**Universitat
Autònoma
de Barcelona**

A Parametric Model for Computational Colour Naming

A dissertation submitted by **Robert Benavente Vidal** at Universitat Autònoma de Barcelona to fulfil the degree of **Doctor en Informàtica**.

Bellaterra, 14th May 2007

Director: **Dra. Maria Vanrell i Martorell**
Universitat Autònoma de Barcelona
Dept. Ciències de la Computació & Computer Vision Center

A la meva família

Abstract

Colour is an important visual cue widely used in computer vision in the last years. Most of the methods developed so far aim to extract low-level colour features from images and such information do not have a direct link to the high-level semantics that humans use. The lack of this direct link, known as the semantic gap, is even more significant for some applications such as image retrieval where users require systems able to support queries in natural language or, at least, in a high-level language. Hence, the main goal of this thesis is to make a further step in the way of reducing the semantic gap in the task of giving names to colours in images.

The main contribution of this thesis is a parametric colour-naming model for images. The problem is framed on the fuzzy set theory where each one of the 11 basic colour categories (white, black, red, green, yellow, blue, brown, purple, pink, orange, and grey) is characterized by a membership function.

Since the goal of such a model is to obtain the same name assignments as a human observer would provide, a set of human judgements is needed as starting point for the modelling process. To obtain such a data set we propose a methodology for fuzzy psychophysical experiments and a set of fuzzy judgements are obtained. The methodology and the results are validated by computing some usual statistics which are compared to previous experiments to show the equivalence of the results obtained with the new methodology and the ones from former experiments. The data set obtained has been made available online on the Internet for the research community.

The analysis of the results from the experiment allows defining the properties that the membership functions should fulfil. Several functions are proposed and evaluated to achieve the final Triple-Sigmoid with Elliptical centre (TSE) model which provides good fitting to the learning data and a categorization of the Munsell colour space which is consistent with previous works. The result of the fitting process is the set of parameters of the model which allows computing the membership of any given colour sample to the 11 colour categories considered, with all the advantages of a parametric implementation.

The last part of the thesis is devoted to analyse the conditions needed to apply the model on real images under uncalibrated conditions where no information about the acquisition conditions is known. This analysis shows that the model can work with an acceptable error on computer vision applications where a perceptual representation of colour information is needed. The model is tested in one of these applications on a real problem where automatic image annotation is used for image retrieval. The results on the experiments show the potentiality of the colour-naming model for different future applications and open new research possibilities in this field.

Resum

El color és una font d'informació visual molt important i àmpliament utilitzada en els darrers anys en la visió per computador. La majoria dels mètodes desenvolupats fins ara tenen com objectiu extreure característiques de baix nivell de les imatges i aquesta informació no té un lligam directe amb la semàntica d'alt nivell que utilitzem els humans. La falta d'aquesta relació directa, que es coneix com 'semantic gap', és encara més significativa per algunes aplicacions com la recuperació d'imatges de bases de dades on els usuaris requereixen sistemes que permetin realitzar cerques en llenguatge natural o almenys en un llenguatge d'alt nivell. Per tant, el principal objectiu d'aquesta tesi és fer un avanç en la línia de reduir el 'semantic gap' en la tasca de donar noms als colors de les imatges.

La principal contribució d'aquesta tesi és un model paramètric d'assignació de noms de colors en imatges. El problema s'ha emmarcat en la teoria dels conjunts difusos en la que cadascuna de les 11 categories bàsiques de color (blanc, negre, vermell, verd, groc, blau, marró, morat, rosa, taronja i gris) està caracteritzada per una funció de pertinença.

Donat que l'objectiu d'aquest model és obtenir els mateixos noms que proporcionaria un observador humà, es necessita un conjunt de judicis fets per persones com a punt de partida pel procés de modelat. Per obtenir aquest conjunt de dades es proposa una metodologia basada en lògica difusa per experiments psicofísics que ha permès obtenir un conjunt de judicis difusos. La metodologia i els resultats són validats a partir del càlcul d'alguns estadístics habituals que són comparats amb experiments previs per mostrar l'equivalència entre els resultats obtinguts amb la nova metodologia i els d'experiments anteriors. El conjunt de dades obtingut s'ha posat a disposició de la comunitat científica a través d'Internet.

L'anàlisi dels resultats de l'experiment permet definir les propietats que les funcions de pertinença haurien de complir. Proposarem i avaluarem diverses funcions per arribar finalment al model Triple Sigmoid amb centre El·líptic (TSE) que proporciona un bon ajust al conjunt d'aprenentatge i una classificació de l'espai de color Munsell que és consistent amb els treballs previs. El resultat del procés d'ajust és el conjunt de paràmetres del model que permeten calcular la pertinença de qualsevol mostra de color a les 11 categories de color considerades, amb tots els avantatges d'una implementació paramètrica.

La darrera part de la tesi està dedicada a l'anàlisi de les condicions necessàries per aplicar el model en imatges reals sota condicions no calibrades on no es coneix cap informació sobre les condicions d'adquisició. Aquesta anàlisi mostra que el model pot funcionar amb un error acceptable en aplicacions de visió per computador en les que és necessària una representació perceptiva de la informació de color. El model

iv

s'avalua en una d'aquestes aplicacions per un problema real d'anotació automàtica de bases de dades d'imatges. Els resultats en els experiments mostren la potencialitat del model d'assignació de noms de color per diferents aplicacions futures i obren noves possibilitats de recerca en aquest camp.

Contents

List of Tables

List of Figures

List of Acronyms

BO	Boynton and Olson
BVB	Benavente, Vanrell and Baldrich
CBIR	Content Based Image Retrieval
CC	Colour Constancy
CCT	Correlated Colour Temperature
CI	Colour Induction
CNH	Colour Naming Histogram
CNS	Colour-Naming System
CS	Circular-Sigmoid
DS	Double-Sigmoid
DSM	Double-Sigmoid Model
ES	Elliptical-Sigmoid
HVS	Human Visual System
ISCC	Inter-Society Color Council
LGM	Lammens Gaussian Model
LGN	Lateral Geniculate Nucleus
MAE	Mean Absolute Error
MES	MacLaurie's English Speaker
MGM	Multivariate Gaussian Model
MMV	Mean Membership Vector
MSE	Mean Squared Error
NBS	U.S. National Bureau of Standards
PC	Predominant Colour
PLSA	Probabilistic Latent Semantic Analysis

SG	Sigmoid-Gaussian
SGM	Sigmoid-Gaussian Model
SW	Sturges and Whitfield
TSC	Triple-Sigmoid with Circular centre
TSCM	Triple-Sigmoid with Circular centre Model
TSE	Triple-Sigmoid with Elliptical centre
TSEM	Triple-Sigmoid with Elliptical centre Model

Chapter 1

Introduction

The goal of this thesis is to automate the colour-naming task in the computer vision framework. In this chapter we introduce the topic of colour in computer vision and, more specifically, we focus on the colour-naming task. We will define the problem and will consider related topics, such as colour constancy and colour induction, that influence the automation of the colour-naming task. Finally, the fuzzy nature of the problem will be taken into account to define the framework in which the problem is posed.

1.1 Colour

Colour is the perceptual phenomenon caused by the light that reaches the retina. The human visual system only processes the electromagnetic radiation of wavelengths between 380nm and 780nm, that is, the human visible spectrum of light. The assignment of a colour name to each stimulus is the last stage of human colour vision.

Colour is a very important visual cue in human perception. Among the several visual tasks involving colour, one of the most commonly done by humans is colour naming. The aim of this task is, given a region from a scene with a more or less homogeneous colour, to take a decision in natural language about what is the hue or colour that best describes the region. Humans are able to do this task since early ages, it is done frequently and colour names are used without effort. However, the perceptual mechanisms that rule this process are still not completely known as it will be seen in the next chapter.

Colour naming has been studied from very different points of view. For many years philosophers, anthropologists, linguists, physiologists, psychologists and colour scientists have researched on this topic. However, the anthropological study of Berlin and Kay [27] was the starting point of a lot of research about the topic in the subsequent decades. They studied colour naming in different languages and stated the existence

1

of universal colour categories. They also defined the set of 11 basic categories that have the most evolved languages. These are white, black, red, green, yellow, blue, brown, purple, pink, orange and grey. The research on colour naming in the different fields mentioned above is reviewed on Chapter 2.

Colour has also been a very important visual cue on computer vision. Many applications such as object recognition, image retrieval, tracking, or segmentation have used colour information. In computer vision, colour has been normally represented in different spaces that provide a complete representation, but do not easily derive information about how colour is named by humans. Although the high level information that colour names provide would be very useful for computer vision applications, the automatic assignment of names to image regions has not been widely considered up to now.

One of the applications in computer vision where the automation of the colour-naming task may be very useful is Image Retrieval. The amount of multimedia contents to which users have access has incredibly grown in the last years. This fact has brought associated the need of managing such amount of information in a practical and efficient way. To this purpose many image database systems have been developed [106, 110, 116, 62]. Early systems were based on textual annotations associated to images and queries were done by giving a verbal description of the image content. This approach implies a considerable amount of human work to annotate all the images in a database. Moreover, the human subjectivity can cause inconsistencies in annotations. To overcome these problems, a different approach was soon proposed: content-based image retrieval (CBIR).

In CBIR information about image contents, in terms of colour, texture, shape and others, is used to index images. CBIR queries are normally done by providing an image similar to the one the user is searching (i.e. query by example) or by making an sketch of the image content and its spatial distribution (i.e. query by sketch). However, semantic queries are still needed. Users may not have a similar image to the one they are searching, or may want to find images of specific semantic categories (e.g. find images of 'flowers').

Hence, a problem appears at this point. While humans normally use semantic concepts, the information that can be extracted from images is normally based on numerical features. The distance between the high-level semantics of humans and the low-level image features is known as the semantic gap. Therefore, techniques for reducing the semantic gap are needed to achieve CBIR with high-level semantics. Interesting reviews about CBIR can be found in [126] and [95].

At this point, we are able to state the main goal of this thesis. We aim to obtain a computational model of human colour naming able to reduce the semantic gap. A model of colour naming will be very useful to describe image colour contents in terms of natural language. To automate colour naming obtaining the same judgements of a human being, we have to use a learning set of data obtained from human observations. To this end, we present a computational colour-naming model where the eleven basic colour categories proposed by Berlin and Kay are modelled by fuzzy sets. Hence, after defining a fuzzy set framework for colour naming, we develop a psychophysical colour-

naming experiment to obtain an adequate learning set for our purpose. This learning set will be used to fit the membership functions of the fuzzy sets corresponding to the colour categories. Finally, the model will be tested on a real computer vision application of automatic image annotation.

1.2   Colour Spaces

The importance of colour in industry and engineering brought the need of having a way to measure colour. Colorimetry is the field of physics that has mathematically specified colour perception.

The first attempts to describe colour were based on matching experiments with colour lights. In colour matching experiments, a single-wavelength test light, Q, is displayed in one half of a split field on a black background and human observes are asked to adjust the intensities, w_1, w_2 and w_3, of three linearly independent monochromatic primaries, P_1, P_2 and P_3, that are mixed on the other half of the split field to match the test light. These experiments are based on the trichromatic theory [151] (i.e. colour can be described in terms of only three primaries) and the Grassman Laws [63]. The Grassman Laws are:

1. **First Law.** Any light, Q, can be matched by a linear combination of three linearly independent primary lights, P_1, P_2 and P_3:

$$Q = w_1 P_1 + w_2 P_2 + w_3 P_3$$

   where w_1, w_2, and w_3 are the amounts of each primary needed to match the test light.

2. **Second Law.** If two test lights, Q_1 and Q_2, are mixed, the sum of the individual weights that match the test lights matches the mixed light:

$$Q_1 = w_{1,1} P_1 + w_{1,2} P_2 + w_{1,3} P_3 \quad \wedge \quad Q_2 = w_{2,1} P_1 + w_{2,2} P_2 + w_{2,3} P_3$$

$$\Downarrow$$

$$Q_1 + Q_2 = (w_{1,1} + w_{2,1}) P_1 + (w_{1,2} + w_{2,2}) P_2 + (w_{1,3} + w_{2,3}) P_3$$

3. **Third Law.** Colour Matching is linear, that is,

$$Q = w_1 P_1 + w_2 P_2 + w_3 P_3$$

$$\Downarrow$$

$$kQ = (kw_1) P_1 + (kw_2) P_2 + (kw_3) P_3$$

Colour matching experiments were the basis on the definition of the standard colorimetric observers of the CIE (*Comission Internationale de l'Eclairage*).

1.2.1   CIE Standard Colorimetric Observer

In the context described above, where industry needed colour standards, the CIE determined the colour-matching functions, $r(\lambda)$, $g(\lambda)$ and $b(\lambda)$, for a standard observer in 1931.

These functions were obtained by adjusting each primary ($P_1 = 700.0nm$, $P_2 = 546.1nm$ and $P_3 = 435.8nm$) to match a monochromatic source at every wavelength. Hence any light source, $Q(\lambda)$, can be expressed in terms of the three primaries as:

$$Q(\lambda) = RP_1 + GP_2 + BP_3 \qquad (1.1)$$

where R, G and B are the tristimulus coordinates and they are computed from equations (1.2), (1.3) and (1.4):

$$R = \int_\omega Q(\lambda)r(\lambda)d\lambda \qquad (1.2)$$

$$G = \int_\omega Q(\lambda)g(\lambda)d\lambda \qquad (1.3)$$

$$B = \int_\omega Q(\lambda)b(\lambda)d\lambda \qquad (1.4)$$

where ω is the range of wavelengths within the visible spectrum from 380nm to 780nm, and $r(\lambda)$, $g(\lambda)$ and $b(\lambda)$ are the colour matching functions that are shown in figure 1.1.
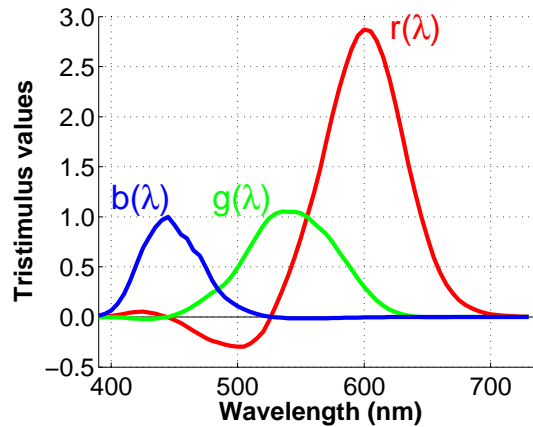
Figure 1.1: Colour-matching functions of the 2-degree CIE standard colorimetric observer in the RGB system.

The computation and use of the RGB colour space present some problems:

- The $\bar{r}(\lambda)$ function has negative values for some wavelengths, which can be an impediment in the design of colour measurement instruments.

- There is no axis directly related to luminosity response function of the visual system.

- The RGB values are not easily associated with perceptual colour attributes such as hue, chroma and lightness.

- Euclidean distances in the space do not correlate with perceptual dissimilarity.

In order to solve the first two of these problems the CIE defined another colour space referred as XYZ, which is a linear trasform of the RGB space. This new system is based on the use of three imaginary primitives that are positive for all λ.

In this new space, the tristimulus values for a given colour, $Q(\lambda)$, are obtained from equations (1.5), (1.6) and (1.7):

$$X = \int_\omega Q(\lambda)x(\lambda)d\lambda \qquad (1.5)$$

$$Y = \int_\omega Q(\lambda)y(\lambda)d\lambda \qquad (1.6)$$

$$Z = \int_\omega Q(\lambda)z(\lambda)d\lambda \qquad (1.7)$$

where ω is the range of wavelengths inside the range of visible spectrum, $x(\lambda)$, $y(\lambda)$ and $z(\lambda)$ are the colour-matching functions for the XYZ space, which define an ideal observer referred to as the CIE Standard Colorimetric Observer (Figure 1.2).
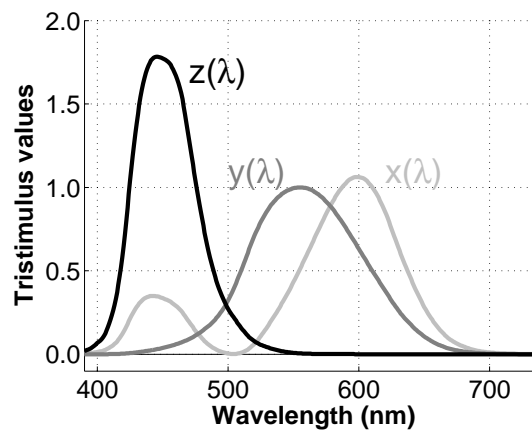
Figure 1.2: Colour-matching functions of the 2-degree CIE 1931 standard colorimetric observer in the XYZ system.

In this new space there is not a perceptual interpretation of the coordinates, with the exception of the approximated relationship between Y-coordinate and colour intensity. Nevertheless, this colour system has been widely accepted as a standard space in colorimetry.

1.2.2 CIE Uniform Colour Spaces

To overcome the problems that XYZ colour space presents, the CIE has defined perceptually uniform spaces, such as the CIELuv and the CIELab systems, where distances correlate with perceptual dissimilarity. Despite it has been shown not to be absolutely perceptually uniform [164], the CIELab space has been widely used in applications where a correlation with human perception is needed. In this thesis the CIELab space will be the work space where the colour-naming model is defined. CIELab colour space is defined as a non-linear transform of the XYZ space ruled by the following equations [157]:

$$L^* = 116 \left(\frac{Y}{Y_n} \right)^{\frac{1}{3}} - 16 \tag{1.8}$$

$$a^* = 500 \left[\left(\frac{X}{X_n} \right)^{\frac{1}{3}} - \left(\frac{Y}{Y_n} \right)^{\frac{1}{3}} \right] \tag{1.9}$$

$$b^* = 200 \left[\left(\frac{Y}{Y_n} \right)^{\frac{1}{3}} - \left(\frac{Z}{Z_n} \right)^{\frac{1}{3}} \right] \tag{1.10}$$

with the constraint that $X/X_n, Y/Y_n, Z/Z_n > 0.01$. X_n, Y_n, and Z_n are the XYZ coordinates of a reference white.

For values of $X/X_n, Y/Y_n, Z/Z_n < 0.01$ the values of L^*, a^*, and b^* can be computed by replacing previous equations by the modified equations:

$$L_m^* = 903.3 \left(\frac{Y}{Y_n} \right) \qquad for \; \frac{Y}{Y_n} \leq 0.008856 \tag{1.11}$$

$$a_m^* = 500 \left[f \left(\frac{X}{X_n} \right) - f \left(\frac{Y}{Y_n} \right) \right] \tag{1.12}$$

$$b_m^* = 200 \left[f \left(\frac{Y}{Y_n} \right) - f \left(\frac{Z}{Z_n} \right) \right] \tag{1.13}$$

where

$$f\left(\frac{X}{X_n}\right) = \begin{cases} \left(\frac{X}{X_n}\right)^{\frac{1}{3}} & \frac{X}{X_n} > 0.008856 \\ \\ 7.787\left(\frac{X}{X_n}\right) + \frac{16}{116} & \frac{X}{X_n} \leq 0.008856 \end{cases} \tag{1.14}$$

$$f\left(\frac{Y}{Y_n}\right) = \begin{cases} \left(\frac{Y}{Y_n}\right)^{\frac{1}{3}} & \frac{Y}{Y_n} > 0.008856 \\ \\ 7.787\left(\frac{Y}{Y_n}\right) + \frac{16}{116} & \frac{Y}{Y_n} \leq 0.008856 \end{cases} \tag{1.15}$$

$$f\left(\frac{Z}{Z_n}\right) = \begin{cases} \left(\frac{Z}{Z_n}\right)^{\frac{1}{3}} & \frac{Z}{Z_n} > 0.008856 \\ \\ 7.787\left(\frac{Z}{Z_n}\right) + \frac{16}{116} & \frac{Z}{Z_n} \leq 0.008856 \end{cases} \tag{1.16}$$

1.2.3 Munsell Colour Space

The Munsell colour system [157] is a perceptually uniform colour space that has been widely used in colour naming research. Although our colour naming model will be defined on the CIELab space, the Munsell notation will be used to test the model and to compare the results to the ones from previous works. Munsell system is represented in a cylindrical coordinate system in which colour is described by three attributes: Hue, Value and Chroma.

Hue is related to the dominant wavelengths of the colour spectrum and it is represented as the angular position in the cylindrical coordinate system. Five principal hues (Red (R), Yellow (Y), Green (G), Blue (B) and Purple (P)) and five half-way hues (Yellow-Red (YR), Green-Yellow (GY), Blue-Green (BG), Purple-Blue (PB), and Red-Purple (RP)) are defined in the system. Each hue is divided in a scale of 10 units.

Value represents lightness in a scale of 10 equally spaced units from 0 (black) to 10 (white). Value is represented in the central vertical axis of the cylindric system.

Chroma corresponds to saturation, that is, the amount of a neutral grey of the same lightness that has the colour. Chroma is represented as the perpendicular distance from the central vertical axis of the cylindrical system. Chroma ranges from C=0 for achromatic colours in the central axis to the maximum producible with pigments. Hence, the maximum possible chroma can be different for different colours.

Colour naming has often been evaluated and studied by means of a representation of the outer surface of the colour solid of the Munsell space which is referred as the Munsell array. In the Munsell array used in colour naming, 320 chromatic chips are represented. Hue scale is sampled at units 2.5, 5.0, 7.5, and 10.0 for each hue. Value is represented in a range from 2 to 9 and each presented chip has its maximum chroma

available. An additional ten-chip column is presented on the left of the chart with the achromatic scale (value from 1 to 10 and chroma equal to 0).

1.3   Image Formation

To deal with the colour-naming task in the frame of computer vision we will consider a simplification of the dichromatic model of Shafer [125] for image formation. It makes us to assume images are formed from an ideal scene where surfaces are Lambertian and uniformly illuminated. Thus the colour of a point is represented as a vector, $\vec{s}$, with components

$$s_i = \int_\omega E(\lambda)S(\lambda)R_i(\lambda)d\lambda \tag{1.17}$$

where s_i represents the ith component of the vector, that integrates over the range of wavelengths, ω, within the visible spectrum. The number of components depends on the number of sensors which are usually based on three primaries: red, green and blue. They depend on each specific device and their sensitivity responses are given by $R_i(\lambda)$ that provide an image in a RGB device-dependent colour space. Finally, $E(\lambda)$ is the spectral power distribution of the illuminant and $S(\lambda)$ is the reflectance function of the point on a surface of the scene.

On this given image we want to simulate the human ability of the colour-naming task, which obviously depends on how colour is perceived by the human visual system. Thus, there is a gap, known as the sensory gap [126], between colour provided by a digital image, and colour perceived by the human visual system. The usual way to fill this gap is to apply an image transform through standard spaces which are based on the spectral sensitivities measured in the human visual system, and which are referred as the standard observer [157], as we detailed in the previous section. The parametric model we present in this thesis will be given in one of these standard spaces, the CIELab space. Hence, to use the colour-naming model on a given image, we will need to know some information on the acquisition system or we will need to perform a calibration process. Depending on what we know about the elements of the acquisition conditions, that is, $R_i(\lambda)$, $S(\lambda)$ and $E(\lambda)$, then we can perform several calibration procedures with different error amounts.

To overcome this problem a standard RGB colour space, denoted as sRGB, has been proposed [7]. This space, as other RGB spaces [137], aims to be a default colour space for multimedia applications by defining the relationship between sRGB values and CIE 1931 XYZ values for a reference display and in reference viewing conditions. The goal is to define a standard for interchange in multimedia. Hence, if the sensor used provides values in the sRGB space, then the transform to CIE standard spaces is straightforward. Otherwise, a calibration process will be needed. In the last years, some devices that provide values on sRGB colour space have been commercialized.

Apart from performing a correct calibration process, other considerations must be taken into account for a naming task on a specific image. The model we present in this thesis has been trained on data derived from psychophysical experiments where a

homogeneous colour area is shown to an observer that has been adapted to the scene illuminant. Therefore, the name assignment has been done under ideal conditions where neither influences from the illuminant nor the surround of the observed area have any effect on the naming process. However, in practice, colour name assignment is a content-dependent task and therefore perceptual considerations about the surround influence must be taken into account.

In figure 1.3, some examples are shown. In the upper part of the figure (images 1.3(a) and 1.3(b)) the same surface can be represented by different RGB values due to an illumination change on the scene. However, in a real scene we would perceive both surfaces pointed by the arrows as being yellow, and we will require an automatic colour-naming system to assign this name to both surfaces. The lower images (1.3(c) and 1.3(d)) present the opposed situation. Although the regions of each image pointed by the arrows have the same RGB values, we perceive them as being of different colours and, thus, we require the model to assign different colour names to the same stimulus depending on the surround of the region.
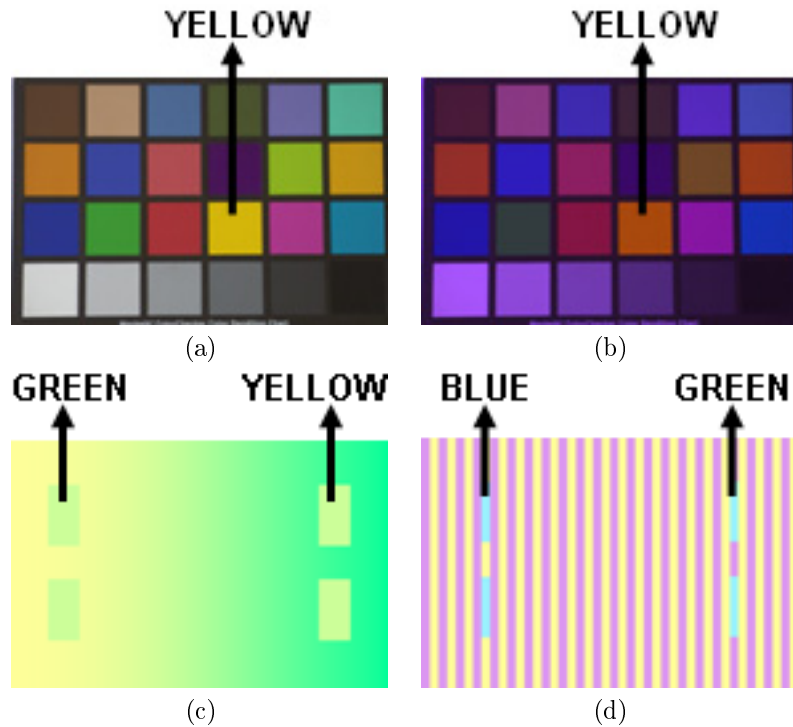
Figure 1.3: Examples of context dependence of the colour-naming task. The labels shown correspond to the names that should be assigned after perceptual considerations. (a) and (b) The blue cast in (b) is removed to assign the same name as in (a). In (c) and (d) the same stimuli are assigned different names due to the surround influences.

