels and 150 blue pixels are chosen from signs in one weather condition and are used to calculate the range of colour in one weather condition. Those pixel values are stored as *RGB* colour space and converted to *XYZ* colour space, using formula (4-7). By using the *XYZ* tristimulus value and input parameters described above, we can obtain the perceptual colour values from CIECAM97s.

In the second stage, the mean and standard deviation of perceptual colour of samples will be calculated under different luminance level. The luminance of the adapting field is determined by the weather conditions and time of day (e.g. reddish at sunrise and bluish at noon), luminance of the adapting field are different. These values can be found in [66]. For adapting luminance in the normal viewing range, the values of luminance change between approximately 10 and 2000 candela square meters (cd/m^2) [14]. Expanding the luminance value to low value 5 and high value 2150, the mean and standard deviation of hue, chroma and lightness of samples in different luminance levels can be obtained. Finally, mean ± standard deviation of perceptual colour values can be used to express the ranges of the colour for each weather condition.

In summary, ranges of hue, chroma and lightness values are used as thresholds for segmentation of traffic signs. When segmenting, each pixel in an image is converted to colour attributes expressed using CIECAM97s colour terms. Then hue, chroma and lightness values are compared to the thresholds to segment the region of interest that is the region containing potential traffic signs. The segmentation results are described in Chapter 5.

4.5 Recognition based on BMV model

In our study, BMV model is applied to identify the traffic signs. This model is created and modified by Russia group [5, 67]. The recognition based on BMV has two procedures to deal with: *memorizing* and *recognition*. *Firstly*, memorizing means a database of each standard traffic signs is created by using BMV model, called model-specified database. This database includes two aspects. One is feature description including features of standard images expressed by using AW at points and scan path by using this model. The other one is the standard images. *Secondly*, recognition is comparing features of the potential object calculated by BMV model to the database. Once features of this object could be found from this database, it means that this potential object is matched one of signs in the database and the corresponding standard image will be shown in screen.

4.5.1 Feature description and model-specific database creation

The following section will describe features of traffic signs by using BMV model firstly.

4.5.1.1 Feature description of traffic signs using BMV model

BMV provides a compressed and invariant representation of each image fragment by space-variant features extracted in the fragment by the Attention Window (AW). To describe the features of an object in traffic sign recognition, the Attention window (AW) (described in Chapter 3) is utilized. Sometimes this is called Input Window (IW) [104], or Sensor Window (SW) [105, 106]. They have similar structures below.

1. An image is presented by 49-dimensional vector of orientation extracted in vicinity of each of 49 nodes of IW or SW.
2. The IW/SW is located at the intersections of sixteen radiating lines and three concentric circles.
3. Orientation of segments in the vicinity of each IW/SW node is determined by means of Gaussian convolution with spatially shifted centres with the step of 22.5°.
4. Representation of space-invariant image is emulated by Gaussian convolutions with different kernels.

To increase a specificity of traffic sign, the algorithm of context description of oriented elements in the vicinity of each of 48 peripheral nodes (except for the central node) of the IW/SW has been developed. The size of context, for sixteen nodes, is equal to 3x3 pixels on the central SW circle, 5x5 for the immediate circle, and 7x7 for the peripheral circle. An example of oriented elements detected in the context area of the indicated node of a sign is shown in the Figure 4-17 below.

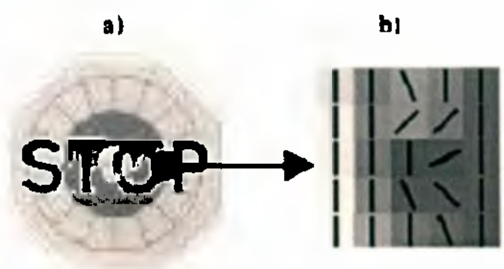


Figure 4-17: Example of AW (IW) feature description [106]

a) Schematic of the SW located in the centre of the informative part of a sign. Circles with different grey levels represent different resolutions in the SW structure, 16 lines means every 22.5 degree have a fix point.

(b) Oriented elements detected in the context area of a SW node (indicated by a small black spot in (a))

Figure 4-17 above shows the feature descriptions of sign 'stop'. Based on such feature description of signs, the standard feature database of traffic signs for UK and Russia can be created.

4.5.1.2 Creation of model-specific database

The model-specific database consists of two components. One is standard traffic sign image and the other is their features respectively. There are two sets of signs considered in our study. One is British standard traffic signs (n=105), images are acquired by scanning them from the book of Highway Code of UK or downloading from the web [9]. The other is Russian traffic signs (n=158), also obtained from the web site [107]. The features of traffic signs are obtained by the following processes. Firstly, the standard traffic signs are normalized to 40×40 pixels and converted into grey-level representation using Matlab6.5 image toolbox. Then, each image's features are described using the methods mentioned above. The standard traffic sign images, with their features, form a model-specific database of traffic signs.

4.5.2 Recognition based on BMV

After database creation, the next step is to compare features of possible signs with the features of the database stored within database to obtain a successful match. After segmentation, the segmented images are resized to 40*40 size images and converted into grey scale image. The features of grey image will be described by AW (IW, SW) of BMV (described in section 4.5.1.1). Following, those features are compared to features of standard database. The results should output the matching signs from standard traffic signs database if it is matched, otherwise it shows no result. The results of traffic sign recognition in real environment will be described in Chapter 5.

In this chapter, a new approach of traffic sign recognition, derived from human cognitive processing has been represented. Firstly, the weather status of images is classified based on the local features of sky and road. Secondly, the procedures of segmentation based on colour appearance model are detailed. Finally, recognition based on the human behaviour model of vision is described in detail. The experimental results and analysis will be given in the next chapter.

Chapter 5

Results and Analysis

In this research, three sequences of processing steps leading to a new approach to traffic sign recognition under different viewing conditions have been developed. These are weather classification, segmentation of traffic signs from rest of scene based on CIECAM97s, and recognition of traffic signs based on the BMV model. The weather identification should afford the viewing conditions of weather from an image. CIECAM97s would provide good segmentation results under different viewing conditions to increase recognition rate and reduce the redundant calculation of traffic recognition. BMV model will give high recognition with invariance to scale, rotation and perspective distortion *etc.*

5.1 Study of weather classification

Participants analyze global and local features to asses the weather conditions of an image. At first sight, most participants always consider a whole image as either bright or colourful. Then, the detailed features such as wet road, umbrella for a rainy day, blue sky for sunny day, are used to distinguish the weather. The following Figure 5-1 gives a schematic of weather analysis of an image by participants.

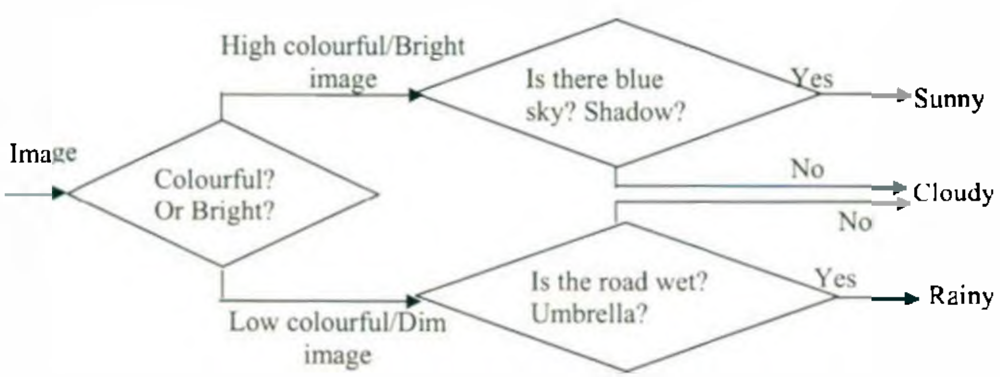


Figure 5-1: Weather classification by participant

From the schema above, Figure 5-1, it can be said that participants judge the weather of an image in two stages: global analysis and local analysis. Firstly, the whole picture is judged as either colourful or less colourful images. Within colourful images group, it has been found that blue sky and shadow are used to determine whether the day is classified as sunny, and inversely whether it can be classified as cloudy. Within less colourful images it was found that a wet road and umbrella(s) are used as the visual features that indicate a rainy day. Otherwise it will be classified it as a cloudy day. The examples of global analysis and local analysis by participants are given in Figure 5-2, Figure 5-3 and Figure 5-4 below.

(a) Sunny day

(b) Cloudy day

(c) Rainy day

Figure 5-2: Three examples of pictures under different weather conditions

The three pictures above illustrate that the top image of a sunny day appears significantly more colourful than the others. The last image, of a rainy day, seems greyer and less colourful. The colourfulness changes from high to low through sunny day. cloudy day to rainy day.

The two pictures below in Figure 5-3 and Figure 5-4 below demonstrate that local features that are used to identify weather status. The upper set of pictures in Figure 5-3 have similar colourfulness but different results for cloudy day and sunny day, determined by the local feature of blue sky. The lower set of pictures in Figure 5-4 have similar colourfulness but are classified as rainy and cloudy respectively. this is determined by the presence of a wet road.

(a) Cloudy day

(b) Sunny day

Figure 5-3: Similar colourfulness or brightness of picture, different weather for cloudy, sunny day judging by blue sky

(a) Cloudy day

(b) Rainy day

Figure 5-4: Similar colourfulness or brightness of picture, different weather for cloudy, rainy day judging by wet road

This preliminary study shows that the global features, brightness or colourfulness, of an image are used for the weather classification firstly. After considering global features, most people then move to analyzing the local features. In this study the local features are blue sky for sunny day and wet road for rainy days. These features pave the way for implementing a computerized weather identification system.

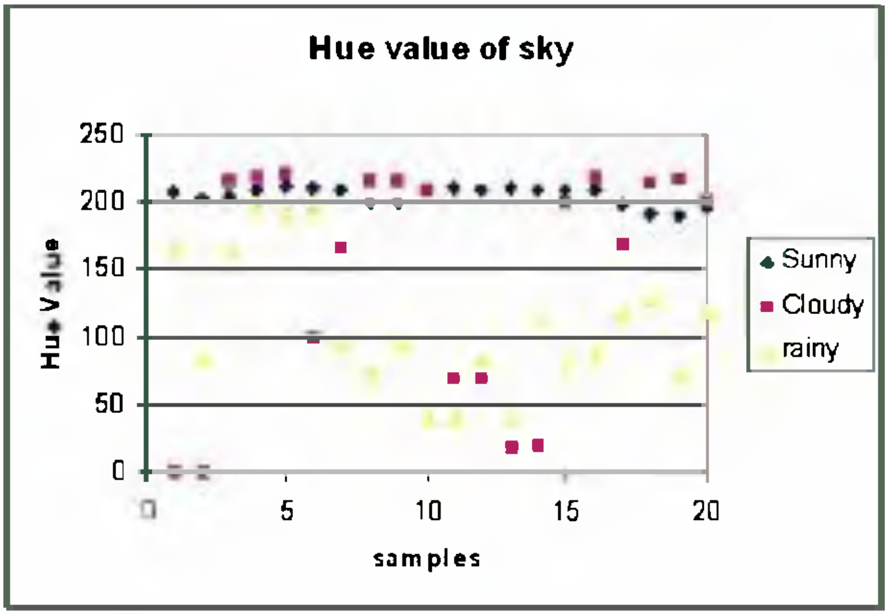
5.2 Weather classification by computer

The previous study on weather classification gives us a clue that local features are important to identify the weather conditions from pictures. Therefore, in this study blue sky is used to detect the sunny day and wet road are used to detect the rainy day. These features are adopted to classify weathers by computer.

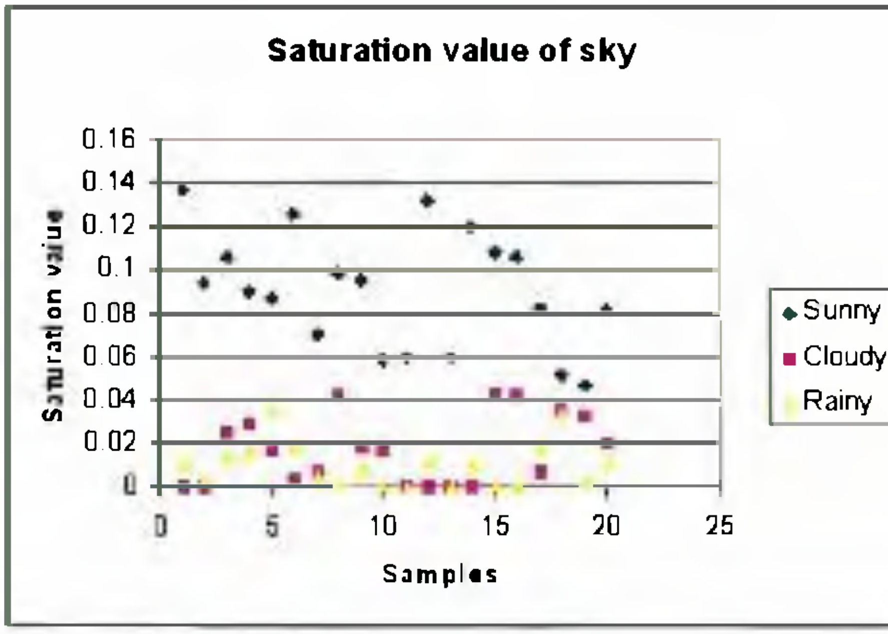
In total, 48 pictures of sunny day, 27 pictures of cloudy day and 53 pictures of rainy day are picked up as sample images for this study. According to the procedures detailed in Chapter 4, results of classification are described below.

5.2.1 Sunny day identification based on sky colour

According to the preliminary study, humans distinguish weather status using sky colour. Hue and saturation of *HSI* colour space are therefore used to describe the sky colour features. 20 blocks of blue sky are selected from 10 samples of sunny days and 20 blocks of sky are also selected from 10 pictures of cloudy days. The mean values of both hue and saturation for each block are represented in Figure 5-5 below.



(a) Hue



(b) Saturation

Figure 5-5: Hue and saturation value of sky examples

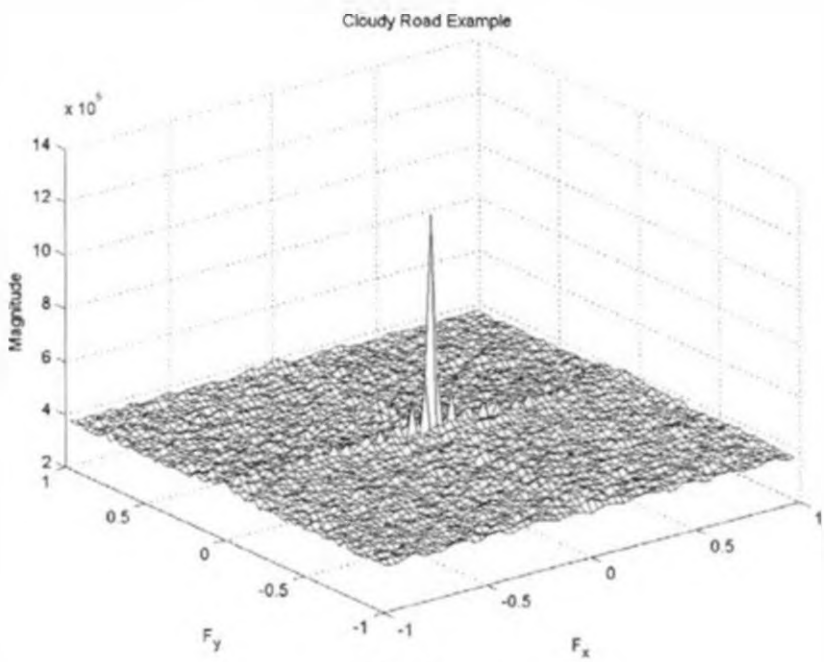
The figures above provide the hue and saturation results of sky examples which are cropped from the original images. Figure 5-5 (a) describes the hue value of sky examples. The horizontal axis represents the sample and the vertical axis represents hue value. Figure 5-5 (b) describes the saturation value of example sky. The horizontal axis expresses the sample and vertical axis expresses the saturation. (◆) expresses the sunny day. (■) represents cloudy day and (△) represents the rainy day.

It is shown, in Figure 5.5, that sunny day has hue value near to 200 and saturation value that is larger than 0.04. Conversely the hue values of sky from cloudy day change from 0, to near 200; and the saturation values are smaller than 0.04. In addition Figure 5.5 above shows that hue values of sky under rainy conditions varies between 40 and 200 and the saturation values are also smaller than 0.04. The values from rainy and cloudy days are comparable; therefore hue and saturation can be used to differentiate a sunny day from both rainy and cloudy.

5.2.2 Rainy day and cloudy day classification based on FFT

To identify rainy and cloudy days the texture of road is quantified, and then utilized as the determinant. A method of texture quantification based on the *Fourier* transform has been employed in this study and was previously detailed in Chapter 4.

The results obtained from applying the *Fourier* transform to two different examples in Figure 4-11 are presented in Figure 5-6 below. The figure at the top is the *Fourier* transform of road in cloudy day, and the bottom one is the *Fourier* transform of road in rainy day.



Cloudy

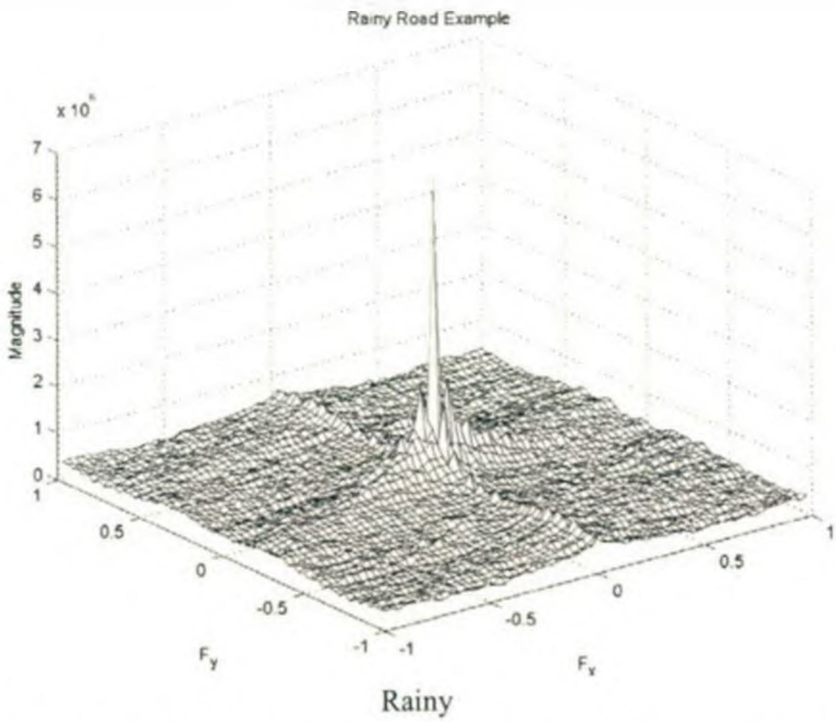


Figure 5-6: Examples of Fourier transform of two road textures in rainy day and cloudy day

These two pictures apparently demonstrate that the texture pattern on road of rainy day is coarser than from a cloudy day. It illustrates the difference of road features between a cloudy and a rainy day. In the top picture of Figure 5-6, the amplitude distribution of *Fourier* transform of cloudy day, is flat. However, the rainy day shown in the bottom picture of Figure 5-6 has significantly increased variation. The average magnitude (AM) of the *Fourier* transform is used to describe the road texture feature. 15 blocks of road are selected from 10 pictures of cloudy day and 12 blocks of road are selected from 10 pictures of rainy day. The average magnitude (AM) of each block is represented in Figure 5-7 below.

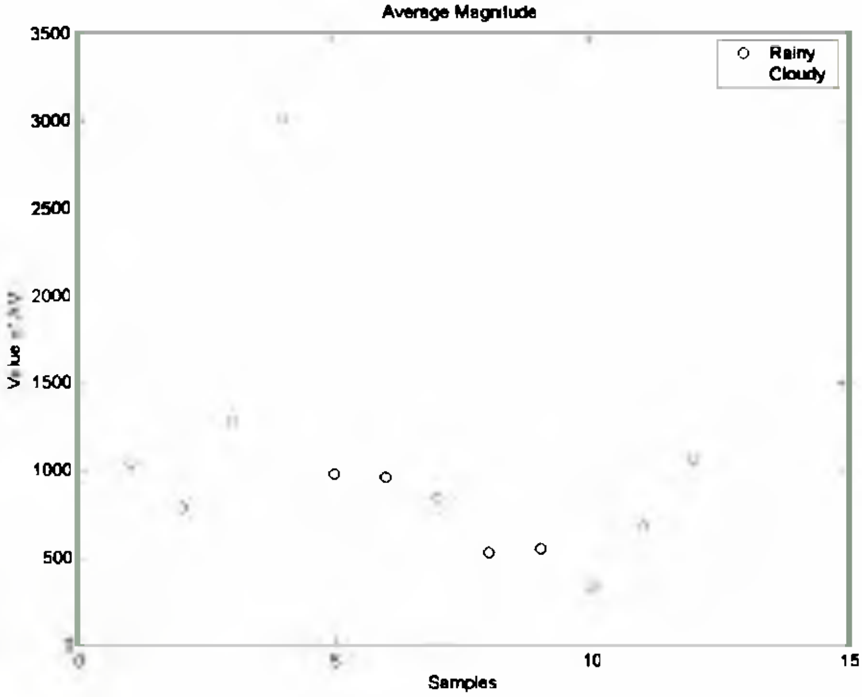


Figure 5-7: Average magnitude value of road examples in rainy day and cloudy day

Figure 5-7 above shows the average magnitude of the *Fourier* transform for road examples. The horizontal axis and vertical axis represent the samples and average magnitude respectively. From the figure above it can be seen that the majority of AM values from a rainy day (○) are larger than 500. However, the majority of AM values from a cloudy day (◇) are smaller than 500. It provides evidence for the postulation that the average magnitude of *Fourier* transform can be employed to distinguish between a rainy or cloudy day. In fact the evidence presented suggests that 500 is the threshold value between cloudy and rainy day

Based on the previously described the local features detailed above, images can be identified into 3 different weather conditions, sunny, cloudy or rainy. Final classification results based on this method are shown in Table 5-1. Finally, 43 of 48 sunny day pictures, 23 of 27 cloudy day pictures and 48 of 53 rainy day pictures can be identified as correct weather condition, which gives 90%, 85% and 91% success classification rates respectively.

Weather	Pictures	Correct classification
Sunny	48	43
Cloudy	27	**23**
Rainy	53	48

Table 5-1: Classification results of three weather cnnditinns

The weather status of an image is a fundamental input of the CIECAM97s colour appearance model. After correctly classify the weather type of an image, the next step is to segment potential traffic signs from the images, by utilizing the CIECAM97s.

5.3 Segmentation based on CIECAM97s

To segment traffic signs from images it is necessary to obtain the range of perceptual colour attributes of traffic signs under different weather conditions based on CIECAM97s. Then the images can be segmented by comparing to those thresholds.

5.3.1 Thresholds---Ranges of perceptual colour

To segment images the thresholds of perceptive colour, of traffic signs, under different weather conditions are obtained. This procedure has been described in Chapter 4.

The mean and standard deviation of perceptual colour values when exposed to different luminance levels are listed in the tables below. Table 5-2 and Table 5-3 show the results of mean and standard deviation values from red samples in sunny day. Table 5-4 and Table 5-5 show the results of mean and standard deviation values from blue samples in sunny day.

Luminance	Hue	Lightness	Chroma	Brightness	Saturation	Colourfulness
5	389.12	62.18	36.39	23.30	69.54	30.26
50	392.15	64.2	38.37	33.58	74.64	35.80
100	392.92	65.12	38.51	37.40	74.84	37.20
300	393.56	66.82	37.97	44.10	72.99	38.75
500	393.76	67.72	37.50	47.48	71.52	39.27
700	393.90	68.35	37.14	49.79	70.41	39.54
900	394.01	68.85	36.84	51.56	69.50	39.72
1100	394.08	69.25	36.59	52.99	68.74	39.84
1300	394.15	69.59	36.37	54.20	68.08	39.93
1500	394.21	69.89	36.17	55.25	67.49	40.01
1700	394.26	70.16	35.99	56.17	66.96	40.06
1900	394.31	70.40	35.83	57.00	66.48	40.10
2150	394.3	70.67	35.64	57.92	65.94	40.14

Table 5-2: Mean value of perceptual colour attributes of red samples in each luminance in sunny day

Luminance	Hue	Lightness	Chroma	Brightness	Saturatinn	CuluurTulness
5	20.08	9.96	5.05	2.50	14.84	4.20
50	18.05	9.82	5.4	3.44	16.40	5.08
100	17.58	9.71	5.53	3.74	16.68	5.35
300	17.37	9.48	5.67	4.20	16.87	5.79
500	17.36	9.35	5.73	4.40	16.86	6.00
700	17.36	9.2	5.75	4.52	16.81	6.13
900	17.37	9.17	5.77	4.61	16.76	6.22
1100	17.37	9.1	5.78	4.67	16.71	6.30
1300	17.38	9.04	5.79	4.73	16.66	6.36
1500	17.38	8.99	5.80	4.77	16.61	6.41
1700	17.38	8.94	5.80	4.81	16.56	6.46
1900	17.39	8.90	5.80	4.84	16.51	6.50
2150	17.39	8.85	5.81	4.87	16.4	6.54

Table 5-3: Standard deviation of perceptual colour attributes of red samples in each luminance in sunny day

Luminance	Hue	Lightness	Chroma	Brightness	Saturation	Colourfulness
5	296.64	42.63	48.09	18.03	113.29	39.98
50	296.8	44.29	48.82	26.10	115.47	45.55
100	296.78	45.16	48.70	29.17	114.79	47.04
300	296.38	46.91	48.53	34.68	113.51	49.52
500	296.12	47.88	48.34	37.52	112.51	50.61
700	295.93	48.58	48.18	39.48	111.67	51.29
900	295.77	49.13	48.02	41.00	110.94	51.78
1100	295.64	49.58	47.88	42.23	110.30	52.15
1300	295.53	49.97	47.76	43.28	109.72	52.44
1500	295.44	50.32	47.64	44.20	109.19	52.69
1700	295.35	50.63	47.53	45.00	108.71	52.89
1900	295.27	50.90	47.42	45.74	108.26	53.07
2150	295.19	51.22	47.30	46.54	107.75	53.27

Table 5-4: Mean value of perceptual colour attributes of blue samples in each luminance in sunny day

Luminance	Hue	Lightness	Chroma	Brightness	Saturation	Colourfulness
5	6.66	10.01	9.39	2.82	33.81	7.81
50	6.69	10.33	10.52	4.07	38.23	9.81
100	6.71	10.42	10.80	4.50	39.12	10.43
300	6.77	10.53	11.13	5.21	40.02	11.36
500	6.79	10.56	11.25	5.54	40.24	11.78
700	6.81	10.58	11.32	5.76	40.31	12.06
900	6.82	10.59	11.37	5.92	40.33	12.26
1100	6.83	10.60	11.41	6.05	40.32	12.42
1300	6.83	10.60	11.43	6.15	40.36	12.56
1500	6.84	10.60	11.46	6.24	40.27	12.67
1700	6.85	10.61	11.47	6.32	40.24	12.77
1900	6.85	10.61	11.49	6.38	40.20	12.86
2150	6.86	10.61	11.50	6.46	40.15	12.96

Table 5-5: Standard deviation of perceptual colour attributes of blue samples in each luminance in sunny day

Tables (Table 5-2 to Table 5-5) above show that the hue, chroma and lightness of blue and red change exhibit little variance, whilst luminance changes widely from 5 to 2150. Therefore, the mean and standard deviation of hue, chroma and lightness in Table 5-2 and Table 5-3 of red, in Table 5-4 and Table 5-5 of blue can be averaged to express mean and standard deviation of perceptual colour in sunny day. These two values can be used to calculate the ranges of perceptual colour. The values are listed in the Table 5-6 below.

Colours	Hue		Chroma		Lightness	
	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation
Red	393	18	37	6	68	9
Blue	296	9	48	11	48	11

Table 5-6: Mean and standard deviation of hue, chroma and lightness in sunny day

Brightness and saturation of CIECAM97s are not used as the thresholds because these 3 values change hugely with luminance changes. According to the Table 5-2 and Table 5-4, the mean value of brightness of colour red changes from 23 to 60 and blue changes from 18 to 48. The highest value of them is as around 3 times as the lowest value. Furthermore, the mean of saturation of colour red changes from near 65 at luminance 2150 to 75 at luminance 100, but the value of standard deviation is around 19. For the saturation of colour blue the mean changes from near 110 to 118, but the value of standard deviation is around 40. Although mean value of the saturation of colour red and blue change small, the standard deviation value is big. Therefore, brightness and saturation of CIECAM97s are not utilized in our study.

Finally, mean ± standard deviation of these perceptual colour values can be used to express the range of colour for a sunny day, listed in Table 5-7.

Colours	Ranges of hue (Mean± Standard deviation)	Ranges of Chroma(Mean± Standard deviation)	Ranges of Lightness (Mean± Standard deviation)
Red	375-411	31-43	59-77
Blue	287-305	37-59	37-59

Table 5-7: Range of Hue, chroma and lightness of colour red and blue in sunny day

The same procedure is applied to data from cloudy and rainy days. The results are shown in the Table 5-8 below.

Weather conditions	Hue		Chroma		Lightness	
	Red	Blue	Red	Blue	Red	Blue
Sunny day	375-411	287-305	31-43	37-59	59-77	37-59
Cloudy day	370-413	275-295	28-45	32-65	55-73	36-67
Rainy day	345-405	270-298	30-44	35-57	50-70	33-57

Table 5-8: The ranges of perceptual colour of traffic signs under different weather conditions

5.3.2 Segmentation results

This section shows the results of segmentation. A screenshot of the graphical user interface for the traffic sign recognition system is shown in Figure 5-8 below.

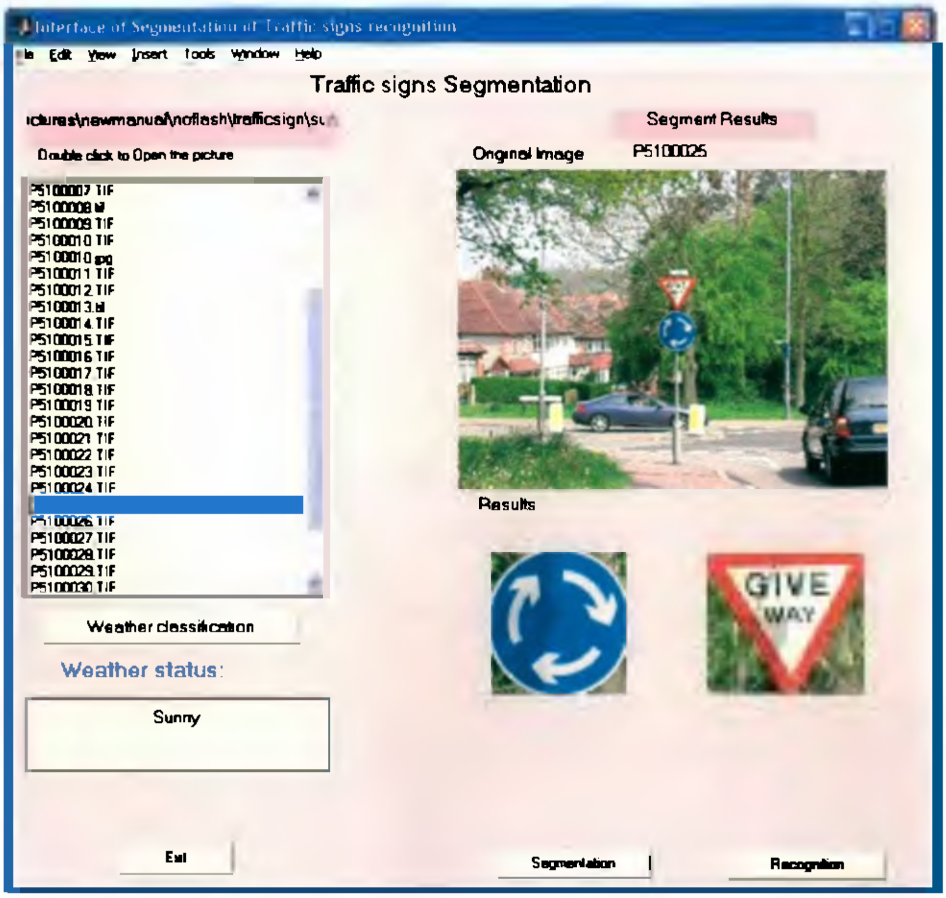


Figure 5-8: Interface of segmentation

From the interface above, those images being processed are listed in the left roller box. When image is selected and opened, it will be shown in right top of the interface. After opening image, the weather condition is classified firstly. The result of weather will be shown in the left bottom box and segmentation results are shown in the right bottom of interface.

The examples of segmentation under three weather conditions are listed in the below Figure 5-9.

Sunny	Example	
	Results	
Cloudy	Example	
	Results	

Rainy	Example	
	Results	

Figure 5-9: Examples of segmentation based on CIECAM97s under three weather conditions

Figure above shows examples of segmentation under 3 weather conditions. Though colour appearance of traffic signs presents difference under different weather condition, we still can segment them by using CIECAM97s. In some cases, for example, the middle of Figure 5-9 shows an incorrect potential sign. This erroneous potential sign will be discarded in the recognition process, this was discussed in Chapter 4 and results will be presented in the following section.

To evaluate the results of segmentation, two terms are used. One is the *probability of correct detection,* denoted by P_c and the other is the *probability of false detection* denoted by P_f. [108]. They are defined as:

P_c = sub regions with traffic signs/ total signs.

P_f = sub regions with no traffic signs/total signs.

Three groups of images under three weather conditions are input to this model for segmentation. Total 128 pictures including 48 pictures of sunny day, 53 pictures of rainy day and 27 pictures of cloudy day are used. From these images a total of 142 traffic signs are manually detected, 53 signs in sunny days, 32 signs in cloudy days and 57 signs in rainy days. The results of segmentation are listed in Table 5-9 below.

Weather condition	Total signs	Correct segmentation	False segmentation	P_c	P_f
Sunny	53	50	15	94%	28%
Cloudy	32	29	11	90%	33%
Rainy	57	48	18	85%	32%

Table 5-9: Segmentation results based on CIECAM97s

The above table shows that around 94% correct segmentation and 28% false segmentation for sunny day. 90% correct segmentation and 33% false segmentation for cloudy day, and around 85% correct segmentation and 32% false segmentation for rainy day. Totally, correct segmentation rate is near 90%.The segmentation is around 7 to 10 seconds at standard Pentium 3 400 MHz computer.

Although, the segmentation based on CIECAM97s demonstrate reliability and validity it is not 100% accurate. Some traffic signs were not extracted from the image and false segmentation sometimes occurs: these include objects such as rear car light. Two examples of this are shown in the following Figure 5-10.

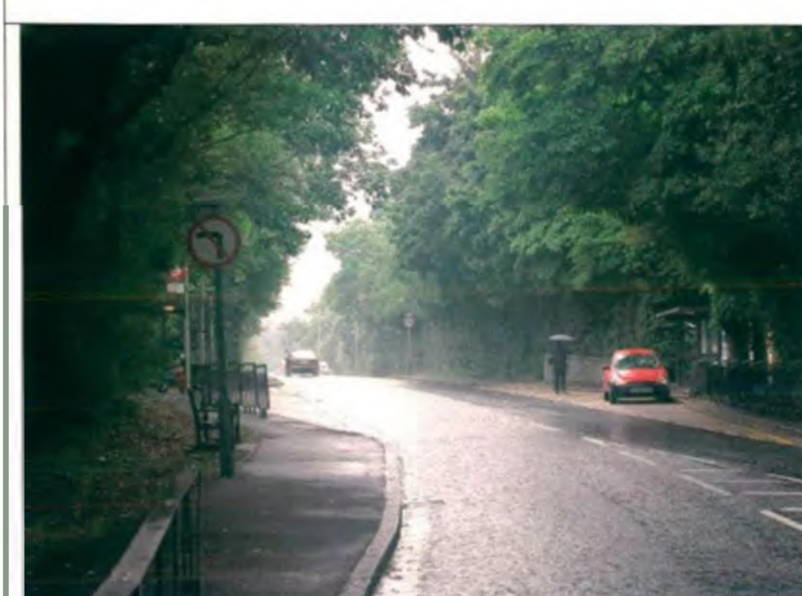
Example	Segmentation Results
	None

Figure 5-10: Examples of false segmentation

The top picture in Figure 5-10 gives an example of signs not being segmented. The reason is that viewing condition/reference white, which was obtained, can not represent this situation. The traffic sign in Figure 5-10 is under the tree in rainy day, however, it is much darker than the situation of rainy day in Chapter 4. Hence, the reference white of this is different. Therefore, the reference white obtained in Chapter 4 can not represent some special conditions: for example, the sign is under a big tree and the environment is very dark. Under those special conditions, using the reference white obtained above will lead to perceptive colour calculation inaccurately. The lower picture in Figure 5-10 gives an example of false detection from other objects. In this case, the rear car light is being segmented. It will be discarded by recognition.

Regions of interest, which have been segmented from the original images, are then passed to the next stage of recognition based on the BMV model; the results from this stage are described in detail in the following section 5.4.

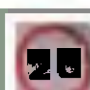
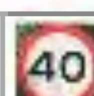
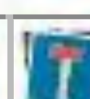
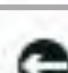
5.4 Recognition based on BMV model

After segmentation, it is necessary to recognize and identify those segmented signs. The interface of traffic sign recognition is in Figure 5-11 below.

Figure 5-11: Interface of BMV recognition

Figure 5-11 above shows that when recognizing a sign, the potential traffic sign is loaded into the program and displayed in left side of interface. If it is matched to one sign stored within database the matching sign will be displayed in right picture of interface. An error is caused if a matching sign is not found. The points in pictures are 49 features points which described in Chapter 4.

Some examples of the recognition results that have been obtained are listed in Table 5-10 below.

Segmented Images								
Results					Non identify image			
	a	*b*	*c*	*d*	*e*	*f*	*g*	*h*
Segmented Images								
Results		Non identify image					Non identify image	
	i	*j*	*k*	*l*	*m*	*n*	*o*	*p*

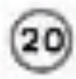
Segmented Images				
Results			Non identify image	Non identify image
	q	*r*	*s*	*t*

Table 5-10: Examples of recognition results

From the table above, it can be seen that this model correctly recognizes potential signs that have been segmented from the real environment. Images *a*, *k* and *p* show that this model can recognize signs that posses a large degree of perspective distortion. Image *h* demonstrates that this model can recognize rotated signs, and images *i*, *k* and *l* show that this model can recognize imperfectly segmented signs. Image *p* provides evidence that the model can recognize small-scale images. The above results also show that the method can recognize traffic signs of any shape, whether they are rectangular as in image *h*, triangles (up/down) as in images *g* and *i*, octangular as in image *q*, and also round as in images *a*, *b*, *m*, *n*. However, some segmented signs are not recognized, such as *e*, *j* and *o*. Image *e* gives an unrecognized example of imperfectly segmented sign. Image *j* shows an unrecognized example of segmented sign with a large amount of noise. Image *o* shows an unrecognized example of a scale transform. Image *s*, *t* show the discarded images by BMV model. The results of recognition based on BMV model is listed in the Table 5-11.

Weather	Potential signs after segmentation	Correct recognition	Recognition rate
Sunny	50	48	96%
Cloudy	29	27	94%
Rainy	48	44	93%

Table 5-11: Recognition rate under three weather conditions

Table 5-11 shows that 50 signs of sunny day, 29 signs of cloudy day and 48 signs of rainy day are identified separately. It shows BMV model correctly identified 48 out of 50 potential traffic signs images for sunny weather conditinns, 27 out of 29 for cloudy weather cnnditions and 44 out of 48 for rainy weather conditions, which gives 96% recognition rate in sunny day and 94% recognition rate of cloudy day, also 93% recognition rate of rainy day respectively. In total, this BMV model can identify 119 out of 127 potential traffic signs, which gives nearly 94% recognition rate. Recognition time is varied from 0.35 seconds to 0.6 seconds per image on standard Pentium 3 - 400 MHz machine.

In this chapter, results of segmentation and recognition based on human visual perception have been detailed. According to the description in chapter 4, an image undergoes three processing steps. First is weather identification, secondly segmentation based on CIECAM97s is carried out and finally the BMV model is used to recognize those segmented sub images. Results of weather classification are first described, and then the segmentation results under these three weather conditinns are detailed. Listed finally are the traffic sign recognition results.

Chapter 6

System Evaluation

The last chapter describes the results, of the three separate stages, of traffic sign recognition based upon human visual perception. In this chapter, three steps are used to evaluate the whole system, firstly, segmentation based on CIECAM97s is compared to segmentation based on other two colour spaces *HSI* and CIELUV. Then, quantitative estimations of recognition based on BMV model invariance ranges are analysed. Finally, the total recognition rate of this system will be calculated.

6.1 Comparison of segmentation using two colour spaces

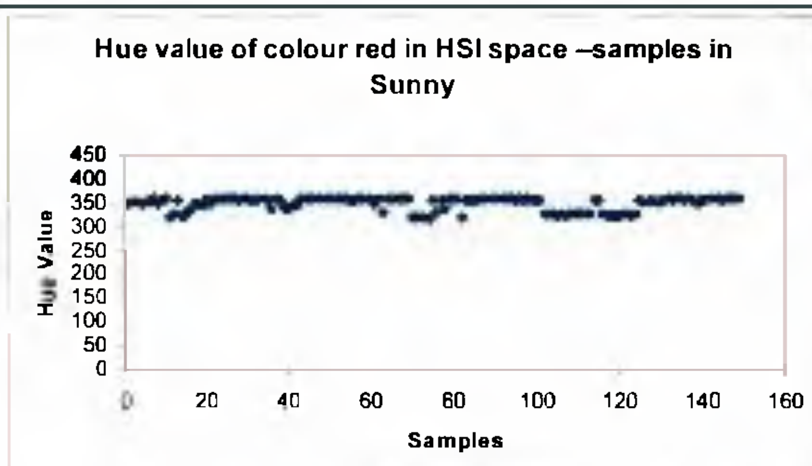
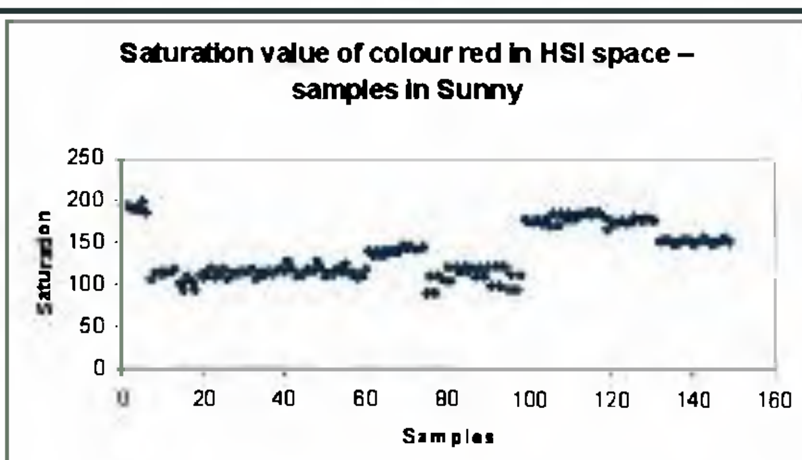
Two colour spaces are chosen, *HSI* and CIELUV, which are widely applied to colour image segmentation in the field of imaging research. *HSI* is based on the human perceptual attributes of hue, saturation and intensity. CIELUV was introduced by the CIE group and is used to predict colour under one standard viewing condition. In a process similar to the segmentation method based on CIECAM97s model introduced in chapter 4, thresholds of *HSI* and CIELUV colour space under three weather conditions are first obtained.

6.1.1 Thresholds ---Colour ranges

The same pixels that were employed in the segmentation based on CIECAM97s model to obtain thresholds are also utilized to create thresholds for the two colour spaces, *HSI* and CIELUV. Colour ranges, or thresholds, are again calculated using the mean and standard deviation of colour perceptual values calculated by using these two colour spaces. For example, hue, saturation and intensity value of sample pixels are obtained by using *HSI* colour space first. Then, the mean and standard deviation of hue, saturation and intensity value of sample pixels are calculated. Finally mean ± standard deviation of hue, saturation and intensity are used as thresholds.

6.1.1.1 Thresholds based on *HSI* colour space

To segment the traffic signs from the images based on *HSI* colour space, the formula (2-2) is used. Those pixels selected from sample images, under the three different weather conditions are calculated to obtain thresholds respectively. An example of, in this case, the colour red range of hue, saturation and intensity based on *HSI* colour space in sunny day is shown below in the Figure 6-1.