The model we propose assumes to work on perceived images, that is, images where

the effects of perceptual adaptation to the illuminant and to the surround have been previously considered. In the following section, we explain in detail these effects and review some of the approaches that have been proposed in the computer vision literature to deal with them.

1.4 Colour Constancy and Colour Induction

If we aim to obtain the same name assignments a human observer would provide from an automatic colour-naming system, the perceptual effects cited above, namely, colour constancy and colour induction, must be taken into account.

In this thesis we focus on the name assignment decision on the perceived images and therefore, the perceptual considerations detailed in this section must be considered in a preprocessing step before applying the model on real images. The solutions that have been adopted to solve these perceptual issues in order to apply the model on real computer vision problems will be explained later in Chapter 5.

1.4.1 Colour Constancy

As it has been presented in the previous section, the RGB values on a colour image are related to the illuminant spectral composition. Thus, the same surface may present very different appearances under different illuminants or under different intensities of the same illuminant. Figure (1.4) shows an example of the variation that the same scene can suffer when there are changes in the illuminant characteristics. However, a human observer will still be able to infer that the colours of the two kind of peppers in figure 1.4 are yellow and red. That is because the human visual system has an adaptive mechanism that allows us to avoid the spectral variations of the scene light and assign stable colour names to the surfaces. This perceptual ability is called colour constancy.

Figure 1.4: Two images of the same scene under different illuminants. The RGB values acquired by a camera will be completely different in the two images.

The way the human visual system performs its colour constancy ability is still

unknown. The most accepted hypothesis is Von Kries model [156], also called the coefficient rule, which suggests that a change on the chromatic adaptation is caused by the sensitivity reduction or extension in each cone type and without affecting their relative spectral sensitivity. Since standard cameras do not have colour constancy ability, the illumination variability is one of the main problems in computer vision. Small changes in the illuminant intensity or colour may dramatically decrease the performance of RGB-based colour algorithms.

The colour constancy ability presented by the human visual system has been modelled in computer vision by different colour constancy methods. These methods have focussed on recovering the illuminant of the scene when the image was acquired. There are many algorithms to solve this problem [35, 86, 98, 54, 46, 37, 50, 52, 140],

Hence, the application of a colour constancy algorithm can provide images under a reference illuminant, thus simulating an adaptation process to the illuminant. Interesting comparatives between colour constancy methods can be found in the literature [56, 10, 11, 71].

1.4.2 Colour Induction

Human perception of colour at a certain region is influenced by context and texture in the scene. Significative changes on colour naming for different surround conditions have been found on psychophysical experiments [111]. The mechanisms of colour induction that the human visual system has, include colour assimilation and colour contrast. Both effects depend on the colour surrounding a certain stimulus, but the chromatic change is opposed. See figure 1.5 for examples of these phenomena.

(a)   (b)
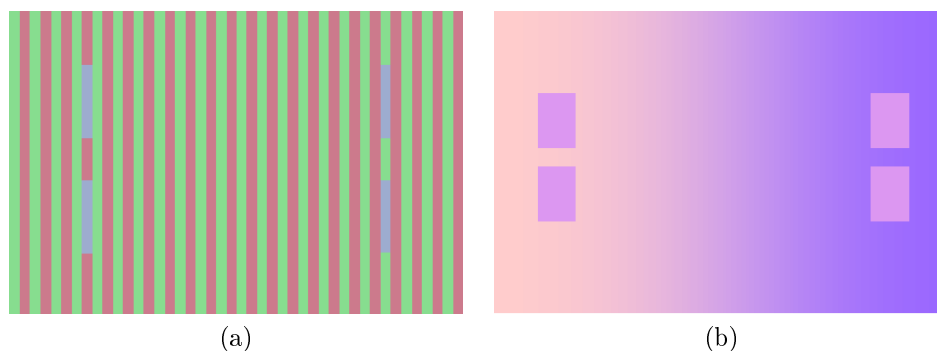
Figure 1.5: Examples of colour induction. (a) Colour assimilation. (b) Colour contrast.

Colour assimilation occurs when our perception of the chromaticity of a stimulus changes towards the chromaticity of the colour surrounding it. This phenomenon will normally happen when we are observing a textured surface with a high spatial

frequency. In computer vision, colour assimilation has been modelled as a blurring effect [149, 117].

Colour contrast happens when the chromaticity of the stimulus changes away from the chromaticity of the surround. This effect appears when we are observing a surface with low spatial frequencies. In computer vision, colour contrast has been modelled as a sharpening operation [149, 112].

Once the most important colour induction phenomena have been explained, we assume that before we apply our model of colour naming, the input image will have been pre-processed with any of the algorithms that model these phenomena, that is, we assume that the input image in our model will take into account these perceptual issues.

1.5   Problem Definition: Colour Naming as a Fuzzy-set Problem

The basis of most of the works on colour naming has been the study of Berlin and Kay [27] in which they stated the existence of a unique and common set of eleven basic colour terms in different languages. In the way of explaining the colour-naming process, Kay and McDaniel [78] proposed a general model of colour naming which attempted to find the relationship between the neuro-physiological mechanisms involved in colour naming and the semantic categories of basic colour terms. The model is based on fuzzy set theory [81] where each colour category has a characteristic function which defines a membership degree to the category. The most interesting of the model is that it considers the colour-naming problem as something more than the assignment of a colour term to a stimulus, since the fuzzy approach takes into account the non-discrete nature of the problem.

In this thesis we propose to model the colour-naming task on the frame of fuzzy-sets theory. A fuzzy set is described by its membership function. In colour naming, we can consider that any colour category, C_k, is a fuzzy set with a membership function, μ_{C_k}, which assigns, to any colour sample $\vec{s}$ represented in a certain colour space, a membership value $\mu_{C_k}(\vec{s})$ within the $[0,1]$ interval. This value represents the certainty we have that $\vec{s}$ belongs to category C_k, and therefore has to be named with its corresponding linguistic term t_k. The essential contribution of this thesis will be the proposal of a parametric model based on sigmoid functions to represent the membership of these fuzzy sets.

According to the fuzzy sets theory, if n categories are considered, the membership functions must fulfil the constraint:

$$\sum_{k=1}^{n} \mu_{C_k}(\vec{s}) = 1 \qquad \text{with} \qquad \mu_{C_k}(\vec{s}) \in [0,1], \quad k = 1, \ldots, n \qquad (1.18)$$

Hence, for any given colour sample $\vec{s}$ it will be possible to compute a colour

descriptor, $\mathcal{CD}$, such as

$$\mathcal{CD}(\vec{s}) = (\mu_{C_1}(\vec{s}), ..., \mu_{C_n}(\vec{s})) \tag{1.19}$$

where each component of this n-dimensional vector describes the membership of $\vec{s}$ to a specific colour category.

The information contained in such a descriptor can be used by a decision function, $\mathcal{N}(\vec{s})$, to assign the colour name of the stimulus $\vec{s}$. The most easy decision rule we can derive is to choose the maximum from $\mathcal{CD}(\vec{s})$:

$$\mathcal{N}(\vec{s}) = t_{k_{\max}} \quad | \quad k_{\max} = \arg\max_{k=1,...,n} \{\mu_{C_k}(\vec{s})\} \tag{1.20}$$

where t_k is the linguistic term to name the colour category C_k.

In our case the categories considered are the basic categories proposed by Berlin and Kay, that is, $n = 11$ and the set of categories is:

$$C_k \in \{Red, Orange, Brown, Yellow, Green, Blue, Purple, Pink, Black, Grey, White\} \tag{1.21}$$

For implementation issues, chromatic categories[1] in equation (1.21) are anticlockwise ordered according to their location on a chromaticity plane starting on Red and ending at Pink, and then, achromatic categories[2] are ordered according to their lightness, from dark to bright.

1.6 Thesis Outline

This thesis has been organized in six chapters. In Chapter 1 we have presented the goals of the thesis and have defined the framework in which this thesis will be developed.

In Chapter 2 a bibliography review about the topic of colour naming is presented in order to give the necessary background to study this task. The main contributions from different fields are discussed, giving special attention to those which are nearer to our goal of colour-naming automation.

Chapter 3 is devoted to psychophysical experimentation. In the first part, a review of the different methodologies used so far is presented to show the lack of an adequate previous experiment for fuzzy modelling of colour naming. In the second part of the chapter, a new methodology is proposed and the results of a psychophysical experiment following the proposed methodology are presented. These results are analyzed and compared to previous works to show its validity for our purpose.

Chapter 4 begins with a discussion on the advantages that parametrical models present in front of non-parametrical models to model the colour-naming task. In the second part of this chapter, four approaches to colour-naming modelling are proposed

---

[1] Red, Orange, Brown, Yellow, Green, Blue, Purple, and Pink
[2] Black, Grey, and White

before presenting the main contribution of this thesis, the Triple-Sigmoid with Elliptical centre (TSE) model, which is developed and discussed in the last part of the chapter.

Chapter 5 considers the issues related to the application of the proposed TSE model to real images. These issues are colour space transforms, colour constancy and colour induction effects. A modular scheme to take these topics into account is proposed and colour-naming descriptors for images are proposed. In the last part of the chapter, the model is tested on a real problem of automatic annotation for image retrieval to show its usefulness for real applications in computer vision.

Finally, in Chapter 6, the results of this thesis are discussed and the contributions of the thesis are summarized. To end, the future research lines that can be derived from the thesis are explained.

Chapter 2

Colour Naming Background

Colour naming, or colour categorization, has been studied from different disciplines. For many years, psychologists, linguists, anthropologists, physicians and engineers have worked about this topic providing very different points of view about the problem. In this chapter we summarize the main contributions to the topic from these different fields.

2.1 The Human Visual System and Colour Naming

Colour perception phenomenon begins when a luminous stimulus reaches the retina. However, only a part of the whole radiant energy that receives the retina causes a visual stimulus. This is what is called the visible spectrum referring to those signals containing wavelengths between 380nm and 780nm. Outside this band of frequencies the human visual system has no sensitivity. The way in which the human visual system (HVS) processes these visual stimuli that reach at the eye has been the topic of extensive research during the last two centuries.

In 1802, the physician Thomas Young was the first in suggesting that colour perception involved only three fundamental mechanisms with different sensitivities [161]. It was only an hypothesis, but it was the start point of the trichromatic theory. Some years later, von Helmholtz [151] proposed the existence of three different channels to process colour.

Another important contribution to colour theory was done by Grassmann who in 1853 proposed that colour matches were based on linear operations [63]. Since then, his laws have been the basis of psychophysical research about colour matches.

These first contributions to the understanding of the mechanisms involved in colour perception helped to develop the trichromatic theory of colour perception. According to it, colour perception is based on the information provided by only three sensors with different spectral sensitivity.

At present, we know the retina contains two kinds of receptors: rods and cones. While rods are sensitive to low intensity stimulus, cones give response to higher intensities and also provide the information used by the HVS to perceive colour.

Colour perception is related to the presence in the cones of three different pho-

topigments. Depending on these photopigments there are three types of cones. Each
of them presents different sensitivity and concentrates its response on certain wave-
lengths. Thus, the three cone types are called L (*large*), M (*medium*) and S (*small*)
related to the spectral wavelength where they present maximum sensitivity. Cones
have also been called *red*, *green* and *blue* respectively, although the maximum sensi-
tivity wavelength does not exactly correspond to these colours.

The spectral sensitivity curves of the three cone types in the retina were first
measured by König and Dieterici [82] in 1886. Since then the three cone absorptions
have been measured by different authors and with different techniques [152, 127]. In
all cases, results were very similar [88]. The most recent estimates were measured by
Stockman and Sharpe [133, 132]. Figure 2.1 shows these sensitivity curves.
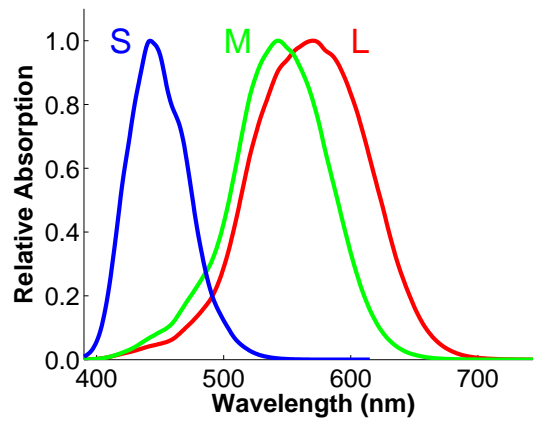
Figure 2.1: Sensitivity curves of the three cone types of the human visual system
estimated by Stockman and Sharpe [133, 132].

Hence, the information contained by the spectral distribution of any input stimulus
is reduced to a three values representation by the HVS. The output of the cones has
no wavelength information since each cone type has response to a wide range of wave-
lengths. The magnitude of the cone responses does not bring wavelength information
either. An increase in the output of a cone can be due to a change of the stimulus
to a wavelength for which the cone has more sensitivity or due to an increase in the
intensity of the stimulus. Since human beings are able to distinguish colours at dif-
ferent wavelengths it seems obvious that the HVS must perform additional processes
at later stages.

There are psychophysical [69] and physiological [41] evidences that there is a sec-
ond level of processing based on colour differences. These evidences include the fact
that while we are able to perceive some colours as a mixture of two other, e.g. orange
as a mixture of red and yellow, we are not able to perceive any colour as a mixture
of certain colours, e.g. there is no colour perceived as a mixture of red and green.
An opponent colour mechanism was proposed by Hering [69]. According to Hering's
theory, there are three opponent channels with opposed responses to pairs of stimulus.
These pairs are red-green, blue-yellow and white-black. His opponent theory was first

considered as antagonistic and excluding to the trichromatic theory. However, some years later Hurvich and Jameson [73] revised the opponent theory and proposed the opponent mechanism as a second stage of the colour processing. They also measured the responses of the three opponent channels by psychophysical experimentation. The opponent mechanisms were also quantified by Ebenhöh and Hemminger some years later [45].

According to several studies [42, 128, 129], this opponent behavior takes place in the retinal ganglion cells and the Lateral Geniculate Nucleous (LGN). Two main interactions between the cones responses are involved in this stage. De Valois et al. [41] identified four different types of spectrally opponent cells in the LGN of the macaque monkey. One type is excited by a red signal and is inhibited by a green signal, while a second type has the opposite behaviour. This pair is normally referred as the "RG" system. The two other types of opponent cells have the same behavior as the "RG" system but with yellow and blue signals. This pair is referred as the "YB" system. The system is completed by two types of spectrally non-opponent cells which are only sensitive to intensity changes and that are referred as the "$Wh-Bl$" system. Each one of the opponent channels would receive different inputs from the cones in the retina. Several possibilities for the connections from cones to opponent cells have been proposed [88].

Krauskopf et al. [83] also described three independent channels in colour vision by means of psychophysical experimentation. They called these channels cardinal directions, which included a luminance direction and two colour-opponent directions. The novelty of this work is that the directions found were different from the red-green and blue-yellow axis of the Hering theory. De Valois and De Valois [43] proposed a multi-stage model that integrates both opponent axis proposals. According to this model, whereas the cardinal directions of Krauskopf et al. are directly related to the cone outputs on a second processing stage, Hering's opponent axis correspond to a higher stage.

The last step in colour processing is performed in the visual cortex. The knowledge about this stage of colour perception is considerably smaller than about the first stages explained previously [87]. In fact, as De Valois and De Valois stated in [43], "one can thus cite some cortical study in support of (or against) almost any suggestion about cortical color processing".

De Valois and De Valois [42] proposed that information in the cortex goes in two directions. Whereas one would include colour-specific channels that separate luminance and colour information, the other would go into multiple-colour channels with cells that use colour to extract form information. The existence of a specific area in the visual cortex, frequently referred to as the "Colour Center", exclusively devoted to colour processing has been the topic of many research in the last decades. Zeki [162] reported an area in the macaque monkey cortex, termed V4, formed only by colour-specific cells that responded to limited spectral regions. However, Schein et al. [121] reduced the estimate of colour-specific cells in V4 to only the 20% of the total population of cells. In humans, brain imaging studies [100] have revealed an analogue area, human V4, with high activity in different colour vision tasks. Hadjikhani et al. [67] also found such area with high activity for colour tasks, but they considered it was different to monkey V4 and called V8 in humans. Zeki [163] insisted on that

both areas were the same and should be called V4. However, although these works maintain the existence of the "colour center" in the cortex, other works argue that such center processing only colour information does not exist [58].

Colour naming is one of the visual tasks involving colour that is supposed to be done in the last stages of processing. The way in which the visual system manages the information from the opponent-colour system to codify the colour names and label each stimulus with its corresponding name is not known yet. Although, it seems reasonable that the "$Wh - Bl$" system provides the luminance information, and that the chromaticity information from "RG" and "YB" systems is used to codify hue and saturation, the mechanisms that rule this process are unknown. However, there have been some attempts to link the responses of the opponent system to the colour names assigned by humans [42]. A recent psychophysical experiment [111] found that this relationship could be modelled with a network based on very simple operators.

After this brief review on the mechanisms of human colour vision the main conclusion to which we can come is that there is still a long way to know the mechanisms of the whole process. Although much progress has been done on the understanding of the first stages, a lot of the last stages is still unknown. A good revision of the last advances in this area can be found in [60, 59, 129].

2.2   Basic Colour Terms

Colour naming, and all the semantic fields in general, have been involved for many years in a discussion between two points of view in linguistics. On the one hand, relativists support the idea that semantic categories are conditioned by experience and culture, and therefore, each language builds its own semantic structures in a quite arbitrary form. On the other hand, universalists defend the existence of semantic universals shared across languages. These linguistic universals would be based on the human biology and directly linked to the neurophysiological mechanisms. Colour has been presented as a clear example of relativism since each language has a different set of terms to describe colour.

Although some works had investigated the use of colour terms in English [97], the anthropological study of Berlin and Kay [27] about colour naming in different languages was the start point of many works about this topic in the subsequent years.

Berlin and Kay studied the use of colour names in speakers of a total of ninety-eight different languages (20 experimentally and 78 through literature review). With their work, Berlin and Kay wanted to support the hypothesis of semantic universals by demonstrating the existence of a set of colour categories shared across different languages. To this end, they first defined the concept of "basic colour term" by setting the properties that any basic colour term should fulfil. These properties are:

- It is monolexemic, i.e. its meaning can not be obtained from the meaning of its parts.

- It has a meaning which is not included in that of other colour terms.

- It can be applied to any type of objects.

- It is psychologically salient, i.e. it appears at the beginning of elicited lists of colour terms, it is consistently used along time by speakers and across different speakers, and it is used by all the speakers of the language.

In addition, they defined a second set of properties for the terms that might be doubtful according to the previous rules. These properties are:

- The doubtful form should have the same distributional potential as the previously established basic terms.

- Basic colour terms should not be also the name of an object that has that colour.

- Foreign words that have recently been incorporated to the language are suspect.

- If the monolexemic criterion is difficult to decide, the morphological complexity can be used as a secondary criterion.

The work with informants from the different languages was divided in two parts. In the first part, the list of basic colour names in each informant's language, according to the previous rules, was verbally elicited. This part was done in the absence of any colour stimuli and using as little as possible of any other language. In the second part, subjects were asked to perform two different tasks. First, they had to indicate on the Munsell colour array all the chips that they would name under any condition with each of their basic terms, i.e. the area of each colour category. Second, they had to point out the best example (focus) of each basic colour term in their language.

Data obtained from the 20 informants was completed with information from published works in other 78 languages. After the study of these data, Berlin and Kay extracted three main conclusions from their work:

1. *Existence of Basic Colour Terms.* They stated that colour categories were not arbitrary and randomly defined by each language. The foci of each basic colour category in different languages were all in a close region of the colour space. This finding led them to define the set of eleven basic colour terms. These terms for English are white, black, red, green, yellow, blue, brown, pink, purple, orange and grey.

2. *Evolutionary order.* Although languages can have different number of basic colour terms, they found that the order in which languages encoded colour terms in their temporal evolution was not random, but it followed a fixed order that defined seven evolutionary stages:

   - Stage I: Terms for only white and black.
   - Stage II: A term for red is added.
   - Stage III: A term for either green or yellow (but not both) is added.
   - Stage IV: A term for green or yellow (the one that was not added in the previous stage) is added.
   - Stage V: A term for blue is added.

- Stage VI: A term for brown is added.
- Stage VII: Terms for pink, purple, orange and grey are added (in any order).

These sequence can be summarized with the expression:

$$[white, black] < [red] < [green, yellow] < [blue] < [brown] < [pink, purple, orange, grey]$$

where symbol '$<$' indicates temporal precedence, that is, for two categories A and B, $A < B$ means that A is present in the language before B, and order between terms inside '[ ]' depends on each language.

3. *Correlation with cultural development.* They noticed a high correlation of colour vocabulary of a language with technological and cultural evolution. Languages from developed cultures were all in the last stage of colour terms evolution, while languages from isolated and low-developed cultures were at lower stages of colour vocabulary evolution.

Figure 2.2 shows the boundaries in the Munsell space of the eleven basic colour categories for American English that were obtained by Berlin and Kay. This categorization of the Munsell array has been used as a reference in later colour-naming studies [31, 135] and will also be used to evaluate the model we propose later in Chapter 4.
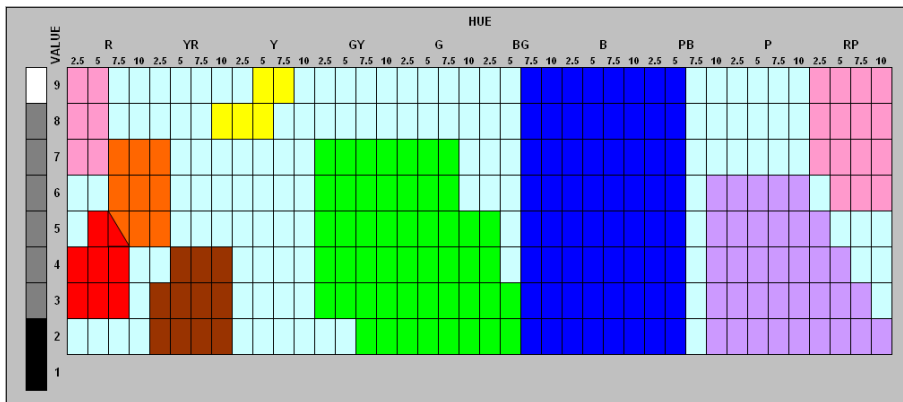
Figure 2.2: Categorization of the Munsell colour array obtained by Berlin and Kay in their experiments for American English.

The findings of Berlin and Kay have been questioned on some works. Schirillo [122] accepts the universality of the six Hering primaries (Red, Green, Blue, Yellow, Black and White) but questions the validity of Berlin and Kay theory for the other five basic colour categories (Brown, Orange, Purple, Pink, and Grey). Schirillo suggests that these categories might be more culturally determined, thus accepting up to a certain point, the relativist theory. A complete opposition to the universality of

colour categories is defended by Roberson et al. in [120], where colour categorization is affirmed to be totally conditioned by cultural experience. For Roberson and her colleagues the organization of colours into categories is based on language instead of on the neurophysiology of the visual system.

Despite these discrepancies, the universalist theory of Berlin and Kay about the existence of semantic universals for colour names is the most accepted one and has been supported by many later results that will be reviewed in the following section.

2.3   Colour Naming in Psychophysics

After the important study of Berlin and Kay about colour naming, many contributions in the field of psychophysics have been done. Most of this research has been focused on the discussion about the existence and the location of the basic colour terms. In this section we review the most relevant works on the colour-naming task and discuss on the results and contributions of these works. However, we will not analyze the methodologies used to obtain the results we summarize in this section because they are explained in the next chapter.

The results of the work of Berlin and Kay have been analysed in several later works. Boynton and Olson performed an experiment [31] to locate the regions occupied by the basic colour categories in the OSA space [107]. In this experiment, subjects were presented with a wide set of colour samples and they were asked to give a monolexemic colour name to the sample. The results of the experiment showed that basic colour terms were used by the subjects of the experiment more frequently than non-basic colour terms and that basic colour terms had a lower mean response time than non-basic colour terms. With their experiment, Boynton and Olson confirmed the existence of the eleven basic colour terms proposed by Berlin and Kay. Moreover, they considered that these results are the confirmation that the link between basic colours and their names was congenital and physiologically based. These results were confirmed in a later study by the same authors [32] where three measures on the results of a similar experiment showed the saliency of basic colour names over non-basic colours. The analysis was focused on chromatic basic colour names and showed that all chromatic basic colour names were used faster, more consistently and with more consensus than non-basic colour names.

The previous experiments were replicated in the Munsell space by Sturges and Whitfield in [135] and [136]. The results agreed with Berlin and Kay theory and confirmed the existence and location of the eleven basic colour terms proposed by Berlin and Kay. Again, the eleven basic colour terms were used faster, more consistently and with more consensus than the rest of colour terms. The similarity of the results in the two colour spaces, the OSA space and the Munsell space, was also shown in a comparative study [30]. A later experiment with a completely different technique [105] located the centroids of basic colour categories in similar positions of the colour space. All these results seem to indicate that the location of basic colour categories does not depend on neither the colour space nor the technique used in the experiments.

The prevalence of basic colour terms in front of non-basic colour terms was also

confirmed by Guest and Van Laar [65]. They made a further step by applying the Principal Component Analysis (PCA) technique from three simple measures obtained from the experiment (response times, confidence ratings and consistencies). The result was a single measure of the naming easiness, which they called "nameability". Hence, each point of the colour space can be assigned with a nameability value and thus obtain a map of the colour space that represents which areas were areas of no confusion about the name to be assigned and which areas were hard to name. In a later experiment, Guest and Van Laar [66] studied the correlation between nameability and performance on a naming task. They found that observers performed better when a set of highly nameable colours was used than when a different set of colours with very similar metric characteristics but without the nameable property was used.

A different group of experiments were designed to study colour naming across different languages and cultures. Uchikawa and Boynton [144] repeated with Japanese speakers the experiment done with American speakers [31]. The comparison of the results showed that both languages had the eleven basic colour terms proposed by Berlin and Kay, and also showed that the colours these terms described were the same in both languages, which agreed with the thesis of Berlin and Kay. Lin et al. made comparative experiments [89, 90] with English and Taiwanese Mandarin speakers in order to evaluate the differences between the use of colour names in both languages. The results demonstrated again that there were small differences between these languages and that the use of basic colour terms was equivalent in such different languages. Once again, the results of Berlin and Kay were reaffirmed. However, in these experiments, the authors found some differences in the use of modifiers (light, deep, etc.) which they attributed to cultural differences.

As we have seen, the existence and universality of the eleven basic colour terms has been widely supported from psychophysics. However, some results in the opposed direction have been reported. In [74] and [75], although the existence and importance of the eleven basic colour terms is accepted, the linkage of these colour sensations and the neurophysiology of the visual system is questioned.