	Mean: 349 Std: 14
	Mean: 138 Std: 30

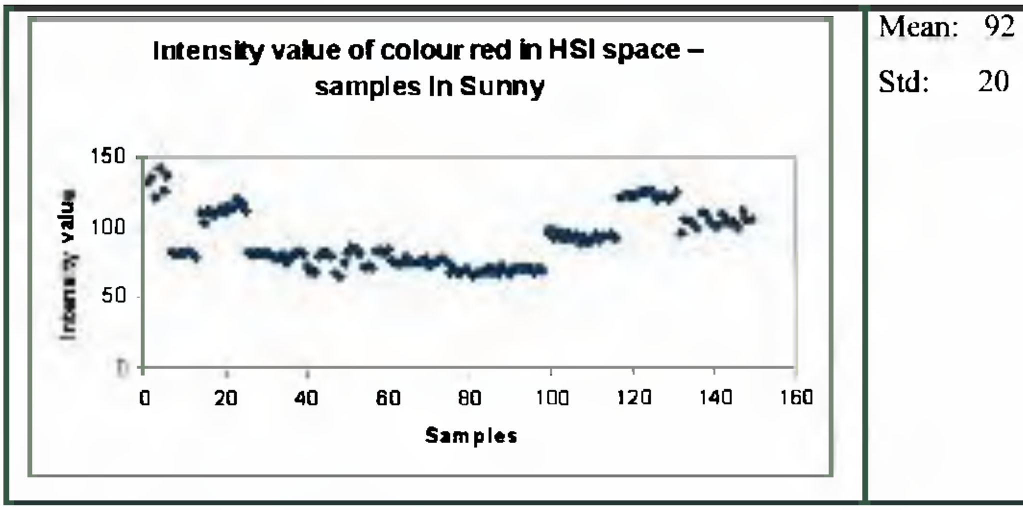


Figure 6-1: Example of red colour ranges of HSI colour space in sunny day

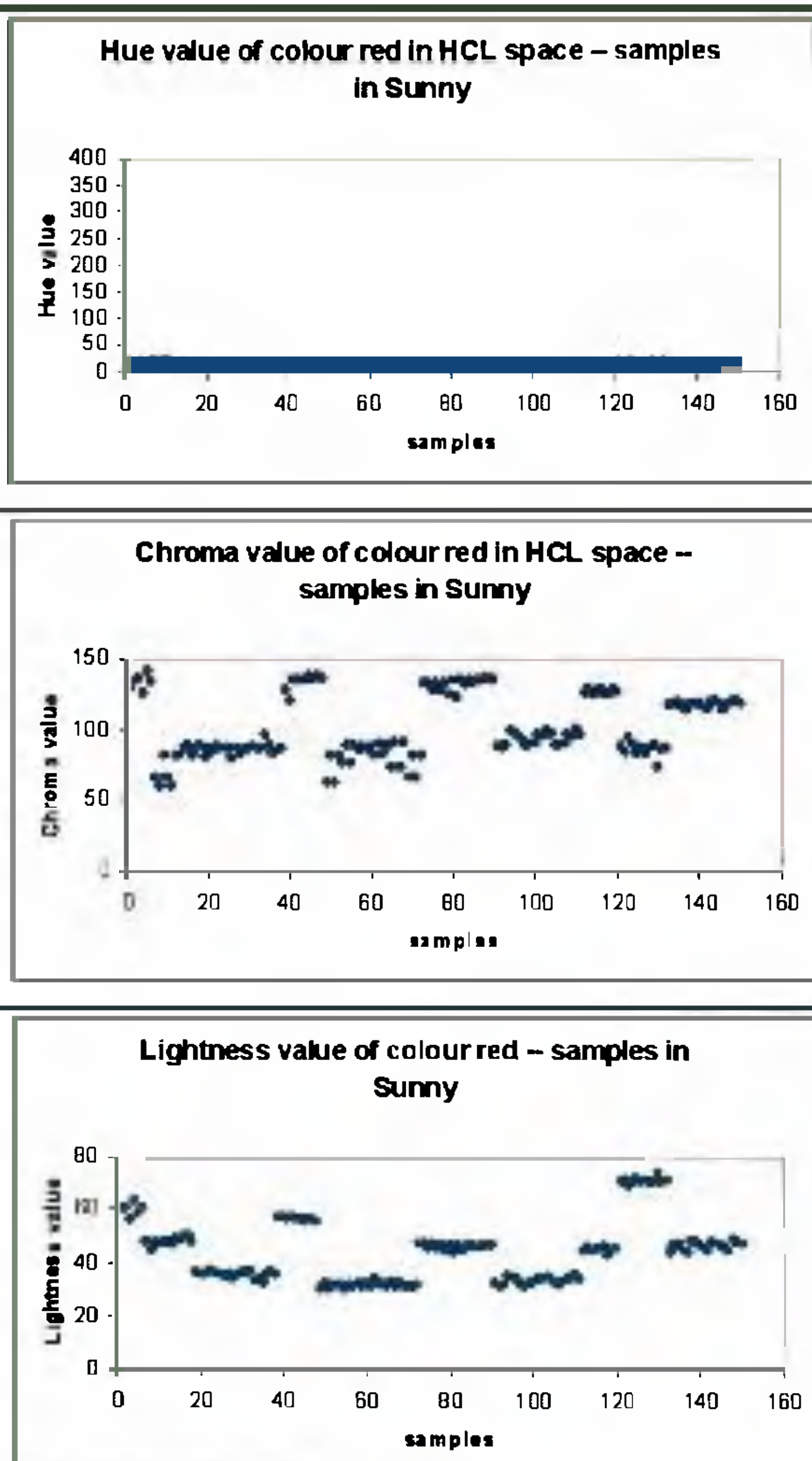
Figure 6-1 above shows that saturation changes from near 50 to 200 and intensity changes from near 50 to 150. To compare this *HSI* segmentation procedure to the method based on CIECAM97s, the mean ± standard deviation is again used to calculate colour range. The range is then used as the threshold to segment traffic signs from the images. It is therefore necessary to obtain thresholds of traffic signs based on *HSI* colour space under three weather conditions and these are listed in the Table 6-1 below.

Weather condition	Hue		Saturation		Intensity	
	Red	Blue	Red	Blue	Red	Blue
Sunny	335-363	190-215	108-168	80-155	72-112	55-105
Cloudy	332-361	190-205	105-175	150-205	85-155	100-125
Rainy	321-355	195-215	70-165	55-165	60-100	50-120

Table 6-1 : Thresholds of traffic signs based on *HSI* colour space

6.1.1.2 Thresholds based on *HCL* (CIELUV) colour space

Another colour space is considered. It is based upon three components of colour the Hue, Chroma and Lightness and is called *HCL*. *HCL* was created by CIELUV, and a detailed introduction is given in Chapter 2. The same pixels that were used in both the CIECAM97s and *HSI* models are again utilized to obtain thresholds under the three different weather conditions. An example of hue, chroma and lightness values, for the colour red in the *HCL* space, in sunny day viewing conditions is shown in Figure 6-2 below.



Mean: 13
Std: 6

Mean: 103
Std: 23

Mean: 45
Std: 18

Figure 6-2: Example of red colour ranges of *HCL* colour space in sunny day

The figures show that hue value of *HCL* is stable. However, lightness and chroma changes from near 50 to 150 and near 30 to 70 respectively. The mean ± standard deviation is used to calculate colour ranges for the thresholds of this colour space. Therefore, the thresholds of traffic signs based on *HCL* colour space under the three different weather conditions are obtained and listed in the Table 6-2 below.

Weather condition	Hue		Chroma		Lightness	
	Red	Blue	Red	Blue	Red	Blue
Sunny	7-23	230-255	80-126	30-80	27-63	25-75
Cloudy	6-27	230-245	75-125	55-80	35-60	45-60
Rainy	5-25	230-255	60-125	25-70	25-65	20-75

Table 6-2: Thresholds of traffic signs based on *HCL* (CIELUV) colour space

6.1.2 Segmentation results comparison

After thresholds are created, segmentation is performed. The examples of segmentation based on these three colour spaces are shown in the following figures from Figure 6-3 to Figure 6-5 under sunny, cloudy and rainy weather conditions respectively.

Example

Segmentation result

HCJ colour space (CIECAM97s)	 
HSI Colour space	  
HCL Colour space (CIELUV)	   

Figure 6-3: Example of segmentation in sunny day based on three colour spaces respectively

The Figure 6-3 above shows that in sunny day, segmentation based on CIECAM97s can detect the traffic signs "give way" and "roundabout". However, segmentation based on *HSI* colour space can detect those two signs and also the red roof. Segmentation based on CIELUV identified the two potential signs which are not traffic signs (the red roof and blue car decal).

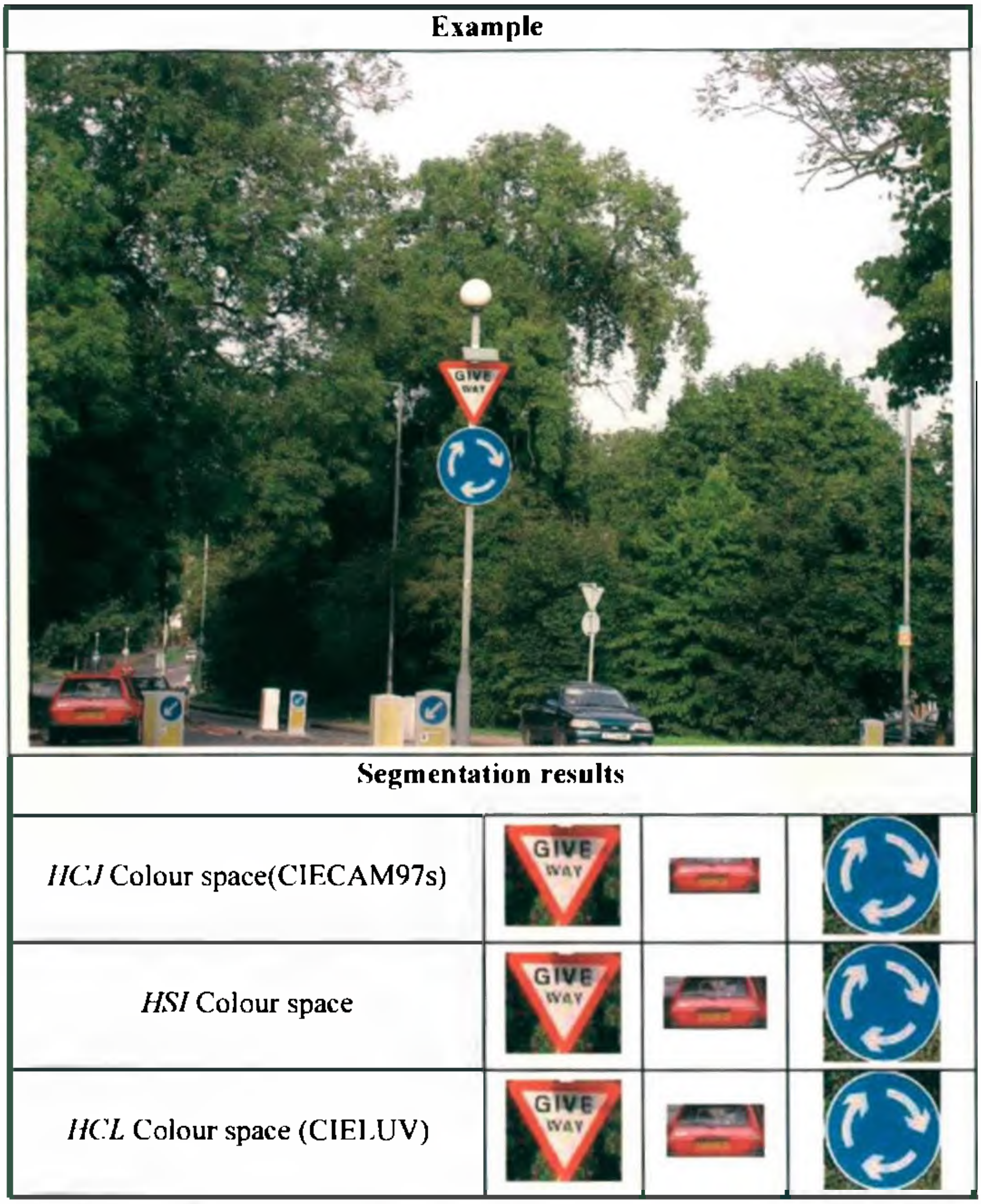


Figure 6-4: Example of segmentation in cloudy day based on three colour spaces respectively

The above Figure 6-4 shows that segmentation based on CIECAM97s, *HSI* and CIELUV have identical results for this picture, in cloudy day viewing conditions. The three methods detect the two traffic signs and also the rear of the red car.

Example	
	
Segmentation results	
HCJ colour space(CIECAM97s)	
HSI Colour space	
HCL Colour space (CIELUV)	

Figure 6-5: Example of segmentation in rainy day based on three colour spaces respectively

Figure 6-5 above shows that segmentation based on CIECAM97s can recognize two traffic signs "keep way" and "no entry" in rainy day viewing conditions. However, *HSI* colour space can only detect "keep way" traffic sign and the front of red car. Segmentation based on CIELUV only detects "Keep way" blue sign.

As well, segmentation results of those 128 pictures can be listed in tables. The following Table 6-3 make a conclusion of segmentation based on these three colour spaces under three weather conditions.

Weather condition	Total signs	Colour space	Results			
			Correct Segmentation	False Segmentation	*Pc*	*Pf*
Sunny	53	*HCJ*(CIECAM97s)	50	15	94%	28%
		HSI	46	19	88%	35%
		HCL(CIELUV)	46	17	88%	32%
Cloudy	32	*HCJ*(CIECAM97s)	29	11	90%	33%
		HSI	24	14	77%	46%
		HCL(CIELUV)	26	12	82%	38%
Rainy	57	*HCJ*(CIECAM97s)	48	18	85%	31%
		HSI	41	26	73%	45%
		HCL(CIELUV)	43	24	76%	42%

Table 6-3: Segmentation results comparison under three weather conditions

The Table 6-3 above gives the total results of segmentation based on these three colour spaces under sunny, cloudy and rainy weather conditions. In sunny day, the results show the segmentation based on CIECAM97s can segment 50 of 53 traffic signs, which give 94% probability of correct detection. Also, the probability false detection is 28%. However, the probability of correct detection based on *HSI* and *HCL* are 88%. The false detection is 35% and 32% respectively.

The Table 6-3 above also gives the segmentation results in cloudy day. From the table, the segmentation based on *HCJ* can achieve 90% correct detection rate. The segmentation based on *HCL* has 82% correct detection rate. The correction detection rate of segmentation based on *HSI* is 77%. However, the false detection rate of segmentation based on *HCJ* is 34%. The false detection rate of segmentation based on *HSI* and *HCL* are 46% and 38% separately.

The Table 6-3 above proves that segmentation based on CIECAM97s performs better results than other two colour spaces under rainy weather conditions. Segmentation based on CIECAM97s can reach 85% correct detection rate, which is 12% and 9% higher than *HSI* colour space and *HCL* colour space. The false detection is 31% lower than those two colour spaces, which gives 14% and 11% lower than *HSI* and CIELUV colour spaces.

These results in the table prove *HCJ* of CIECAM97s performs better than *HSI* and *HCL* (CIELUV) colour spaces in colour image segmentation in real environments. The theory and data analysis as following will testify this point.

6.1.3 Analysis

Theoretically, *HCJ* colour space can create a more accurate colour range than the *HSI* and CIELUV colour space. This may in some part be due to the fact that CIECAM97s, and therefore *HCJ*, not only considers colour stimulus itself, but also considers the background and surrounding conditions. It can predict the change of colour appearance when viewing conditions are altered by utilizing the basic perceptual attribute of hue. Moreover, it also predicts colour by the relative perceptual attributes of lightness and chroma. Through judging colour to reference white under same illuminant and environment, lightness and chroma would tend to remain stable for a given colour regardless view conditions [14]. However, CIELUV colour space is limited to being used under fixed viewing conditions [17]. CIELUV colour space does not consider colour changing affected by surround, and background. Considering *HSI* colour space, it does not consider any viewing conditions and surround either. It is apparently found from the equation (2-2) and (2-5) too. From equation (2-2), the hue, saturation and intensity are obtained from *R*, *G*, and *B* directly. It does not consider any viewing conditions. From equation (2-5), it can be found that the L, U and V are judged to the reference white, which normally is set to D65. In addition, it does not consider the viewing conditions parameters. Therefore, perceptual attributes of the two spaces, *HCL* and *HSI*, can be said to vary greatly for a single given colour under different view conditions.

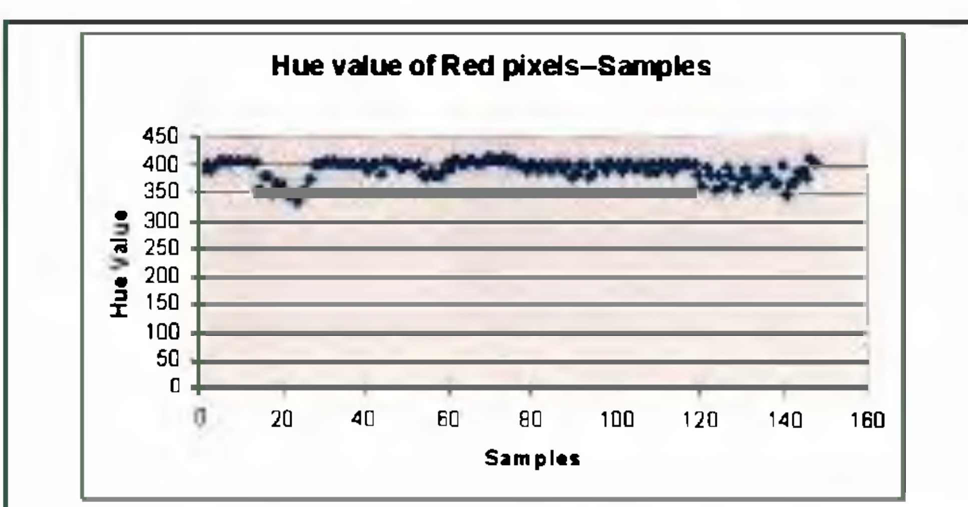
The previous examples of thresholds that were acquired in chapter 5 and chapter 6 in sunny day also provide evidence that segmentation based on CIECAM97s has better results than other two colour spaces. From example listed in Table 5-6, thresholds of colour red acquiring in sunny day, the standard deviation for hue, chroma and lightness of *HCJ* (CIECAM97s) is 18, 6 and 9 respectively. In Figure 6-1, the standard deviation for hue, saturation and intensity of *HSI* is 14, 30 and 20 respectively. From Figure 6-2, the standard deviation for hue, chroma and lightness of *HCL* are 5, 25 and 15 respectively. To compare those standard deviation, the standard deviation value are normalized to [0, 1] by divided to the scale of corresponding perceptual colour attributes values, which is the maximum value – minimum value of colour attributes value. The Table 6-4 below summarises the variation of standard deviation (truncated) of red colour.

Colour space	Scale of perceptual colour attribute	Weather	Standard deviation	Variation
HCJ	Hue: [0,400] Chroma:[0,125] Lightness :[0,100]	sunny	Hue: 18 Chroma: 6 Lightness: 9	Hue: 4% Chroma: 5% Lightness: 9%
		cloudy	Hue: 21 Chroma: 8 Lightness: 9	Hue: 5% Chroma: 6% Lightness: 9%
		rainy	Hue: 30 Chroma: 10 Lightness: 10	Hue: 8% Chroma: 8% Lightness: 10%
HSI	Hue: [0,360] Saturation:[0,255] Intensity:[0,255]	sunny	Hue: 14 Saturation: 30 Intensity: 20	Hue: 4% Saturation: 11% Intensity: 7%
		cloudy	Hue: 14 Saturation: 35 Intensity:30	Hue: 4% Saturation: 14% Intensity:12%
		rainy	Hue: 16 Saturation: 43 Intensity: 20	Hue: 8% Saturation: 17% Intensity: 8%
HCL	Hue: [0,360] Chroma: [0,260] Lightness: [0,100]	sunny	Hue: 5 Chroma: 25 Lightness 15	Hue: 1% Chroma: 9% Lightness: 15%
		cloudy	Hue: 11 Chroma: 25 Lightness: 13	Hue: 3% Chroma: 10% Lightness: 13%
		rainy	Hue: 10 Chroma: 33 Lightness: 15	Hue: 3% Chroma: 13% Lightness: 15%

Table 6-4: Variation of standard deviation of red colour to the scale of perceptual colour attributes

Table 6-4 above illustrates that *HCJ* colour space provides the more accurate threshold than other two colour spaces. By taking sunny day value as an example, the variation value of standard deviation of hue, chroma and lightness of colour red based on *HCJ* are 4%, 5% and 9% respectively. The variation value of standard deviation of hue, saturation and intensity based on *HSI* are 4%, 11% and 7% separately. The variation value of standard deviation of hue, chroma and lightness based on *HCL* are 1%, 9% and 15% separately. The hue value of *HSI* colour space changes bigger than other spaces. The intensity value of *HSI* colour space changes up to 11%. That means hue and intensity of *HSI* colour space have bigger range than hue and chroma value of *HCJ* and *HCL* colour space. From value above, it can be found that the range of chroma and lightness based on *HCL* colour space are bigger than chroma and lightness value based on *HCJ* colour space although the hue value has smaller variation than *HCJ* colour space. That means lightness and chroma in *HCJ* remain more stable than other colour space for a given colour with view condition change. That is to say, the values of hue, chroma and lightness created by CIECAM97s represents more accurate range of colour than other two colour spaces. It will provide more accurate segmentation results than other colour spaces. The values under cloudy and rainy day in Table 6-4 also prove that the variation of standard deviation of colour perceptual values based on *HCJ* are more accurate than other two colour spaces. For example, variation of hue, and saturation of *HSI* colour space in rainy day are 8% and 17%. However, the variations of hue and chroma value of *HCJ* are 8% and 10%.

Also the red colour samples values of *HCJ* for sunny weather conditions under one luminance value can be drawn in the following Figure 6-6.



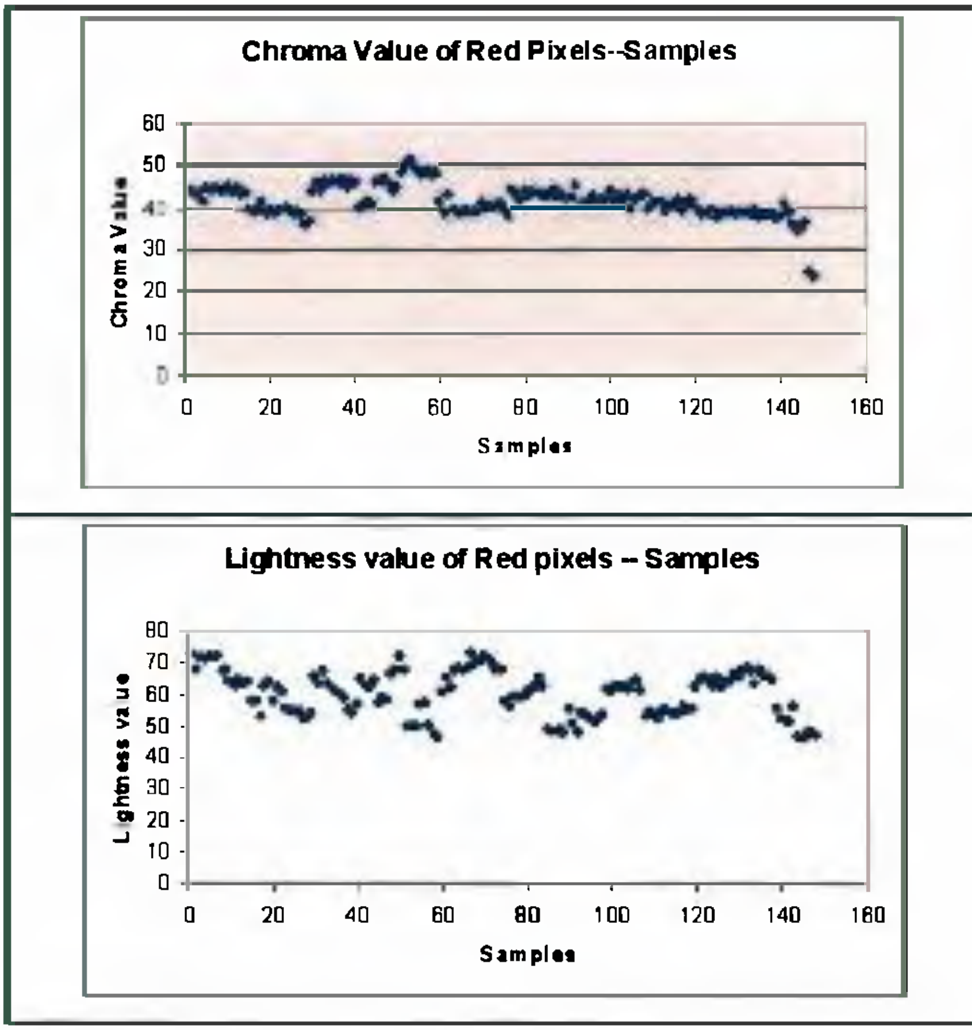


Figure 6-6: Samples values of *HCJ* of colour red under sunny weather conditions

By comparing Figure 6-1 and Figure 6-2, Figure 6-6 above proves that Hue, chroma and lightness of *HCJ* colour space created by CIECAM97s have smaller changes than *HSI* and *HCL* colour spaces. Although change of hue value of *HCL* is smaller than that of *HCJ* colour space, lightness and chroma changes bigger than *HCJ*. From figures above, value of Hue, intensity and saturation of *HSI* colour space changes bigger than hue, chroma and lightness of *HCJ* colour space.