Nonetheless, Berlin and Kay results are widely accepted and, as we have seen in this section, they are the basis of present studies on colour naming.

2.4   A Neurophysiological Model of Colour Naming

Kay and McDaniel [78] proposed a model for colour naming based on the neurophysiology of the human visual system. In their work, they defended that colour naming is not a discrete process, but a continuous process. This proposal was in contradiction with a widely extended theory in linguistics that stated that the semantic primes were discrete entities.

Kay and McDaniel were the first in posing the problem of colour naming as a fuzzy-set problem and defined each colour category as a fuzzy set with a membership function. These fuzzy sets were directly derived from neural response functions of the four opponent colours proposed by Hering. To do this, they identified the neural responses of opponent cells defined by De Valois et al. [41] with the semantic categories. Hence, they defined four fundamentals (red, green, blue and yellow) which would

have a neuronal basis. Membership to these fundamental categories were computed as the proportion on the total chromatic response [155] contributed by each opponent state. The values of membership for each wavelength define the fuzzy sets for the four fundamental categories. Two additional categories (white and black) are modelled in terms of percentage of reflectance. The membership degree to White is computed as the Munsell value (brightness) divided by ten (Munsell value scale runs from one to ten). Membership degree for black is modelled as unity minus the membership to white

After defining these six fuzzy sets for the primary colour categories, they state that the semantic categories corresponding to all the basic colour terms can be defined as:

- one of the four fundamentals (Red, Green, Blue, and Yellow).

- one of the two additional fundamentals (White and Black).

- fuzzy unions among these six primary categories.

- simple functions of fuzzy intersections among these six primary categories.

Basic colours of languages with few basic colours, i.e. languages in the first stages defined by Berlin and Kay [27], are defined as fuzzy unions. Some of these languages have a colour category corresponding to the *cool* areas of the colour space, i.e. green and blue for a language with 11 basic colour terms. This category, which is sometimes referred as *Grue*, can be defined as the fuzzy union of Blue and Green fuzzy sets. Similarly, a category referred as *Warm* can be defined as the fuzzy union of red and yellow, *Light-warm* as the union of White, Red and Yellow, and *Dark-cool* as the union of Black, Blue and Green.

The rest of the basic colour terms are defined as fuzzy intersections of the six fundamentals. Hence, Brown is defined where Yellow and Black overlap, Pink where Red and White overlap, Purple in the region where Red and Blue overlap, Orange where Red and Yellow overlap and Grey where Black and White overlap. The fuzzy intersections used to define these categories are sometimes completed with non-fuzzy operations (multiplications and subtractions) in order to obtain fuzzy sets coherent with the fuzzy framework defined. Table 2.1 shows a summary of the fuzzy sets defined by the model of Kay and McDaniel.

The most important point of this work is that it changes the vision of colour naming as a discrete problem, where colour categories were defined by focus and boundaries, to a fuzzy vision that takes into account the continuous nature of colour naming. This point of view has been the framework for our work and for many other approaches on the field of computer automation of the colour-naming task.

2.5   Computational Models of Colour Naming

A completely different point of view to the problem of colour naming is that of computer vision. In the last two decades, many attempts to automate this task have been done. Although the high level information that colour names provide has not been widely used in computer vision until recently, the interest for colour naming

---

Fundamental categories: fuzzy sets based from neural responses

f_{Black}
f_{White}
f_{Red}
f_{Yellow}
f_{Green}
f_{Blue}

First-Stages Categories: fuzzy sets based on fuzzy union

$f_{Dark-cool}$	$=$	$f_{Black\ OR\ Green\ OR\ Blue}$
$f_{Light-warm}$	$=$	$f_{White\ OR\ Red\ OR\ Yellow}$
f_{Warm}	$=$	$f_{Red\ OR\ Yellow}$
f_{Cool}	$=$	$f_{Green\ OR\ Blue}$

Basic Colour Categories: fuzzy sets based on fuzzy intersection

f_{Brown}	$=$	$f_{Black\ +\ Yellow}$
f_{Purple}	$=$	$f_{Red\ +\ Blue}$
f_{Pink}	$=$	$f_{Red\ +\ White}$
f_{Orange}	$=$	$f_{Red\ +\ Yellow}$
f_{Grey}	$=$	$f_{White\ +\ Black}$

Table 2.1: Summary of the fuzzy sets defined in the model of Kay and McDaniel. Symbol '+' indicates a fuzzy intersection, sometimes along with one or more non-fuzzy operations.

is increasing. In this section we review the most important contributions from the computer vision field.

This review is divided in two parts. The first one is concerned with those models that have not considered the degree of uncertainty that colour naming involves. The second part is devoted to the methods that have used the fuzzy framework to take into account the non-discrete nature of the problem.

2.5.1   Non-fuzzy Colour-Naming Models

In the decade of the thirties some fields, such as industry, arts and science, brought the need of unifying the vocabulary used to describe colour and soon standards for this field were required. The ISCC-NBS system [79, 80] was a first attempt to standardize the use of colour names. It was proposed by the U.S. National Bureau of Standards (NBS) in 1939 following the recommendations of the Inter-Society Color Council (ISCC). The system defines a set of valid terms and modifiers which can be combined to obtain a final name. Five degrees of lightness are allowed (*very light, light, medium, dark* and *very dark*). Saturation can be described with four adjec-

tives (*grayish, moderate, strong* and *vivid*). Three additional terms substitute certain lightness-saturation combinations (*pale* for *light grayish*, *brilliant* for *light strong*, and *deep* for *dark strong*). The preceding terms are used as modifiers to the set of valid hues: *pink, red, orange, brown, yellow, olive, green, blue, violet, purple*, and combinations of them formed by two hues or by adding the *-ish* suffix to one of them. Finally, the three achromatic hues *white, grey*, and *black* are also included. Although the main principle of the system is quite simple, the rules to combine the different elements of the lexicon are confusing and the lexicon results a bit difficult to use. Hence, for example, not all of the possible hues, cover the full range of lightness and saturations. The complete system includes a set of 312 possible colour names whose representatives in the Munsell space are listed in [108].

Many years later, Berk et al. [26] proposed the Colour-Naming System (CNS) which aimed to solve the problems presented by the ISCC-NBS lexicon. In CNS, the syntax of the system is simplified by defining a set of rules that can be summarized as a grammar. Lightness and saturation terms are the same as in the ISCC-NBS system, with the exception of the additional terms *brilliant, deep* and *pale* that are not included in CNS. Hue description includes seven chromatic terms (*red, orange, brown, yellow, green, blue*, and *purple*) and three achromatic terms (*white, grey*, and *black*). The system can generate 627 different colour names, but only 340 distinct colours of these 627 can be found in the the Munsell Book of Colour [6]. The system is tested with humans on a task of colour description in three colour notation systems (CNS, RGB, and HSV). Although humans achieve better results in CNS, i.e. colour is easier to describe in CNS than in the other two, this colour-naming system is still not much intuitive. For example, some of the possible colour names in CNS are not realizable and some names (e.g. very light grayish greenish-blue) can result a bit difficult to understand and interpret.

Although neither of both systems were originally implemented for image applications, they could be easily used for this purpose. First, a representative should be defined for each colour term of the system, and then any arbitrary sample from an image could be assigned with a colour name by applying a nearest neighbour classifier similarly as it has been done in a later system [103].

In the early 80's, with the development of CRT displays and graphical interfaces, the importance of using the high level information from colour names in computer applications was soon appreciated. The first computational models of colour naming that were developed were based on colour vocabularies and the division in blocks of the colour space. The classification of stimuli was done according to that division, that is, the sample to be named was mapped on the colour space and was assigned with the colour name of the block that included the sample. These models were not fuzzy and therefore they were only a tessellation of the colour space.

Following this approach, Tominaga [139] defined a colour-naming system in which the Munsell space is divided into a set of blocks. Each block represents a colour name and its boundaries are specified by the coordinates that limit each block. The system defines four levels of naming accuracy and the rules for each level in order to obtain a vocabulary that overcomes the problems of the ISCC-NBS and the CNS systems. The levels of the system and the valid terms for each level are:

Level 1 encodes 16 colour terms: *red, yellow, green, blue, white, black, grey, pink,*

orange, brown, purple, olive, yellow-green, blue-green, violet, and *red-purple.*

Level 2 includes nine additional terms such as *beige, sky, lilac* or *lavender,* that are added to the 16 terms of level 1.

Level 3 includes nine tone modifiers: *pale, light, bright, vivid, deep, dark, dull, light grayish,* and *grayish.*

Level 4 includes 3 additional adjectives for tone modifiers (*dark grayish, soft,* and *strong*) and 8 hue shades (*reddish, yellowish, greenish, bluish, purplish, pinkish, brownish,* and *olive*).

The basic order to create colour names in this system is:

[modifier on tone]+[modifier on hue]+[basic name of Level 2]

and the exact rules of application were defined as a grammar.

Although the system was proposed to overcome some of the problems of the previous colour-naming vocabularies, it has the same drawbacks. The resulting set of colour names is too extensive and some of the possible names are difficult to interpret. Moreover, the division of the colour space in the 236 colour blocks seems a bit arbitrary. Anyway, the system is interesting since it was tested in images to describe its colour content, and an algorithm with the necessary steps to use the system was defined.

The same idea of dividing the colour space and mapping a sample to assign a colour name is proposed by Lin et al. in [91]. They consider the 11 basic colour categories and their boundaries in the CIELab space were adjusted to minimize the number of wrongly named samples from data obtained in psychophysical experiments [89, 90]. Later, Wang et al. [154] improved the model with data from a more controlled experiment and the model was extended to the CIECAM02 space. The interesting point of this approach is that the division of the space is based on data from psychophysical experiments.

A different approach to the problem was proposed by S.N. Yendrikhovskij [159, 160]. This model assumes that colour categories depend on both the experience of observations of the real world and the physiology of the human visual system. Colour categories are modelled by applying the k-means algorithm on a sample of 10.000 pixels from a wide set of natural images. To name an arbitrary sample, the nearest neighbour rule with the centers obtained from the k-means algorithm is used. An interesting result of this method is that it allows to simulate the order of appearance of colour categories which agrees quite well with the order proposed by Berlin and Kay [27]. The formation of colour categories was also simulated on a community of agents by Belpaeme [15, 14].

2.5.2   Fuzzy Colour-Naming Models

In a completely different line, we can consider the works that have taken the fuzzy nature of the problem into account and have developed fuzzy models for colour naming. From this point of view, given a sample, it is not assigned a unique colour name but it is assigned a membership value to each colour category. In this section, the most important contributions from this point of view are reviewed.

Univariate Gaussian Model

Lammens [85] modelled the 11 basic colour categories with a variant of the Gaussian normal distribution.

$$G_n(\vec{s};\vec{m},\sigma) = e^{-\frac{1}{2}\left(\frac{\sqrt{\frac{\sum_{i=1}^{N}(s_i-m_i)^2}{\sigma}}}{\sigma}\right)^2} \tag{2.1}$$

where N is the dimension of the colour space, $\vec{m}$ is the mean and σ is the variance.

For each colour category, the parameters of the model ($\vec{m}$ and σ) are fitted such that the foci of Berlin and Kay results have membership 1 and the boundaries have membership 0.5. The model is fitted in three different colour spaces (CIELab, CIE XYZ and a neuro-psychophysical (NPP) space) and the best results are obtained in the CIELab space. Figure 2.3 shows the categorization of the Munsell array obtained with Lammens' method.
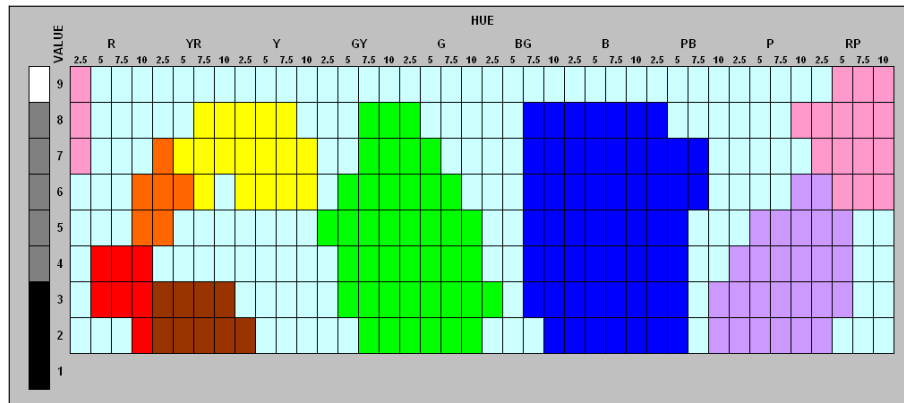
Figure 2.3: Results of Lammens Gaussian model on the CIELab colour space which is the one where the model obtained the best results. Chips in light blue were not assigned any name by the model.

The most interesting of the model is that it considers psychophysical data and the fact of being the first approach to the problem with a parametric statistical model with the advantages that this implies and that will be explained in detail in Chapter 4. Another important point of Lammens' work is that it takes the colour constancy problem into account by applying a colour constancy method to the images previously to apply the colour-naming model. The model is applied to real images on tasks of naming sample regions in the image and pointing out examples of colours.

Colour-Naming Metric

Mojsilović [103] defines a metric for the CIELab colour space which is more perceptual than the Euclidian distance to measure colour similarity in terms of colour naming. The set of colour names considered was the one from the ISCC-NBS system but with some modifications on the names of some colours to agree with the results of

a preliminary colour-naming experiment with human subjects. The syntax of the
model is very similar to that of ISCC-NBS and is defined by means of a grammar.
Given a sample to be named, the distances using the defined metric to the centroids
of each colour category are computed and the nearest neighbour rule is applied to
assign a colour name. As the distance to representatives of all the colour names
are computed, membership to any colour name can be easily computed. Mojsilović
considers the problems of colour constancy and spatial averaging previously to apply
the colour-naming model. For the first problem, the method used by Lammens in
[85] is adapted to the metric defined. The spatial averaging effect is modelled as an
adaptive low-pass filtering operation. Although the solutions proposed to these two
problems are not very complex, it is interesting the fact of considering those problems
that condition the perception of colour names.

Fuzzy Colour Category Map

The model of Seaborn et al. [124] uses the data from the psychophysical experiment
of Sturges and Whitfield [135] as the basis to create a colour-naming map of colour
categories on the Munsell colour space. The model is based on the fuzzy k-means
algorithm to generate the membership of any sample on the Munsell space.

Consensus samples from Sturges and Whitfield experiment are considered the
representatives of each category. The convex hull of such samples are assigned with
membership 1 and to compute the membership of any other point of the Munsell
space the fuzzy k-means algorithm is applied. This is a fully fuzzy model that allows
assigning the membership to the eleven basic categories for any sample of the Munsell
space. Figure 2.4 shows the colour category map obtained.
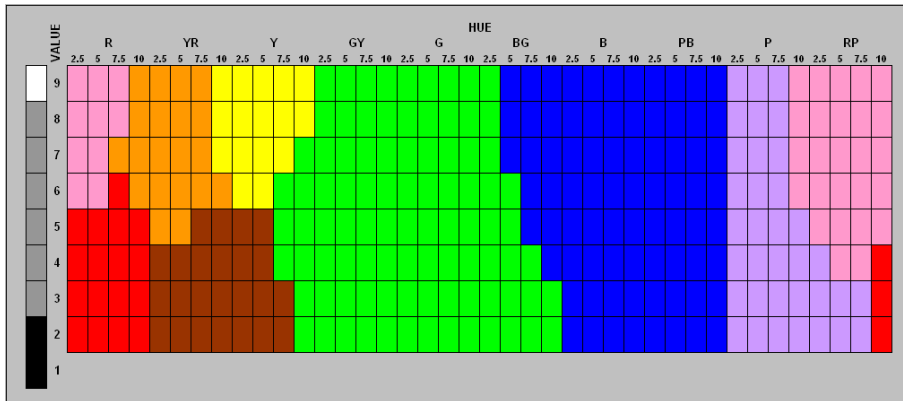
Figure 2.4: Categorization of the Munsell array obtained by the Fuzzy Colour
Category Map.

The main drawback of this model is that it is non-parametric (the disadvantages
that this implies are reviewed in Chapter 4). Another important drawback is the fact
of being a model defined on the Munsell space, since images are usually represented
on RGB spaces and there are not direct and exact translation from RGB to Munsell

system. Anyway, the fact of using psychophysical data and being a fully fuzzy model makes it a very interesting approach to the problem.

Linear Interpolation

Menegaz et al. propose in [102] a fuzzy model based on linear interpolation of the membership values obtained in a psychophysical experiment. Their experiment used the set of 424 samples of the OSA colour space [107] and memberships to each of the 11 basic colour categories of Berlin and Kay were computed as the relative frequency of responses. The OSA samples were represented in the CIELab colour space that was partitioned on 3D tetrahedrons by three-dimensional Delaunay triangulation. To name an arbitrary sample, the memberships to the 11 colour categories are computed by linear interpolation of the memberships from the four vertex of the tetrahedron that includes the sample in the CIELab space.

Probabilistic Latent Semantic Analysis

An alternative approach to the previous ones is proposed by van de Weijer et al. [145]. In this case, colour names are not modelled from data obtained or derived from psychophysical experiments, but from real-world images. The data to model each one of the 11 basic colour terms of Berlin and Kay is obtained from Google by making searches of the style "red+color".

Colour names are learned from the images by Probabilistic Latent Semantic Analysis (PLSA) which is a generative model previously used for document analysis [70] and adapted by van de Weijer et al. to the purpose of learning colour names. After learning, given a pixel, the model provides a probability value (i.e. a membership value) for each colour category and the name assigned is the one with maximum probability.

Although the initial vocabulary includes the 11 basic colour terms of Berlin and Kay, it can be easily increased by just adding images labelled with the name to be learnt to the learning set. The model is tested on image retrieval and pixel labelling applications in which it obtains good results.

2.6   Summary

In this chapter a review of the most important bibliography about colour naming has been presented. In the first part, the basic references about the research on the human visual system have been reviewed to conclude that the state of the art in the physiology of the visual system does not explain yet the mechanisms that rule the assignment of colour names. In this field, there is still much research to do.

In the second part, the main contributions to the problem from the field of psychophysics have been reviewed. The work of Berlin and Kay has been given special attention since it has been the basis of most of the subsequent research. In fact, the most accepted theories about colour naming were proposed in their pioneering study.

Finally, we have reviewed the works that have tried to computationally automate the task of colour naming. These works have been grouped in two blocks according

to whether they consider the fuzzy nature of the problem or not. In the following chapters the contributions of this thesis to the last two fields will be presented.

Chapter 3

Psychophysical Experiments on Colour Naming

An automatic colour-naming system should assign colour names in the same way a real human observer does. To achieve this, the system must have a learning step based on human judgements. Therefore, the success on the automation of the colour-naming task mainly relies on an appropriate learning step. In computer vision, a learning process can be the procedure of fitting a mathematical model to known data, in our case psychophysical data, to simulate a specific task. For colour naming, the psychophysical judgements to be fitted must be obtained from a colour-naming experiment. Hence, considering colour naming as a fuzzy process implies defining a model able to assign to a given colour sample, in our case a RGB triplet from a digital image, a membership value to all colour categories according to the psychophysical data.

As we have seen in the previous chapter, many psychophysical experiments on colour naming have been done in the last decades. However, most of them have not taken into account the fuzzy nature of the problem and the data they provide are not suitable for our purpose. For this reason, in this chapter we present a colour-naming experiment to obtain the fuzzy judgements we need as a previous step to automate the colour-naming task.

The fuzzy data set obtained from the experiment we present, is the main contribution of this chapter. This data set has been made available online to the scientific community. Furthermore, the methodology proposed to obtain the fuzzy judgements and the analysis performed to validate the method is another contribution of the chapter. Finally, a last contribution of the chapter is the set of fuzzy judgements for four more illuminants that can be useful for future research on colour naming and colour constancy.

3.1 A Review on Colour-Naming Experiments

As we have seen in the previous chapter, there is an active psychophysical research on colour naming. The goals and methodologies of all these works are diverse. In the

previous chapter we focused on the results and the contributions to colour naming of all these works. In this section we are going to review and analyse the different methodologies used in psychophysical experiments on colour naming.

In the pioneering work of Berlin and Kay, informants from different languages were shown the Munsell array with 329 colour chips and were asked to select the focal point and the boundary of their basic categories. These basic categories had been previously elicited from the informant in a first step.

Since then, a lot of psychophysical research on colour naming has been done. The basic principle of a colour-naming experiment is quite simple. Usually, these experiments work as follows. A set of colour samples are presented to a group of human observers who must assign a colour name and/or more information to each of the samples presented.

Although the main idea is quite simple, a colour-naming experiment can follow very different procedures. The methodology of a colour-naming experiment is conditioned and constrained by the final goal of the experiment. Hence, colour-naming experiments can be classified according to several criteria.

In the following paragraphs we will review the variables included in a colour-naming experiment and the different proposals that can be found in the literature. The review is divided in four topics and we will focus on these points that are the most important for the final result of a colour-naming experiment.

Subjects

The first variable in a colour-naming experiment is the **number of subjects** that participate on it. The most usual number of subjects is between 10 and 20 approximately (9 in [32], 10 in [144, 65, 154], 20 in [135, 136], and 22 in [66]). However many experiments have used less subjects (7 in [31], 6 in [30, 33], 4 in [158]), or even 2 in [130, 111]. Some experiments [143] have included more subjects (30 people). In cross-cultural experiments the number of subjects is normally higher. For example, Lin et al. worked with 90 subjects in [89] and 40 subjects in [90], and Jameson and Alvarado [75] made their experiment with 92 subjects.

A special case is the experiment of Moroney [105] that used the concept of **distributed psychophysics** and instead of having a small group of subjects and a wide set of samples, a reduced set of samples were presented to a wide set of subjects via the Internet. Thus, more than 700 subjects participated on this experiment

Subjects are normally tested to assure they have normal colour vision. To this purpose, Ishihara test has been the most frequently used one [135, 136, 65, 158, 66, 89, 90, 75, 154]. Farnsworth D-15 and F-100 hue tests have also been used in [65] and [144, 30] respectively. However, in some experiments [31] it is not reported whether subjects were tested for normal colour vision or not, or it is not mentioned which test was used [33, 32, 143, 130, 111]. Finally, in [105], due to the specific characteristics of the experiment, it was not possible to test subjects for normal colour vision.

Stimuli

Decisions about the experimental stimuli are also important. The variables related with stimuli include the number of samples, the colour space where the samples are

represented and the source of these stimuli.

The first of the variables referring to stimuli is the **number of samples**. While some experiments use a large set of samples (292 in [158], 300 in [130], 330 in [30], 348 in [65, 66], 424 in [31, 144, 32], 446 in [135, 136], 729 in [154], or even 1526 in [90]), other experiments work with a lower number of samples (78 in [111], 110 in [75], 144 in [143], 200 in [89], and 215 in [33]). Moroney [105], used a total number of 216 different samples, but each subject only evaluated a reduced subset of seven samples.

About the **colour spaces** that have been used, perceptually uniform colour spaces have been preferred to select and represent the samples used in colour-naming experiments. Hence, Munsell space was used in [30, 135, 136, 158], OSA space in [144, 31, 33, 32, 75], CIELuv in [143, 65, 66], CIELab in [158, 154], and CIE appearance models in [158] (CIECAM97) and in [154] (CIECAM02). However, other colour spaces have been used to select the samples of the experiments. Such spaces include CIE xy in [130, 158, 111], RGB in [105], ISCC-NBS in [89], and Natural Color System [68] in [90].

Finally, there are several possibilities about the **source of the stimuli** used in a experiment. Most experiments work with physical samples [31, 144, 30, 33, 32, 135, 136, 158, 89, 90, 75, 154]. More recent experiments have used CRT displays to show the samples to subjects [65, 66, 105], although this support was used before in [143] . Other experiments to study aspects of colour naming have also used lights [130, 111].

Apparatus

Psychophysical colour-naming experiments have normally been done on **calibrated conditions**, that is, the experiment is done with the same controlled conditions for all the subjects in the experiment. In most cases, experiments are developed on a closed room or a booth with fixed illuminant and geometry, although less restrictive experiments have used different environments for subjects [30, 89] or even have been done under completely **uncalibrated conditions** as it is the case of distributed psychophysics developed by Moroney [105].

The **illuminant** used is one of the most important condition in experiments for the implications it may have in the final results. In the colour-naming literature a wide set of illuminants used in experiments can be found. These illuminants include CIE D50 [154], CIE D65 [89, 90], CIE C [75], incandescent lights of different colour temperature (3000K [31], 3200K in [144, 32], 6500K in [135, 136]), or natural daylight [30, 89]. In CRT based experiments, illumination has been simulated on the display [65, 66] or no illumination apart from that of the display has been included [111]. Experiments developed to study aspects of colour constancy have used multiple illuminants [143, 130, 158] or specific illumination, as in [33] where mixtures of low-pressure sodium and tungsten illuminants were used.

Method

As we said at the beginning of this section, a colour-naming experiment normally is done by presenting a set of colour samples to a group of subjects that evaluate them. This process can be done on one or more **trials**. In experiments including two trials [31, 144, 30, 32, 135, 136], samples are normally presented in random order in

the first trial, and in the second the reverse order is used to avoid memory effects and systematic errors. However, many experiments have been done on a single trial [30, 143, 130, 66, 89, 90, 111, 75, 154] or in more than two, as in [158] and [65] where each sample was evaluated three times.

The vocabulary of colour names includes a high variety of terms: single names (e.g. red, yellow), compound names (e.g. blue-green,yellow-green), derived names (e.g. purplish blue, greenish yellow), modifiers (e.g. light, deep), etc. Hence, the information asked to the subjects in a experiment is a very important decision in the design of the experiment. Results of the experiment are highly conditioned by the rules imposed. Therefore, and depending on the goal of the experiment, this point must be carefully treated.

Hence, experiments having general goals, such as analysing aspects of nameability [65, 66] or cross-cultural studies [89, 75] have used **unconstrained** methods, that is, no restrictions on the set of names that subjects can use are imposed. On the other hand, those experiments having as goal a specific aspect of colour naming have used **constrained** methods, that is, the subject is only allowed to use a set of predefined terms.

The most general constraint is to allow only one-word names. These experiments are referred as **monolexemic**. The degree of restriction varies between experiments: all monolexemic names allowed [31, 144, 30, 32, 135, 136], only some monolexemic are allowed [143], or only the eleven basic colour terms defined by Berlin and Kay are allowed [130, 158, 111, 154]. **Polylexemic** experiments, that is, experiments allowing the use of compound colour terms and modifiers, are normally unconstrained [30, 65, 89, 105, 75]. However, polylexemic experiments can also be constrained [90]

Although in most of the experiments the only information that subjects must provide is a colour name, some variations of the usual method have been proposed and some experiments have used **ratings** to evaluate colour samples. Guest and Van Laar [65, 66] asked subjects to provide a confidence rating on the scale from 1 to 5 according to their certainty about the name they provided. The same scale was used in [75]. In [111] a scale from 0 to 10 was used.

Troost and De Weert [143] introduced the idea of describing colour samples with a naming vector where each dimension corresponds to a colour term. The value of each position of the vector, was the percentage of answers with that term that the sample had received. The same approach was used in [154]. This representation was also adopted by Speigle and Brainard [130] by using an 11-dimensional vector where each dimension corresponds to one of the basic colour terms. In that experiment, the ratings ranged from 0 to 9 and were provided by the subjects who were asked to "rate how good an X the stimulus is", where X was each of the eleven basic colour names.

As we have seen in this section, the design of a colour-naming experiment incorporates many variables and the selection of the exact methodology for an experiment is conditioned by the goal of the experiment. Table 3.1 summarizes the analysis of the different methodologies that we have done in this section.

	SUBJECTS		STIMULI			APPARATUS		METHOD			
	Number	Test	Number	Space	Source	Conditions	Illuminant	Trials	Constraints	Names	Ratings
Boynton & Olson (1987)	7	No	424	OSA	Physical	Cal	3000K	2	Con	Mono	No
Uchikawa & Boynton (1987)	10	F100	424	OSA	Physical	Cal	3200K	2	Con	Mono	No
Boynton et al.) (1989)	6	F100	330	Munsell	Physical	Uncal	Outdoor	1	Unc	All	No
Boynton & Purl (1989)	6	NR	215	OSA	Physical	Cal	Low-press Na-W	2	Con	Mono	No
Boyton & Olson (1990)	9	NR	424	OSA	Physical	Cal	3200K	2	Con	Mono	No
Troost & de Weert (1991)	30	NR	144	CIELuv	CRT	Cal	Multiple	1	Con	12 mono	% Resp.
Sturges & Whitfield (1995)	20	Ishihara	446	Munsell	Physical	Cal	6500K	2	Con	Mono	No
Sturges & Whitfield (1997)	20	Ishihara	446	Munsell	Physical	Cal	6500K	2	Con	Mono	No
Speigle & Brainard (1997)	2	NR	300	CIExyY	Lights	Cal	Multiple	1	Con	B&K	0-9
Guest & Van Laar (2000)	10	Ishi/F-D15	348	CIELuv	CRT	Cal	Simulated D65	3	Unc	All	1-5
Yaguchi (2001)	4	Ishihara	292	Multiple	Physical	Cal	Multiple	3	Con	B&K	No
Lin et al., part I (2001)	90	Ishihara	200	ISCC-NBS	Physical	Uncal	D65/Daylight	1	Unc	All	No
Lin et al., part II (2001)	40	Ishihara	1526	NCS	Physical	Cal	D65	1	Con	Poly	1-5
Guest & Van Laar (2002)	22	Ishihara	348	CIELuv	CRT	Cal	Simulated D65	1	Unc	All	1-10
Okajima et al. (2002)	2	NR	78	CIExy	Lights	Cal	No	1	Con	B&K	1-10
Jameson & Alvarado (2003)	92	Ishihara	110	OSA	Physical	Cal	CIE C	1	Unc	All	1-5
Moroney (2003)	>700	No	216	RGB	CRT	Uncal	—	—	Unc	All	No
Wang et al (2006)	10	Ishihara	729	Multiple	Physical	Cal	D50	1	Con	B&K	% Resp.
Benavente et al (2006)	10	Ishi/F-D15	387	Munsell	Physical	Cal	5955K	2	Con	B&K	0-10

Table 3.1: Summary of methodologies used in colour-naming experiments. In the last row the experiment presented in this thesis is also included. Although it will be explained in detail in next section, it is included in the table in order to make it possible to compare our experiment with the previous ones. Abbreviations used: NR:Not reported, Cal:Calibrated, Uncal:Uncalibrated, Con:Constrained, Unc:Unconstrained, Na-W:Sodium-Tungsten, Mono:Monolexemic, Poly:Polylexemic, B&K:Berlin and Kay, % Resp.:Percentage of responses

3.2 Building a New Data Set for Fuzzy Colour Naming

In this section we present a colour-naming experiment we have developed to obtain a data set of fuzzy colour-naming judgements that will be the basis of the subsequent modelling process that will be presented in the next chapter.

As we have seen in the previous section, data from most of the previous colour-naming experiments are not adequate to model the colour-naming space as a fuzzy process. These experiments are usually done by only asking subjects to assign a colour name to a given stimulus without any other judgement. Although some experiments have introduced variations on the usual method, these experiments were not thought to be used for fuzzy colour-naming modelling but were part of colour constancy studies and the corresponding data sets are directed to that purpose.

To provide a sufficient data set of colour judgements adequate for the computational modelling of the colour-naming task, we have performed a colour-naming experiment using a methodology similar to the used by Speigle and Brainard, and that was previously tested on a preliminary experiment [21]. In the experiment, subjects were asked to assign a membership value to all the colour categories considered. The experiment is restricted to the eleven basic colour terms proposed by Berlin and Kay because previous experiments [32, 136, 65] showed that basic names are used more consistently and with more consensus than non-basic names. This fact is essential for computational approaches due to the problems caused by outliers, that is, low consistent and low consensus samples. In addition, non-basic colour terms correspond to small regions on colour space and this is difficult to model from a computer vision point of view.

The final goal of this experiment is to obtain the adequate data set suitable to be used as learning set on a fitting process to model colour naming as a fuzzy task. In addition, we make two more contributions with this experiment. First, we make available to the scientific community a complete dataset derived from a fuzzy colour-naming experiment. Such dataset will be the basis of a mathematical model of the colour-naming task by means of a fuzzy model. Second, we have computed common statistics used to evaluate colour-naming experiments to demonstrate the validity of these data. We have shown that data obtained from a scoring methodology present similar statistics to data obtained in previous monolexemic experiments. After our analysis we can assure that our data is valid in a perceptual sense, that is, data is not affected by the method used which a priori could introduce more subjectivity to the judgements obtained.

3.2.1 Method

Subjects

The subjects that took part in the experiment were 10 researchers (5 male and 5 female) from our lab. They were between 24 and 30 years old and all of them were volunteers. All the subjects were tested with the Ishihara and Farnsworth D-15 tests to guarantee they had normal colour vision. All the subjects were bilingual Catalan

and Spanish speakers with an advanced level of English since the experiment was developed using the English terms for the eleven basic colour categories. The fact of having Catalan and Spanish speakers doing the experiment in English should not be a problem for our final purpose of computationally modelling colour categories as shown in [118].

Stimuli

The stimuli used were a total of 387 samples which included 36 achromatic and 351 chromatic samples. Each sample had a size of 24×16 cm to assure a wide visual angle. All the samples were printed by a HP DesignJet 2500CP Plotter, and their reflectance functions were measured with a PhotoResearch PR-650 spectro-radiometer.

The selection of samples was done according to three criteria: to cover as much as possible the colour space, to avoid overlapping between samples and to have a reasonable number of samples, that is, a number of samples that did not make the experiment too hard for subjects. To this end, the Munsell colour space was sampled at each 2.5 units of hue, at each value unit (from 3 to 9) and at the highest chroma value available. The selected samples were printed with a HP DesignJet 2500CP Plotter. Obviously, the printed samples suffered a deviation from the real Munsell due to the use of the plotter. To have the real reflectances of the printed samples, they were measured with a PhotoResearch PR-650 spectro-radiometer.

Since the spectra of the selected samples are known, the coordinates of the samples set at any colour space can be computed. Thus, we provide the CIELab values computed according to standard equations [157] using the CIE Illuminant D65 and the Two Degree Standard Observer. The corresponding Munsell values have been computed with the 'Munsell conversion - Version 4.01' software from Gretagmacbeth [5].

To cover some gaps that appeared in the Munsell space we printed some additional surfaces to have the final set of 351 samples that accomplished the three criteria mentioned above. The set of spectra as well as the corresponding CIELab and Munsell values of the samples used in the experiment have been made available online to the scientific community [3].

Apparatus

The experiment was developed in a dark room. The subjects were sat in an adjustable chair in front of a booth where the samples were presented. They were at a viewing distance of 50 cm. from the sample that was presented on a support painted with a neutral grey corresponding to Munsell N7. Because the final goal of our work is to define a model that assigns colour names in the same way as a human observer, we must consider the effects of the colour constancy mechanisms of the human visual system. Thus, the inside of the booth was white to assure that psychophysical data would be acquired considering the colour constancy mechanisms of the human subjects doing the experiment. The samples were illuminated from the top of the booth by an illuminant with a correlated colour temperature (CCT) of 5955K and a Luminance of 150 cd/m^2. Figure 3.1 shows a diagram of the experimental conditions.
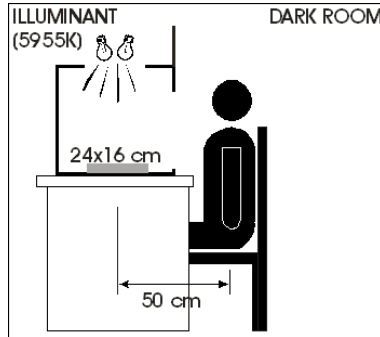
Figure 3.1: Scheme of the experiment conditions. The experiment was developed under controlled conditions in a dark room to assure that samples were only illuminated from the top of the booth.

Procedure

The procedure followed was very similar to the one followed by Speigle and Brainard. Subjects were instructed to use only the eleven basic colour terms. These are white, black, red, blue, green, yellow, purple, pink, orange, brown and grey. Any other colour name was not allowed.

For each one of the samples presented, the observer was asked to distribute a total score of 10 points among the 11 possible colour names according to the certainty they had about the sample belonging to the different categories. Thus, if the subject was absolutely sure about the colour name of a sample, then the 10 points had to be assigned to the category corresponding to that name. Otherwise, if there was a doubt between two or more names, the 10 points had to be distributed between the categories corresponding to those names (e.g. blue - 4 and green - 6). Hence, the result of the naming for each sample is a colour descriptor of 11 components, one for each basic colour term. No time limitations were set to give a response.

The 387 samples were presented one at a time, twice each, to the ten subjects. This means a total number of 7740 observations. The samples were first presented in random order, and the reverse was used in the second trial. For each sample, the scores from the 10 subjects were normalized to the [0,1] interval and averaged to obtain the mean colour descriptor of the sample. The set of colour judgements obtained is available at [3].

3.2.2   Results

The goal of the colour-naming experiment presented in this chapter is to provide a set of fuzzy colour-naming judgements (i.e. the membership values of a wide set of samples) for the eleven basic colour categories. Hence, we do not try to prove the existence of the eleven basic colour categories nor their location in the colour space, as some previous studies have done. Moreover, and in order to prove the feasibility of the methodology followed, in this section we show that usual statistics computed on monolexemic colour-naming experiments agree with the ones we can

derive from our data. The statistics analysed are consistency, consensus, position of focal colours and centroids, and the confusion matrix in the use of colour terms. The analysis of consistency and consensus was done over the whole set of samples while the analysis of the other three statistics only considered the chromatic categories due to the reduced number of samples that were considered achromatic by the subjects. After this analysis, we will be able to state that the methodology followed in the experiment is a generalization that comprises the results of monolexemic experiments.

Consistency

Consistency is a measure of the degree of coincidence in the two evaluations that each subject makes for each sample. In previous works, consistency has been calculated by counting the number of times that a colour sample has been given the same name by the same subject on the two observations of the sample. Due to the differences between these works and our experiment, some variations of the consistency measure are proposed. In our case, we do not have a colour name for each sample, but a colour descriptor of membership values with 11 components, one for each of the 11 basic colour names. The proposed measures are the following:

- *Consistency based on identical membership values:* Judgements for a sample are consistent when the same subject gives exactly the same values to all the categories in the two observations of the sample.

- *Consistency based on the highest membership value:* Judgements for a sample are consistent when the same subject gives the highest membership value to the same category in the two observations of the sample, no matter how the values are distributed among the 11 categories.

- *Consistency based on the city-block metric with a threshold (τ):* Judgements for a sample are consistent when the city-block distance (equation (3.1)) between the two colour descriptors given by the same subject is lower than the value of the threshold τ.

$$d = \sum_{k=1}^{11} \left| CD_k(\vec{x}_i)^{j,1} - CD_k(\vec{x}_i)^{j,2} \right| \tag{3.1}$$

  where $CD_k(x_i)^{j,1}$ is the kth component of the colour descriptor, that is the score given to the colour category C_k, of the ith sample for the jth subject in the first observation, $CD_k(x_i)^{j,2}$ is the same component for the second observation.

- *Global Consistency:* The city-block distance between the two colour descriptors given by the same subject is normalized to one and the global consistency is calculated according to equation (3.2):

$$Consistency = \frac{\sum_{i=1}^{i=ns} \sum_{j=1}^{j=np} 1 - \frac{\sum_{k=1}^{11} \left| CD_k(\vec{x}_i)^{j,1} - CD_k(\vec{x}_i)^{j,2} \right|}{2}}{ns \cdot np} \tag{3.2}$$

where $CD_k(\vec{x_i})^{j,1}$ is the *kth* component of the colour descriptor of the *ith* sample for the *jth* subject in the first observation, $CD_k(\vec{x_i})^{j,2}$ is the same component for the second observation, *ns* is the number of samples, and np is the number of subjects in the experiment. Notice that the sum of the differences is divided by the highest possible difference which is 2. This measure has been designed to take into account the nature of the fuzzy methodology of our experiment.

The consistency values were calculated for the four measures proposed. Obviously, the results for the less restrictive measures (global consistency and highest membership value) are the best, but the values obtained by the other measures are also acceptable. According to the measure based on identical membership values, subjects scored the samples consistently in 2501 of the 3870 evaluations. This means that 64.63% of the samples were scored identically. If we consider a threshold of 0.4 in the calculation of the consistency (i.e. if the city-block distance between the two colour descriptors is lower than 0.4, we consider the naming of the sample is consistent), the number of samples scored consistently increases up to 2762, which means that consistency was reached 71.37% of the time. If we only consider the colour name with the highest membership value to study consistency, the number of samples scored consistently is 3327 (85.97%). Finally, the application of the global consistency measure provides a result of 85.43% of consistent use of the colour names in the experiment. These results are summarized in table 3.2.

Consistency method	Coincidences	%
Highest membership value	3327	85.97
Global consistency	3306.2[a]	85.43
City-block consistency with $\tau = 0.4$	2762	71.37
Identical membership values	2501	64.63

[a]Result obtained from equation (3.2)

Table 3.2: Consistency results obtained according to the different proposed measures.

If we assume that in a monolexemic colour-naming experiment the name assigned to any sample would have been the one corresponding to the category with the highest membership value in our experiment, we can compare the consistency results from Boynton & Olson [32] and Sturges & Whitfield [136] experiments with our measure based on the highest membership value. In those works the consistency values for basic colour terms were 75% and 84.4% respectively. As can be seen, our consistency value (85.97%) is very similar to the results obtained in these previous experiments. Moreover, if we consider the consistency measure based on identical membership values, the value obtained (64.63%) is not bad if we take into account that an exact coincidence of the membership values is much harder to obtain than a coincidence of only the highest value.

Consensus

Consensus is a measure of the agreement in the judgements between subjects. In the previous experiments, the consensus in a particular sample was computed by counting the number of times that the most used name for that sample had been given by all subjects in all the trials. Hence, perfect consensus for a sample was achieved when all subjects used the same colour name on both trials of the sample.

In our experiment, we have 20 11-dimensional colour descriptors for each sample, obtained from the two trials of the 10 subjects. From these 20 colour descriptors, we have computed three different consensus measures.

- *Consensus based on the membership values:* Consensus is reached when all subjects have assigned the same values to the eleven categories in all the evaluations of the sample. Hence, in our case, a sample will have consensus if it has 20 identical colour descriptors.

- *Consensus based on the highest membership value:* Consensus is reached when all subjects have assigned the highest membership value to the same category in all the evaluations of the sample.

- *Consensus based on the highest membership value with a threshold (τ):* Consensus is reached when the highest membership value has been assigned to a certain category in at least τ evaluations of the sample. For example, if τ is set to 15, consensus will be reached if the category corresponding to the highest membership value is the same for, at least, 15 of the 20 colour descriptors.

Consensus was computed for the three measures defined above. The first measure gave as a result that consensus was obtained for 43 of the 387 samples; this is, for 11.11% of the samples. As happened with consistency, the second measure can be considered the equivalent to the measure computed in previous experiments. In the present experiment, total consensus based on the highest membership value was obtained for 136 of the 387 samples. This means that 35.14% of the samples were given the highest value to the same category by all the subjects. If we compare these results to the consensus obtained in previous works, we can see that perfect consensus in the experiments of Boynton and Olson, and Sturges and Whitfield was obtained on 30% and 23% of the samples respectively. In our case, we obtain a slighter higher consensus probably due to the fact that we do not allow non-basic colour names which are used with lower consensus than basic colours.

Using the third statistic, the percentage of consensus increases gradually as the threshold is reduced. Thus, when only 15 coincidences on the category which has been assigned the highest membership value are required, consensus is of 75.97%, and this value increases up to 98.19% when the threshold is set to 11 coincidences, which is the minimum possible majority. In figure 3.2, the evolution of consensus value in terms of the threshold (number of coincidences required) is presented.

Focal Colours

Focal colours are defined as the best examples for each colour category. In previous experiments, focal colours have also been defined as the fastest named samples with
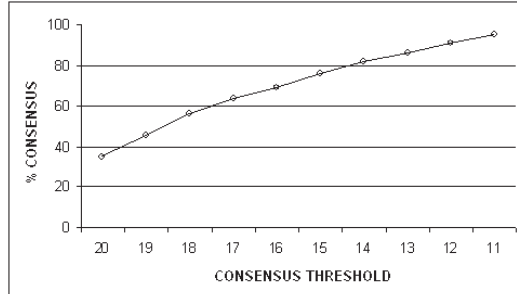
Figure 3.2: Percentage of consensus when the criterion for defining consensus is relaxed from 20 coincident responses (of 20 possible) to the smallest possible majority of 11 coincidences.

consensus for each colour category. In our experiment, response time was not considered. Hence, we have selected as focal candidates all the samples which have been assigned 10 points to the same colour category by all the subjects and in the two evaluations of the sample. For two categories, orange and yellow, there is no sample with total consensus and the samples with the highest mean membership value in those categories have been selected as foci. In figure 3.3, our focal candidates and the foci from Berlin and Kay, and Sturges and Whitfield experiments are shown in terms of hue and value on Munsell colour space.
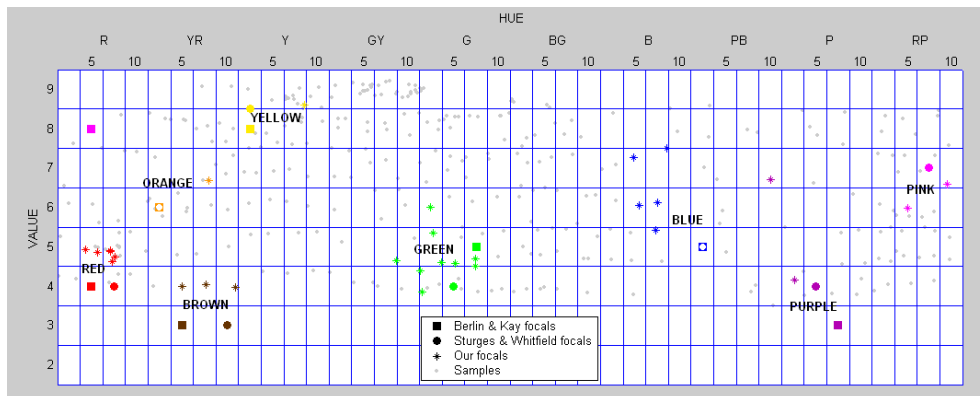
Figure 3.3: Location of candidate foci obtained in the experiment and foci from two previous works (Berlin and Kay, and Sturges and Whitfield). Samples used in the experiment are shown as grey dots.

As can be seen in figure 3.3, the consensus samples from our experiment lie, in general, near the focal colours found by Berlin and Kay, and Sturges and Whitfield. The focal samples from Sturges and Whitfield have, in most of the cases, a candidate focal of our experiment at a distance lower than one Munsell chip (2.5 hue units or 1 value unit). This distance is higher for three categories: orange, yellow and blue. In the case of orange, the shift of the focus might be due to cultural reasons (it had been

detected in a preliminary experiment [23] that Catalan and Spanish speakers located the orange region at a different area from English). In the case of blue, the problem could be due to the different criteria of selection of samples in the experiments, since the focal locations are highly dependent on the stimulus set.

Centroids

Centroids are a measure of the central tendency of the location of the categories in the colour space. In previous experiments the centroids of each category have been computed by averaging the values of all the samples named with the corresponding term, and weighted according to whether the term was used once or twice. In our case, we have computed the centroids as the average of the values of the samples, but weighted according to the membership values provided by the subjects for each colour category (equation (3.3)).

$$Centroid_k = \frac{\sum_{i=1}^{ns} m_k^i \vec{x_i}}{\sum_{i=1}^{ns} m_k^i} \qquad (3.3)$$

where ns is the number of samples, $\vec{x_i}$ is the ith sample and m_k^i is the membership value of the ith sample for category C_k.

Table 3.3 shows the centroids obtained in our experiment and in two previous monolexemic works [31, 135]. Table 3.4 allows a comparison to the previous results by giving the CIELab differences between experiments. In figure 3.4, the centroids for the chromatic categories of the three experiments are presented in terms of Munsell Hue and Value.

	Boynton & Olson			Sturges & Whitfield			Benavente et al.		
	H	**V**	**C**	**H**	**V**	**C**	**H**	**V**	**C**
Red	3.50R	3.97	9.19	6.01R	3.91	11.72	6.55R	5.02	10.82
Orange	1.73YR	5.91	10.10	3.88YR	6.21	11.45	7.42YR	6.77	7.90
Brown	3.85YR	4.41	4.52	7.69YR	4.32	5.78	6.88YR	6.01	4.72
Yellow	2.51Y	7.56	8.28	3.35Y	7.86	10.44	7.89Y	7.96	6.82
Green	8.21GY	6.07	5.07	1.73G	5.35	7.52	1.21G	6.38	4.16
Blue	8.44B	5.42	5.12	8.23B	4.87	7.33	2.67B	6.46	4.45
Purple	6.37P	4.53	5.22	4.75P	4.18	7.86	8.69P	5.65	5.82
Pink	2.04R	6.04	5.68	2.15R	6.14	8.59	8.38RP	6.72	6.73

Table 3.3: Centroids obtained in our experiment and in two previous studies.

As can be seen in table 3.3 and in figure 3.4, the centroids found in our work have, in general, a difference of one value unit to the centroids of previous works. Again, this is due to the different criteria in the samples selection which has been previously explained. The fact that the samples were printed with a plotter has brought a lack of samples in the low value areas of the Munsell space. The value difference is considerably smaller for those categories that are located in the areas of high value in the Munsell system, such as pink and yellow. The bigger differences are found for the categories located at the low value areas (brown, purple and red) because they

	BO vs. SW	BO vs. BVB	SW vs. BVB
Red	15.75	12.56	17.69
Orange	14.05	20.83	19.81
Brown	12.17	18.81	19.02
Yellow	16.03	24.55	14.61
Green	17.19	21.27	10.47
Blue	10.93	22.89	14.38
Purple	12.78	20.03	12.08
Pink	12.45	12.86	8.99
Mean diff.	13.92	19.23	14.63

Table 3.4: Comparison of centroids obtained in our experiment (Benavente, Vanrell and Baldrich [24] (BVB)) and in previous studies (Boynton and Olson [31] (BO), and Sturges and Whitfield [135] (SW)), in terms of the CIELab differences.
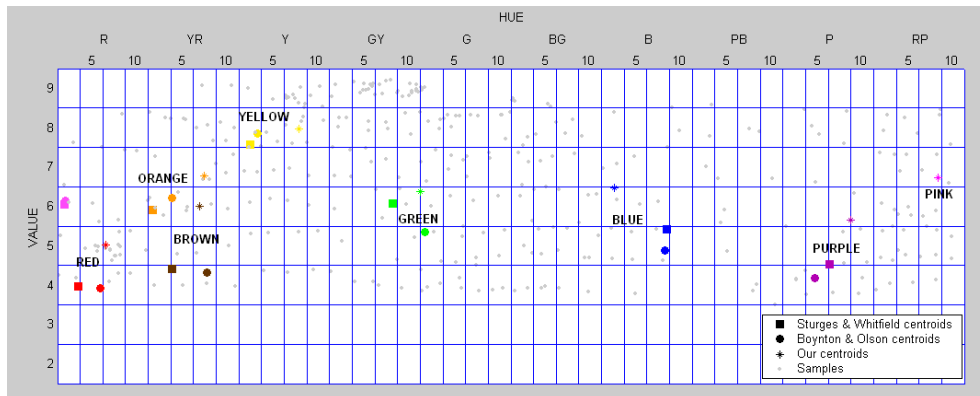
Figure 3.4: Location of centroids obtained in the experiment and in two previous works (Boynton & Olson and Sturges & Whitfield). Samples used in the experiment are shown as grey dots.

are more affected by the lack of low value samples in the experiment. The mean difference to Boynton and Olson study is 19.23 CIELab units and the difference to Sturges and Whitfield is 14.63, while the mean difference between the centroids found by Boynton and Sturges was of 13.92. Another possibility for these differences in the results could be the different illumination conditions in the experiments. Boynton and Olson used an illuminant with a temperature of 3200K and Sturges and Whitfield used an illuminant of 6500K while our experiment was done under a 5955K illuminant. The mean difference between centroids is smaller compared to Sturges and Whitfield who used an illuminant similar to ours. However, the difference between Sturges and Boynton is also small, but the illuminants are quite different. Hence, the influence of illumination conditions in the results is not clear and needs further study. Anyway, we can see that the differences between our centroids and the ones from previous studies are reasonable. Moreover, the higher difference found in our case seems to be due to the high dependence of the centroids locations on the set of stimuli selected as was concluded by Speigle and Brainard [130].

Linked Colours

In previous studies the confusion matrix in the use of basic colour terms was analysed to study the way in which basic colour terms were related. Boynton and Olson defined two colours as linked when a majority of subjects applied the same different basic colour terms to describe a particular colour sample on the two observations of the sample. According to Sturges and Whitfield this inconsistency indicated that the sample was perceived as a mixture of colours that contained elements of the two terms used. With the information from the confusion matrix, it is possible to build a three-dimensional model of the relationships between the basic colours.