The analysis above gives the evidence that *HCJ* should have more accurate data reflected a given colour in different view conditions than the other two colour spaces. Therefore, the segmentation results based on CIECAM97s performs better than other two colour spaces in real environment.

6.2 Quantitative estimations of recognition invariance range

In the natural environment, traffic signs are often affected by noise, rotation and viewing angles. To evaluate recognition ability of the BMV model, quantitative estimations of the recognition invariance range are given in this section. Graduated artificial transformations of standard traffic sign images have been employed to evaluate the traffic sign recognition by BMV model. Firstly, the method of artificial transformation will be described. Then, the results of recognition by BMV model will be detailed.

6.2.1 Artificial transformation of traffic signs

To imitate possible sign transformations in real road conditions and obtain the quantitative estimations of recognition invariance range, graduated artificial transformations (noise, scale, rotation, perspective distortions and occluded shapes) of standard traffic sign images have been performed by means of the Adobe Photoshop 7. Then, the transformed images are tested for recognition.

Noise

The qualities of real images are often degraded by some random errors (artefacts). This degradation is usually called *noise*. Noise can occur during image capture, transmission, or processing and may be dependant on, or independent of, image content [34]. For example:

If the image is scanned from a photograph made on film, the film grain is a source of noise. Noise can also be the result of damage to the film.

If the image is acquired directly in a digital format, the mechanism for gathering the data (such as a CCD detector) can introduce noise.

Electronic transmission of image data can introduce noise.

Usually, there are two noise types being considered. One is white Gaussian noise and the other is salt-pepper noise [11]. Noise has been simulated by adding graduated noise (5%, 10%, 20% and 50%) to the images from the standard database. The examples are in Figure 6-7 and Figure 6-8.

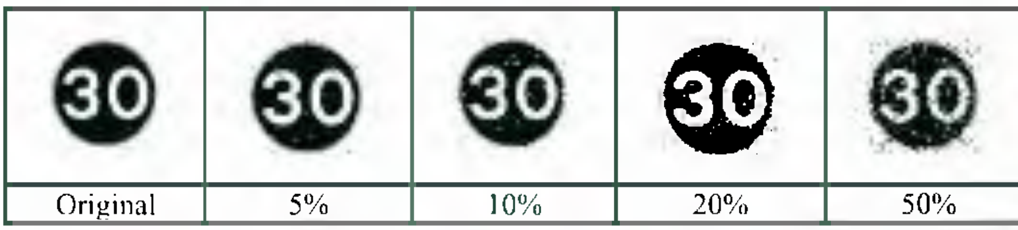


Figure 6-7: Gaussian Noise gradually added to signs

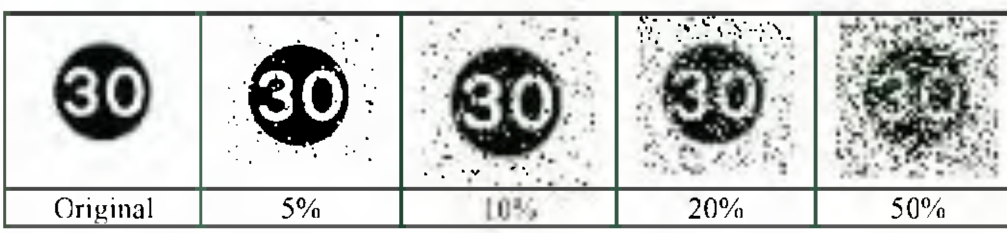


Figure 6-8: Salt and Pepper noise gradually added to signs

Scale

Scale transformations reflect the change of signs' size in images relative to the distance change between signs and viewing position. As described in picture collection in chapter 4, 10, 20, 30, 40, and 50 meters are considered. The following Table 6-5 and Figure 6-9 give the size of traffic signs corresponding to the distance.

Size of traffic signs	Distance between sign and Camera
15*15	50 meters
20*20	40 meters
25*25	30 meters
35*35	20 meters
55*55	10 meters

Table 6-5: Size of traffic signs in images corresponding to distance

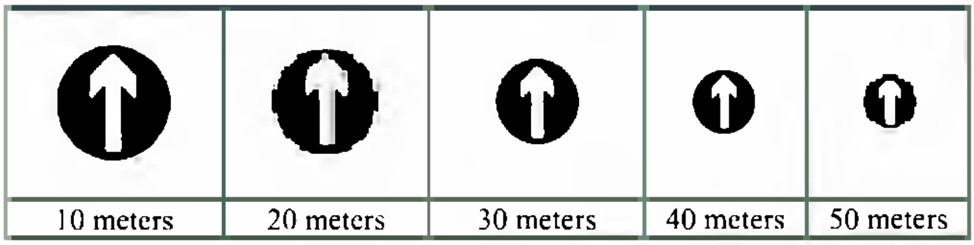


Figure 6-9: Size of traffic signs in images changing with distance

Rotation

Rotation is another factor to leading to misrecognition or unrecognisable. In real environment, traffic signs may be rotated by installation or altered later. Here, we do not consider the extreme condition. For example, if the "Ahead Only" sign above is rotated 90 degree clockwise or anti-clockwise, people will misidentify them to "turn right" or "turn left" sign. Therefore, rotation was simulated by changing angle clockwise and anti-clockwise direction from 5 degree to 20 degree listed in the following Figure 6-10.

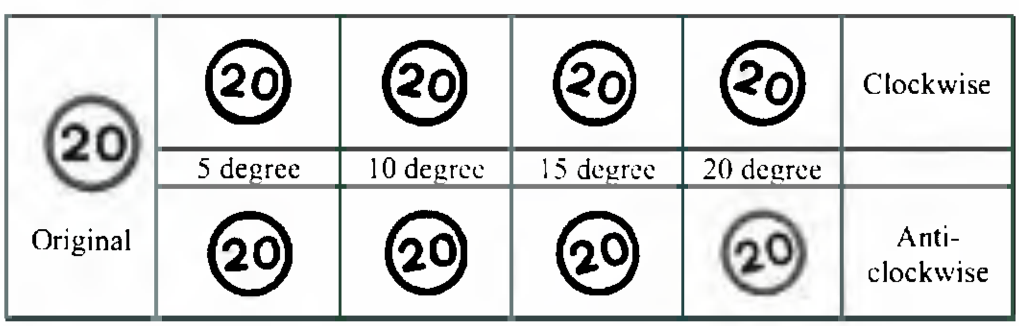


Figure 6-10: Rotation examples of traffic signs

Perspective Distortion

Perspective distortion imitates the differences caused by car movement and the consequent viewing position alterations. The examples are listed in the following Figure 6-11. Considered distortions are 5 and 10 degrees at two different viewing positions. High level viewing position represents driving a lorry whilst lower level viewing position relates to driving a car.

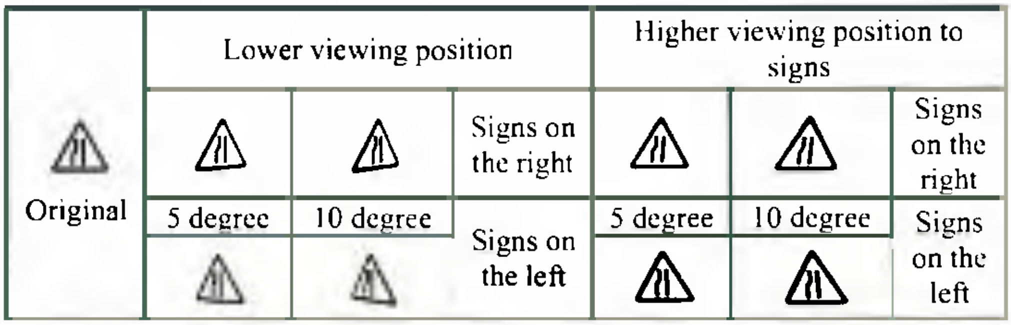


Figure 6-11: Original and perceptual distortion examples

Signs on the right mean perspective from left, Signs on left means perspective from right

Occluded Shapes

Occluded shape is when some parts of the shapes are covered or invisible. For example, some parts of signs are hidden by houses or trees, or cars, then, the segmented signs are not perfect. As well, some potential traffic signs can not be fully described after segmentation. Hence, some parts of original traffic sign will be lost. By cropping parts of traffic signs, these segmentation imperfection and occlusions are simulated. Examples of cropping 5, 10, and 15 pixels in 4 directions are listed in the following Figure 6-12.

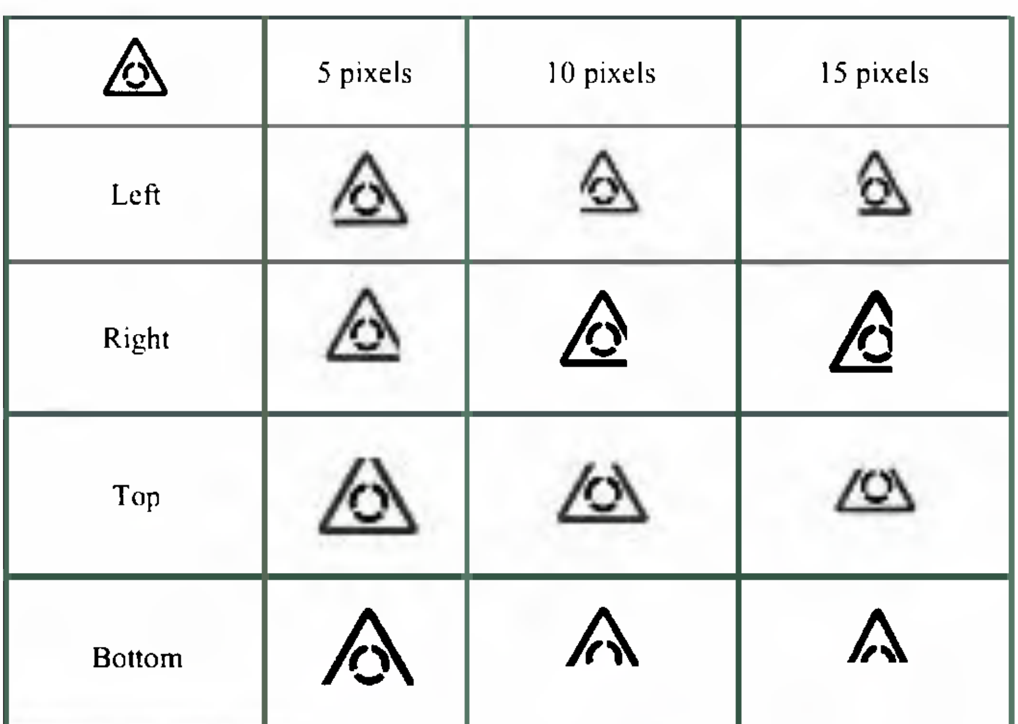


Figure 6-12: Cropped traffic signs

Artificial transformations of traffic signs caused by noise, scale, rotation, perspective distortions and occluded shape in real environments have been detailed in this section. The results of recognition, using the BMV model, on artificially transformed images will be given in the following section.

6.2.2 Results of quantitative estimations of recognition

To estimate the recognition invariance for BMV model, 18 standard blue signs and 50 standard red signs have had various levels of artificial transformations applied. The transformed images are then tested with the BMV model.

Noise

There are 2 types of noise used to test BMV model. One is Gaussian noise and the other is "salt and pepper". The noise is been added gradually from 5% to 50%. The correction results are listed in the Table 6-6 and Table 6-7 below.

Noise Level	Sign Groups	
	Blue	Red
5%	100%	100%
10%	100%	100%
20%	100%	97%
50%	93%	71%

Table 6-6: Gaussian Noise results

The table above shows that BMV model can effectively recognize traffic signs with 20% white Gaussian noise affected. It reaches 100% recognition rate of blue and 97% of red at 20% Gaussian noise. At 50% Gaussian noise level, the recognition rate of red is decreasing to 71% and blue to 93%.

Noise Level	Sign Groups	
	Blue	Red
5%	100%	100%
10%	100%	96%
20%	94%	83%
50%	50%	37%

Table 6-7: Salt and Pepper Noise results

The table above shows recognition rate when "salt and pepper" noise is applied. It reaches 100% recognition rate of blue and red at 5% noise. At 50% noise level, the recognition rate of red is decreasing to 37% and blue to 50%.

Scale

Recognition results when scaling transforms have been applied are listed in the Table 6-8 below. The size of traffic signs is changing from 15*15 to 55*55 as the distance changing from 50 meters to near 10 meters.

Size of traffic signs	Distance to signs	Sign Groups	
		Blue	Red
55*55	10 meters	100%	91%
35*35	20 meters	100%	100%
25*25	30 meters	100%	96%
20*20	40 meters	97%	90%
15*15	50 meters	94%	88%

Table 6-8: Scale modelling from 10 meters to 50 meters

The above table illustrates that recognition based on BMV is relatively invariant when scale alterations are considered. Both groups of signs obtain 100% recognition for 35*35 sized images. The minimal recognition rates are 94% and 88% for blue and red respectively, both obtained with 15*15 size images.

Rotation

Table 6-9 below summarizes the recognition results for images that have been rotated. Rotation varies from 5° to 20° (with increments of 5°) in both clockwise and anti-clockwise directions.

Degree		Sign Groups	
		Blue	Red
Clock wise	5	100%	100%
	10	100%	100%
	15	100%	100%
	20	92%	88%
Anti-clockwise	5	100%	100%
	10	100%	95%
	15	98%	90%
	20	88%	83%

Table 6-9: Rotation of recognition results

The results show 100% recognition rate for 5° rotation. However, it decreases to around 85% rate with 20° rotation in both directions.

Perspective distortion

Perspective distortion includes 2 types, lower and higher position. The recognition results are shown in the following Table 6-10.

Condition	Degree		Signs	
			Blue	Red
Lower	Signs on the right	5	100%	100%
		10	94%	90%
	Signs on the left	5	100%	100%
		10	94%	90%
Parallel near to sign	Signs on the right	5	100%	100%
		10	100%	96%
	Signs on the left	5	100%	100%
		10	100%	96%

Table 6-10: Perspective distortion results

The table above illustrates that BMV model has high recognition rate for images that have perspective distortion. It reaches more than 90% even when distorted by 10°.

Occluded shape

Table 6-11 represents the recognition results of imperfect shapes, cropping 5-15 pixels from 4 the different edge.

Crop pixels		Sign Groups	
		Blue	Red
Bottom	5	100%	100%
	10	98%	96%
	15	80%	70%
Left	5	100%	100%
	10	100%	100%
	15	87%	85%
Right	5	100%	100%
	10	100%	100%
	15	87%	85%
Top	5	100%	100%
	10	100%	90%
	15	90%	82%

Table 6-11: Imperfect shape results

The table above shows high recognition rate (100%) when only 5 pixels are cropped at any side. 90% recognition rate is achieved when 10 pixels cropped and reduced further to 70% when 15 pixels cropped.

The quantitative estimations of recognition demonstrate the ability of BMV model to recognize traffic signs effectively and invariantly, with respect to scale, plain rotation, perspective distortion, imperfect shape and noise. The BMV model can therefore be applied to recognize traffic signs in real environments within a certain ranges of shape transformations. It has been outlined in section 5.3.

6.3 System evaluation results

A new approach to traffic sign recognition based upon human visual perception has been presented. Results (in chapter 5) show that it can effectively detect and recognize traffic signs under different viewing conditions. This has been shown by comparing segmentation based on CIECAM97s colour appearance model to other perceptual colour spaces. Quantitative estimation of recognition based on BMV model has been performed separately. To assess validity of whole system, the recognition results of whole system are given in this section.

Total 142 traffic signs in 128 images are used to test our system. Every image is input to system and is classified as one of three weather conditions. After weather classification, the image is segmented through CIECAM97s by comparing the ranges of colour according to different weather status. The BMV model is then applied to recognize those potential segmented signs.

The whole system can recognize 119 traffic signs of total 142 correctly, which reaches 84% recognition rate. This system can detect different traffic signs under three weather and difference viewing conditions. The result proves that this approach can effectively overcome the drawback of other traffic sign recognition methods stated in chapter 2. The processing speed for one image is approximately 11 seconds using P 3-400 computer.

In this chapter, the complete system has been evaluated. By comparing to the other perceptual colour spaces, segmentation based on CIECAM97s is shown to produce good results in real environments. The recognition based on BMV model can identify a signs with a degree of distortion. Finally, the recognition results of whole system are given. Those results provide evidence of the validity of this new approach to traffic sign recognition in real environments.

Chapter 7

Conclusions and Future Work

The evaluation result shows that this approach based on human visual perception for traffic sign recognition under different viewing conditions is an effective recognition method. This chapter will draw a conclusion for the whole project. The contribution to scientific knowledge is presented and finally the direction of possible future work is discussed.

7.1 Summary

This dissertation presents a new approach for traffic sign recognition under different viewing conditions. This set of algorithms is based on human visual perception and consists of two human vision models. The CIECAM97s colour appearance model, which is based upon human visual perception, is used to segment images under different weather conditions to get potential traffic signs. After this segmentation a human vision behavioural model, the BMV model, is used to recognize those segmented images and obtain a match to traffic signs. Some work have been published in the conferences [104, 105, 109, 110].

In summary, the following results have been achieved during this study; they are:

1. Weather classification by people is investigated. Results are used to provide the knowledge of classifying weather automatically by computer

2. Weather classifications by computer are carried out. The method described here provides a system to identify weather from an image according to the local features of image.

3. Parameter input for the CIECAM97s is selected according to the real environment.

4. Reference white detection in real environment is performed under different weather conditions. The method provided the experimental way of reference white decision by using luminance meter to measure white hoard and objects.

5. Conversion between camera output RGB and CIE XYZ is done by using practical experimental methods and linear regression. It is used to transfer the colour data in real environment taken by a digital camera to the corresponding tristimulus XYZ value.

6. Image segmentation based on CIECAM97s is realized by threshold. The mean ± standard deviation is used as threshold. The result is shown that around 94% correct segmentation rate in sunny day, 90% correct segmentation rate in cloudy day and 85% correct detection rate in rainy day.

7. Image segmentation based on CIECAM97s is compared to segmentation based on two different colour spaces. The results show the segmentation based on the CIECAM97s is better than the other two colour spaces.

8. Quantitative estimations of recognition invariance range for BMV model proves that recognition based on BMV model can detect the traffic signs in a wide range of artificial transforms successfully.

9. Potential segmented images are recognized by using BMV model. The result demonstrates that this approach can distinguish objects effectively in real environment.

The research has demonstrated that a high recognition rate, of traffic signs located at a moderate distance, is achieved under different weather and viewing conditions.

7.2 Contribution

This dissertation presents the approach of traffic sign recognition based on human visual perception. The approach consists of two human vision models: colour perceptual CIECAM97s and behaviour vision model BMV. This set of algorithms provides a novel method for object recognition in real environments. This project contributes to the fields of computer vision and image processing in the following aspects:

A new approach for colour image segmentation in a real environment. This is a significant improvement to colour knowledge and CIECAM97s application in the fields of computer vision and image processing. It includes the:

- A new algorithm for identifying three weather conditions: sunny, cloudy and rainy day. This algorithm is implemented by analysing local features of an image.

- Viewing conditions setting for CIECAM97s colour appearance model in image processing. The importance of them is that reference white value setting in difference weather conditions is provided.

- Camera characterization method is presented. This method is used to transform *RGB* colour space, which is the output of camera response, to *XYZ* colour space, which is colour input of CIECAM97s.

- Test CIECAM97s efficiency in image segmentation under three weather conditions by comparing to the other two colour perceptual space CIELUV and *HSI*.

An accurate recognition method invariance with noise, scale, rotation, and distortion used in the field of traffic sign recognition. This method, by implementing the BMV model, has demonstrated successful traffic recognition.

7.3 Future work

The current work has paved a road to image segmentation and object recognition in real environments based on human visual perception. There is however still room to extend and improve this area.

Currently, segmentation based on CIECAM97s is only applied to 3 types of weather conditions: sunny, cloudy and rainy. However, there are more than 3 types of weather. It is more investigation to how to classify other weather conditions such as foggy day and snowing day.

Also, at present only day time situation is considered. However, in the evening colour appearance are affected by different illuminants. This brings another challenge for application of CIECAM97s in multi-illuminant situations in real environments.

The viewing condition/reference white, which we obtained, represents normal situations of different weather status not for special conditions such as multi-illuminant situations. It may lead to inaccurate perceptive colour calculation in some special conditions. How to select reference white/viewing condition parameters that can represent complex environments or special viewing conditions, automatically is a challenge.

The setting of camera will be taken into account in the future. Currently, weather conditions are considered firstly. Those pictures are discarded by human manually if the pictures can not reflect the weather condition, when the pictures are taken under unsuitable camera setting. However, the camera setting parameters affect the quality of images. It will be considered whether the camera setting changes create similar results as the weather changes.

Furthermore, the processing speed of segmentation is not fast and an execution time of around 7 to 10 seconds is achieved. The whole system processing speed for one image is around 11 seconds. According to the complicated model of

CIECAM97s, computation takes a lot of time. To increase the speed, a higher performance computer can be used or an integrated hardware system can be built.

Another important area of future work is to develop this approach into an application compatible with video sequences. Currently, this vision-based approach has succeeded in traffic sign recognition for static images. However video sequences taken from a moving car would reveal interesting topics.

In this dissertation the scope of the study was limited to the specific application of traffic sign recognition. However, it is believed that the methods developed in this research can be extended to other image segmentation and object recognition problems in complex real environments.

Bibliography

1. M. Lalonde and Y. Li, *"Road Sign Recognition"*, in *CRIM/IIT*. 1995.

2. Stitles and Wyszecki, *"Color Science: Concepts and Methods, Quantitative Data and Formulae"*. 2 ed. 2000: Wiley Classics Library.

3. T. Gevers, *"Color in Image Database"*. 2001, Faculty of Mathematics, Compuer Science, Physics and Astronomy, University of Amersterdam: Kruislaan, Amsterdam, Netherlands.

4. International Committee on Illumination/Commission Internationale de l'Eclairage, *"The CIE 1997 Interim Colour Appearance Model (Simple Version), CIECAM97s"*. Vol. CIE TC1-34. 1998.