In our case, the experiment has provided more information about these relationships than previous experiments and we can use the whole set of membership values to build the confusion matrix. Hence, our confusion matrix has been built by counting the number of times that any subject has scored two particular colour names for the same sample. This definition of the confusion matrix agrees with the idea of Sturges and Whitfield that when a sample is named with two different basic names on the two observations of the sample, it is perceived as a mixture of the two colours. In our experiment, we assume that if subjects give values to two or more categories for the same sample, it is because none of the basic colour terms describes exactly the sample and they doubt about which name must be given to the sample. Table 3.5 shows the confusion matrix for our experiment (white and black are omitted due to lack of sufficient data). To avoid some outliers detected on data, values less than 2 points have not been considered.

From the confusion matrix a three-dimensional model of the relationships between the basic colour categories has been built. Figure 3.5 shows a schematic representation of this model. The criteria to decide whether two colours are linked are similar to the used by Sturges and Whitfield. As in their work, we represent all the relationships of the confusion matrix (excluding confusions from a unique subject) and the width of the line linking two colours indicates the confusion between the two colours.

The analysis of the linkages in Figure 3.5 shows that the colours that are linked

	Red	Green	Yellow	Blue	Orange	Purple	Grey	Pink	Brown
Red	—				38 (10)	14 (6)		39 (9)	38 (9)
Green		—	90 (9)	274 (10)	1 (1)		55 (9)		44 (9)
Yellow	90 (9)		—		32 (8)		5 (4)	1 (1)	59 (10)
Blue		274 (10)		—		29 (7)	24 (7)		
Orange	38 (10)	1 (1)	32 (8)		—			25 (6)	70 (10)
Purple	14 (6)			29 (7)		—	13 (5)	57 (9)	5 (3)
Grey		55 (9)	5 (4)	24 (7)		13 (5)	—	6 (4)	30 (5)
Pink	39 (9)		1 (1)		25 (6)	57 (9)	6 (4)	—	19 (3)
Brown	38 (9)	44 (9)	59 (10)		70 (10)	5 (3)	30 (5)	19 (3)	—

Table 3.5: Confusion matrix computed from the results of our experiment. Each cell of the matrix shows the number of times that a sample has been given values to the categories intersecting on the cell. The number of subjects who scored the two categories for at least one sample is given in parentheses.
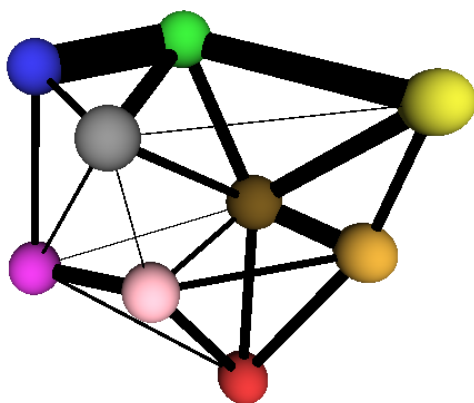
Figure 3.5: Schematic representation of the linkages between basic colour categories (excluding black and white). The spheres represent the position of colours on the chromatic plane of the CIELab space. The model is viewed in perspective from L=160.

and unlinked are the same as in previous works, except for yellow-grey and pink-grey which were unlinked in both previous studies. However, the linkages observed in these two cases are very weak and confusions between those colours also appear in the confusion matrices of the previous studies.

Hence, these results mean that although the method employed seems to include more subjectivity, the rating of samples is done in a perceptual sense. Thus, for example, nobody has given values to red and green or blue and yellow at the same time, which would be contradictory to the opponent theory of colour perception. Hence, the comparison of the results in terms of linked colours also supports the validity of the method and the equivalence to previous studies about colour naming.

3.3 Fuzzy Colour-Naming Data for Different Illuminants

As we have seen in the review of colour-naming experiments, there has been research on the relationship between colour naming and colour constancy [29, 33, 143, 130, 158]. To provide fuzzy data for future research on this topic, the colour-naming experiment presented in the previous section has been repeated with four more different illuminants.

3.3.1 Method

The ten participant subjects, the stimuli used, and the dark room where the replications of the experiment were developed, were the same as in the previous experiment.

Illuminants

The four additional illuminants used in the repetitions of the experiment were selected along the Planckian locus to obtain a range of illuminants from a strong reddish 1765K illuminant to a strong bluish 26285K. These illuminants were obtained by adding colour filters to the original illumination maintaining the Luminance of $150 cd/m^2$. Hence, the four illuminants obtained had Correlated Colour Temperatures (CCT) of 26285K, 3435K, 2320K and 1765K. Figure 3.6 shows the location of the five illuminants (the original 5955K illuminant and the four for the repetition) in the CIE xy chromaticity plane.

The spectra of the five illuminants are available at [3]. Table 3.6 summarizes the information about the illuminants used.

Procedure

Each repetition of the experiment with a different illuminant was done a week later than the trial with the previous one. Therefore the five repetitions of the experiment were done in a period of five weeks.

The experiment using the illuminant with a CCT of 5955K was done first for all the participants. The order of the rest of the illuminants was random for each subject in order to avoid systematic errors. As we had done in the original experiment with
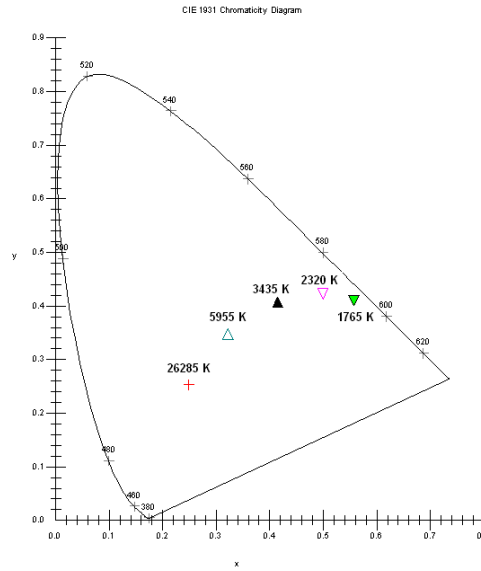
Figure 3.6: Locations on the CIE xy chromaticity plane of the five illuminants used in the experiments.

CCT	x	y
26285K	0.2486	0.2541
5955K	0.3220	0.3472
3435K	0.4151	0.4085
2320K	0.4991	0.4232
1765K	0.5567	0.4106

Table 3.6: Correlated Colour Temperature (CCT) and CIE xy chromaticity coordinates of the illuminants used in the experiment.

the first illuminant, the experiment was repeated twice for each illuminant and the order of the samples in the second trial was the reverse from the first trial.

3.3.2 Results

The repetitions of the experiment provided the set of fuzzy colour-naming judgements for the four new illuminants. As the number of subjects and samples was the same as in the first experiment, the total number of observations for each illuminant was of 7740, which means a total number of 38700 observations. As we did in the first experiment, memberships for each sample were averaged and normalized to the [0,1] interval. Hence, we finally provide a total of 1548 additional fuzzy judgements (387 for each of the four additional illuminants). These data for the four new illuminants is also available at [3].

The set of fuzzy jugdements for five different illuminants can be very useful to study the relationship between colour naming and colour constancy. As we explained before, the inside of the booth where the experiment was developed was painted white in order to allow the colour constancy mechanisms of the human visual system to act. However, some of the illuminants used in the repetitions of the experiment were very extreme and it would be interesting to evaluate if the colour-naming perception of the subjects varied or the colour constancy mechanisms of the visual system could avoid the effects of the illuminant to leave colour naming unaffected. The data provided can be used for this goal, but it is out of the scope of this thesis.

3.4 Discussion

In this chapter, we have presented the colour-naming experiment developed to obtain a set of colour judgements useful to be used as the basis for the fuzzy modelling of the colour-naming task. Once we have the set of fuzzy colour-naming judgements the next step is to find a mathematical model that fits the data from the experiment. Using such data as learning set will allow us to obtain a model of colour naming that provides the same colour-naming judgements as a real human observer.

This data set is the main contribution of this chapter. The set of fuzzy judgements, the reflectances and, the CIELab and Munsell values of the samples used in the experiment are available online at [3]. Making psychophysical data available to the research community is an important step, at least in the computer vision field, because it can contribute to achieve faster progress in this field. Moreover, it will allow testing and comparing different models over the same data sets. This fact is also very important in our field due to the complexity of obtaining that kind of data sets for testing.

The methodology used in this experiment and the subsequent analysis is the second contribution of the chapter. The results of our experiment have been compared to the ones from previous monolexemic experiments to show that although the experimental methods are considerably different, the results in terms of some usual measures in colour naming are similar. This fact allows converting the results of the present experiment to a monolexemic experiment by just assigning to each sample the colour

name with the highest membership value in the colour descriptor provided by the subjects from the experiment.

The analysis of the results also supports the statement of Speigle and Brainard about the heavy influence of the stimuli gamut in the location of focal colours and centroids. Hence, some deficiencies in the selection of the samples used in the experiment have caused that our foci and centroids are not exactly in the same locations as in some previous works. Although these differences are not important for our purpose, it would be interesting to find a way to analyse colour-naming results that is not so dependent on the set of stimuli used as the computation of focal colours and centroids.

As a final contribution of this chapter, we have repeated the colour-naming experiment with four additional illuminants to provide a set of fuzzy judgements under different illumination conditions. These data can be very useful for future studies on the relationship between colour naming and colour constancy. As in the case of the initial experiment, the resulting set of fuzzy judgements and illuminant spectra used in the experiments are available online at [3].

Chapter 4

A Parametric Model for Colour Naming

Once we have obtained a set of psychophysical data, the next step to build a colour-naming model is to define the functions that will provide the membership values for any given colour sample. The fuzzy judgements from the experiment will be used as learning set to build the final model that should assign colour names in the way the human observers from the experiment did. In this chapter we propose different possible membership functions. Observation of the psychophysical data from the experiment allows us to define the set of properties that the membership functions should fulfil. Hence, several functions are proposed, each one improving the results of the preceding. The last function that is proposed, the Triple-Sigmoid with Elliptical centre (TSE), is the basis of the final colour-naming model which is the most important contribution of this thesis.

4.1 Colour Data Modelling

To automate a visual task implies to build a computational model that reproduces the behaviour of humans when doing such task. Many colour tasks, amongst them colour naming, have been modelled following different approaches. In Chapter 2 we reviewed and detailed the different methods that have been used to model the colour-naming task. In general pattern recognition [44, 28], models are normally classified in three groups:

Parametric models
  Parametric models are characterized by a function with a set of tuneable parameters that determine the form of the function. Usually the number of parameters will determine the complexity of the function, which can vary from a trivial function with few parameters to a complex function with a higher number of parameters.
  Parametric models have been previously used to model colour information [9]. A parametric model for colour naming was proposed by Lammens [85]. Such model

51

was based on a variant of the Gaussian normal function. Gaussian models have been widely used for different tasks involving colour [61, 134, 77].

Non-Parametric models

In non-parametric models, the structure of the distribution is defined by the available sample. This does not necessary imply that these models completely lack parameters, but that the number and nature of the parameters are flexible and not fixed in advance. Hence, as few or none parameters have to be estimated, the building step is normally faster than in parametric models. However, the use of the model is more complicated since it normally implies calculations with the full set of samples that define the model.

Amongst non-parametric models, the nearest neighbour and the k-means algorithm have been used for colour naming. The nearest neighbour classifier based on a perceptual metric defined on the CIELab space from the results of psychophysical experimentation is used by Mojsilović in [103]. A similar approach is used by Menegaz et al. in [102] where memberships for an arbitrary sample are linearly interpolated from the four nearest points of the initial sample. The k-means algorithm was used by Yendrikhovskij in [159, 160] to model colour categories. Seaborn et al. [124] used the fuzzy definition of the k-means algorithm to define a fuzzy model. Other non-parametric colour-naming models are based on the definition of a look-up-table (LUT) that assigns any point of the space with a colour name as in [139, 91] and sometimes also a membership value as in [145].

Semi-Parametric models

Semi-parametric models have parametric and non-parametric components. Usually, these models are composed of multiple parametric components. If the number of parametric components is elevated, the model is able to fit almost any distribution of samples as happens with non-parametric models. Hence, semi-parametric models have the flexibility of non-parametric models but with some of the advantages of parametric models such as the interpretability of its parameters. Typical semi-parametric models are mixture models and neural networks.

An example of the use of a Gaussian Mixture model in a colour task can be found in [101] where it is applied to skin detection. Neural networks have been applied to solve the colour constancy problem [36, 37]. In the field of colour naming Okajima used a network model to map elemental colour responses into colour names in [111].

In this work, we will focus on parametric models. Similarly, as it has been done in previous works, such as Mojsilovic in [103] or Seaborn et al in [124], we present the colour-naming task as a decision problem formulated in the frame of the fuzzy-set theory [81]. The essential difference of our proposal relies on the definition of a parametric model, that is, we propose a set of tuneable parameters that analytically define the shape of the fuzzy sets representing each colour category. The suitability of such approach can be summed up in the following points:

- **Inclusion of prior knowledge**: Prior knowledge about the structure of the data allows choosing the best model on each case. However, this could turn into

a disadvantage if a non-appropriate function for the model is selected.

- **Compact categories**: Each category is completely defined by few parameters and training data do not need to be stored after an initial fitting process. This implies lower memory usage and lower computation when the model is applied.

- **Meaningful parameters**: Each parameter has a meaning in terms of the characterization of the data, which allows modifying and improving the model by just adjusting the parameters.

- **Easy Analysis**: As a consequence of the previous one, the model can be analysed and compared by studying the values of its parameters.

The parameters of a model must be estimated from a set of examples. This learning process can be done in different ways [44]. Hence, learning methods are classified in three groups according to whether the set of examples, i.e. the learning set, is labelled or not. These three groups are the following:

Supervised learning

In supervised learning the samples used to estimate the parameters of a model have been previously assigned with a category label or cost and the learning process consists on minimising a cost function (i.e. the mean squared error or the number of misclassified samples).

Unsupervised learning

In unsupervised learning or clustering, the learning sample is not previously labelled and the process leads to the formation of clusters according to a criteria or cost function defined in the learning algorithm.

Reinforcement learning

In reinforcement learning a labelled learning sample is not provided but the process has feedback information. After assigning a category, the process has information whether the category assigned is correct or not, but information about how wrong it is, is not provided.

In colour-naming modelling, learning has normally been supervised [26, 85, 103, 124, 145]. However, unsupervised learning has also been used as in the approach of Yendrikhovskij [159, 160] that derived colour categories from the distribution of samples obtained from a wide set of images of natural scenes.

In our case, learning will be supervised and, we will use a set of psychophysical data, D, composed by a set of samples from the colour space and their membership values to the eleven categories,

$$D = \{< \vec{s}_i, m_1^i, ..., m_{11}^i >\}, \qquad i = 1, \ldots, n_s \qquad (4.1)$$

where $\vec{s}_i$ is the ith sample of the learning set, n_s is the number of samples in the learning set and m_k^i is the membership value of the ith sample to the kth category.

Such data will be the knowledge basis for a fitting process to estimate the model parameters. The experiment presented in the previous chapter had the goal of building an adequate psychophysical fuzzy dataset for that purpose. Hence, the fitting process will use these data as learning set.

In the following sections, we propose different parametric models to automate the colour-naming task. In all the cases, the procedure to build the model has been the following:

1. **Model definition.**
   The analysis of the psychophysical data from the experiment allows defining the properties that the colour-naming model should fulfil. This prior knowledge must be considered to find the analytic form of the membership functions of the model.

2. **Parameter estimation.**
   Estimating the model parameters implies to solve a non-linear data-fitting problem in the least squares sense, that is, finding an estimation of the parameters, $\hat{\theta} = (\hat{\theta}_{C_1}, ..., \hat{\theta}_{C_{11}})$, that minimises the mean squared error (MSE) between the membership values provided by the model and the set of colour-naming judgements used as learning set:

   $$\hat{\theta} = \arg\min_{\theta} \frac{1}{n_s} \sum_{i=1}^{n_s} \sum_{k=1}^{11} (\mu_{C_k}(\vec{s_i}; \theta_{C_k}) - m_k^i)^2 \qquad (4.2)$$

   where:
   $\hat{\theta}$ is the estimation of the parameters of the model
   n_s is the number of samples in the learning set
   μ_{C_k} is the membership function of the colour category C_k
   $\vec{s_i}$ is the ith sample of the learning set
   θ_{C_k} is the set of parameters of category C_k
   m_k^i is the membership value of the ith sample of the learning set to the kth category.

   Minimisations for the first approaches that will be presented were done by using a subspace trust region method which is based on the interior-reflective method proposed in [38, 39]. In the final model, minimisations were performed by using the Nelder-Mead simplex search method proposed in [84].

3. **Evaluation.**
   The estimated model can be evaluated in terms of different measures. Alexander in [9] proposes three classes of measurement that can be made on models to evaluate their performance:

   - *Goodness of fit to the sample distribution.*
     The evaluation of the fitting process will be done in terms of the mean absolute error (MAE_{fit}) between the learning set memberships and the memberships obtained from the parametric membership functions:

     $$MAE_{fit} = \frac{1}{n_s} \frac{1}{11} \sum_{i=1}^{n_s} \sum_{k=1}^{11} |m_k^i - \mu_{C_k}(\vec{s_i})| \qquad (4.3)$$

where:

n_s is the number of samples in the learning set

m_k^i is the membership of $\vec{s_i}$ to the kth category

$\mu_{C_k}(\vec{s_i})$ is the parametric membership of $\vec{s_i}$ to the kth category provided by the model

The value of MAE_{fit} is a measure of the accuracy of the model fitting to the learning dataset and is used here instead of the usual MSE because MAE is more intuitive and easily interpretable in terms of the membership error for membership values.

The MAE_{fit} measure will also be computed for any individual category, C_k, by considering only the membership function, μ_{C_k}, and the membership values, m_k, of the category being evaluated:

$$MAE_{fit} = \frac{1}{n_s} \sum_{i=1}^{n_s} |m_k^i - \mu_{C_k}(\vec{s_i})| \tag{4.4}$$

- *Goodness of fit to the parent distribution.*
  The evaluation of the fitting of the model to the real distribution can be done by computing error measures after applying the model to data that was not used in the learning step.

  If the available sample is abundant, it can be divided in two sets: one for learning and one for testing. In our case, these measure will only be computed for the final proposed model because for the first proposed models the set of available data did not suffice to build two different sets with enough data on each.

  However, to evaluate, up to a certain degree, the fitting of the estimated models to the parent distribution we define a second measure to evaluate the degree of fulfilment of the fuzzy unity-sum constraint. According to the fuzzy framework in which our colour-naming model is formulated, the sum of the memberships to the eleven colour categories for any point of the colour space should be the unity, that is:

$$\sum_{k=1}^{11} \mu_{C_k}(\vec{s}; \theta_{C_k}) = 1, \qquad \forall \vec{s} \tag{4.5}$$

Considering as error the difference between the unity and the sum of all the memberships at a point, the measure proposed is:

$$MAE_{unitsum} = \frac{1}{n_s} \sum_{i=1}^{n_s} |1 - \sum_{k=1}^{11} \mu_{C_k}(\vec{s_i}))| \tag{4.6}$$

where n_s is the number of samples considered and μ_{C_k} is the membership function of category C_k.

- *Classification performance.*
  In many previous works on colour naming [27, 96, 85, 124], results have
  been evaluated in terms of the categorization of the Munsell array. To
  evaluate classification performance we will also categorize the Munsell ar-
  ray by computing, for each chip of the Munsell dataset, the colour-naming
  descriptor, $CD(\vec{s})$, proposed in equation (1.19) and by applying the maxi-
  mum criteria (equation (1.20)) as decision rule to assign a colour name to
  each chip of the Munsell dataset.
  The results of the works of Berlin and Kay [27] and Sturges and Whitfield
  [135] will be used as ground truth to evaluate the models. The error mea-
  sure will be the number of chips inside the boundaries proposed by Berlin
  and Kay (210 of the 329 chips on the Munsell array) that are named dif-
  ferently by models. In the case of Sturges and Whitfield data, the error
  measure will be the number of chips from their consensus areas that are
  named differently by models.

Once the model is defined, it will be possible to apply it to any given colour sample
by computing the membership of the sample to the eleven colour categories considered
that will compose the colour descriptor defined in 1.19. The information contained
in such descriptor can be used in different ways as it will be shown in Chapter 5. In
the following sections, the different proposed membership functions are presented.

4.2   A Preliminary Approach: The 3D-Gaussian Function

This first approach follows the idea of the work of Lammens [85] that assumes uni-
variate Gaussian membership functions given by:

$$G_n(\vec{s}; \vec{m}, \sigma) = e^{-\frac{1}{2}\left(\frac{\sqrt{(\sum_{i=1}^{N}(s_i-m_i)^2)}}{\sigma}\right)^2} \tag{4.7}$$

where N is the dimension of the colour space, $\vec{m}$ is the mean and σ is the variance.

Using univariate Gaussian membership functions on a three-dimensional colour
space implies to assume that colour categories are spherical. This is a strong assump-
tion. In our first approach we propose to use a multivariate Gaussian as membership
function defined as a three-dimensional function, $G : \mathbb{R}^3 \to \mathbb{R}$, given by:

$$G(\vec{s}; \theta_G) = e^{-\frac{1}{2}(\vec{s}-\vec{m})^T \Sigma^{-1}(\vec{s}-\vec{m})} \tag{4.8}$$

where $\theta_G = (\vec{m}, \Sigma)$ is the set of parameters of the Gaussian function, $\vec{m}$ is the
mean and Σ is the covariance matrix.

Hence, each colour category, C_k will have membership function μ_{C_k}:

$$\mu_{C_k}(\vec{s}; \theta_k) = G(\vec{s}; \theta_k), \quad k = 1, \ldots, 11 \tag{4.9}$$

where θ_k is the set of parameters of the kth colour category C_k (see equation (1.21)
for the correspondence of colour categories and values of k).

Results

For each colour category, the mean, $\vec{m} = (m_1, m_2, m_3)$ and the covariance matrix, Σ, must be estimated. As Σ is a symmetric 3×3 matrix only one semi-matrix must be estimated. This implies that a total of 9 parameters must be estimated for each category.

Hence, estimating the parameters of the eleven colour categories by minimising the expression of equation (4.2) implies solving a minimisation problem of 99 variables. For implementation issues (i.e. complexity and convergence problems) estimation for each colour category, C_k, was done individually:

$$\hat{\theta}_{C_k} = \arg\min_{\theta_{C_k}} \frac{1}{n_s} \sum_{i=1}^{n_s} (\mu_{C_k}(\vec{s_i}; \theta_{C_k}) - m_k^i)^2 \tag{4.10}$$

where θ_{C_k} is the set of parameters of the category C_k, n_s is the number of samples in the learning set, μ_{C_k} is the membership function of the colour category C_k, $\vec{s_i}$ is the ith sample of the learning set, and m_k^i is the membership value of the ith sample of the learning set to the kth category.

The estimation process was done by using the learning set defined in previous section (see equation (4.1)) and obtained from the experiment presented in Chapter 3. The parameters obtained are presented in table 4.1. The MAE_{fit} measure defined by equation (4.4) was computed for each colour category. The values obtained and the total sum are presented in table 4.2.

The total error obtained means that each membership value provided by the model has, in average, an absolute difference of 0.0393. This global value is acceptable, but for some of the categories, such as *Green*, the error is too high to consider that the model provides correct membership values.

The categorization of the Munsell space obtained by the Gaussian model can be seen in figure 4.1 where each chip is painted with a colour representing the assigned category (the colour palette used has no special meaning). Chips with a highest membership lower than 0.5 are coloured in light blue meaning that they do not have a definite colour name.

In the figure, the boundaries defined by Berlin and Kay [27] for American English (figure 2.2) are superimposed. The chips inside the boundaries that are labelled differently by our model are marked with a cross. A total of 46 chips inside Berlin and Kay boundaries are wrongly labelled by the Gaussian model, which means a 21.90% of the total number of chips inside the boundaries. The same computation was done for the results of two previous works. These are from a 35 years old English speaker presented by MacLaury (MES) in [96] (see figure 4.2) and from Lammens Gaussian model (LGM) [85] (see figure 2.3). Results are summarized in table 4.3.

As can be seen in the table, our Multivariate Gaussian Model obtains similar results than previous computational Lammens model but it is far from MacLaury's English speaker. At this point, it is easy to see the subjectivity of the problem, since the categorization of MacLaury's English speaker has very important differences with Berlin and Kay division for English of the colour space. See, for example, that *Red* in the subject's categorization overlaps the *Purple* region in Berlin and Kay results.

Category	$\tilde{\mathbf{m}} = (\mathbf{L}, \mathbf{a}, \mathbf{b})$	$\mathbf{\Sigma}$		
Red	(50.23, 54.40, 31.36)	236.87	223.96	36.84
		223.96	545.05	346.59
		36.84	346.95	607.27
Orange	(68.30, 22.88, 44.92)	252.77	58.33	-175.80
		58.33	412.09	-244.37
		-175.80	-244.37	308.05
Brown	(43.95, 17.39, 27.38)	302.91	-25.87	-206.27
		-25.87	309.43	18.68
		-206.27	18.68	544.98
Yellow	(87.13, -4.91, 50.21)	59.65	-8.65	-20.16
		-8.65	606.57	-218.77
		-20.16	-218.77	190.51
Green	(59.40, -23.60, 19.12)	238.77	82.51	34.58
		82.51	979.89	925.42
		34.58	925.42	16608.00
Blue	(62.65, -12.63, -24.41)	466.71	-372.94	-861.42
		-372.94	592.12	1354.20
		-861.42	1354.20	4165.60
Purple	(47.90, 27.23, -21.19)	1041.50	231.18	-2189.50
		231.18	199.16	-469.55
		-2189.50	-469.55	5828.70
Pink	(66.85, 38.20, -0.08)	662.07	-259.73	-546.66
		-259.73	353.14	202.03
		-546.66	202.03	244.32
Black	(32.77, -1.25, -3.61)	0.77	0.40	-2.53
		0.40	0.53	-2.37
		-2.53	-2.374	11.80
Grey	(75.66, 0.35, -3.86)	18.59	-7.34	19.99
		-7.34	56.78	32.39
		19.99	32.39	351.44
White	(96.39, 3.13, -6.14)	0.18	0.81	0.55
		0.81	1.00	0.94
		0.55	0.94	0.88

Table 4.1: Parameters of the Gaussian model for the eleven colour categories.