5. I.A. Rybak, V. I. Gusakova, A.V. Golovan, L. N. Podladchikova, and N. A. Shevtsova, *"Behavioral Model of Visual Perception and Recognition"*. 1994-present, Rostov State University: Rostov-on-Don, Russia. http://www.rybak-et-al.net/vnc.html.

6. International Committee on Illumination/Commission Internationale de l'Eclairage, *"Road signs"*, C.B.o.T. CIE, Editor. 1988.

7. International Committee on Illumination/Commission Internationale de l'Eclairage, *"Recommendations for surface colours for visual signalling"*. CIE No. 39-2 (TC-1.6) ed. 1983.

8. R. C. Moeur, *"The Manual of Traffic Signs"*. 2003. http://members.aol.com/rcmoeur.

9. Driving Standards Agency, *"The Highway Code"*. 1999, London, England: The Stationery Office.

10. P. Zahradnik and M. Vlcek, *"Road Sign Recognition Survey"*. 1999http://euler.fd.cvut.cz/research/rs2/files/skoda-rs-survey.html.

11. K. R. Castleman, *"Digital Image Processing"*. 1995: Prentice Hall.

12. Quantitative Imaging Group, *"Digital Image Definitions"*. 2004, Department of Imaging Science and Technology: The Netherlands. http://www.ph.tn.tudelft.nl/Courses/FIP/frames/fip.html.

13. B. Zhao, *"Colour Space"*. 2002, Computer Vision Laboratory, Department of Electrical and Computer Engineering, Stony Brook University: New York.

14. R.W.G. Hunt, *"Measuring Colour"*. 2 ed. 1992: Ellis Horwood Limited.

15. B. Fraser, Chris Murphy, and Fred Bunting, *"Real World Color Management"*. 2003: Peachit Press.

16. A. Davis and P. Fennessy, *"Digial Imaging for Photographers"*. 3 ed. 1998: Focal Press.

17. S.J. Sangwine and R.E.N. Horne, *"The Colour Image Processing Handbook"*. 1998: Chapman&Hall.

18. Kodak, *"Digital Color"*. 2002-2003 http://www.kodak.com/US/en/digital/dlc/book3/chapter2/index.shtml.

19. C. R. (Rod) Nave, *"HyperPhysics"*. 2000, Department of Physics and Astronomy, Georgia State University: Atlanta, Georgia.

20. Technical Advisory Service for Images, *"Glossary of Technical Terms"*. T.J.I.S.C. (JISC), Editor. 2002, Institute for Learning and Research Technology, University of Bristol: Bristol. http://www.tasi.ac.uk/glossary/glossary_technical.html.

21. A.N. Venetsanopoulos and K.N. Plataniotis, *"Colour Image Processing and Applications"*. 1 ed. 2000: Springer-Verlag.

22. A. Miller, *"Colour Space Conversion"*. 2001, Xilinx UK: UK. http://www.xilinx.com/xlnx/xweb/xil_tx_display.jsp?sTechX_ID=am_colour&languageID=1.

23. D. Bourgin, *"Color spaces FAQ"*. 1995. http://www.rmbwoc.com/vidpage/color_faq.html.

24. J. Bradley Davis, *"Computer Graphics Color Space"*. http://www.cs.fit.edu/wds/classes/cse5255/cse5255/davis/index.html.

25. A. Ford and A. Roberts, *"Colour Space Conversions"*. 1998, University of Westminster and British, Broadcasting Corporation: London, UK. http://www.biology.duke.edu/johnsenlab/pdfs/tech/colorconversion.pdf.

26. Apple Computer Inc., *"Color Spaces"*. 1999. http://developer.apple.com/documentation/QuickTime/REF/refVectors.21.htm#pgfId=20708).

27. G. Beretta, *"Understanding Color"*. 2000, The Society for Imaging Science and Technology and The International Society for Optical Engineering: San Jose, CA.

28. C. Poynton, *"Color FAQ - Frequently Asked Questions Color"*. 2002 http://www.poynton.com/notes/colour_and_gamma/ColorFAQ.html.

29. S. Westland, *"Frequently asked questions about Colour Physics"*. 2000 http://www.colourware.co.uk/cpfaq.htm.

30. D. Neagu, *"Information on the CIE LUV color space"*. 2002, Department of computing, University of Bradford. http://airlab.elet.polimi.it/imagetox/dneagu/Courses/MSCG/applets/colormodels/infoluv.html.

31. D. Neagu, *"Information on the YIQ color space"*. 2002http://www.nacs.uci.edu/~wiedeman/cspace/me/infoyiq.html.

32. Corel, *"Paintshop"*. 2001.

33. Image Lib, *"UK Road Traffic Signs"*. http://www.bris.ac.uk/imagelib/ts.html. 2000.

34. V. Hlavac, R. Boyle and M. Sonka, *"Image Processing Analysis and Machine Vision"*. 2 ed. 1999: PWS Publishing.

35. D. H. Ballard, *"Generalizing the Hough transform to detect arbitrary shapes"*. Pattern Recognition and Image Analysis, 1981. Vol. 13 (2): p. 111-122.

36. T. Kaneko and M. Okudaira, *"Encoding of arbitrary curves based on the chain code representation"*. IEEE Transaction on Communications, 1985. Vol. COM-33 (7): p. 697-707.

37. H. Freeman, *"Shape description via the use of critical points"*. Pattern Recognition and Image Analysis, 1978. Vol. 10: p. 159-166.

38. M. Jiang, *"Digital Image Processing"*. 2003, Department of Information Science , Peking University: Beijing, China.

39. H. Chang and U. Robles, *"Face Detection"*. 2000, Computer Science: Los Angles, USA. http://www-cs-students.stanford.edu/~robles/ee368/main.html.

40. L. C. Mai, *"Introductin to Computer Vision and Image Processing"*. 1999-2000http://www.netnam.vn/unescocourse/computervision/861.htm.

41. G Piccioli, E.D Michelli, and M Campani, *"A robust method for road sign detection and recognition"*. Image and Vision Computing Journal, 1996. Vol. 14: p. 209-223.

42. P. Douville, *"Real-Time Classification of Traffic Signs"*. Real-Time Imaging, 2000. Vol. 6: p. 185-193.

43. M. Shneier. *"Road Sign Detection and Recognition"*. in IEEE Computer Society International Conference on Computer Vision and Pattern Recognition. 2005.

44. C. Bahlmann, Y. Zhu, V. Ramesh, M. Pellkofer, and T. Koehler. *"A System for Traffic Sign Detection, Tracking, and Recognition Using Colour, Shape, and Motion Information"*. in IEEE Intelligent Vehicles Symposium (IV 2005),. 2005. Las Vegas, USA.

45. A. de la Escalera, J.Mª Armingol, and M.A. Salichs. *"Traffic Sign Detection for Driver Support Systems"*. in International Conference on Field and Service Robotics. 2001. Espoo, Finlandia.

46. A. de la Escalera, J. Mª Armingol, J. M. Pastor, and F. J. Rodriguez, *"Visual Sign Information Extraction and Identification by Deformable Models for Intelligent Vehicles"*. IEEE Transactions on Intelligent Transportation Systems, 2004. Vol. 15 (2): p. 57-68.

47. U. Büker, P. Haasper, and B. Mertsching. *"Efficient Shape Recognition of Hierachically Encoded Structures by Moments from Boundaries"*. in Shape, Structure and Pattern Recognition. 1994 p. 109-117. Singapore.

48. D.M. Gavrila. *"Traffic Sign Recognition Revisited"*. in Procceediong of the 21st DAGM Symposium fur Musterekennung. 1999 p. 86-93.

49. D.M. Gavrila and V. Philomin. *"Real-Time Object Detection For "Smart" Vehicles"*. in International Conference on Computer Vision. 1999 p. 20-25. Corfu, Greece.

50. C.Y. Fang, S.W. Chen, and Chiou-Shann Fuh, *"Road-Sign Detection and Tracking"*. IEEE Transactions on Vehicular Technology,, 2003. Vol. 52 (5): p. 1329-1341.

51. N. Kehtarnavaz, N.C. Griswold, and D.S. Kang, *"Stop-sign recognition based on colour-shape processing"*. Machine Vision and Applications, 1993. Vol. 6: p. 206-208.

52. D.Kellmeyer and H.Zwahlen. *"Detection of highway warning signs in natural video images using color image processing and neural networks"*. in IEE Proceedding of International Conference Neural Networks. 1994. Vol. 7 p. 4226-4231.

53. G. Nicchiotti, E. Ottaviani, P. Castello, and G. Piccioli. *"Automatic road sign detection and classification from color image sequences"*. in Procceeding of 7th International Conference On Image Analysis and Processing. 1994 p. 623-626.

54. P. Paclík, J. Novovicová, P. Pudil, and P. Somol, *"Road Sign Classification using the Laplace Kernel Classifier"*. Pattern Recognition Letters, 2000. Vol. 21 (13-14): p. 1165-1173.

55. M.M. Zadeh, T.Kasvand, and C.Y.Suen. *"Localization and Recognition of Traffic Signs for Automated Vehicle Control Systems"*. in Conference on Intelligent Transportation Systems, part of SPIE's Intelligent Systems & Automated Manufacturing. 1997. Pittsburgh, PA, USA.

56. L. Priese, J. Klieber, R. Lakmann, V. Rehrmann, and R. Schian. *"New results on traffic sign recognition"*. in IEEE Proc. Intelligent Vehichles'94 Symposium. 1994 p. 249-253.

57. V. Rehrmann L. Priese, R. Schian and R. Lakmann. *"Traffic Sign Recognition Based on Colour Image Evaluation"*. in Procceeding of Intelligent Vehicle Symposium. 1993 p. 390-395. Tokyo, Japan.

58. R. Lakmann, L. Priese and V. Rehrmann. *"A parallel system for real-time traffic sign recognition"*. in International Workshop on Advanced Parallel Processing Technologies. 1995. Beijing, China.

59. L. Priese and V. Rehrmann. *"A Fast Hybrid Colour Segmentation Method"*. in DAGM Symposium Symposium for Pattern Recognition. 1993 p. 297-304. Lübeck, Germany.

60. S. Prince, *"Road Sign Detection and Recognition Using Perceptual Grouping"*. 1998, Compute Vision and Systems Laboratory, University of Laval: Laval, Canada.

61. P. Haasper, B. Mertsching, and U. Büker. *"Efficient Shape Recognition of Hierachically Encoded Structures by Moments from Boundaries"*. in Shape, Structure and Pattern Recognition. 1994 p. 109-117. Singapore.

62. D.M.Gavrila and V.Philomin. *"Real-Time Object Detection For "Smart" Vehicles"*. in International Conference on Computer Vision. 1999 p. 20-25. Corfu, Greece.

63. X.. Chen, M. Kusanagi and S. Puntavungkour. *"Automatic recognition and location of road signs from terrestrial color image"*. in Geoinformatics & Dmgis'2001. 2001. Bangkok, Thailand.

64. T. Gevers and A. W. M. Smeulders. *"An Approach to Image Retrieval for Image Databases"*. in Database and Expert Systems Applications (DEXA). 1993 p. 615-626. Prague, Czech Republic.

65. S.W. Chen, C.Y. Fang, and C.S. Fuh, *"Road-Sign Detection and Tracking"*. IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY,, 2003. Vol. 52 (5): p. 1329-1341.

66. M.R. Luo and R.W.G. Hunt, *"The Structure of the CIE 1997 Colour Appearance Model (CIECAM97s)"*. Color Research and application, 1998. Vol. 23: p. 138-146.

67. I.A. Rybak, V.I. Gusakova, A.V. Golovan, L.N. Podladchikova, and N.A. Shevtsova, *"A model of attention-guided visual perception and recognition"*. Vision Research, 1998. Vol. 38: p. 2387~2400.

68. F. Ziliani, *"Spatio-Temporal Image Segmentation: A new Rule-Based Approach"*, in *The Signal Processing Laboratory*. 2000, Swiss Federal Institute of Technology (EPFL).

69. G. Lacey, F. Shevlin, and N. Winters, *"Computer Vision and Robotics"*. 1998, Computer Science Department, University of Dublin, Trinity College: Dublin, Ireland.

70. R.N. Haber and M. Hershenson, *"The psychology of visual perception"*. 1973: Holt, Reinhart and Winston.

71. B. Girod, *"Image and Video Compression"*. 2000, Electrical Engineering and Computer Science, Stanford University: Stanford, CA, USA.

72. J. C. Sabataitis and J. M. DiCarlo, *"CIECAM97 Color Appearance Model"*. 1998, Department of Electrical Engineering. Stanford University.

73. N. Moroney. *"Usage guidelines for CIECAM97s"*. in Image Processing, Image Quality, Image Capture, Systems Conference, (PICS 2000). 2000 p. 164-168. Portland, OR, USA.

74. M. D. Fairchild, *"CIE TC1-34 Testing Colour Appearance Models"*. 1997.

75. M.D. Fairchild, *"A revision of CIECAM97s for practical applications"*. Color Research and Application, 2001. Vol. 26: p. 418-427.

76. M. Jin-Seo and K., Cho, *"Applying CIECAM97s in Color Management System"*. Proceedings of the IEEE Region 10 Conference (TENCON 99), 1999. Vol. 2: p. 1367-1370.

77. J. S. Laine. *"Experiments on Adaptive Soft-Copy Color Reproduction"*. in Proceedings of the IS&T/SID 10th Color Imaging Conference. 2002 p. p. 190-195. Scottsdale, Arizona.

78. D. McG. Squire, *"Model-based Neural Networks for Invariant Pattern Recognition"*. 1996, Curtin University of Technology: Perth, Western Australia.

79. S. Gil R. Milanese, and T. Pun, *"Attentive mechanisms for dynamic and static scene analysis"*. Optical Engineering, 1995. Vol. 34 (8): p. 2428-2434.

80. L. Wiskott, *"How Does Our Visual System Achieve Shift and Size Invariance?"* Problems in Systems Neuroscience, 2004.

81. J. Gibson, *"The Senses Considered as Perceptual Systems"*. 1966, Boston, Massachusetts, USA: Greenwood Publishing Group.

82. I. Biederman and E. E. Cooper, *"Size invariance in visual object priming"*. Experimental Psychology: Human Perception and Performance, 1992. Vol. 18 (1): p. 121-133.

83. I. Biederman and E. E. Cooper, *"Evidence for complete translational and reflectional invariance in visual object priming"*. Perception. 1991. Vol. 20: p. 585-593.

84. C. S. Furmanski and S. A. Engel, *"Perceptual learning in object recognition: Object specificity and size invariance"*. Vision Research, 2000. Vol. 40 (5): p. 473-484.

85. M. J. Tov'ee, E. T. Rolls, and P. Azzopardi, *"Translation invariance in the responses to faces of single neurons in the temporal visual cortical areas of the alert macaque"*. Journal of Neurophysiology, 1994. Vol. 72 (3): p. 1049-1060.

86. R. Ihaka, *"Information Visualisation"*. 2004, Department of Statistics: New Zealand. http://www.stat.auckland.ac.nz/~ihaka/120/notes.html.

87. A. L. Yarbus, *"Eye Movements and Vision"*. 1967, New York: Plenum Press.

88. L Stark D Noton, *"Scanpaths in eye movements during pattern recognition"*. Science, 1971. Vol. 171 (72-75).

89. M Mishkin, and LG Ungerleider *"Two cortical visual systems"*. in Analysis of Visual Behavior. 1982 p. 549-586: MIT Press.

90. LG Ungerleider, M Mishkin, and KA Mackn, *"Object vision and spatial vision: two cortical pathways"*. Trends Neuroscience, 1983. Vol. 6: p. 414-417.

91. D Essen Van. *"Functional organization of primate visual cortex"*. 1985. Vol. 3 p. 259-329.

92. RA Flynn SM kosslyn, JB Amesterdam and G Wang, *"Components of high-level vision: a cognitive neuroscience analysis and account of neurological syndomes"*. Cognition, 1990. Vol. 34: p. 203-277.

93. B.A. Skinner, *"Science and Human Behavior"*. 1953, New York: Macmillan.

94. MA. Arbib and RL. Didday, *"Eye movements and visual perception: a two visual system model"*. Int. J. Man-Machine Study, 1975. Vol. 7: p. 547-569.

95. A.V. Golovan, N.A. Shevtsova, L.N. Podladchikova, S.N. Markin, and D.G. Shaposhnikov, *"Image Preprocessing for Identifying the Most Informative Regions in Facial Images"*. Pattern Recognition and Image Analysis, 2001. Vol. 11 (2): p. 313-316.

96. Mathworks, *"Matlab Guide"*. 2002.

97. European Broadcasting Union (EBU), *"Standard for Chromaticity Torerances for Studio Monitors"*. 1975.

98. P. Scheunders, S. Livens, G. Van de Wouwer, P. Vautrot, and D. Van Dyck, *"Wavelet-based texture analysis"*. Computer Science and Information Management, 1998. Vol. 1: p. 22-34.

99. .M.F.Clements, L.E., and K.A Shaw, *"Performance evaluation of texture measures for ground cover identification in satellite images by means of a neural network classifier"*. IEEE trans. Geoscience and Remote Sensing,, 1995. Vol. 33 (3): p. 616-626.

100. Y. Rubner and C. Tomasi. *"Coalescing Texture Descriptors"*. in Proceedings of the ARPA Image Understanding Workshop. 1996 p. 927-935. Palm Springs, CA.

101. T.L.V. Cheung and S.Westland. *"Colour camera characterization using artificial neural networks"*. in IS&T/SID Tenth Colour Imaging Conference. 2002 p. 117-120. Scottsdale, USA.

102. G.W. Hong, M.Ronnier Luo, and P. A.Rhodes, *"A study of digital camera colorimetric characterization based on polynomial modeling"*. Color Research and Application, 2001. Vol. 26: p. 76.

103. W. Mendenhall and L. Ott, *"Undersdanting Statistics"*. 5 ed. 1990: PWS-KENT Publishing Company.

104. X.W. Gao, N. Shevtsova, K. Hong, S. Batty, Podladchikova L., A. Golovan, Shaposhnikov D., and V Gusakova. *"A New Approach to Traffic Sign Recognition"*. in The 2002 International Conference on Imaging Science, Systems, and Technology (CISST'02). 2002. Las Vegas, USA.

105. X.W. Gao, Golovan A., K. Hong, Podladchikova L.N., Shaposhnikov D.G., and N Shevtsova. *"Road Sign Recognition by Means of the Behavioral Model of Vision"*. in Third All-Russian Conference on Neuroinformatics,. 2002. Moskow, Russia.

106. D.G. Shaposhnikov, L.N. Podladchikova, K. Hong, X.W. Gao, A.V. Golovan, and N.A. Shevtsova. *"Road Sign Recognition by Single Positioning of Space-Variant Sensor Window"*. in 15th International Conference on Vision Interface (VI 2002). 2002. Calgary, Canada.

107. *"Russian traffic signs"*. http://www.domkrat.ru/Laws/rules/znkl.shtml.

108. Pratt W. Adbdou I E, *"Quantitive design and evaluaion of enhancement/thresholding edge detectors"*. Proceeding of IEEE, 1979. Vol. 67: p. 753-763.

109. D. G. Shaposhnikov, K.B. Hong, and L. N. Podladchikova. *"Road Sign Recognition by single Positioning of Space-Variant Sensor Window"*. in 15th International Conference on Vision Interface. 2002 p. 213-217. Calgary, Canada: Canadian Imae processing and Pattern Recognition Society.

110. X.W. Gao, N. Shevtsova, K. Hong, S. Batty, Podladchikova L., A. Golovan, Shaposhnikov D., and V Gusakova. *"Vision Models-based Identification of Traffic Signs"*. in First European Conference on Color in Graphics, Image and Vision. 2002. Poitiers, France.

Appendix 1: Relative Formula of Two Human Vision Models

CIECAM97s Model

Chapter 3 details the knowledge of colour appearance model (CIECAM97s). In this appendix, how to obtain 6 colours perceptual attributes: hue, chroma, lightness, brightness, saturation and colourfulness according to the input data of tristimulus values and viewing conditions, will be described.

According to [4, 66], the input data includes

- XYZ: Relative tristimulus values of colour stimulus
- $X_W Y_W Z_W$: Relative tristimulus values of white
- La: Luminance of the adapting field (cd/m*m)
- Y_b: Relative luminance of the background
- Surround parameters: c, Nc, F_{LL}, F

And the surround parameters are list in Table 3-1, which are re-listed in the table below:

Viewing condition	c	Nc	F_{LL}	F	Examples
Average Surround, samples subtending $>4^\circ$	0.69	1.0	0.0	1.0	Viewing surface colours
Average Surround	0.69	1.0	1.0	1.0	
Dim Surround	0.59	1.1	1.0	0.9	Viewing television
Dark Surround	0.525	0.8	1.0	0.9	Viewing film projected in a dark room
Cut-sheet Transparencies (on a viewing box)	0.41	0.8	1.0	0.9	Viewing cut-sheet films in light boxes

To obtain the colour perceptual values, two steps are adapted. Firstly, chromatic adaptation transformation is calculated. Secondly, the colour appearance correlates are obtained.

Chromatic Adaptation

An initial chromatic adaptation transform is used to go from the source viewing conditions to corresponding colours under the equal-energy illuminant reference viewing conditions. First, tristimulus values for both the sample and white are normalized and transformed to spectrally sharpened cone responses, given in Equation A-1 and A-2. The forward matrix transformation given in A-2 was applied to the spectral tristimulus values of the CIE 1931, which is described in chapter 2. In the formula below RGB represents three cone responses. XYZ is the tristimulus value.