	Red	Orange	Brown	Yellow	Green	Blue
MAE_{fit}	0.91	2.09	5.78	1.65	12.72	4.83
	Purple	Pink	Black	Grey	White	**Global**
MAE_{fit}	5.09	5.65	0.19	3.91	0.39	**3.93**

Table 4.2: Results in terms of the MAE_{fit} ($\times 10^{-2}$) of the Gaussian functions for each of the 11 colour categories. The mean for all the memberships is presented as the global error measure of the model.
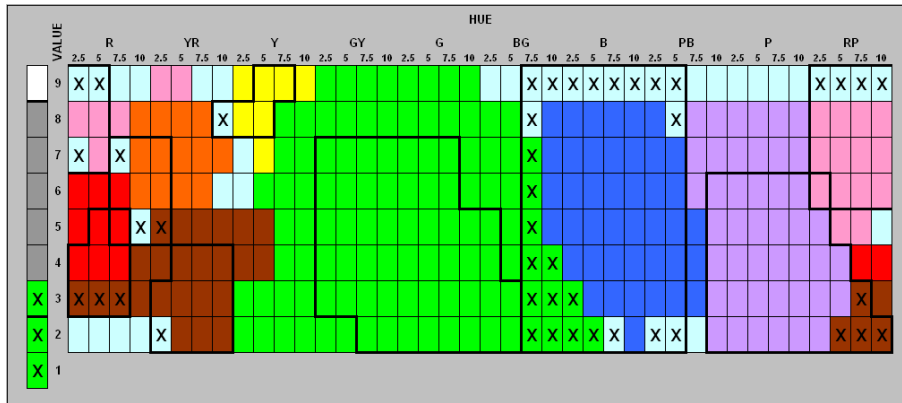
Figure 4.1: Categorization of the Munsell colour array obtained by applying the Gaussian model.
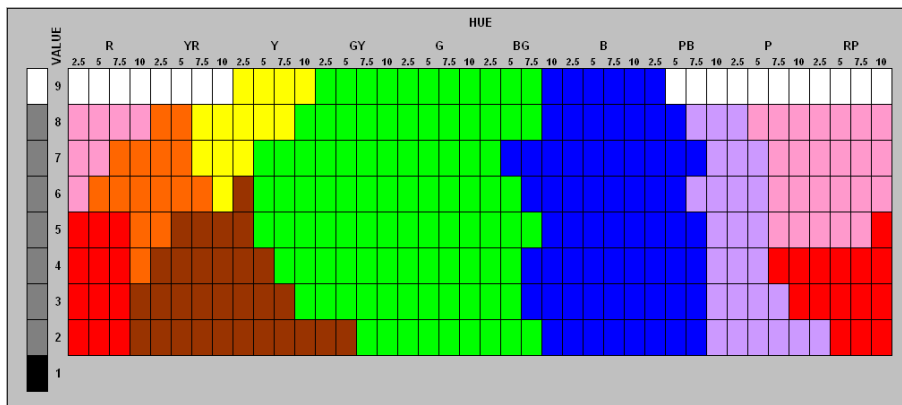
Figure 4.2: Categorization of the Munsell colour array provided by MacLaury's English speaker [96].

Model	Coincidences	Errors	% Errors
LGM	161	49	23.33
MGM	**164**	**46**	**21.90**
MES	182	28	13.33

Table 4.3: Comparison of the proposed Multivariate Gaussian Model (MGM) to other categorizations in terms of the coincidence with Berlin and Kay categorization.

Conclusion

Despite the global error of the model might seem not very high, several categories (*Brown, Green, Blue, Purple* and *Pink*) present high individual errors which implies that the 3D-Gaussian model is not appropriate to model them. In addition, the categorization of the Munsell space shows too many errors when compared to the boundaries derived by Berlin and Kay. These results brought us to conclude that a more complex model was needed to automate the colour-naming task. In the next section, a different approach is proposed.

4.3 Outline of the Colour-Naming Models Based on Sigmoid Functions

If the CIELab space is sliced in a few number of levels along the L axis and all the samples in each level are represented on a chromaticity plane, we are able to see that the membership distributions of chromatic categories[1] are not normal distributions (figure 4.3(a)).

As can be seen in figure 4.3(b), the membership distribution that can be inferred seems to have a triangular form, with values of 1 in the centre and a graded decrease from membership 1 to 0 at the two sides of the triangle that are in contact with the neighbouring categories. This vision that we can have on the two-dimensional chromaticity plane is more difficult to be inferred in the three-dimensional colour space.

For this reason, in our new proposal, the three-dimensional colour space is sliced in a set of N_L levels along the lightness axis (see figure 4.4), obtaining a set of chromaticity planes over which membership functions will be modelled by two-dimensional functions.

Therefore, any specific chromatic category will be defined by a piecewise function, μ_{C_k}, depending on lightness intervals, as it is expressed in equation (4.11).

$$\mu_{C_k}(\vec{s}) = \begin{cases} \mu_{C_k}^1(c_1, c_2) & \text{if } I < I_1, \\ \mu_{C_k}^2(c_1, c_2) & \text{if } I_1 < I < I_2, \\ \vdots & \vdots \\ \mu_{C_k}^{N_L}(c_1, c_2) & \text{if } I_{N_L-1} < I, \end{cases} \tag{4.11}$$

---

[1] Red, Orange, Brown, Yellow, Green, Blue, Purple, and Pink

(a)

(b)
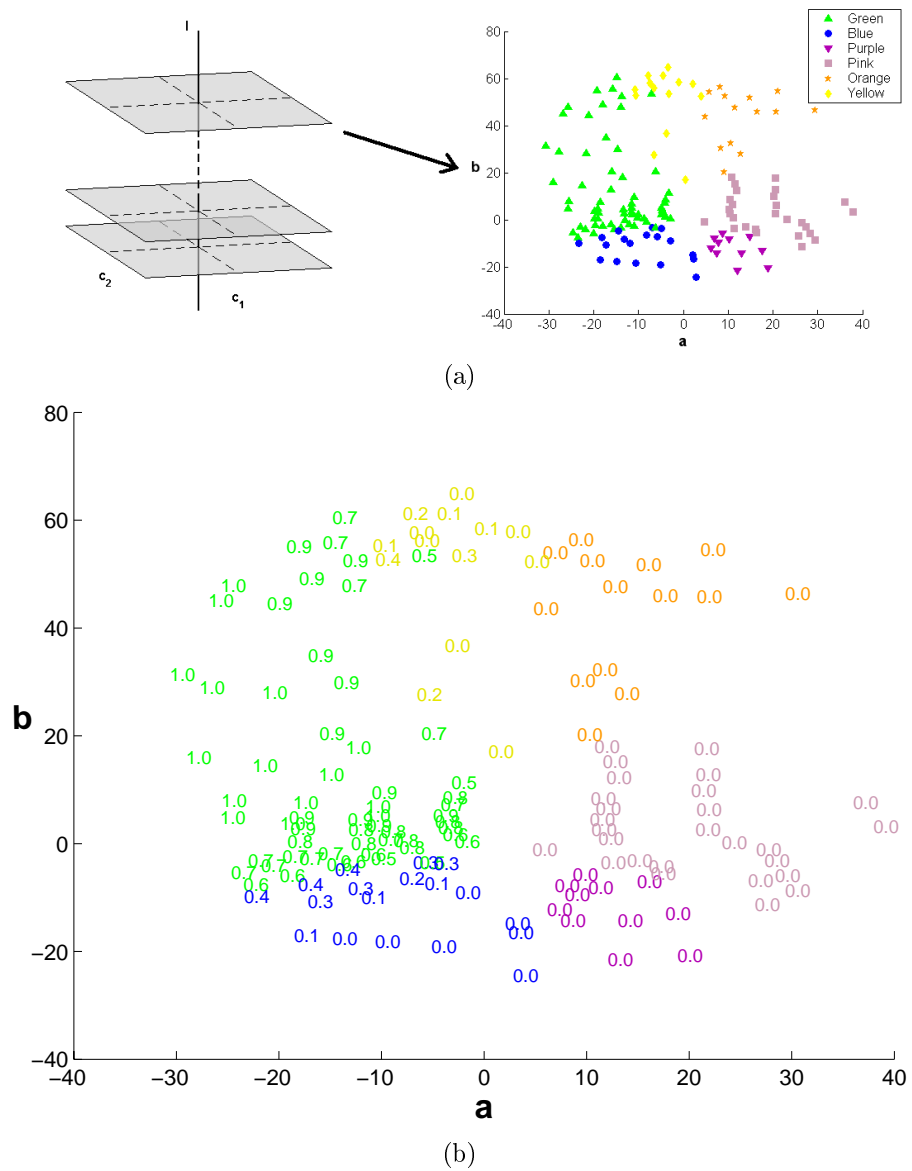
Figure 4.3: The distributions of the membership values of chromatic categories are not a normal distributions. (a) Representation of the samples of a given interval on a chromaticity plane. (b) Membership values to category *Green* of all the samples in the plane. The colour of the numbers indicates the category for which the sample has its maximum membership value.
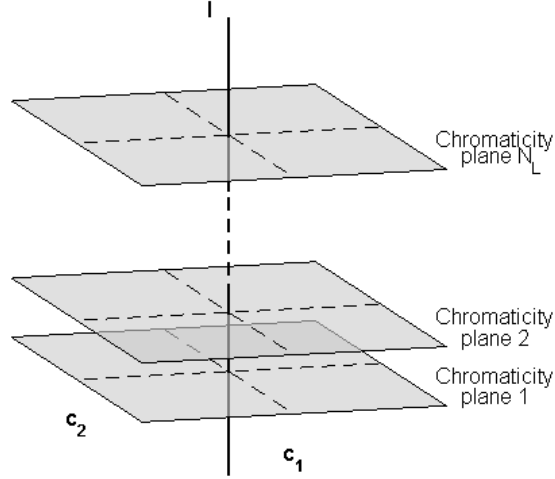
Figure 4.4: Scheme of the model. The colour space is divided in N_L levels along the lightness axis.

where $I_1, \ldots, I_{N_L-1}$ are the values that define the levels and chromaticity planes along the lightness axis. The number of chromaticity planes and the values that define them must be defined before the estimation process and depend on the data set available for the learning process.

As we did in previous section, we propose to work on a perceptual uniform colour space as it is the CIELab. However, other spaces could be suitable whenever one of the dimensions correlates with colour lightness and the other two with chromaticity components.

For generality, while the models are being formulated, we will denote any colour sample as $\vec{s} = (I, c_1, c_2)$, where I is the lightness and, c_1 and c_2 are the chromaticity components. For the case of models fitted on the CIELab space, samples will be denoted as $\vec{s} = (L, a, b)$, that is, L coordinate is identified with I and a and b coordinates from CIELab are idetified with c_1 and c_2 respectively.

After the previous preliminary approach based on Gaussian functions, we studied the membership distributions on the chromaticity planes obtained from the colour-naming experiment of Chapter 3 to find the set of properties that candidates functions should fulfil to be the adequate membership function in the colour-naming model.

The study of the membership values of the psychophysical data over a chromaticity plane (figure 4.3) allowed us to define the desirable properties that should fulfil a membership function, $\mu_{C_k}(c_1, c_2)$, for the chromatic categories in a given chromaticity plane:

- *Bounded:* Membership values provided must be in the $[0, 1]$ interval, i.e. $\mu_{C_k}(c_1, c_2) \in [0, 1]$

- *Triangular basis:* Membership functions present a triangular shaped basis with a principal vertex shared by all the categories. Within this basis, the distribution of the membership values forms a triangular plateau with all the colours that do not present confusion on its name assignment (i.e. membership value equal to one) on top of it.

- *Asymmetry:* Membership functions have parameters controlling the slope of naming certainty towards the neighbouring categories. These slopes can be different on each side of the category.

- *Central shift:* Membership functions have parameters allowing a shift of the triangular plateau of membership one with respect to the bisector of the triangular basis of the category.

In figure 4.5 we show an scheme of the preceding conditions on a chromaticity diagram where the samples of the colour-naming experiment have been plotted.
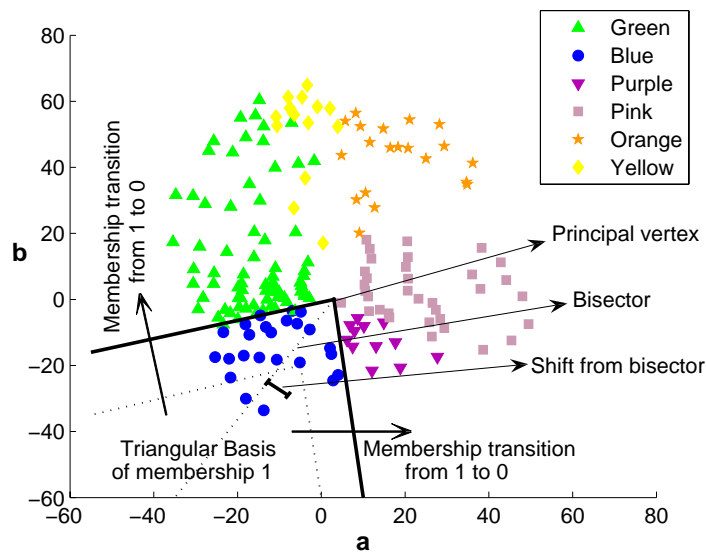
Figure 4.5: Desirable properties of the membership function for chromatic categories. In this case, on the *Blue* category.

Achromatic categories[2] will be modelled as a unique category on each chromaticity plane to differentiate them from chromatic ones. Differentiation between the three achromatic categories will be done only in terms of lightness.

In the following sections, we propose different membership functions for chromatic and achromatic categories, but in all the cases, the outline of the model is the one defined in this section.

---

[2]Black, Grey, White

4.4 Membership Functions for Fuzzy Colour Naming

In this section we detail different parametric functions proposed as membership functions. The results obtained with each proposed function are analysed and the model is evolved to improve its performance. The problems detected in these approaches have allowed us to set the basis of the final proposed model, the Triple-Sigmoid with Elliptical Centre (TSE) function.

4.4.1 The Sigmoid-Gaussian Function

The outline and the properties defined in the previous section are the basis to progressively find adequate functions to model the memberships to colour categories. As we saw in previous section, the functions to model memberships to chromatic and achromatic categories must be considered separately. In the following paragraphs the functions used to model both types are explained in detail.

Chromatic Categories

After defining the properties that the membership function for chromatic categories must fulfil, an adequate function must be defined. The proposed function in this section is a combination of two well-known functions: the sigmoid and the Gaussian functions. The proposed function, named Sigmoid-Gaussian (SG) [20, 23] from now on, is defined as a two-dimensional function, $SG : \mathbb{R}^2 \to \mathbb{R}$.

The definition of the SG starts from the one-dimensional sigmoid function:

$$S(x; \beta) = \frac{1}{1 + e^{-\beta x}} \qquad (4.12)$$

where β controls the slope of the transition from 0 to 1 (see figure 4.6(a)).

This can be extended to a two-dimensional sigmoid function, $S : \mathbb{R}^2 \to \mathbb{R}$, as:

$$S(\vec{p}; \beta, \vec{u}_i) = \frac{1}{1 + e^{-\beta \vec{u}_i \vec{p}}}, \qquad i = 1, 2 \qquad (4.13)$$

where $\vec{p} = (x, y)$ is a point in the plane, and vectors $\vec{u}_1 = (1, 0)$ and $\vec{u}_2 = (0, 1)$ define the axis in which the function is oriented (see figure 4.6(b)).

For simplicity, we will denote $S(\vec{p}; \beta, \vec{u}_1)$ as S_1 and $S(\vec{p}; \beta, \vec{u}_2)$ as S_2. By multiplying S_1 and S_2 we define a function, S_3, which fulfils the first three properties proposed before:

$$S_3(\vec{p}; \beta_x, \beta_y) = S_1(\vec{p}; \beta_y) \cdot S_2(\vec{p}; \beta_x) \qquad (4.14)$$

where β_y is the slope of the sigmoid oriented in the x-axis and β_x is the slope of the sigmoid oriented in the y-axis. Figure 4.7(a) shows an example of this function.

To achieve the last property, we propose to modulate the sigmoid with a Gaussian function in the direction perpendicular to the bisector of the category defined as:

$$G(\vec{p}; m, \sigma) = e^{-\frac{1}{2}\left(\frac{\frac{x-y}{\sqrt{2}} - m}{\sigma}\right)^2} \qquad (4.15)$$
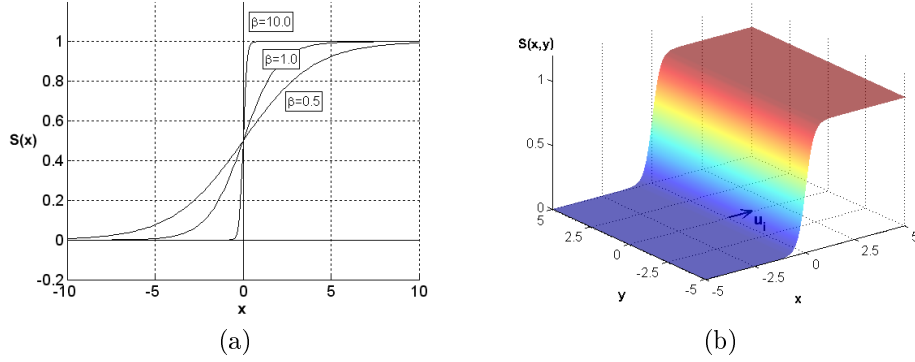
Figure 4.6: Sigmoid functions. (a) Sigmoid function in one dimension. The value of β determines the slope of the function. (b) Sigmoid function in two dimensions. Vector $\vec{u}_i$ determines the axis in which the function is oriented.

where m is the mean and σ is the standard deviation. Note that the value $\frac{x-y}{\sqrt{2}}$ is the distance of a point (x, y) to the line $x = y$. Figure 4.7(b) shows an scheme of the purpose of this function.
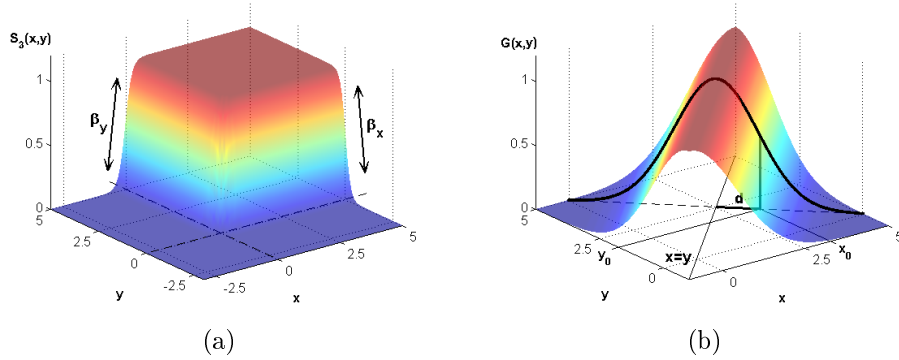
Figure 4.7: Sigmoid-Gaussian components. (a) Function S_3 as a result of the product of two 2D-sigmoids. (b) A 1D-gaussian function is used to modulate the sigmoid in the direction $x = y$. For a point (x_0, y_0) the sigmoid function is modulated by the value of the Gaussian at the position $d = (x_0 - y_0)/\sqrt{2}$, which is the distance of point (x_0, y_0) to line $x = y$.

Hence, the final expression of the function defined on the first quadrant of the space is:

$$SG(\vec{p}; \beta_x, \beta_y, m, \sigma) = S_3(\vec{p}; \beta_x, \beta_y) \cdot G(\vec{p}; m, \sigma) \tag{4.16}$$

By adding a translation, $\vec{t} = (t_x, t_y)$, and a rotation, α, to the previous equation, the function can be fitted to all the categories at the different locations of the chro-

matic plane. In order to represent the formulation in a compact matrix form, we will use homogeneous coordinates [64]. Let us redefine $\vec{p}$ to be a point in the plane expressed in homogeneous coordinates as $\vec{p} = (x, y, 1)$, and let us denote the vectors $\vec{u}_1 = (1, 0, 0)$ and $\vec{u}_2 = (0, 1, 0)$. The resulting function is:

$$SG(\vec{p}; \theta_{SG}) = \frac{1}{1 + e^{-\beta_y \vec{u}_1 R_\alpha T_t \vec{p}}} \cdot \frac{1}{1 + e^{-\beta_x \vec{u}_2 R_\alpha T_t \vec{p}}} \cdot e^{-\frac{1}{2}\left(\frac{\frac{\vec{u}_1 R_\alpha T_t \vec{p} - \vec{u}_2 R_\alpha T_t \vec{p}}{\sqrt{2}} - m}{\sigma}\right)^2} \tag{4.17}$$

where $\theta_{SG} = (\vec{t}, \alpha, \beta_x, \beta_y, m, \sigma)$ is the set of parameters of the Sigmoid-Gaussian function, and T_t and R_α are a translation matrix and a rotation matrix respectively:

$$T_t = \begin{pmatrix} 1 & 0 & -t_x \\ 0 & 1 & -t_y \\ 0 & 0 & 1 \end{pmatrix} \qquad R_\alpha = \begin{pmatrix} \cos(\alpha) & \sin(\alpha) & 0 \\ -\sin(\alpha) & \cos(\alpha) & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{4.18}$$

Figure 4.8, shows the shape of the SG function.
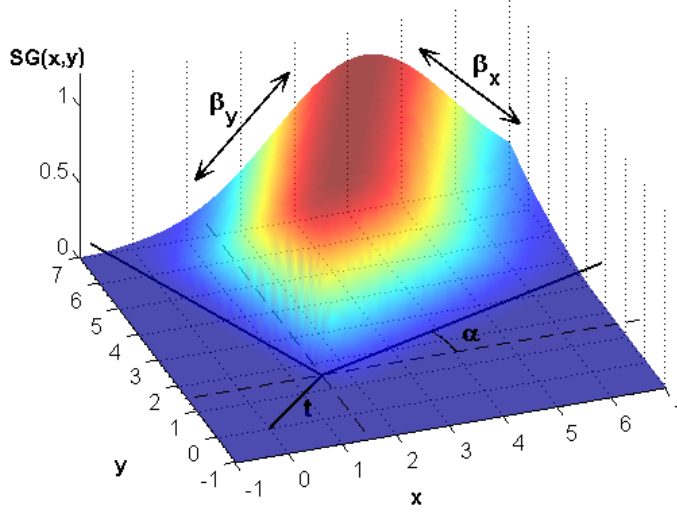
Figure 4.8: Example of the Sigmoid-Gaussian function.

Hence, once we have the analytic form of the chosen function, the membership function, μ_{C_k}, for a chromatic category, C_k, is given by:

$$\mu_{C_k}(\vec{s}) = \begin{cases} \mu_{C_k}^1 = SG(c_1, c_2; \theta_{C_k}^1) & \text{if } I \leq I_1, \\ \mu_{C_k}^2 = SG(c_1, c_2; \theta_{C_k}^2) & \text{if } I_1 < I \leq I_2, \\ \vdots & \vdots \\ \mu_{C_k}^{N_L} = SG(c_1, c_2; \theta_{C_k}^{N_L}) & \text{if } I_{N_L - 1} < I, \end{cases} \tag{4.19}$$

where $\vec{s} = (I, c_1, c_2)$ is a sample on the colour space, N_L is the number of chromaticity planes (i.e. lightness levels), $\theta_{C_k}^i$ is the set of parameters of the category C_k on the ith chromaticity plane and I_i are the lightness values that divide the space in the N_L lightness levels.

Achromatic Categories

Due to their position in the chromatic plane the three achromatic categories (*Black*, *Grey* and *White*) are first considered as a unique category at each chromaticity plane, and the global achromatic membership, μ_A, is modelled for each level with a multivariate two-dimensional Gaussian function:

$$\mu_A^i(c_1, c_2) = e^{-\frac{1}{2}((c_1,c_2)-\vec{m})^T \Sigma^{-1}((c_1,c_2)-\vec{m})} \tag{4.20}$$

where i is the lightness level that contains the sample $\vec{s} = (c_1, c_2, I)$, $\vec{m}$ is the mean and Σ is the covariance matrix.

The differentiation between the three achromatic categories must be done in terms of lightness. To model the fuzzy boundaries between these three categories we use one-dimensional Gaussian functions along the lightness axis:

$$\mu_{A_{C_k}}(I; \theta_{C_k}) = e^{-\frac{1}{2}\left(\frac{I-m_k}{\sigma_k}\right)^2}, \qquad 9 \leq k \leq 11 \tag{4.21}$$

where $\theta_{C_k} = (m_k, \sigma_k)$ is the set of parameters for category C_k with values of k from 9 to 11 which are the values that correspond to the achromatic categories (see equation (1.21)). Figure 4.9 shows an scheme of this division along the lightness axis.
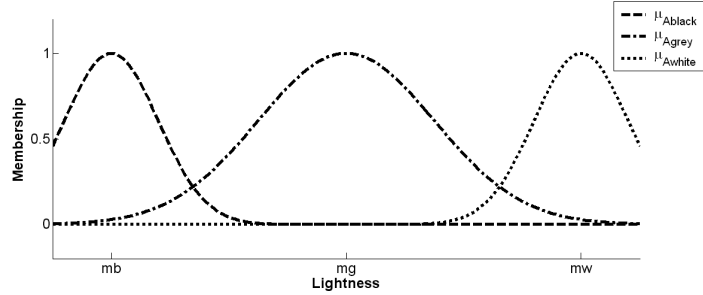
Figure 4.9: Sigmoid functions are used to differentiate between the three achromatic categories.

Hence, the membership to the three achromatic categories of a given sample is computed by weighting the global achromatic membership (equation (4.20)) with the corresponding membership in the lightness dimension (equation (4.21)):

$$\mu_{C_k}(\vec{s}; \theta_{C_k}) = \mu_A^i(c_1, c_2) \cdot \mu_{A_{C_k}}(I; \theta_{C_k}), \qquad 9 \leq k \leq 11, \quad I_i < I \leq I_{i+1} \tag{4.22}$$

where i is the lightness level in which the sample is included and the values of k correspond to the achromatic categories (see equation (1.21)).

Results

Once the Sigmoid-Gaussian colour-naming model has been defined, the next step is to estimate the parameters for each membership function. The number of lightness levels was set to three, i.e. $N_L = 3$, and the values that divide them were set at L=55 and L=75. These values were chosen in order to isolate some categories in only one or two of the intensity levels (i.e. *Yellow* is only present for high intensity and *Orange* does not appear for low intensity), because this fact provided the best results in the fitting process.

At each lightness level, the parameters of the function must be estimated for each colour category by fitting the function to the samples included in that level. The parameters obtained as a result of the estimation process are presented in table 4.4.

The results in terms of the fitting error MAE_{fit} for all the categories (equation (4.4)) are presented in table 4.5. The global fitting error of the model (equation (4.3)) is 0.0281. If these results are compared with the obtained for the Gaussian model, we can see that the improvement is important for some of the categories (e.g. *Green*, *Purple* and *Pink*), although other categories (e.g. *Red* and *Yellow*) have higher fitting error. Globally, the results confirm that the proposed Sigmoid-Gaussian model is modelling the colour categories better than the Gaussian model.