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = M_B \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \tag{A-1}$$

$$M_B = \begin{bmatrix} 0.8951 & 0.2664 & -0.1614 \\ -0.7502 & 1.7135 & 0.0367 \\ 0.0389 & -0.0685 & 1.0296 \end{bmatrix} \quad M_B^{-1} = \begin{bmatrix} 0.9870 & -0.1471 & 0.1600 \\ 0.4323 & 0.5184 & 0.0493 \\ -0.0085 & 0.0400 & 0.9685 \end{bmatrix} \tag{A-2}$$

The chromatic-adaptation transform is a modified von Kries-type transformation [14, 75] with an exponential nonlinearity on the short-wavelength sensitive channel as given in Equation A-3 through A-6. In addition, the variable D is used to specify the degree of adaptation. D is set to 1.0 for complete adaptation or discounting the illuminant (as is typically the case for reflecting materials). D is set to 0.0 for no adaptation. D takes on intermediate values for various degrees of incomplete chromatic adaptation. Equation A-7 allows calculation of such intermediate D values for various luminance levels and surround conditions.

$$R_C = [D(1.0/R_W) + 1 - D]R \tag{A-3}$$

$$G_C = [D(1.0/G_W) + 1 - D]G \tag{A-4}$$

$$B_C = [D(1.0/B_W^P) + 1 - D]|B|^P \tag{A-5}$$

$$P = (B_W/1.0)^{0.0834} \tag{A-6}$$

$$D = F - F/[1 + 2(L_A^{1/4}) + (L_A^2)/300] \tag{A-7}$$

If B happens to be negative, then B_c is also set to be negative. Similar transformations are also made for the source white since they are required in later calculations. Various factors must be calculated prior to further calculations as shown in Equation A-8 through A-12. These include a background induction factor, n, the background and chromatic brightness induction factors, N_{bb} and N_{cb}, and the base exponential nonlinearity, z.

$$k = 1/(5L_A + 1) \tag{A-8}$$

$$F_L = 0.2k^4(5L_A) + 0.1(1-k^4)^2(5L_A)^{1/3} \tag{A-9}$$

$$n = Y_b / Y_W \tag{A-10}$$

$$N_{bb} = N_{cb} = 0.725(1/n)^{0.2} \tag{A-11}$$

$$z = 1 + F_{LL} n^{1/2} \tag{A-12}$$

The post-adaptation signals for both the sample and the source white are then transformed from the sharpened cone responses to the Hunt-Pointer-Estevez cone responses as shown in Equation A-13 and A-14.

$$\begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} = M_H M_B^{-1} \begin{bmatrix} R_C Y \\ G_C Y \\ B_C Y \end{bmatrix} \tag{A-13}$$

$$M_H = \begin{bmatrix} 0.38971 & 0.68898 & -0.07868 \\ -0.22981 & 1.18340 & 0.04641 \\ 0.00 & 0.00 & 1.00 \end{bmatrix} \quad M_H^{-1} = \begin{bmatrix} 1.9102 & -1.1121 & 0.2019 \\ 0.3710 & 0.6291 & 0.00 \\ 0.00 & 0.00 & 1.00 \end{bmatrix} \tag{A-14}$$

The post-adaptation cone responses (for both the sample and the white) are then calculated using Equation A-15 through A-17.

$$R'_a = \frac{40(F_L R'/100)^{0.73}}{\left[(F_L R'/100)^{0.73} + 2\right]} + 1 \tag{A-15}$$

$$G'_a = \frac{40(F_L G'/100)^{0.73}}{\left[(F_L G'/100)^{0.73} + 2\right]} + 1 \tag{A-16}$$

$$B_a' = \frac{40(F_L B'/100)^{0.73}}{\left[(F_L B'/100)^{0.73} + 2\right]} + 1 \qquad \text{(A-17)}$$

Appearance Correlates

Preliminary red-green and yellow-blue opponent dimensions are calculated using Equation A-18 and A-19.

$$a = R_a' - 12G_a'/11 + B_a'/11 \qquad \text{(A-18)}$$

$$b = (1/9)(R_a' + G_a' - 2B_a') \qquad \text{(A-19)}$$

The CIECAM97s hue angle, h, is then calculated from a and b using Equation A-20.

$$h = \tan^{-1}(b/a) \qquad \text{(A-20)}$$

Hue quadrature, H, and eccentricity factors, e, are calculated from the following unique hue data via linear interpolation between the following values for the unique hues:

Red: $h = 20.14$, $e = 0.8$, H = 0 or 400,
Yellow: $h = 90.00$, $e = 0.7$, H = 100,
Green: $h = 164.25$, $e = 1.0$, H = 200,
Blue: $h = 237.53$, $e = 1.2$. H = 300

Equations A-21 and A-22 illustrate calculation of e and H for arbitrary hue angles where the quantities subscripted 1 and 2 refer to the unique hues with hue angles just below and just above the hue angle of interest.

$$e = e_1 + (e_2 - e_1)(h - h_1)/(h_2 - h_1) \qquad \text{(A-21)}$$

$$H = H_1 + \frac{100(h - h_1)/e_1}{(h - h_1)/e_1 + (h_2 - h)/e_2} \qquad \text{(A-22)}$$

The achromatic response is calculated as shown in Equation A-23 for both the sample and the white.

$$A = \left[2R_a^{'} + G_a^{'} + (1/20)B_a^{'} - 2.05\right]N_{bb} \qquad \text{(A-23)}$$

CIECAM97s Lightness, J, is calculated from the achromatic signals of the sample, A, and white, Aw, using Equation A-24.

$$J = 100(A/A_W)^{cz} \qquad \text{(A-24)}$$

CIECAM97s Brightness, Q, is calculated from CIECAM97s lightness and the achromatic response for the white using Equation A-25.

$$Q = (1.24/c)(J/100)^{0.67}(A_W + 3)^{0.9} \qquad \text{(A-25)}$$

Finally, CIECAM97s saturation, s; CIECAM97s chroma, C; and CIECAM97s colourfulness, M; are calculated using Equations A-26 through A-28, respectively.

$$s = \frac{50(a^2 + b^2)^{1/2}100e(10/13)N_c N_{cb}}{R_a^{'} + G_a^{'} + (21/20)B_a^{'}} \qquad \text{(A-26)}$$

$$C = 2.44s^{0.69}(J/100)^{0.67n}(1.64 - 0.29^n) \qquad \text{(A-27)}$$

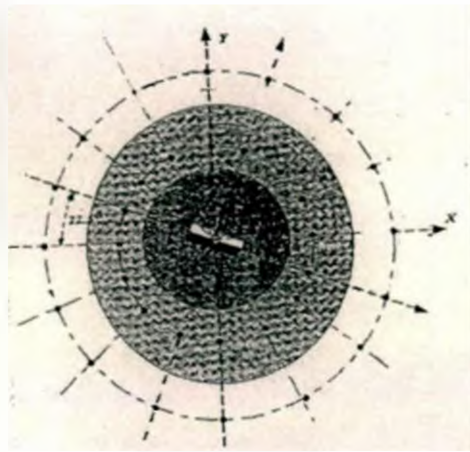
$$M = CF_L^{0.15} \qquad \text{(A-28)}$$

Behaviour model of vision (BMV)

BMV model provides the ability to recognize complex images invariantly with respect to shift, rotation and scale. The main three procedures: primary transform, detecting primary features and invariant representation are described here. In [5, 67], Rybak presents the detail of the model.

Primary transformation: formation of the retinal image within the AW

As described in chapter 3, the schematic of AW is re-drawn in here.



Schematic of the AW

The retinal image results from the initial image $I = \{x_{ij}\}$ by way of a special transformation used to obtain a decrease in resolution from the AW centre to its periphery. To represent a part D of the image $((i, j) \in D)$ at resolution level $l (l \in \{1, 2, 3, 4, 5\})$, the recursive computation of the Gaussian-like convolution at cach point of D.

$$x_{ij}^{(1)} = x_{ij}$$

$$x_{ij}^{(2)} = \sum_{p=-2}^{p=2} \sum_{q=-2}^{q=2} g_{pd} \cdot x_{i-p,j-q}^{(1)}$$

(A-29)

$$x_{ij}^{l} = \sum\sum g_{pq} \cdot x^{(l-1)}_{[illegible]}$$

where the coefficients of convolution belong to the following matrix:

$$[g_{pq}] = \begin{vmatrix} 1 & 4 & 6 & 4 & 1 \\ 4 & 16 & 24 & 16 & 4 \\ 6 & 24 & 36 & 24 & 6 \\ 4 & 16 & 24 & 16 & 4 \\ 1 & 4 & 6 & 4 & 1 \end{vmatrix} \cdot \frac{1}{256} \qquad \text{(A-30)}$$

(p and q = -2, -1, 0, 1, 2)

In the model, the primary image transformation maps the initial image $I = \{x_{ij}\}$ into the retinal image $I^R(n) = \{x^R_{ij}(n)\}$ at each *n*th fixation point. The position of the fixation point $(i_o(n), j_o(n))$ and the resolution level $l_o(n)$ in the vicinity of that point are considered to be parameters of the retinal image. The central point $(i_o(n), j_o(n))$ is surrounded by three concentric circles whose radii are functions of $l_o(n)$:

$$
\begin{aligned}
R_o(l_o) &= 1.5 \cdot 2^{l_o} \\
R_1(l_o) &= 1.5 \cdot 2^{l_o+1} \\
R_o(l_o) &= 1.5 \cdot 2^{l_o+2}
\end{aligned}
\qquad \text{(A-31)}
$$

The retinal image at the *n*th fixation point $I^R(n) = \{x^R_{ij}(n)\}$ is formed from $I = \{x_{ij}\}$ as follows:

$$
x^R_{ij}(n) = \left| \begin{array}{lll} x_{ij}^{l_o(n)}, & \mathit{if} & p_{ij}(n) \le R_o(l_o) \\ x_{ij}^{l_o(n)+1}, & \mathit{if} & R_o(l_o) < p_{ij}(n) \le R_1(l_o) \\ x_{ij}^{l_o(n)+2}, & \mathit{if} & R_1(l_o) < p_{ij}(n) \le R_2(l_o) \end{array} \right| \qquad \text{(A-32)}
$$

where

$$
p_{ij}(n) = \sqrt{(i - i_o(n))^2 + (j - j_o(n))^2} \qquad \text{(A-33)}
$$

Therefore, the initial image is represented in the AW: with the highest resolution $l = l_o(n)$ within the central circle ('fovea'), with lower resolution $l = l_o(n) + 1$ within the first ring surrounding the central circle, and with the lowest resolution $l = l_o(n) + 2$ within the second ring.

Detecting primary features

The module for primary feature detection performs a function similar to the function of primary visual cortex containing orientationally selective neurons. In the model, edges are detected by a network of orientationally selective neurons ad are considered to be the primary features of the image. Each edge is detected with resolution dependent on the position of the edge in the retinal image.

The orientationally selective receptive field (ORF) of the neuron with coordinates (i, j) turned to the orientation α (α=0, 1, 2, ..., 15). The discrete angle step of 22.5° is considered as a unit in all angle measurements. The ORF is described as a difference between Gaussian convolutions with spatially shifted centres. The input signal to the neuron tuned to the orientation α is:

$$Y_{ij\alpha} = \sum_{pq} x^R_{pq} \cdot (G^+_{pqij\alpha} - G^-_{pqij\alpha}), \tag{A-34}$$

where

$$\begin{aligned} G^+_{pqij\alpha} &= \exp\left(-\gamma^2 \cdot ((p - i - m_\alpha)^2 + (q - j - n_\alpha)^2)\right); \\ G^-_{pqij\alpha} &= \exp\left(-\gamma^2 \cdot ((p - i + m_\alpha)^2 + (q - j + n_\alpha)^2)\right) \end{aligned} \tag{A-35}$$

In (A-35), γ is a reciprocal variance. The parameters m_α and n_α depend on the ORF orientation α:

$$\begin{aligned} m_\alpha &= d(l) \cdot \cos(2 \cdot \pi \cdot \alpha / 16) \\ n_\alpha &= d(l) \cdot \sin(2 \cdot \pi \cdot \alpha / 16) \end{aligned} \tag{A-36}$$

where $d(l)$ defines the distance between the centre of each Gaussian and the centre of the ORF and depends on the resolution level l in a given area of the retinal images:

$$d(l) = \max\{2^{l-2}, 1\} \tag{A-37}$$

Sixteen neutrons, whose ORF have the same location but different orientations, interact competitively due to strong reciprocal inhibitory connections:

$$\begin{aligned} \tau \cdot \frac{d}{dt} V_{ij\alpha} &= -V_{ij\alpha} + Y_{ij\alpha} - b \cdot \sum_{\substack{k=0 \\ k \neq \alpha}}^{15} Z_{ijk} - T \\ Z_{ij\alpha} &= f(V_{ij\alpha}); \qquad \alpha = 0,1,2,\ldots,15 \end{aligned} \tag{A-38}$$

where $V_{ij\alpha}$ and $Z_{ij\alpha}$ are the membrane potential and output of the neuron (i, j) with the ORF tuned to the orientation α, respectively; b is the coefficient characterizing the reciprocal inhibition (b>1); T is the neuron threshold; τ is the time constant; $f(V) = V$ if V>0, otherwise $f(V) = 0$.

At each nth fixation (AW position) oriented edges are detected: at the fixation point $((i_o(n), j_o(n))$ (the 'basic edge' in the centre of the AW) and at 48 'context' points which are located at the intersections of sixteen radiating lines (with the angle step of 22.5°) and three concentric circles, each in a different resolution are. The radii of these circles (R_{oo}, R_{o1} and R_{o2}) exponentially increase:

$$
\begin{aligned}
R_{oo}(l_o) &= 2^{l_o} \\
R_{o1}(l_o) &= 2^{l_o+1} \\
R_{o2}(l_o) &= 2^{l_o+2}
\end{aligned}
\qquad \text{(A-39)}
$$

Invariant representation and comparison of the retinal images

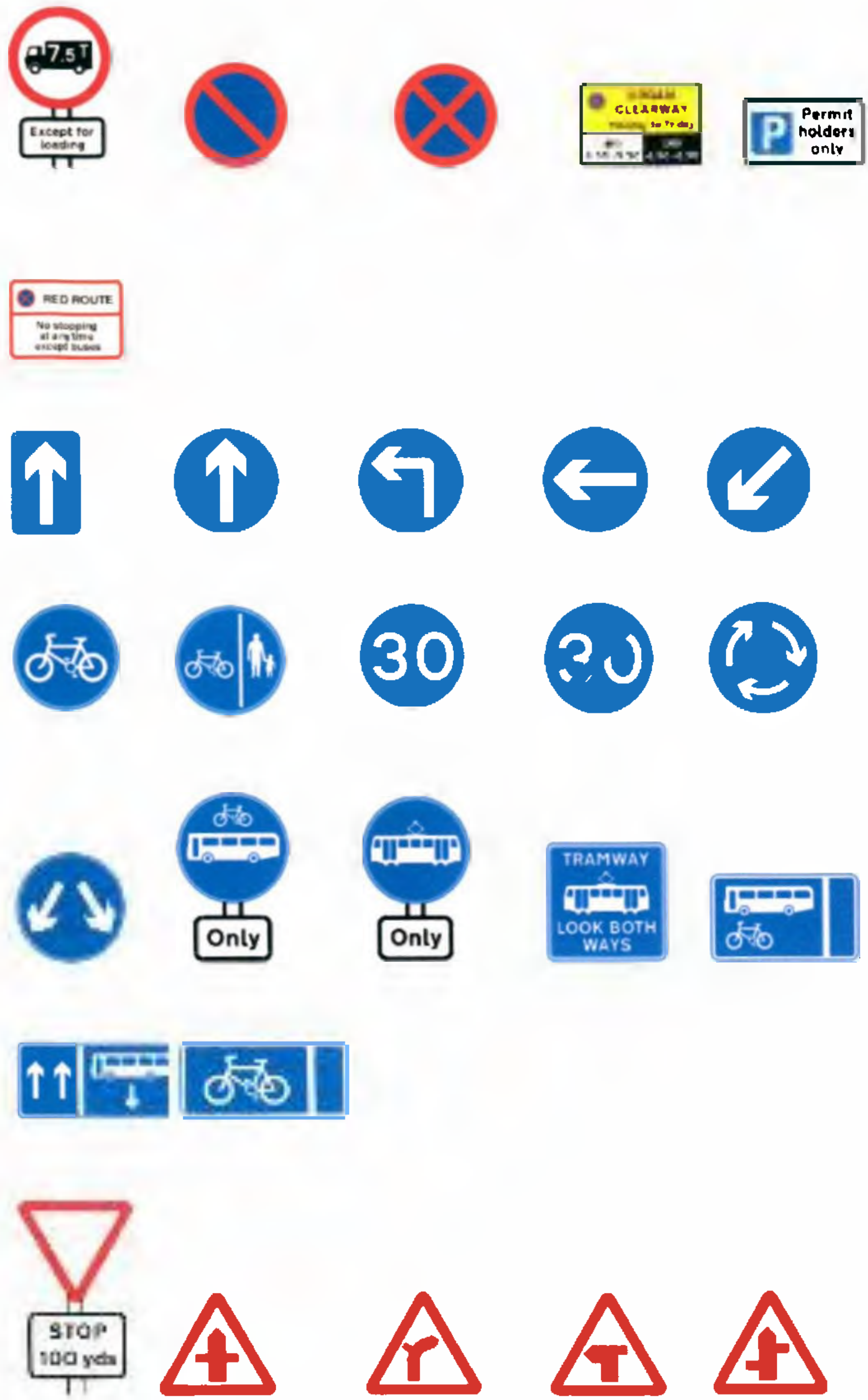
To describe the invariant representation, a coordinate system XOY is used the centre of XOY is the AW centre. The basic edge in the AW centre may be represented by the pair of parameters $((\varphi_o, l_o)$, where: $\varphi_o (\varphi_o \in \{0,1,...15\})$ is the orientation of the basic edge (represented by the angle between the axis $\vec{OX}$ and the vector of brightness gradient of the basic edge) and $l_o (l_o \in \{1,2,3\})$ is the level of resolution in the central area of the AW. Each context edge can be represented in the absolute coordinate system by the three parameters $((\varphi_c, \psi_c, l)$ where: φ_c is the orientation of the context edge; ψ_c characterizes the angular location of the edge in *XOY;* l is the level of resolution in the area of the edge; φ_c and $\psi_c \in \{0,1,...15\}$. Therefore, one of the context edge (φ_c, ψ_c, l) may be represented with respect to the relative coordinate system by the parameters (φ, ψ, λ) where: φ is the relative orientation of the context edge; ψ is its relative angular location; λ characterizes the relative distance from the AW centre. Those parameters are calculated as follows:

$$\begin{aligned}
&\varphi = \mathrm{mod}_{16}(\varphi_c - \varphi_o + 16) \\
&\varphi = \mathrm{mod}_{16}(\psi_c - \varphi_o + 20) \\
&\lambda = l - l_o \\
&\varphi, \psi \in \{0,1,...15\} \\
&\lambda \in \{0,1,2\}
\end{aligned} \tag{A-40}$$

Thus, the retinal image within the AW at the *n*th fixation point can be invariantly represented by three arrays of pairs of numbers

$$\{\varphi_k(n), \psi_k(n)\}_\lambda; k = 0,1,...15; \lambda = 0,1,2 \tag{A-41}$$

Appendix 2: United Kingdom Traffic signs


7.5T
Except for loading
CLEARWAY
P
Permit holders only
RED ROUTE
No stopping at any time except buses
30
Only
Only
TRAMWAY
LOOK BOTH WAYS
STOP
100 yds


GIVE WAY
50 yds
REDUCE
SPEED
NOW
10%
20%
Patrol
No footway
for 400 yds
Queues
likely
Ford
Ice
STOP
when
lights show


14'6"
4.4 m
Safe height
16'6"
Humps for
½ mile
Hidden dip
Soft verges
for 2 miles


Nottingham
23 M 1


Nottingham
A 52
25
½ m


M 1
The NORTH
Sheffield 32
Leeds 59


A 404 Marlow
Birmingham, Oxford M 40


A 46 (M 69) Leicester, Coventry (E)
2 ½ m
The NORTH WEST, Birmingham, Coventry (N) M 6


TURPIN'S CROSSROADS
Lampton (M 11)
Dorfield A 123
Axtley B 1991
Steam railway
PARK STREET ROUNDABOUT
Birmingham
Bourne
M 15 (M 1)
(M 14)
Penderton
Walsham A 1183
Nutfield A 1183
The SOUTH
Nottingham 17
Leicester 32
(M 1 South) 35
Lampton
Axtley
1 mile
Swansea
Abertawe
HANGMAN'S CROSSROADS
(M 11)
Towrley
(A1(M)) 8
Barnes 10
Mackstone 2½
Elkington 1
A 404 (A 41)
Millington Green 3
(A 4011)
Market Walborough 7
B486
WC
150 yds
Wrest Park
300 yds
(A 33) (M 1)
Zoo
4
Public library
Council offices
Northtown
Marton 3
P
Saturday only
HR
M 62
Priority over oncoming vehicles
H
A & E
not 24 hrs

P
Low bridge
2 miles
ahead
4.4 m
14'6"
GOOD FOOD
Puddleworth
services
½ m
Petrol
65p
Tourist
information
Controlled
ZONE
Mon - Fri
8.30 am - 6.30 pm
Saturday
8.30 am - 1.30 pm
Zone
ENDS
local
taxi
Bus lane
Delays
possible
until
Sept 99
1 mile
Sorry
for any
delay
End
SLOW
WET TAR
800 yards
M1
&
A 517
M1 only
ANY
VEH
800 yards
STAY IN LANE
Max
speed
30
800 yds
450 yds


50
mile
ahead

Appendix 3: Publications

Vision Models Based Identification of Traffic Signs[1]

X. Gao, N. Shevtsova*, K. Hong, S. Batty, L. Podladchikova*, A. Golovan*, D. Shaposhnikov*, V. Gusakova*

School of Computing Science, Middlesex University, Bounds Green, London N11 2NQ, UK.

* Laboratory of Neuroinformatics of Sensory and Motor Systems, A.B. Kogan Research Institute for Neurocybemetics, Rostove State University, Rostov-on-Don, Russia.

Abstract

During the last 10 years, computer hardware technology has been improved rapidly. Large memory, storage is no longer a problem. Therefore some trade-off (dirty and quick algorithms) for traffic sign recognition between accuracy and speed should be improved. In this paper, a new approach has been developed for accurate and fast recognition of traffic signs based on human vision models. It applies colour appearance model CIECAM97s to segment traffic signs from the rest of scenes. A Behavioural Model of Vision (BMV) is then utilised to identify the signs after segmented images are converted into gray level images. Two standard traffic sign databases are established. One is British traffic sign and the other is Russian traffic sign. Preliminary results show that around 90% signs taken from the British road with various viewing conditions have been correctly identified.

1. Introduction

Identification of traffic signs correctly at the right time and the right place is very important for car drivers to insure themselves and their passengers' safe journey. However, sometimes, due to the change of weather conditions or viewing angles, traffic signs are difficult to be seen until it is too late. Development of an automatic system to be implemented inside cars for recognition of traffic signs will certainly improve driving safety a great deal.

An automatic real time system requires the identification of traffic signs invariant with respect to various transformations of signs and viewing environment. In particular, in view of the extremely stringent safety requirements for routine use of approaches on public roads, more computing power than was available a few years ago and more robust algorithms will be required in order to provide the necessary accuracy in recognition of traffic signs.

There are broadly 3 major methods applied for traffic sign recognition. They are colour-based, shape-based, and neural network-based recognition. Colour is a dominant visual feature and undoubtedly represents a piece of key information for drivers to handle. Colour is regulated not only for the traffic sign category (red = stop, yellow – danger, etc.) but also for the tint of the paint that covers the sign, which should correspond, with a tolerance, to a specific wavelength in the visible spectrum [1]. Therefore it is widely used in the systems for traffic sign recognition[1], especially for segmentation of traffic sign images from the rest of a scene. The eight colours, red, orange, yellow, green, blue, violet, brown and achromatic, are the most discriminating colours for traffic signs.

Most colour-based techniques in computer vision run into problems if the illumination source varies not only in intensity but also in colour as well. This is the main reason why many researchers have tried to come

[1]First European Conference on Color in Graphics, Image and Vision

up with algorithms for separating the incident illumination from the colour signal perceived by the sensors. So that after this kind of separation, a sign becomes illumination-invariant and is full of characteristics of the surface that reflects the light. As the spectral composition, and therefore the colour, of daylight is known to change depending on weather conditions, e.g., sky with/without clouds, time of day, and night when all sorts of artificial lights are surrounded [2], no method has been widely accepted yet.

In this study, traffic signs are segmented from real world road scenes based on colour contents using a standard colour appearance model CIECAM97 recommended by the CIE (International Committee on Illumination) [3, 4] and are identified after segmentation by the application of a Behavioural Model of Vision (BMV) [5, 6].

2. Methodology

In this student, two standard databases have been established. One is based on British traffic signs (n=142) scanned in from the book of HighWay code. The other is Russian traffic signs (n=158) obtained by the Russian scientists from the web site http://www.domkrat.ru/Laws/rules/znak1.shtml.

2.1 Segmentation

The first step to process the images (still images for the time being) taken from a video camera is to segment the sub-images of traffic-sign-to-be from the rest of scenes. To achieve this, images from the standard databases are firstly utilised to find the range of colour vectors for the colours used in the signs, mainly red, blue, black and white. The ranges for each colour vector, e.g., (red, lightness, chroma) and (blue, lightness, chroma), are found by calculating the corresponding values using the CIECAM97 model.

When an image is downloaded to a computer, it is expressed in a RGB space. To convert RGB space to CIE standard XYZ space, the following equations are applies as shown in Eq(1) for average daylight with CIE standard illuminant D65 as reference white, i.e., [Xw, Yw, Zw] = [0.95045 1.0 1.088754].

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.412453 & 0.357580 & 0.180423 \\ 0.212671 & 0.715160 & 0.072169 \\ 0.019334 & 0.119193 & 0.950227 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (1)$$

Then from CIE XYZ space, the hue, chroma, and lightness are obtained using CIECAM97 model. The reason to apply this model is that it takes viewing condition into account and can predict colours as accurate as an average observer.

During the study, 83 images of British road signs have been taken with variety of viewing and weather conditions using a digital camera, Olympus Digital Camera C-3030. These images are then classified visually according to the viewing conditions, such as cloudy, sunny, etc.. Based on the images in each group, the parameters for each viewing condition are found out from [3] (e.g., direct sun light with colour temperature 5335K and light from overcast sky with colour temperature 6500K) for the application of the colour appearance model. Images taken under real viewing conditions are then transformed from RGB space (the format used in computers to represent an image) to CIE XYZ values and then to LCH (Lightness, Chroma, Hue) using the model of CIECAM97. The range for red sign is between 393-423 that is calculated using mean ± standard deviation, and is between 280 to 290 for blue hue respectively. While for chroma, the range is between 57 to 95. The background also has chroma value ranged between 7 to 50. Figure 1 illustrates the range of blue and red signs plotted on a u'v' chromaticity diagram.

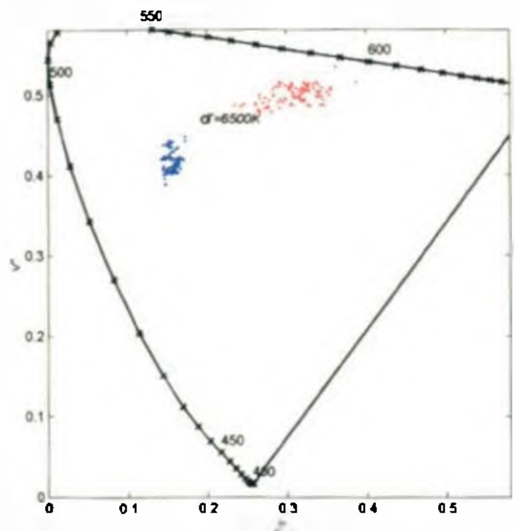


Figure. 1 The ranges of blue and red signs are plotted on a u'v' diagram.

2.2 Recognition

The second step to recognition a sign is to process segmented sub-images. The identification of traffic signs in this study is carried out by the application of the BMV model [5, 6]. This model is developed on the base of biologically plausible algorithms that are representing space-variant images and has demonstrated the ability to recognise complex grey-level images invariantly with respect to shift, plain rotation, and in a certain extent to scale.

To apply the model to the traffic sign identification task, the traffic sign database was transformed into a model-specific form. Firstly, all coloured images from the database are converted into grey-level representation. Then, for each image in the database a specific description is obtained based on trajectory of its viewing according to the most informative regions of the sign [5]. The model provides a compressed and invariant representation of each image fragments along the trajectory of view by space-variant features extracted in the fragment by Attention Window (AW). These descriptions have been stored with the images and form a model-specific database of traffic sign images. The model-specific database for traffic signs needs to be built only once. The descriptions or features for each image are then utilised in all further computer experiments on recognition of real world images of traffic signs.

2.2.1 Feature description of traffic sign

To extract features of signs, or to describe a sign at both memorising (for database images) and recognition (for real world images) stages, a specific structure of the input window (IW) is provided. IW simulates some mechanisms of the vision system, such as space-variant representation of information from the centre (fovea) to the periphery of the retina [7, 8], neuronal orientation selectivity [9], and context encoding of the fovea information [8].

Similar to [6], The following steps are applied for feature description.

(i) an image is presented by 49-dimensional vector of orientation extracted in vicinity of each of 49 nodes of IW;

(ii) the IW are located at the intersections of sixteen radiating lines and three concentric circles, each with a different RL;

(iii) orientation of segments in the vicinity of each IW node is determined by means of Gaussian with spatially shifted centres with the step of 22.5°;

(iv) representation of space-variant image is emulated by Gaussian convolution with different kernels.

Contrary to [6], the IW size is increased to 36 pixels (instead of 16 pixels in the basic BMV) and kernel sizes are changed to process a sign by one fixation of the IW. That is they are equal to 5x5 for the central (fovea one) part of the IW, 7x7 for the immediate (parafovea one), and 9x9 for the peripheral part. An example of oriented elements detected in context area of indicated node of a sign is shown in Figure 2.

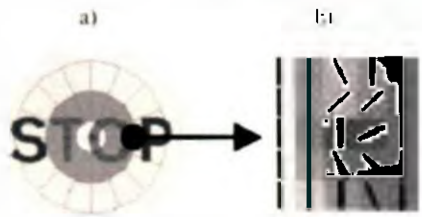


Figure 2. (a) Schematic diagram of the Input Window (IW) located in the centre of informative part of a sign; (b) estimation of orientation context in each of 49 nodes of input window.

Also, estimation of oriented elements in the context area of 48 IW nodes (beside the central node) is used to retrieve a detailed feature description of a sign. The size of context area is varied for different parts of the IW: it is equal to 3x3 for 16 nodes located on the central ring of the IW, 5x5 for the immediate one, and 7x7 for the peripheral ring. Each IW node is described by two values. They are orientation dominating in its context area providing more than 50% of context area points are detected and the density of oriented elements that are detected in the context area as shown in Figure 3.b. Such structure of the IW together with its location in the sign centre provides maximal representation of oriented elements in informative parts of sign at the first and second resolution levels (up to 90%).

2.2.2 Recognition algorithms

The 49-dimension vector for an incoming traffic-sign-to-be image preliminary classified by colour and external form was compared with descriptions of database images of the corresponding subgroup by the formula:

$$K^b = \sum_{i=1}^{49}\left[\operatorname{sgn}\left(Or_i^b - Or_i^{rw}\right)\left(1 - \operatorname{abs}\left(\rho_i^b - \rho_i^{rw}\right)\right)\right] \quad (2)$$

$$\operatorname{sgn}(x) = \begin{cases} 1, & if\ x = 0; \\ 0, & overwise; \end{cases}$$

where *Or* is dominating orientation extracted in the context area of a given IW node (orientations are determined by the step 22.5° and indicated as 1, 2, ..., 16) superscript *b* stands for prototype database images. Also rw is for the incoming real world image, p is the density of the dominating oriented segment in the vicinity of the given IW node. Preliminary testing has shown that minimal value of parameter K^b must be equal to 25 and K^b of the nearest sign in the database must differ from that of the template sign no less than by 4 to provide a correct recognition. A prototype image from the database with maximal K^b was considered as the result of recognition.

3. Results

After a coloured image is segmented using the CIECAM97 colour appearance model, it is firstly converted into a grey-level representation. The BMV model then starts to find representative features from the image-to-be-identified and to search for a hypothesis to be generated about the image in accordance with the model-specific database. During this search the representative description of the query image is compared to the model-specific description of the database of traffic signs. If a successful match occurs the presented image is recognised, and the matched sign image is retrieved.

Initial experimental results show that the majority signs can be segmented correctly by using CIECAM97 colour vision model, up to 90% for sunny days.

After segmentation, the BMV model correctly identified 37 out of 41 potential traffic sign images for sunny weather conditions and 37 out of 42 for cloudy weather conditions, which gives 90% and 88% success rates respectively. The non-identified or falsely identified signs are either of low resolution (taken from very far distance, more than 60 meters) or have very complex information content, for example, the sign "GIVE WAY" with blurry letters, or a complex disturbing background. Recognition time is varied from 0.35 seconds up to 0.6 seconds per image on a standard Pentium III, 400Hz.

Figure 3 illustrates an example of segmentation. The third segmented image (with the rear light from the left-hand-side car) is not traffic sign but within the colour range- Although it is segmented at this stage, it will be recognised as non-sign image during the recognition stage.

Figure 3. An example of segmented using CIECAM97 model. The third segmentation is not traffic sign that will be recognised using BMV model

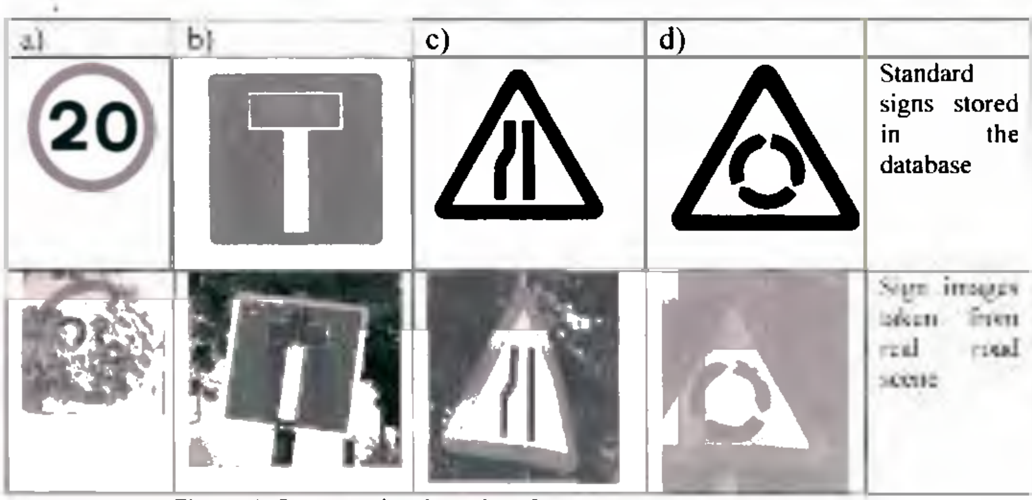


Figure 4. Some retrieval results of recognition using BMV model.

Figure 4 shows some recognition results using BMV models. It can be seen that this model works very well for some images with only partial information (as shown in (a)) and very blur images (picture (d)).

4. Discussion

Overall the models-based approach can give accurate identification for traffic signs located in a moderate distance for still images in various weather conditions and shows a good performance for a wide variety of traffic signs of different colours, forms, and informative content. The use of the CIECAM97 colour vision model allows the segmentation of the majority of traffic signs from the rest of the scenes. Computer experiments with the BMV model indicate that a preliminary separation of traffic signs by shape for each colour (for example, rectangle versus circle for blue traffic sings or triangle versus ring/circle for red ones) can accelerate sign identification. In addition, experimental results demonstrate the importance of AW fixation points chosen while viewing trajectory formation. Also the adequate template image encoding indicates that it is necessary that psychophysical experiments should be conducted to achieve better understanding what attracts a driver's visual attention while driving along the road in order to find the most informative regions in traffic sign images. Modification of the BMV model in accordance with the results of these experiments and the use of special acceleration boards can lead to improvement in its performance and therefore increase its importance for practical applications.

Acknowledgement

This project is founded by The Royal Society, UK, under the International Scientific Exchange Scheme. Their support is gratefully acknowledged.

References

1. M. Lalonde and Y. Li, Road Sign Recognition -- Survey of the State of Art, technique report for Sub-Project 2.4, CRIM-IIT-95/09-35, http://nannetta.ce.unipr.it/argo/theysay/rs2/#publications.

D. Judd, D. MacAdam, and G., Wyszecki, Spectral distribution of typical daylight as a function of correlated color temperature, J. Opt. Soc. Am., 54(8):1031-1040 (1964).

3. M.R. Luo and R.W.G. Hunt, The structure of the CIE 1997 colour appearance model (CIECAM97s), Color Res. Appl., 23:138-146 (1998).
4. CIE, The CIE 1997 Interim Colour Appearance Model (Simple Version), CIECAM97s, CIE TC1-34, April, (1998).
5. L.N. Podladchikova, V.I. Gusakova, D.G. Shaposhnikov, A. Faure, A.V. Golovan, and N.A. Shevtsova, MARR: Active vision model, Proceedings SPIE'97, 3208:418-425 (1997).
6. I.A., Rybak, V.I., Gusakova, A.V., Golovan, L.N., Podladchikova, N.A., Shevtsova, A Model of Attention-guided Visual Perception and Recognition, Vision Research, 38:2387-2400 (1998).
7. E.L. Schwartz, Computational anatomy and functional architecture of striate cortex: A spatial mapping approach to perceptual coding, Brain Research, 20, 645-669 (1980).
8. A.L.Yarbus, Eye Movements and Vision, 1967, New York: Plenum.
9. D.H. Hubel, & T.N. Wiesel, Sequence, regularity and geometry of orientation columns in the monkey striate cortex, J. Comparative Neurology, 158, 267-293 (1974).

Road Sign Recognition by Single Positioning of Space-Variant Sensor Window[2]

D.G. Shaposhnikov, L.N. Podladchikova,
A.V. Golovan, N.A. Shevtsova
A.B.Kogan Research Institute for Neurocybernetics, Rostov State University, Russia
nisms@krinc.ru
K. Hong, X.W. Gao
School of Computing Science, Middlesex University, London, United Kingdom
K.Hong@mdx.ac.uk, X.Gao@mdx.ac.uk

Abstract

A biologically plausible model of traffic sign detection and recognition invariantly with respect to variable viewing conditions is presented. The model simulates several key mechanisms of biological vision, such as space-variant representation of information (reduction in resolution from the fovea to retinal periphery), orientation selectivity in the cortical neuron responses, and context encoding of information. The model was tested on British traffic signs and demonstrated the ability to recognize these signs from a single position of a space-variant sensor window. After performing colour segmentation and classification and finding the sign centre, 85% of the traffic signs tested were identified under various environmental conditions.

1. Introduction

Many methods for automatic traffic sign identification have been developed with some promising results [1, 5, 9, 13]. However, identification of traffic signs invariantly with respect to various natural viewing conditions still remains a challenging task. In particular, with the account of safety requirements on public roads, more robust and fast algorithms are required to provide the necessary accuracy in recognition of traffic signs.

A biologically plausible model of visual recognition BMV (Behavioural Model of Vision) [8] was previously implemented for solution of the task of traffic sign recognition [1]. That model version required visual processing in multiple positions of the space-variant sensor window (SW) and demonstrated the ability of recognition of traffic signs, invariantly with respect to viewing and environmental conditions, with recognition rate about 80%. Several means for the improvement of recognition efficiency (i.e., increasing recognition rate and reduction of computations) have been proposed. In particular, it was suggested that the signs could be quickly recognized from a single position of SW if this position is close to the sign centre, since for most traffic signs the geometrical centre is also the centre of information content.

Choice of informative image fragments for detailed processing is one of the most important problems in the field of image recognition [2, 7]. There are many algorithms for detection of most informative parts of images in the frameworks of both conventional and biologically plausible approaches. Most of them are image-dependent because each type of images has specific important fragments. For example, for face images the most informative fragments are eyes, nose, and mouth, [2, 7, 12]. As mentioned above, for most traffic signs the most informative fragments are concentrated around the sign centre. However, in traffic sign recognition only a few approaches are known that attempt to use a selection of most informative fragments of signs for detailed processing [1, 5, 9].

Colour is a dominant visual feature, which undoubtedly represents a key piece of information used by drivers. Therefore colour is widely used in traffic sign

[2]:15th International Conference on Vision Interface

recognition systems [5, 13], especially for segmentation of traffic sign images from the rest of a scene. The color segmentation can be also used for finding the sign center.

In this study, traffic signs were segmented from road scenes under various environmental conditions by colour contents using a standard colour appearance model CIECAM97 [4, 6, 11]. First, the colour segmentation and classification based on colours and shapes provided the detection of the sign centre. After that, the sign was recognized from a single position of a space-variant sensor window centered in the sign centre.

2. Algorithms and procedures

British traffic sign images (n=105) for standard database were scanned from the book of Highway Code. These images were used for preliminary testing the developed algorithms and procedures. Besides, they served as prototypes for recognition of real world images. Traffic signs (n=97) have been taken in London under various environmental conditions using a digital camera (Olympus Digital Camera C-3030). According to conventional standards, the size of each sign in both the standard database and in real world images was normalized to 40x40 pixels.

2.1. Colour Segmentation

Images taken from real word under different viewing conditions were preprocessed to find the range of colour vectors for the colours usually used in the signs, namely red, blue, black, and white. This preprocessing was performed using model of CIECAM97 (11). CIECAM97 is a standard colour appearance model recommended by CIE in 1997 for measuring colour appearance under various viewing conditions. This model can estimate a colour appearance as accurate as an average observer. For human perception, the most common terms used for colour description or colour appearance are lightness, chroma, and hue. A representative set of traffic signs was classified visually according to the viewing and environmental conditions, such as cloudy, sunny, etc. Based on the images in each group, the parameters for each viewing condition were found from [4] (e.g., direct sun light with colour temperature 5335K and light from overcast sky with colour temperature 6500K) for application of the colour appearance model. Test images taken under real viewing conditions were transformed from RGB space to CIE XYZ values and then to LCH (Lightness, Chroma, Hue) using the model of CIECAM97. The lightness was similar for red and blue signs and background. Therefore, only Hue and Chroma were used for segmentation.

Table 1. **The ranges of Hue and Chroma for red and blue signs.**

Colour	Hue	Chroma
Red	393 - 423	57 - 95
Blue	280 - 290	57- 95

Based on the range of sign colours, traffic-signs-to-be are segmented from the rest of scenes for further identification and classified according to the colour. Only blue and red signs were used in this study (Table 1). All sign images with size more than 10x10 pixels (pictures were taken within 100 meters distance) could be segmented correctly. Sometimes, some other contents, such as the rear red lights of cars were also segmented. However, these non-sign segments could be rejected during recognition stage using BMV.

Information about the colour of the sign was also used for the localization of sign centre based on estimation of the location of colour contour elements.

2.2. Classification of traffic signs according to their shapes

After preliminary color classification, the signs were further classified by shapes (circle, rectangle, or triangle) which were detected using histograms of oriented segments detected on a sign image. Each sign with a certain shape has its own characteristic pattern of oriented segments. In particular, elements of various orientation have nearly equal representation for circle signs in contrary to rectangle signs (Fig. 1, a), which in turn have preferably horizontal and vertical oriented elements (in sum, more than 50% of all oriented elements). This simple method provides classification of signs in frameworks of each colour group into two subgroups according to their shapes: blue rectangle and circular signs and red triangle and circular signs. Quantitative parameters for this classification were obtained for particular groups of signs in the standard database. The same parameters were used for classification of real world images.

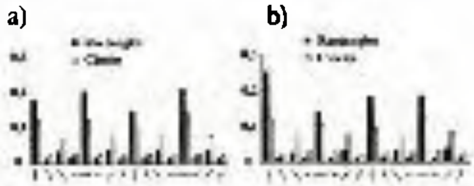


Figure 1. Averaged histograms of oriented element representation for blue traffic signs in (a) standard database (n=76) and (b) real world images (n=56).

2.3. Determination of sign centre for positioning of sensor window

The recognition method the of BMV model [8] is based on the encoding of the image according to the path of image viewing and the description of image fragments at each position of SW by a set of primary features (oriented elements). For traffic sign images, we have found that an effective recognition may be performed from a single position of SW if the latter is placed in about the sign centre. Such location of the SW appears to provide the most specific sign description.

The algorithm for determination of colour contour geometric centre has been developed to find the centre of the internal informative part of signs. In the algorithm, spatial location of colour contour elements in the real world signs was determined on the basis of quantitative estimations of RGB composition for a signs of a given colour in the standard database. This algorithm provided for the exact normalisation of sign size to 40x40 pixels, and extraction of a "pure" (without background) real world sign (Fig. 2). This algorithm provided for determination of the geometric centre of a sign with necessary accuracy (up to 6 pixels).

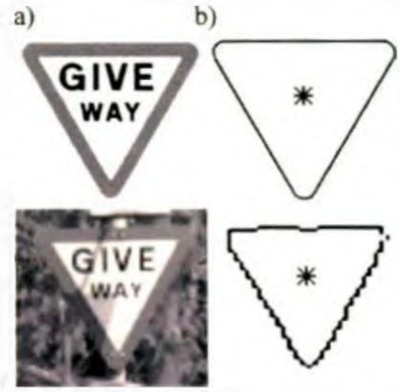
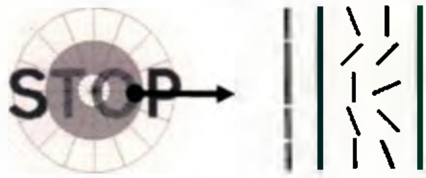


Figure 2. Examples of determination of the sign centre. Images from the standard database (upper row) and a real world picture (lower row) are shown in (a); * - indicates location of the sign centre determined by colour contour (b).

2.4. Feature description of traffic sign

The description of each sign at the memorising stage (for standard database images) and recognition stage (for real world images) was provided by the specific structure of the SW, which imitates some features of the real visual system such as space-variant representation of information from the centre (fovea) to the periphery of the retina [10, 12], neuronal orientation selectivity [3], and context encoding of the foveal information [12].

Basic algorithms of sign processing by space-variant SW were similar to [8], namely: (i) an image was represented by 49-dimensional vector of oriented elements extracted in vicinity of each of 49 nodes of SW; (ii) the SW nodes were located at the intersections of sixteen radiating lines and three concentric circles with increasing radii (Fig.3, a); (iii) orientation of segments in



(a) (b)

Figure 3. (a) Schematic of the SW located in the centre of the informative part of a sign. Circles with different gray levels represent different resolutions in the SW structure. (b) Oriented elements detected in the context area of a SW node (indicated by a small black circle in (a)).

the vicinity of each SW node is determined by means of calculation of the difference between two oriented Gaussians with spatially shifted centres; (iv) space-variant image representation is emulated by Gaussian convolution with different kernels depending on distance from the SW centre. The SW size increased to 36 pixels (instead of 16 pixels in the basic BMV [8]), and kernel sizes were changed to process a sign from a single position of the SW, i.e., they were equal to 5x5 for the central (foveal) part of the SW, 7x7 for the immediate (parafoveal), and 9x9 for the peripheral part. Besides, estimation of oriented elements in the context area of 48 SW nodes (except for the central node) was used to receive a detailed feature description of a sign. The size of context area was varied for different parts of the SW: it was equal to 3x3 pixels for 16 nodes located in the central circle of the SW, 5x5 for the immediate, and 7x7 for the peripheral circle. Each SW node was described by two values, namely, an orientation dominated in the context area and a density of oriented elements detected in this area (Fig. 3, b). Such structure of the SW and its location in the sign centre provided maximal representation of oriented elements (up to 90%) in the informative region of a sign at the first and second resolution levels. An example of detection of oriented elements in the context area of the indicated node of a sign is shown in Fig. 3, b.