To compute the $MAE_{unitsum}$ error measure (equation (4.6)), we have sampled each of the chromaticity planes with values from -80 to 80 at steps of 0.5 units on both a and b axis, which means a number of points $n_p = 76800$. The value obtained of $MAE_{unitsum} = 0.3612$ indicates that the model does not provide a good fulfilment of that constraint. The value of the measure means that, in average, the sum of the memberships for each point is of only 0.6388 and, therefore, some areas of the space are not well modelled.

The categorization of the Munsell array obtained with the Sigmoid-Gaussian model is presented in figure 4.10. As we did in the Gaussian model, each chip is painted with a colour representing the assigned category. Chips without a definite colour term, that is, a highest membership value below 0.5, are coloured in light blue.

The Sigmoid-Gaussian categorization has a total of 35 chips inside Berlin and Kay boundaries named differently. This means an error rate of 16.67%. Table 4.6 shows a summary of the categorization considered up to now (the proposed Sigmoid-Gaussian model (SGM), Lammens Gaussian model (LGM), [85], an English speaker presented by Maclaury (MES) in [96], and the previously proposed Multivariate Gaussian model (MGM)).

The results show that the Sigmoid-Gaussian model improves the results of the previous Multivariate Gaussian model but in some areas there are still too many errors to produce a coherent categorization. Although the errors with respect to Berlin and Kay and to MacLaury's English speaker have been reduced, the differences are still considerable.

Despite the improvements obtained by the model, some problems can be noticed. The effect of the Gaussian, which is useful for some of the categories, is counterproductive for others such as *Green* or *Blue*. As the Gaussian is multiplying the product of the two sigmoids, to obtain a wide plateau we need to have a high value of σ. Anyway, the plateau is never a perfect one since the Gaussian has a smoothing effect

Achromatic axis		
Black	$m = 30.72$	$\sigma = 8.08$
Grey	$m = 74.60$	$\sigma = 28.76$
White	$m = 98.52$	$\sigma = 2.80$

Chromaticity plane 1

Achromatic	$\vec{m} = (-0.87, -4.60)$				$\Sigma = \begin{pmatrix} 26.54 & -2.59 \\ -2.59 & 51.54 \end{pmatrix}$	
	t_a	t_b	α	β_a	β_b	m	σ
Red	0.78	5.53	-0.04	110.03	427.05	15.16	6.47
Brown	0.78	4.53	-0.11	1.10	0.24	-9.01	14.62
Green	1.78	4.61	-1.80	123.30	4.35	34.48	131.26
Blue	0.78	5.53	2.71	0.24	0.68	-81.40	167.35
Purple	1.14	4.56	1.30	18.30	96.97	0.26	17.32
Pink	1.01	5.31	0.21	8.80	165.78	23.99	3.81

Chromaticity plane 2

Achromatic	$\vec{m} = (0.15, -5.49)$				$\Sigma = \begin{pmatrix} 41.93 & 3.56 \\ 3.56 & 30.41 \end{pmatrix}$	
	t_a	t_b	α	β_a	β_b	m	σ
Red	0.32	5.33	0.45	50.23	210.41	-15.83	7.14
Orange	-0.45	5.12	-0.67	6.24	1070.20	19.73	13.41
Brown	-0.47	5.12	-1.00	92.90	195.87	14.25	4.83
Yellow	0.08	5.91	0.90	19.04	51.59	0.50	0.21
Green	0.53	5.68	-1.52	92.74	39.25	0.35	29.58
Blue	-0.21	6.12	2.80	0.45	22.40	-1.88	39.16
Purple	-0.47	6.12	1.33	348.97	0.63	3.65	11.75
Pink	-0.47	5.71	0.68	208.64	31.60	-1.41	12.93

Chromaticity plane 3

Achromatic	$\vec{m} = (1.24, -3.27)$				$\Sigma = \begin{pmatrix} 13.15 & -7.09 \\ -7.09 & 58.55 \end{pmatrix}$	
	t_a	t_b	α	β_a	β_b	m	σ
Orange	-1.84	3.27	-0.71	57.54	38.67	4.25	2.06
Brown	-0.87	2.46	-0.84	0.00	0.79	2.39	1.74
Yellow	-1.57	2.75	-0.94	130.36	53.80	0.46	0.26
Green	-0.87	3.14	-1.81	0.23	41.75	19.14	36.78
Blue	-1.13	3.46	-2.76	173.15	60.72	-8.74	10.60
Purple	-1.87	3.46	1.82	0.61	0.49	-7.22	6.34
Pink	-1.87	2.46	0.35	1.71	1.03	37.62	53.84

Table 4.4: Parameters of the Sigmoid-Gaussian model.

	Red	Orange	Brown	Yellow	Green	Blue
MAE_{fit}	1.62	2.24	4.10	3.79	6.55	3.08
	Purple	Pink	Black	Grey	White	**Global**
MAE_{fit}	2.95	2.92	0.23	2.96	0.48	**2.81**

Table 4.5: Results in terms of the MAE_{fit} ($\times 10^{-2}$) of the Sigmoid-Gaussian functions fitting for each one of the 11 colour categories considered. In the last column, the global measure of error of the model.
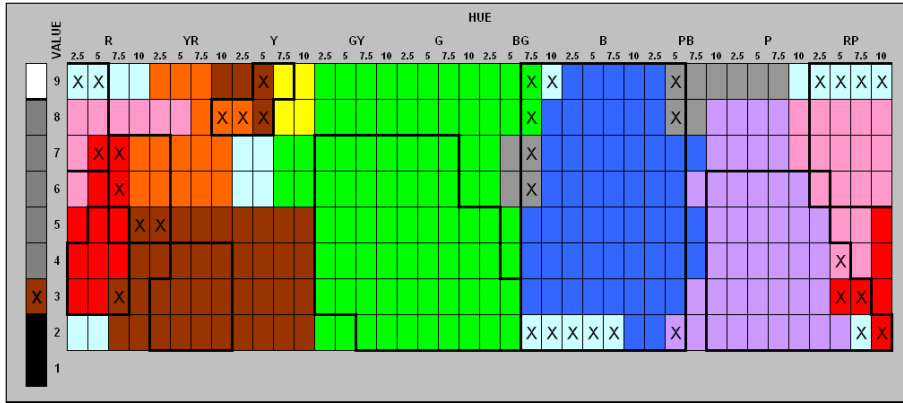
Figure 4.10: Categorization of the Munsell colour array obtained by applying the proposed Sigmoid-Gaussian model. Differences to Berlin and Kay categorization are marked with a cross.

Model	Coincidences	Errors	% Errors
LGM	161	49	23.33
MGM	164	46	21.90
SGM	**175**	**35**	**16.67**
MES	182	28	13.33

Table 4.6: Comparison of the different categorizations in terms of the coincidence with Berlin and Kay categorization.

over the surface formed by the product of the two sigmoids. Another drawback of the Sigmoid-Gaussian function is that it does not allow having membership functions covering a surface wider than an angle of $\frac{\pi}{2}$ and this makes that categories such as *Green* can not be correctly modelled. This lack of flexibility of the function makes that in some regions of the colour space the sum of the membership values for the eleven colour categories is not the unit, as it should be according to the fuzzy framework.

4.4.2 The Double-Sigmoid Function

As we have seen in the previous section, the Sigmoid-Gaussian function present some drawbacks. Firstly, the Sigmoid-Gaussian used as membership function was not suitable for some of the categories (e.g. *Green* and *Blue*) due to its lack of flexibility to adopt the form presented by these categories. Secondly, there were some regions of the colour space where the sum of the membership values for the eleven colour categories was not the unit, as it should be according to a fuzzy framework. To overcome the problems of the Sigmoid-Gaussian model, we propose a new membership function for chromatic categories based only on the product of two sigmoids.

The new proposed function, the Double-Sigmoid (DS) function [22], is a two-dimensional function, $DS : \mathbb{R}^2 \rightarrow \mathbb{R}$, and its definition is based on the functions defined by equation (4.13).

Let us add a translation, $\vec{t} = (t_x, t_y)$, and a rotation, α, to the function of equation (4.13). Thus, we redefine S_1 as a function with orientation α with respect to axis y and S_2 as a function with orientation α with respect to axis x:

$$S_i(\vec{p}; \vec{t}, \alpha, \beta) = \frac{1}{1 + e^{-\beta \vec{u}_i R_\alpha T_t \vec{p}}}, \qquad i = 1, 2 \tag{4.23}$$

where T_t and R_α are a translation matrix and a rotation matrix respectively (equation (4.18)).

By multiplying S_1 and S_2 we define the DS function, which fulfils the properties proposed before in 4.4.1, but avoiding the use of the Gaussian function:

$$DS(\vec{p}; \vec{t}, \theta_{DS}) = S_1(\vec{p}; \vec{t}, \alpha_y, \beta_y) \cdot S_2(\vec{p}; \vec{t}, \alpha_x, \beta_x) \tag{4.24}$$

where $\theta_{DS} = (\alpha_x, \alpha_y, \beta_x, \beta_y)$ is the set of parameters of the Double-Sigmoid function. Functions S_1, S_2 and DS are plotted in figure 4.11.

Hence, once we have the analytic form of the chosen function, the membership function for a chromatic category, μ_{C_k}, is given by:

$$\mu_{C_k}(\vec{s}) = \begin{cases} \mu^1_{C_k} = DS(c_1, c_2; \theta^1_{C_k}) & \text{if } I \leq I_1, \\ \mu^2_{C_k} = DS(c_1, c_2; \theta^2_{C_k}) & \text{if } I_1 < I \leq I_2, \\ \vdots & \vdots \\ \mu^{N_L}_{C_k} = DS(c_1, c_2; \theta^{N_L}_{C_k}) & \text{if } I_{N_L-1} < I, \end{cases} \tag{4.25}$$

where $\vec{s} = (I, c_1, c_2)$ is a sample on the colour space, N_L is the number of chromaticity planes, $\theta^i_{C_k}$ is the set of parameters of the DS function of the category C_k on the ith

(a) $S_1(x, y)$

(b) $S_2(x, y)$

(c) $DS(x, y)$
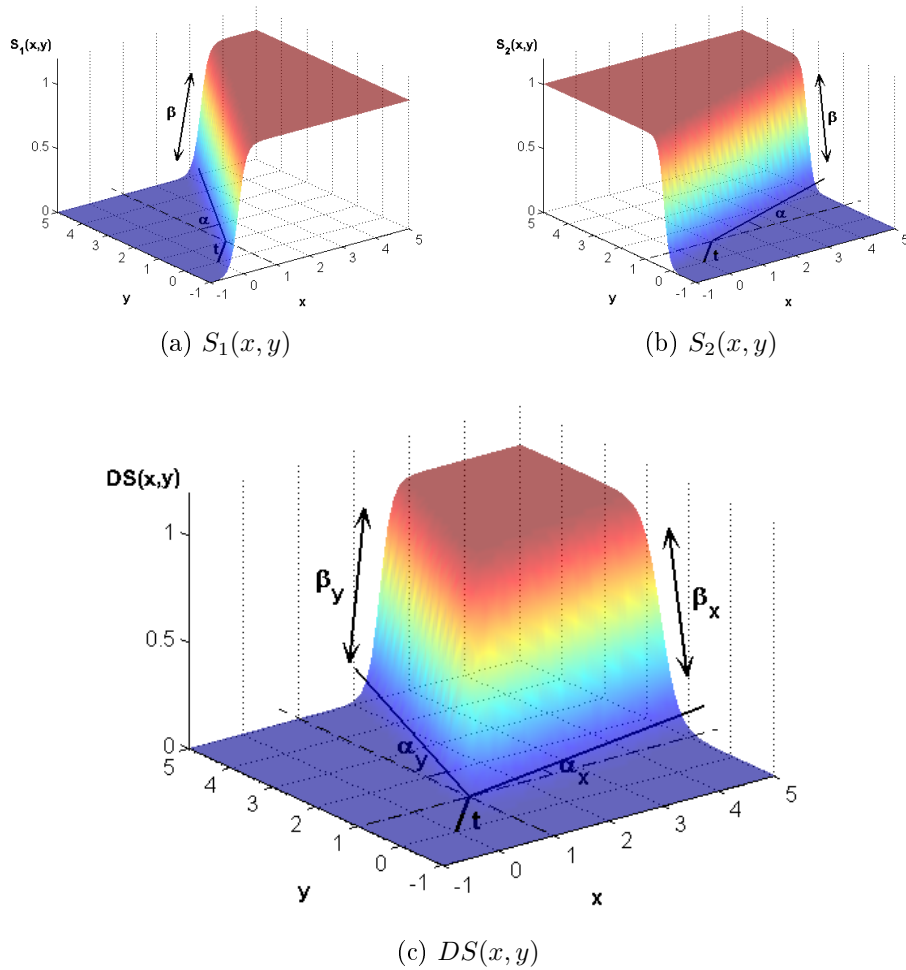
Figure 4.11: Two-dimensional sigmoid functions. (a) S_1: Sigmoid function oriented in axis y direction (b) S_2: Sigmoid function oriented in axis x direction (c) DS: The product of two differently oriented sigmoid functions generates a plateau with some of the properties needed for the membership function.

chromaticity plane and I_i are the lightness values that divide the space in the N_L lightness levels.

The fact that each sigmoid of the function has a different rotation angle allows varying the wideness of the plateau. Hence, this new function solves one of the problems detected on the Sigmoid-Gaussian model. The effect of the two angles is similar to the product by the Gaussian function, but it does not have the smoothing effect of the Gaussian that did not allow wide plateaus in the previous model. Therefore, the new function is more flexible and allows a wider range of shapes by just varying the parameters.

The modelling of achromatic memberships is done in the same way as in the model based on the Sigmoid-Gaussian (see section 4.4.1).

Results

The parameter estimation was done following the same procedure as in the case of the Sigmoid-Gaussian function. The parameters obtained for the Double-Sigmoid model after this fitting process are presented in table 4.7.

As in the previous models, the MAE_{fit} was computed for each colour category and in global. The values obtained are shown in table 4.8.

The value of MAE_{fit} of the model is 0.0304. This value is slightly higher than for the Sigmoid-Gaussian model. Instead, the value of $MAE_{unitsum}$[3] was of 0.1219 and shows a great improvement with respect to the Multivariate Gaussian and Sigmoid-Gaussian models for which $MAE_{unitsum}$ was 0.8206 and 0.3612 respectively. Despite this improvement, the fulfilment of the fuzzy constraint is still not satisfactory enough. Furthermore, one area where we observed that the fuzzy constraint of unit sum was far from being achieved, was the one corresponding to the achromatic categories. Table 4.9 summarizes these results.

The evaluation in terms of the categorization of the Munsell array can be seen in figure 4.12. The number of mislabelled chips inside Berlin and Kay boundaries is reduced to 28, which is the same value of the English speaker of [96] as can be seen in table 4.10.

Although the function proposed in this section, the DS function, improves the results in terms of the Munsell categorization of the previous approaches, the model still presents some drawbacks. The most important is that the fuzzy constraint of unity sum is not fulfilled in several areas of the colour space. The most evident as we mentioned above was in the areas close to achromatic categories. We try to solve this problem with the next proposed funcion, the Triple-Sigmoid with Circular centre function.

4.4.3   The Triple-Sigmoid with Circular Centre Function

The next function we propose, which we will refer as Triple-Sigmoid with Circular centre (TSC) [22], is a modification of the previous model (see equation (4.24)). The TSC function is defined as a two-dimensional function, $TSC : \mathbb{R}^2 \to \mathbb{R}$.

---

[3]The sampling of the colour space used to compute the measure was explained in section 4.4.1 and was the same for the three models

Achromatic axis

Black	$m = 30.72$	$\sigma = 8.08$
Grey	$m = 74.60$	$\sigma = 28.76$
White	$m = 98.52$	$\sigma = 2.80$

Chromaticity plane 1

Achromatic	$\vec{m} = (-0.87, -4.60)$	$\Sigma = \begin{pmatrix} 26.54 & -2.59 \\ -2.59 & 51.54 \end{pmatrix}$

	t_a	t_b	α_a	α_b	β_a	β_b
Red	-1.28	-5.03	0.23	0.90	50.00	50.00
Brown	-0.78	-4.53	0.74	-0.08	67.46	0.27
Green	-1.78	-4.53	1.61	-2.07	0.46	2.39
Blue	-1.78	-5.53	-2.65	2.84	16.44	0.67
Purple	-1.02	-5.01	-1.02	1.45	50.00	11.87
Pink	-1.28	-5.03	0.00	1.28	56.76	3.43

Chromaticity plane 2

Achromatic	$\vec{m} = (0.15, -5.49)$	$\Sigma = \begin{pmatrix} 41.93 & 3.56 \\ 3.56 & 30.41 \end{pmatrix}$

	t_a	t_b	α_a	α_b	β_a	β_b
Red	-0.13	-5.23	0.54	0.79	0.36	34.52
Orange	-0.03	-5.56	0.86	0.10	62.78	57.91
Brown	-0.05	-5.12	1.32	0.01	52.22	3.20
Yellow	-0.01	-5.60	1.62	-0.05	78.48	49.73
Green	-0.11	-5.12	1.60	-1.86	15.71	17.62
Blue	-0.01	-6.12	-2.98	2.91	42.41	7.81
Purple	-0.03	-5.62	-0.93	1.96	60.79	59.31
Pink	0.05	-5.66	-0.25	1.06	57.17	59.50

Chromaticity plane 3

Achromatic	$\vec{m} = (1.24, -3.27)$	$\Sigma = \begin{pmatrix} 13.15 & -7.09 \\ -7.09 & 58.55 \end{pmatrix}$

	t_a	t_b	α_a	α_b	β_a	β_b
Orange	1.87	-3.06	1.23	0.19	11.59	49.96
Brown	1.87	-2.46	1.63	-0.07	26.16	1.36
Yellow	1.32	-2.62	1.52	-0.25	53.46	47.12
Green	0.87	-3.18	1.81	-1.61	0.18	38.08
Blue	0.87	-3.46	-3.08	3.03	39.88	0.42
Purple	1.87	-2.50	-1.16	1.91	238.55	2.73
Pink	1.37	-2.96	-0.32	0.42	50.00	45.89

Table 4.7: Parameters of the Double-Sigmoid model.

	Red	Orange	Brown	Yellow	Green	Blue
MAE_{fit}	1.82	2.26	4.92	4.55	5.44	3.05
	Purple	Pink	Black	Grey	White	**Global**
MAE_{fit}	3.67	4.06	0.23	2.96	0.48	**3.04**

Table 4.8: Results in terms of the MAE_{fit} ($\times 10^{-2}$) of the Double-Sigmoid functions for each one of the 11 colour categories considered. In the last column, the global measure of error of the model.

Model	**MAE$_{fit}$**	**MAE$_{unitsum}$**
MGM	0.0393	0.8206
SGM	0.0281	0.3612
DSM	**0.0304**	**0.1219**

Table 4.9: Results obtained by the three models (Multivariate Gaussian(MGM), Sigmoid-Gaussian(SGM) and Double-Sigmoid(DSM)).
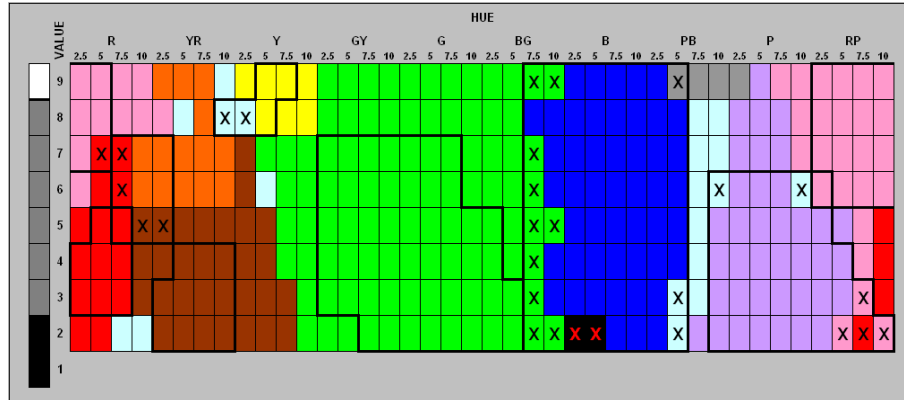
Figure 4.12: Munsell categorization obtained by the Double-Sigmoid model.

Model	**Coincidences**	**Errors**	**% Errors**
LGM	161	49	23.33
MGM	164	46	21.90
SGM	175	35	16.67
MES	182	28	13.33
DSM	**182**	**28**	**13.33**

Table 4.10: Comparison of the different categorizations in terms of the coincidence with Berlin and Kay categorization. The categorizations evaluated are Lammens Gaussian Model (LGM), MacLaury's English speaker (MES), the Multivariate Gaussian Model (MGM), the Sigmoid-Gaussian Model (SGM) and the Double-Sigmoid Model (DSM).

To formulate the TSC function, first let us define the Circular-Sigmoid (CS) function by including the circumference equation in the sigmoid formula:

$$CS(\vec{p}; \vec{t}, \theta_{CS}) = \frac{1}{1 + e^{-\beta_c \left(\left(\vec{u}_1 T_t \vec{p} \right)^2 + \left(\vec{u}_2 T_t \vec{p} \right)^2 - r^2 \right)}} \qquad (4.26)$$

where $\vec{p} = (x, y)$ is a point in the plane, $\theta_{CS} = (\beta_c, r)$ is the set of parameters of the Circular-Sigmoid function, β_c is the slope of the sigmoid curve that forms the circumference boundary, and r is the radius of the circumference. The function obtained is a circular plateau if β_c is negative and a circular valley if β_c is positive. The functions obtained can be seen in figure 4.13.

(a)   (b)
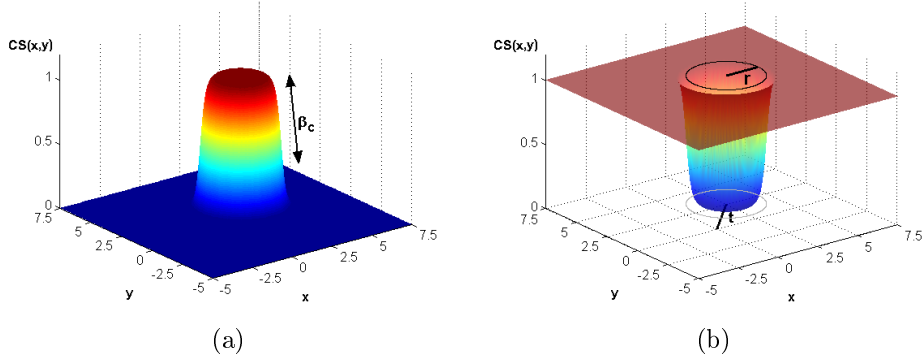
Figure 4.13: Circular-Sigmoid function $CS(\vec{p}; \vec{t}, \theta_{CS})$. (a) CS for $\beta_c < 0$ (b) CS for $\beta_c > 0$

Hence, we define the Triple-Sigmoid with Circular centre(TSC) by multiplying the Double-Sigmoid and the Circular-Sigmoid with a positive β_c, as:

$$TSC(\vec{p}; \theta_{TSC}) = DS(\vec{p}; \vec{t}, \theta_{DS}) \cdot CS(\vec{p}; \vec{t}, \theta_{CS}) \qquad (4.27)$$

where $\theta_{TSC} = (\vec{t}, \theta_{DS}, \theta_{CS})$ is the set of parameters of the TSC.

The TSC function fulfils the properties defined at the beginning of section 4.3. Figure 4.14 shows the form of the TSC function.

The effect of the third sigmoid allows improving the modelling of the central area of the colour space, that is, the boundary between chromatic and achromatic categories.

Hence, once we have the analytic form of the chosen function, the membership function for a chromatic category, μ_{C_k}, is given by:

$$\mu_{C_k}(\vec{s}) = \begin{cases} \mu_{C_k}^1 = TSC(c_1, c_2; \theta_{C_k}^1) & \text{if } I \leq I_1, \\ \mu_{C_k}^2 = TSC(c_1, c_2; \theta_{C_k}^2) & \text{if } I_1 < I \leq I_2, \\ \vdots & \vdots \\ \mu_{C_k}^{N_L} = TSC(c_1, c_2; \theta_{C_k}^{N_L}) & \text{if } I_{N_L-1} < I, \end{cases} \qquad (4.28)$$
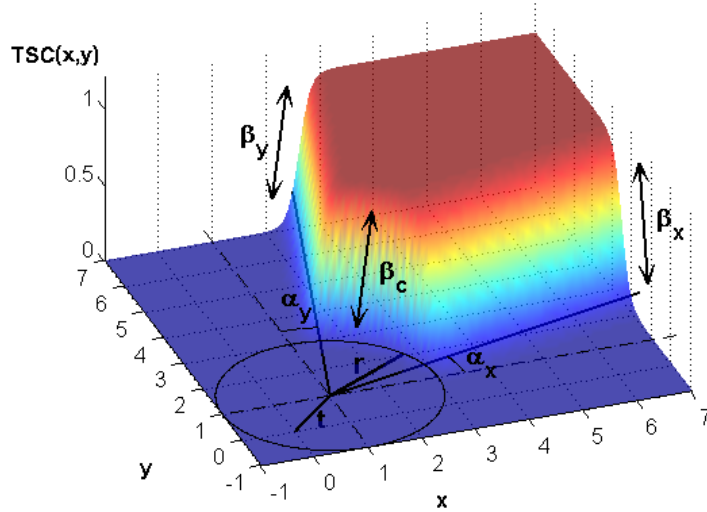
Figure 4.14: Triple-Sigmoid with Circular centre (TSC).

where $\vec{s} = (I, c_1, c_2)$ is a sample on the colour space, N_L is the number of chromaticity planes, $\theta^i_{C_k}$ is the set of parameters of the category C_k on the ith chromaticity plane and I_i are the lightness values that divide the space in the N_L lightness levels.

Results

The parameter estimation was done following the same procedure used for the previous models. The parameters obtained for de TSC model after this fitting process are presented in table 4.11.

As in the previous models, the MAE_{fit} was computed for each colour category and in global. The values obtained are shown in table 4.12.

The TSC model improves the results of the DS model in terms of the MAE_{fit} but the improvement is small. In fact, while some categories such as *Blue* are better modelled with the TSC model, other categories such as *Brown* are worse modelled. The value of $MAE_{unitsum}$, which for the TSC model is 0.1135, is also improved but, again, the improvement is not significant. Table 4.13 summarizes these results.

The categorization of the Munsell array obtained with the TSC model is shown in figure 4.15.

The comparison with the previous categorizations (Table 4.14) shows that the TSC model improves the results and has even less Munsell chips labelled different with respect to Berlin and Kay boundaries than a real English speaker. However, some of the errors, such as some chips at the areas of highest or lowest value which are labelled with achromatic names, must be improved.