2.5. Recognition algorithms

The 49-dimensional vector for an incoming traffic-sign-to-be image preliminary classified by colour and shape was compared with descriptions of database images of the corresponding subgroup by the formula:

$$K^b = \sum_{i=0}^{i<49}\left[\mathrm{sgn}\left(Or_i^b - Or_i^{rw}\right)\cdot\left(1 - abs\left(\rho_i^b - \rho_i^{rw}\right)\right)\right]$$

$$\mathrm{sgn}(x) = \begin{cases} 1, & if \;\; x = 0; \\ 0, & overwise; \end{cases}$$

where *Or* is dominating segment orientation in the context area of a given SW node (orientations are determined by the step 22.5° and indicated as 1, 2, ..., 16); superscript *b* stands for prototype database images, *rw* - for the incoming image; ρ is the density of the dominating oriented segment in the context area (see Section 2.4.) of the given SW node. A prototype image from the database with maximal K^b was considered as the result of recognition.

3. Computer simulation

During the stage of memorising, each traffic sign from the database was preliminary classified according to colour and shape. Thus, the database was divided into 4 subgroups. Then for each subgroup, each sign was transformed into a model-specific form (i.e., presented by a 49-dimensional vector of oriented elements extracted in vicinity of each of 49 nodes of SW) and a specific description for each image in each subgroup was obtained. These descriptions were stored in a model-specific prototype database. The model-specific database for traffic sign images needs to be built only once. The descriptions or features for each database image are then utilised in all further computer experiments on recognition of traffic signs.

During recognition, first a test colour image is extracted from the scene using the CIECAM97 colour appearance model. Then the image is classified according to its colour and shape. After that, the sign centre is determined. Then the recognition algorithm searches for a prototype image in the model-specific database. During this search the representative description of the incoming image is compared to the model-specific description of the database traffic signs. If a successful match occurs the presented image is recognized, and the matched database sign image is retrieved.

Our experimental studies have shown that the majority of signs (more than 90%) can be segmented correctly by using CIECAM97 colour vision model. After segmentation and classification according to colour and shape, the model identified 83 out of 97 potential traffic sign images, which gives 85% success rate. Similar results were obtained for different viewing and environmental conditions (87% and 84% for sunny and cloudy weather respectively). The non-identified or falsely identified signs were either of low resolution (taken from very far distance, more than 60 meters) or have information content similar to other signs (for example, the triangle sign "Roundabout") (Fig. 4), or a complex disturbing background. Recognition time varied from 0.25 seconds up to 0.4 seconds per image on a standard Pentium 233.

a) b) c)

Figure 4. The examples of recognised (a, b) and non-recognised (c) signs (lower row). Retrieved prototype images are shown in the upper row.

4. Conclusion

In this work we tested our suggestion that the BMV model would be able to recognize traffic signs from a single position of the space-variant sensor window [1]. The presented results confirm this suggestion and even demonstrate an increase of recognition efficiency from a single SW position as compared to the version of the model based on multiple positions of the SW while processing the same signs. Task-oriented modifications of the SW, including increase of the SW size, determination of context in the vicinity of each SW node, setting the SW in the sign centre, etc. – all together allow a detailed feature description of the informative part of a sign. This description is sufficiently compressed while quite enough for effective recognition. In addition, this description is stable to local image disturbances in a certain range. Overall, the described model-based approach provides an accurate identification of the traffic signs located at a moderate distance under various environmental conditions. The resultant recognition system shows a good performance for a wide variety of traffic signs of different colours, forms, and informative content.

Acknowledgement

This work is supported by The Royal Society, UK, under the International Scientific Exchange Scheme.

References

[1] X.W. Gao, A.V. Golovan, K. Hong, L.N. Podladchikova, D.G. Shaposhnikov, N.A. Shevtsova. Road Sign Recognition by Means of the Behavioral Model of Vision. *Proc. of Conf. on Neuroinformatics*, January 2002, Moscow, 63-69.

[2] A.V. Golovan, N.A. Shevtsova, L.N. Podladchikova, S.N.Markin, D.G.Shaposhnikov. Image Preprocessing for Identifying the Most Informative Regions in Facial Images. *Pattern Recognition a. Image Analysis*, 2001, 11, 2, 313-316.

[3] D.H. Hubel, T.N. Wiesel. Sequence, Regularity and Geometry of Orientation Columns in the Monkey Striate Cortex. *J. Comparative Neurology*, 1974, 158, 267-293.

[4] D. Judd, D. MacAdam, and G. Wyszecki. Spectral Distribution of Typical Daylight as a Function of Correlated Colour Temperature. *J. Opt. Soc. Am.*, 1964, 54(8), 1031-1040.

[5] M. Lalonde, Y. Li. Road Sign Recognition - Survey of the State of Art. *Tech. rep for Sub-Project 2.4*, CRIM-IIT-95/09-35, http://nannetta.ce.unipr.it/argo/theysay/rs2/#publications.

[6] M.R. Luo, R.W.G. Hunt, The Structure of the CIE 1997 Colour Appearance Model (CIECAM97s). *Color Res. Appl.*, 1998, 23, 138-146.

[7] W. Osberger and Maeder A.J. Automatic Identification of Perceptually Important Regions in an Image. *Proc. 14th Int. Conf. on Pat. Rec.*, Brisbane, Australia, 1998, 701-704.

[8] A. Rybak, V.I. Gusakova, A.V. Golovan, L.N. Podladchikova, N.A. Shevtsova I. A Model of Attention-Guided Visual Perception and Recognition. *Vision Research*, 1998, 38, 2387-2400.

[9] G. Salgian, D. Ballard. Visual Routines for Vehicle Control. In D. Kriegman, G. Hager, and S. Morse, editors, *The Confluence of Vision and Control*, Springer Verlag, 1998.

[10] E.L. Schwartz. Computational Anatomy and Functional Architecture of Striate Cortex: A Spatial Mapping Approach to Perceptual Coding. *Brain Research*, 1980, 20, 645-669.

[11] CIE, The CIE 1997 Interim Colour Appearance Model (Simple Version), *CIECAM97s*, CIE TC1-34, April 1998.

[12] A.L.Yarbus. *Eye Movements and Vision*, Plenum, New York, 1967.

[13] M. M. Zadeh, T. Kasvand, C. Y. Suen. Localization and Recognition of Traffic Signs for Automated Vehicle Control Systems. In *Conf. on Intelligent Transportation Systems, part of SPIE's Intelligent Systems & Automated Manufacturing*, Pittsburgh, PA, October 1997.

A New Approach for Traffic Sign Recognition[3]

X. Gao[1], N. Shevtsova[2], K. Hong[1], S. Batty[1], L. Podladchikova[2], A. Golovan[2], D. Shaposhnikov[2], V. Gusakova[2]

1 School of Computing Science, Middlesex University, Bounds Green, London N11 2NQ, UK. {x.gao, k.hong, s.batty}@mdx.ac.uk

2 Laboratory of Neuroinformatics of Sensory and Motor Systems, A.B. Kogan Research Institute for Neurocybernetics, Rostove State University, Rostov-on-Don, Russa, nisms@krinc.rc

Abstract

A new approach has been developed for accurate and fast recognition of traffic signs based on human vision models. It applies colour appearance model CIECAM97s to segment traffic signs from the rest of scenes. This model takes viewing conditions into account and can recognise colours as accurate as an average observer invariant of lighting sources, background, and surrounding colours. A Behavioural Model of Vision (BMV) is then utilised to identify the signs after segmented images are converted into gray level images. The BMV model is transform invariant and can recognise most signs with distorted shapes. Two standard traffic sign databases are established. One is British traffic signs and the other is Russian traffic signs. Preliminary results show that around 90% signs taken from the British road with various viewing conditions have been correctly identified.

1. Introduction

For car drivers, correct identification of traffic signs at right time and right place plays a crucial part in insuring themselves and their passengers' safe journey. Sometimes, due to changing weather conditions or viewing angles, traffic signs are not easily to be seen until it is too late. Development of automatic systems for identification of traffic signs is therefore an important approach to improve driving safety.

Many methods for automatic traffic sign identification have been developed with some promising results. However, an automatic real time system requires the identification of traffic signs invariant with respect to various transformations of signs and viewing environment, which still remains a challenging issue. In particular, in view of the extremely stringent safety requirements for routine use of approaches on public roads, more computing power than was available a few years ago and more robust algorithms will be required in order to provide the necessary accuracy in recognition of traffic signs.

Colour is a dominant visual feature and undoubtedly represents a piece of key information for drivers to handle. Colour is regulated not only for the traffic sign category (red = stop, yellow = danger, etc.) but also for the tint of the paint that covers the sign, which should correspond, with a tolerance, to a specific wavelength in the visible spectrum [1]. Therefore it is widely used in traffic sign recognition systems [1], especially for segmentation of traffic sign images from the rest of a scene. The eight colours, red, orange, yellow, green, blue, violet, brown and achromatic, are the most discriminating colours for traffic signs.

[3]: The 2002 International Conference on Imaging Science, Systems, and Technology (CISST'02)

Most colour-based techniques in computer vision run into problems if the illumination source varies not only in intensity but also in colour as well. This is the main reason why many researchers have tried to come up with algorithms for separating the incident illumination from the colour signal perceived by the sensors. So that after this kind of separation, a sign becomes illumination-invariant and is full of characteristics of the surface that reflects the light. As the spectral composition, and therefore the colour, of daylight is known to change depending on weather conditions, e.g., sky with/without clouds, time of day, and night when all sorts of artificial lights are surrounded [2], no method has been widely accepted yet.

In this study, traffic signs are segmented from real world road scenes based on colour contents using a standard colour appearance model CIECAM97 recommended by the CIE (International Committee on Illumination) [3, 4] and are identified after segmentation by the application of a Behavioural Model of Vision (BMV) [5, 6].

2. Methodology

Two standard databases have been established in our study. One is based on British traffic signs (n=142) scanned in from the book of HighWay code. The other is Russian traffic signs (n=158) obtained by the Russian scientists from the web site http://www.domkrat.ru/Laws/rules/znak1.shtml.

2.1 Segmentation

The advantages of using CIECAM97 is that most of visual colours are illuminance invariant, which is quite similar to human vision's reaction. After adaptation to the day light, we see the colours of signs are more or less similar.

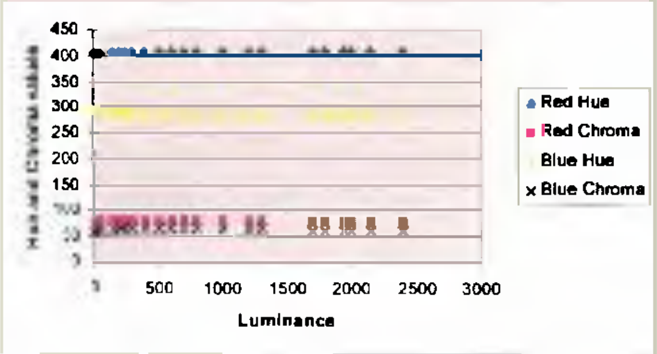


Fig. 1. The change of chroma and hue values with the change of luminance in cd/sqm for red signs.

Figure 1 illustrats the change of hue and chroma values for the typical red with average hue value of 407 and blue having hue value of 290 signs with the change of luminance values. The luminance is changed from 5 to 2500 candela square meters (cd/sqm). It can be seen that after the calculation from the model, the hue and chroma values are hardly changed. The standard deviations for red hue and red chroma are 0.62 and 1.10 respectively, leading to less than 1% variations. Therefore, it is reliable to segment signs based on hue and chroma values. The vector range for segmenting signs are the average values plus/minus standard deviations. Table 1 gives the range vectors for red and blue signs.

Table 1. Range vectors for red and blue signs.

Colour	Hue	Chroma
Red	393 - 423	57 - 95
Blue	280 - 290	57- 95

So that if any pixels with hue and chroma values falling within these ranges, these pixels are classified as potential traffic sign pixels and are to be processed further. Otherwise, those pixels outside these ranges will be categorised as part of background.

In this study, the range vectors are found out using the colours taken from the signs, mainly red, blue, black and white. Figure 2 displays the ranges of traffic signs taken in sunny day and plotted on a u'v' chromaticity diagram after conversion.

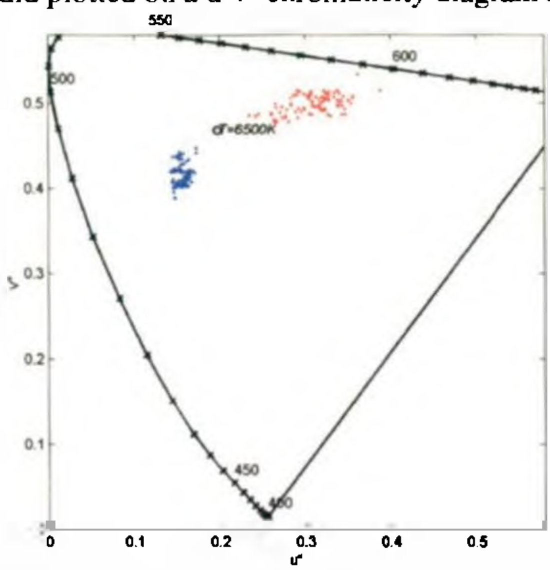


Fig. 2. Traffic signs are plotted on the u'v' diagram.

Based on the range of sign colours, traffic-signs-to-be are segmented from the rest of scenes for further identification.

2.2 Recognition

The identification of traffic signs in this study is carried out by the application of the BMV model [5, 6]. This model is developed on the base of biologically plausible algorithms of representation of space-variant images with different image viewings and has demonstrated the ability to recognise complex grey-level images invariantly with respect to shift, plain rotation, and in a certain extent to scale. The basic version of the BMV model has been modified to adjust it for traffic sign recognition task.

2.2.1 Classification of traffic signs according to their external forms

For all signs, both from standard databases and from real world images, preliminary classified according to the colour, their external form (circle, rectangle.

or triangle) can be determined by means of histograms of orientations detected at resolution level 3 (RL 3). RL 3 is emulated by Gaussian convolution (kernel size is equal 9). Each sign with a certain external form (in spite of its inner content) has characteristic relationship of horizontally, vertically, and obliquely oriented segments at RL 3. In particular, all oriented elements had nearly equal representation for circle signs contrary to rectangle signs (Fig. 3) that had preferable horizontal and vertical orientations (in sum, more than 50% of all oriented segments). For each external form, quantitative estimations were obtained for classification into particular groups of signs from British and Russian standard databases and from real world images.

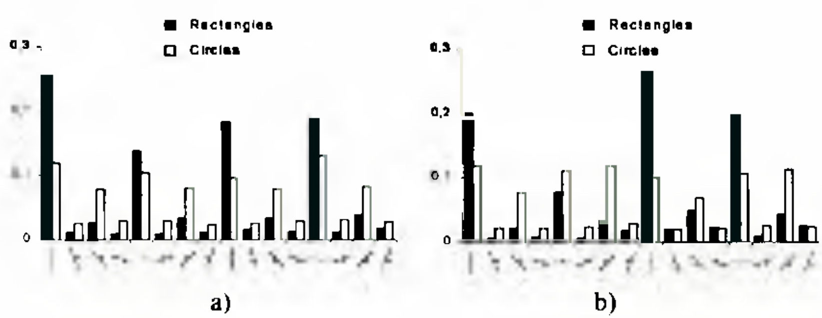


Fig. 3. Averaged histograms of orientations for Russian blue traffic signs in (a) standard database (n=66) and (b) real world images (n=19).

2.2.2 ***Recognition algorithms.***

Traffic sign recognition task was solved by the BMV developed earlier for invariant memorising and recognition of complex grey-level images such as human faces [5,6]. The basic properties of the BMV consist in:

(i) space-variant representation of image fragment in each fixation (a change of resolution level from the centre to the periphery of the Attention Window (AW));

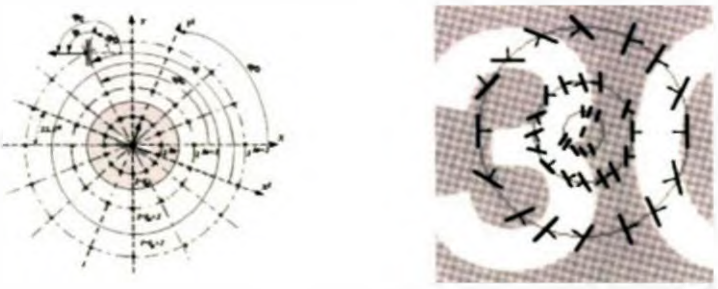
(ii) description of image oriented elements in each fixation by 49-dimensional vector;

(iii) transition description from a fixation point to the next one, and

(iv) image viewing trajectory formation.

The basic BMV demonstrates invariance to scale, rotation, noise, shift, and in part to point of view. The peculiarities of the traffic sign recognition task demand some modifications of the model. First, these modifications are determined by conventional format and relative simplicity of traffic sign images. Second, the basic model invariance to plain rotations should be eliminated because the same geometrical objects in traffic signs rotated in plain (such as arrows) have different meaning for driver. The modifications were performed in the (iii) – (iv) algorithms.

Similar to the basic version of the model, an image fragment in each fixation point is presented by 49-dimensional vector of oriented elements in relative coordinate system which is illustrated in Figure 4a. Orientation of segments near each of 49 AW point is then determined by means of the Gaussians with spatially shifted centres. Such procedures allow receiving sign fragment description that is relatively stable to any local image disturbances in some ranges. Contrary to the basic model version, transition description from a fixation point to the next one is realised in absolute coordinate system to exclude rotation invariance mentioned above. Example of oriented elements detected in a sign fragment is shown in Figure 4. b.



a) b)

Fig. 4. (a) Schematic of the AW; the areas of different resolutions are separated by shadings. The context points are located at the intersections of sixteen radiating lines and three concentric circles, each in a different resolution area. XOY is the absolute coordinate system. The relative coordinate system X^1OY^1 is attached to the basic edge in the centre of the AW; (b) example of oriented elements detected in an initial fixation point of the AW at different resolution levels for a traffic sign image.

To apply the model to the traffic sign identification task, the traffic sign database is transformed into a model-specific form. As the first step, all coloured images from the database are converted into grey-level representation. Then, for each image in the database a specific description is obtained based on trajectory of its viewing according to the most informative regions of the sign [7]. The model provides a compressed and invariant representation of each sign image fragments along the trajectory of viewing by representative space-variant features extracted in the fragment by Attention Window (AW). These descriptions have been stored with the images and form a model-specific database of traffic sign images. The model-specific database for traffic sign images needs to be built only once. The descriptions or features for each image are then utilised in all further computer experiments on recognition of real world images of traffic signs.

After a coloured image is segmented using the CIECAM97 colour appearance, it is firstly converted into a grey-level representation. The BMV model then starts to find representative features from the image-to-be-identified and to search for a hypothesis to be generated about the image in accordance with the model-specific database. During this search the representative description of the query image is compared to the model-specific description of the database traffic signs. If a successful match occurs the presented image is recognised, and the matched sign image is retrieved.

3. Results

During the study, 83 images of British road signs have been taken with variety of viewing and weather conditions using a digital camera, Olympus Digital Camera C-3030. These images are then classified visually according to the viewing conditions, such as cloudy, sunny, etc.. Based on the images in each group, the parameters for each viewing condition are found out from [3] (e.g., direct sun light with colour temperature 5335K and light from overcast sky with colour temperature 6500K) for the application of the colour appearance model. Images taken under real viewing conditions are then transformed from RGB space (the format used in computers to represent an image) to CIE XYZ values and then to LCH (Lightness, Chroma, Hue) using the model of CIECAM97.

Figure 5 shows an example of segmentation using CIECAM97 model and Figure 6 illustrates traffic recognition using BMV model.

Fig.5 An example of segmentation using CIECAM97 model.

Fig. 6. Sign recognition using the BMV.

Initial experimental results show that the majority signs can be segmented correctly by using CIECAM97 colour vision model.

After segmentation, the BMV model, correctly identified 37 out of 41 potential traffic sign images for sunny weather conditions and 37 out of 42 for cloudy weather conditions, which gives 90% and 88% success rates respectively. The non-identified or falsely identified signs are either of low resolution (taken from very far distance, more than 60 meters) or have a very complex information content, for example, the sign "GIVE WAY" with blurry letters, or a complex disturbing background. Recognition time is varied from 0.35 seconds up to 0.6 seconds per image on a standard Pentium III, 250 MHz.

4. Discussion

Overall the models-based approach can give accurate identification for traffic signs located at a moderate distance for still images in various weather conditions and shows a good performance for a wide variety of traffic signs of different colours, forms, and informative content. The use of the CIECAM97 colour vision model allows the segmentation of the majority of traffic signs from the rest of the scenes. Computer experiments with the BMV model indicate that a preliminary separation of traffic signs by shape for each colour (for example, rectangle versus circle for blue traffic sings or triangle versus ring/circle for red ones) can accelerate sign identification. In addition, experimental results demonstrate the importance of AW fixation points chosen while viewing trajectory formation. Also the adequate template image encoding indicates the necessity of psychophysical experiments to better understand what attracts a driver's visual attention while viewing traffic signs with various complexity in real world conditions and to find the most informative regions in traffic sign images. Modification of the BMV model in accordance with the results of these experiments and the use of special acceleration boards can lead to improvement in its performance and therefore increase its importance for practical applications.

Acknowledgement

This project is founded by The Royal Society, UK, under the International Scientific Exchange Scheme. Their support is gratefully acknowledged.

Reference:

1. M. Lalonde and Y. Li, Road Sign Recognition – Survey of the State of Art, technique report for Sub-Project 2.4, CRIM-IIT-95/09-35, http://nannetta.ce.unipr.it/argo/thevsav/rs2/#publications.
2. D. Judd, D. MacAdam, and G.. Wyszecki, Spectral distribution of typical daylight as a function of correlated color temperature, J. Opt. Soc. Am., 54(8):1031-1040, 1964.

3. M.R. Luo and R.W.G. Hunt, The structure of the CIE 1997 colour appearance model (CIECAM97s), Color Res. Appl., 23:138-146, 1998.
4. CIE, The CIE 1997 Interim Colour Appearance Model (Simple Version), CIECAM97s, CIE TC1-34, April, 1998.
5. L.N. Podladchikova, V.I. Gusakova, D.G. Shaposhnikov, A. Faure, A.V. Golovan, and N.A. Shevtsova, MARR: Active vision model, Proceedings SPIE'97, 3208:418-425
6. I.A., Rybak, V.I., Gusakova, A.V., Golovan, L.N., Podladchikova, N.A., Shevtsova, A Model of Attention-guided Visual Perception and Recognition, Vision Research, 38:2387-2400, 1998.
7. M. M. Zadeh, T. Kasvand, C. Y. Suen. Localization and Recognition of Traffic Signs for Automated Vehicle Control Systems. In Conf. on Intelligent Transportation Systems, part of SPIE's Intelligent Systems & Automated Manufacturing, Pittsburgh, PA, October 1997.