As a conclusion of this section, we must say that although the global results obtained with the TSC function are slightly better than the ones obtained with the

Achromatic axis		
Black	$m = 30.72$	$\sigma = 8.08$
Grey	$m = 74.60$	$\sigma = 28.76$
White	$m = 98.52$	$\sigma = 2.80$

Chromaticity plane 1

Achromatic		$\vec{m} = (-0.87, -4.60)$			$\Sigma = \begin{pmatrix} 26.54 & -2.59 \\ -2.59 & 51.54 \end{pmatrix}$		
	t_a	t_b	α_a	α_b	β_a	β_b	β_c	r
Red	-1.28	-5.03	0.23	0.90	50.00	50.00	50.00	0.01
Brown	-0.98	-5.33	0.74	-0.07	0.65	0.30	35.86	14.14
Green	-1.18	-5.13	1.65	-2.03	0.46	0.58	20.97	14.13
Blue	-0.78	-5.53	-2.68	2.69	0.27	0.57	37.18	12.82
Purple	-1.02	-5.01	-1.02	1.45	50.00	13.31	50.00	0.01
Pink	-1.28	-5.03	0.01	1.28	54.28	113.90	50.00	0.01

Chromaticity plane 2

Achromatic		$\vec{m} = (0.15, -5.49)$			$\Sigma = \begin{pmatrix} 41.93 & 3.56 \\ 3.56 & 30.41 \end{pmatrix}$		
	t_a	t_b	α_a	α_b	β_a	β_b	β_c	r
Red	-0.04	-5.21	0.54	0.80	5.57	26.05	44.74	2.35
Orange	-0.03	-5.62	0.86	0.10	50.05	57.48	49.96	0.06
Brown	-0.05	-5.32	1.32	0.01	54.04	3.38	48.85	-0.70
Yellow	-0.02	-5.63	1.62	-0.05	50.02	49.87	50.00	0.03
Green	-0.03	-5.52	1.61	-1.78	0.21	109.19	48.05	1.18
Blue	-0.04	-6.11	-2.98	2.90	40.26	5.66	6.06	0.42
Purple	-0.03	-5.62	-0.93	1.96	70.71	50.05	49.99	0.08
Pink	-0.02	-5.63	-0.25	1.06	61.86	50.00	47.31	0.08

Chromaticity plane 3

Achromatic		$\vec{m} = (1.24, -3.27)$			$\Sigma = \begin{pmatrix} 13.15 & -7.09 \\ -7.09 & 58.55 \end{pmatrix}$		
	t_a	t_b	α_a	α_b	β_a	β_b	β_c	r
Orange	1.87	-3.46	1.24	0.11	17.46	49.86	50.00	0.01
Brown	1.67	-2.66	1.62	-0.07	4.88	1.16	50.00	0.01
Yellow	1.18	-2.46	1.64	-0.25	64.84	1.12	50.00	0.01
Green	1.27	-2.86	1.82	-1.64	0.22	0.65	50.00	0.01
Blue	0.87	-3.46	-3.08	3.03	39.82	0.42	50.00	0.01
Purple	1.87	-2.50	-1.16	1.91	237.39	2.73	50.00	0.01
Pink	1.37	-2.96	-0.32	0.42	50.00	45.89	50.00	0.01

Table 4.11: Parameters of the Triple-Sigmoid with Circular centre model.

	Red	Orange	Brown	Yellow	Green	Blue
MAE_{fit}	1.92	2.24	5.26	4.15	5.40	2.38
	Purple	Pink	Black	Grey	White	**Global**
MAE_{fit}	3.67	4.07	0.23	2.96	0.48	**2.98**

Table 4.12: Results in terms of the MAE_{fit} ($\times 10^{-2}$) of the Triple-Sigmoid with Circular centre functions for each one of the 11 colour categories considered. In the last column, the global measure of error of the model.

Model	$\mathbf{MAE_{fit}}$	$\mathbf{MAE_{unitsum}}$
MGM	0.0393	0.8206
SGM	0.0281	0.3612
DSM	0.0304	0.1219
TSCM	**0.0298**	**0.1135**

Table 4.13: Results obtained by the four models proposed up to now (Multivariate Gaussian (MGM), Sigmoid-Gaussian (SGM), Double-Sigmoid (DSM) and Triple-Sigmoid with Circular centre (TSCM)).
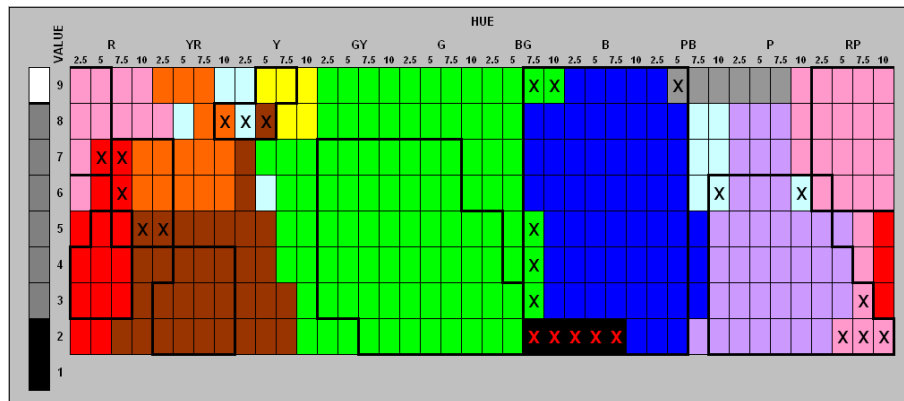
Figure 4.15: Munsell categorization obtained by the Triple-Sigmoid with Circular centre model.

Model	Coincidences	Errors	% Errors
LGM	161	49	23.33
MGM	164	46	21.90
SGM	175	35	16.67
MES	182	28	13.33
DSM	182	28	13.33
TSCM	**185**	**25**	**11.90**

Table 4.14: Comparison of the different categorizations in terms of the coincidence with Berlin and Kay categorization. The categorizations evaluated are Lammens Gaussian Model (LGM), MacLaury's English speaker (MES), the Multivariate Gaussian Model (MGM), the Sigmoid-Gaussian Model (SGM), the Double-Sigmoid Model (DSM), and the Triple-Sigmoid with Circular centre Model (TSCM).

DS function, the improvement is not significant because as can be seen comparing tables 4.8 and 4.12 the error decrease is concentrated in only two of the categories (*Yellow* and *Blue*), while for the rest of the categories the error increases or the decrease is minimal. Therefore, the TSC function has the same drawbacks noticed on the evaluation of the DS results. As we saw before, although the fitting error, MAE_{fit}, is acceptable, the value of $MAE_{unitsum}$ is still high which means that in some areas of the colour space the sum of the memberships for all the categories is far from the unity. This fact can be noticed on the obtained Munsell categorization where some areas do not have any category with membership above the threshold of 0.5. In addition, we have noticed that there are some errors on the Munsell categorizations which are concentred in the extremes of the lightness scale, that is, the samples with low and high value in the Munsell system. This problem is due to some deficiencies of the learning set with these samples that, in some cases, were unavailable. The fact that the fitting error is not high but the unity-sum error and the Munsell categorization are not satisfactory could also indicate some deficiencies of the learning set used or in the estimation step performed up to now.

4.5   The Final Colour-Naming Model

As we have seen in the previous section, the Triple-Sigmoid with Circular centre is the function among the proposed that obtains the best results in terms of the error measures defined and in terms of the categorization of the Munsell space. However, we also pointed out some problems that should be solved if a successful model for colour naming is desired.

In this section we will propose some variations on the previous models to obtain our final proposal, the Triple-Sigmoid with Elliptical centre function [25].

4.5.1   The Triple-Sigmoid with Elliptical Centre Function

The analysis of the results obtained with the previous functions, brought us to propose some modifications on the colour-naming models proposed in the previous sections:

1. The modelling of chromatic categories will be improved by redefining some of the properties that membership functions for chromatic categories should fulfil. To this end, a variant of the TSC function will be proposed.

2. Achromatic categories[4] are given as the complementary function of the chromatic ones but weighted by the membership function of each one of the three achromatic categories. These functions are also improved by using one-dimensional sigmoid functions instead of the Gaussian functions used before.

3. A wider learning set will be used to solve the problems presented by the learning set used in the previous models.

4. The parameter estimation process will be changed to consider the fuzzy constraint (i.e. the sum of all memberships at a given point must be one) and thus reduce the unity-sum error ($MAE_{unitsum}$).

In the following sections, the new proposed model is detailed.

Chromatic Categories

The fuzzy framework defined previously, the observation of the membership values of psychophysical data obtained from the colour-naming experiment, and the feedback from the previous proposed models (SG, DS, and TSC) made us to redefine the set of necessary properties that membership functions for the chromatic categories should fulfil:

- *Bounded*: Membership values must be in the [0,1] interval, i.e. $\mu_{C_k}(\vec{s}) \in [0, 1]$

- *Triangular basis*: Chromatic categories present a plateau, or area with no confusion about the colour name, with a triangular shaped basis and a principal vertex shared by all the categories.

- *Asymmetry*: For a given chromatic category the slope of naming certainty towards the neighbouring categories can be different on each side of the category (e.g. transition from *Blue* to *Green* can be different to that from *Blue* to *Purple*).

- *Central notch*: The transition from a chromatic category to the central achromatic one has the form of an elliptical notch around the principal vertex.

In figure 4.16 we show an scheme of the preceding conditions on a chromaticity diagram where the samples of the colour-naming experiment have been plotted.

The Triple-Sgimoid with Elliptical centre function (TSE) is a generalization of the TSC proposed in the previous section. The functions proposed previously, fulfilled the first three properties defined above. The TSC function included a Circular-Sigmoid function to model the boundary between chromatic and achromatic categories. Now, we propose to change this circular shape. To obtain the central notch shape needed to

---

[4]Black, Grey and White
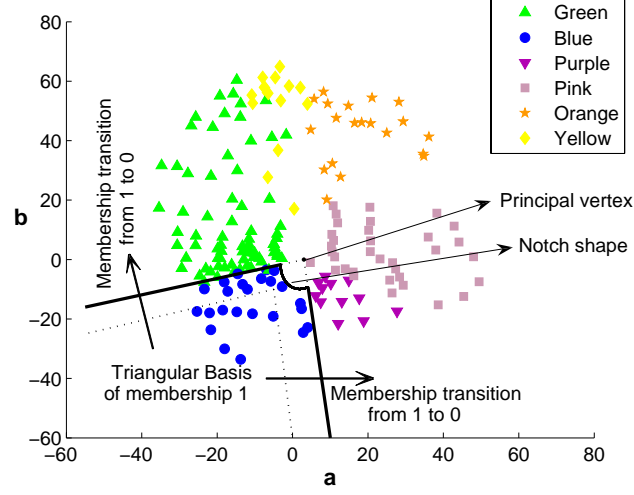
Figure 4.16: Desirable properties of the membership function for chromatic categories. In this case, on the *Blue* category.

fulfil the fourth proposed property, let us define the Elliptical-Sigmoid (ES) function by including the ellipse equation in the sigmoid formula:

$$ES(\vec{p};\vec{t},\theta_{ES}) = \frac{1}{1 + e^{-\beta_e \left(\left(\frac{\vec{u}_1 R_\phi T_t \vec{p}}{e_x} \right)^2 + \left(\frac{\vec{u}_2 R_\phi T_t \vec{p}}{e_y} \right)^2 - 1 \right)}} \qquad (4.29)$$

where $\theta_{ES} = (e_x, e_y, \phi, \beta_e)$ is the set of parameters of the Elliptical-Sigmoid function, e_x and e_y are the semiminor and semimajor axis respectively, ϕ is the rotation angle of the ellipse, and β_e is the slope of the sigmoid curve that forms the ellipse boundary. The function obtained is an elliptical plateau if β_e is negative and an elliptical valley if β_e is positive. The surfaces obtained can be seen in figure 4.17.

Finally, by multiplying the Double-Sigmoid and the Elliptical-Sigmoid with a positive β_e, we define the Triple-Sigmoid with Elliptical centre (TSE) as:

$$TSE(\vec{p};\theta) = DS(\vec{p};\vec{t},\theta_{DS}) \cdot ES(\vec{p};\vec{t},\theta_{ES}) \qquad (4.30)$$

where $\theta = (\vec{t}, \theta_{DS}, \theta_{ES})$ is the set of parameters of the TSE.

The TSE function defines a membership surface that fulfils the properties defined above. Figure 4.18 shows the form of the TSE function.

Hence, once we have the analytic form of the chosen function, the membership function for a chromatic category, μ_{C_k}, is given by:

$$\mu_{C_k}(\vec{s}) = \begin{cases} \mu_{C_k}^1 = TSE(c_1, c_2; \theta_{C_k}^1) & \text{if } I \leq I_1, \\ \mu_{C_k}^2 = TSE(c_1, c_2; \theta_{C_k}^2) & \text{if } I_1 < I \leq I_2, \\ \vdots & \vdots \\ \mu_{C_k}^{N_L} = TSE(c_1, c_2; \theta_{C_k}^{N_L}) & \text{if } I_{N_L-1} < I, \end{cases} \qquad (4.31)$$
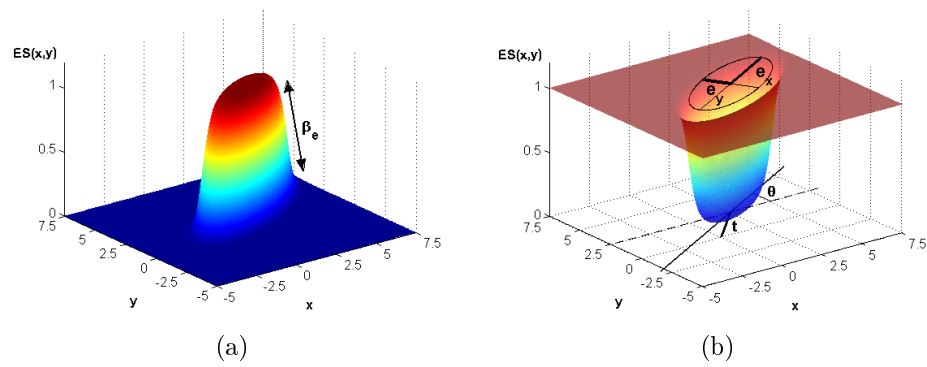
Figure 4.17: Elliptical-Sigmoid function $ES(\vec{p}; \vec{t}, \theta_{ES})$. (a) ES for $\beta_e < 0$ (b) ES for $\beta_e > 0$
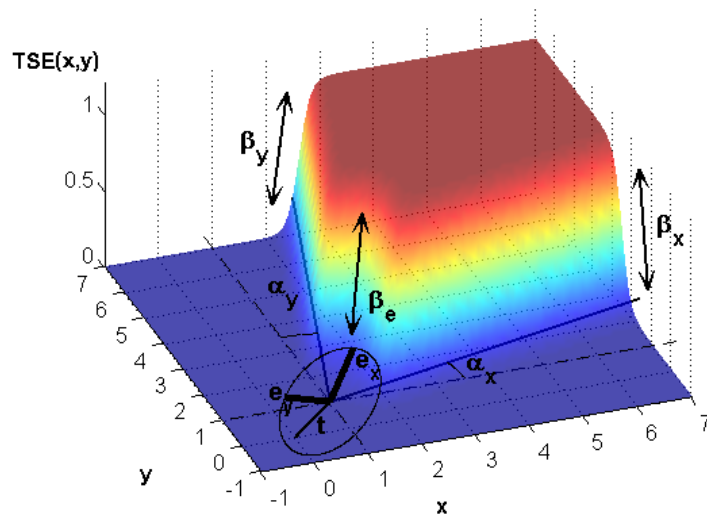
Figure 4.18: Triple-Sigmoid with Elliptical centre (TSE).

where $\vec{s} = (I, c_1, c_2)$ is a sample on the colour space, N_L is the number of chromaticity planes, $\theta^i_{C_k}$ is the set of parameters of the category C_k on the ith chromaticity plane and I_i are the lightness values that divide the space in the N_L lightness levels.

By fitting the parameters of the functions, it is possible to obtain the variation of the chromatic categories through the lightness levels. Doing this for all the categories, it will be possible to obtain membership maps, that is, for a given lightness level we have a membership value to each category for any colour point $\vec{s} = (I, c_1, c_2)$ of the level. Figure 4.19 shows an example of the membership map provided by the TSE functions for a given lightness level.
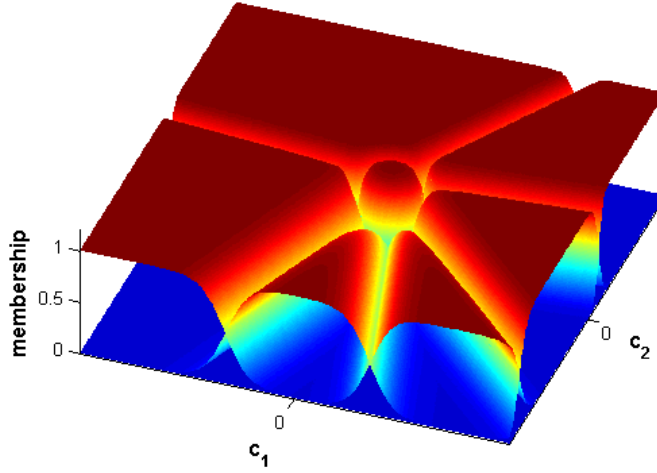
Figure 4.19: TSE function fitted to all the chromatic categories on a given lightness level.

Achromatic Categories

The three achromatic categories (*Black*, *Grey* and *White*) are first considered as a unique category at each chromaticity plane. To ensure that the fuzzy constraint is fulfilled (i.e. the sum of all memberships must be one) the global achromatic membership, μ_A, is computed for each level as:

$$\mu^i_A(c_1, c_2) = 1 - \sum_{k=1}^{n_c} \mu^i_{C_k}(c_1, c_2) \tag{4.32}$$

where i is the chromaticity plane that contains the sample $\vec{s} = (c_1, c_2, I)$ and n_c is the number of chromatic categories (in our case, $n_c = 8$). The differentiation between the three achromatic categories must be done in terms of lightness. To model the fuzzy boundaries between these three categories we use one-dimensional sigmoid functions

along the lightness axis:

$$\mu_{A_{Black}}(I; \theta_{Black}) = \frac{1}{1 + e^{-\beta_b(I - t_b)}} \tag{4.33}$$

$$\mu_{A_{Grey}}(I; \theta_{Grey}) = \frac{1}{1 + e^{\beta_b(I - t_b)}} \cdot \frac{1}{1 + e^{-\beta_w(I - t_w)}} \tag{4.34}$$

$$\mu_{A_{White}}(I; \theta_{White}) = \frac{1}{1 + e^{\beta_w(I - t_w)}} \tag{4.35}$$

where $\theta_{Black} = (t_b, \beta_b)$, $\theta_{Grey} = (t_b, \beta_b, t_w, \beta_w)$, and $\theta_{White} = (t_w, \beta_w)$ are the set of parameters for *Black*, *Grey*, and *White* respectively. Figure 4.20 shows an scheme of this division along the lightness axis.
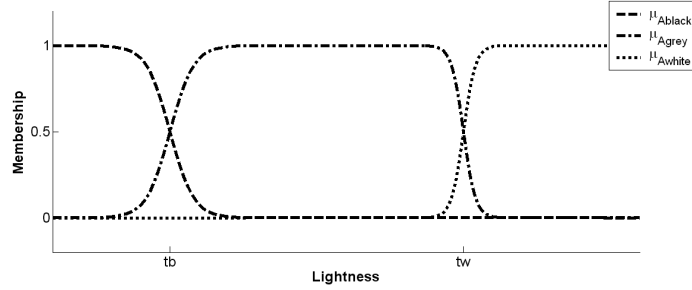
Figure 4.20: Sigmoid functions are used to differentiate between the three achromatic categories.

Hence, the membership of the three achromatic categories on a given chromaticity plane is computed by weighting the global achromatic membership (equation (4.32)) with the corresponding membership in the lightness dimension (equations (4.33)-(4.35)):

$$\mu_{C_k}(\vec{s}; \theta_{C_k}) = \mu_A^i(c_1, c_2) \cdot \mu_{A_{C_k}}(I; \theta_{C_k}), \qquad 9 \le k \le 11, \quad I_i < I \le I_{i+1} \tag{4.36}$$

where i is the chromaticity plane in which the sample is included and the values of k correspond to the achromatic categories (see equation (1.21)). In this way we can assure that the fuzzy constraint is fulfilled on each specific chromaticity plane. For the ith chromaticity plane we have:

$$\sum_{k=1}^{n} \mu_{C_k^i}(\vec{s}) = \sum_{k=1}^{n_c} \mu_{C_k^i}(\vec{s}) + \sum_{k=n_c+1}^{n} \mu_{C_k^i}(\vec{s}) \tag{4.37}$$

where n is the total number of colour categories (in our case, $n = 11$) and n_c is the number of chromatic categories (in our case, $n_c = 8$).

The second summation can be replaced by the membership functions of the three achromatic categories according to equation (4.36):

$$\begin{aligned}
\sum_{k=1}^{n} \mu_{C_k^i}(\vec{s}) = \sum_{k=1}^{8} \mu_{C_k^i}(\vec{s}) + \mu_A^i(c_1, c_2) \cdot \mu_{A_{C_9}}(I; \theta_{C_9}) + \\
+ \mu_A^i(c_1, c_2) \cdot \mu_{A_{C_{10}}}(I; \theta_{C_{10}}) + \mu_A^i(c_1, c_2) \cdot \mu_{A_{C_{11}}}(I; \theta_{C_{11}})
\end{aligned} \tag{4.38}$$

By definition,

$$\mu_A^i(c_1, c_2) = 1 - \sum_{k=1}^{8} \mu_{C_k}^i(c_1, c_2) \qquad \text{(see equation (4.32))}$$

$$\mu_{A_{C_9}}(I; \theta_{C_9}) + \mu_{A_{C_{10}}}(I; \theta_{C_{10}}) + \mu_{A_{C_{11}}}(I; \theta_{C_{11}}) = 1 \quad \text{(see equations (4.33)-(4.35))}$$

Then,

$$\sum_{k=1}^{n} \mu_{C_k^i}(\vec{s}) = \sum_{k=1}^{8} \mu_{C_k^i}(\vec{s}) + \left(1 - \sum_{k=1}^{8} \mu_{C_k}^i(c_1, c_2) \right) \cdot 1 = 1 \tag{4.39}$$

Hence, we assure that for any chromaticity plane we have:

$$\sum_{k=1}^{n} \mu_{C_k^i}(\vec{s}) = 1 \qquad i = 1, \dots, N_L \tag{4.40}$$

where N_L is the number of chromaticity planes on the model.

4.5.2   Fuzzy Sets Estimation

Once we have defined the membership functions of the model, the next step is to estimate their parameters. The parameter estimation will be done taking into account the basic fuzzy property of unity-sum given in equation (4.40). To be able to do this, the data set for the fitting process must be perceptually significant, that is, the judgements should be coherent with results from psychophysical colour-naming experiments and the samples should cover all the colour space. As we saw in the previous sections, the data set from our colour-naming experiment provided the set of fuzzy judgements, but the sampling of the colour space is not large enough to fit the presented model. Hence, we need to build a new learning set.

Learning Set

To build a wider learning set, we have used the fuzzy colour category map proposed by Seaborn et al in [124]. This colour map has been built by making some considerations on the consensus areas of the Munsell colour space provided by the psychophysical data from the experiments of Sturges and Whitfield [135]. Using such data and the fuzzy k-means algorithm, this method allows deriving the memberships of any point in the Munsell space to the eleven basic colour categories.

In this way, we have obtained the memberships of a wide sample set and, afterwards we have converted this colour sampling set to their corresponding CIELab representation. Our data set was initially composed of the 1269 samples of the Munsell Book of Color [6]. Their reflectances and CIELab coordinates, calculated by using the CIE D65 illuminant, are available at the web site of the University of Joensuu in Finland [2].

In order to avoid problems in the fitting process due to the reduced number of achromatic and low chroma samples, the set was completed with 18 achromatic samples (from Value=1 to Value=9.5 at steps of 0.5), 320 low chroma samples (for Values

from 2 to 9, Hue at steps of 2.5 and Chroma=1), and 10 samples with Value=2.5 and Chroma=2 (Hues 5YR, 7.5YR, 10YR, 2.5Y, 5Y, 7.5Y, 10Y, 2.5GY, 5GY, and 7.5GY). The CIELab coordinates of these additional samples were computed with the Munsell Conversion software available at [5].

Therefore, the total number of samples of our learning set is 1617. Hence, with such data set we accomplish the perceptual significance required for our learning set. First, by using Seaborn's method we include the results of the psychophysical experiment of Sturges and Whitfield and, in addition, it covers an area of the colour space that suffices for our purpose.

Parameter Estimation

Before starting with the fitting process, the number of chromaticity planes and the values that define the lightness levels (equation (4.11)) must be set. These values depend on the learning set used and must be chosen taking into account the distribution of the samples from the learning set. In our case, the number of planes that delivered best results was found to be six and the values that define the levels were selected by choosing some local minima in the histogram of samples along the lightness axis. Figure 4.21 shows the samples histogram and the values selected. However, if a more extensive learning set was available a higher number of levels would possibly deliver better results.
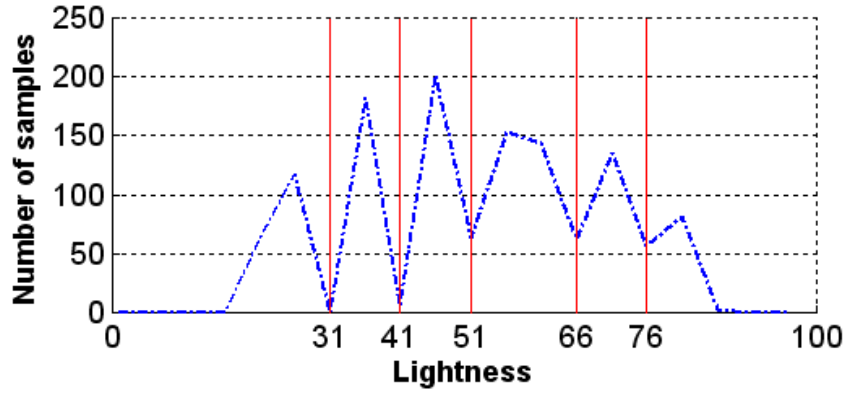
Figure 4.21: Histogram of the learning set samples used to determine the values that define the lightness levels of the model.

For each chromaticity plane, the global goal of the fitting process is finding an estimation of the parameters, $\hat{\theta}^j$, that minimises the mean squared error between the memberships from the learning set and the values provided by the model:

$$\hat{\theta}^j = \arg\min_{\theta^j} \frac{1}{n_{cp}} \sum_{i=1}^{n_{cp}} \sum_{k=1}^{n_c} (\mu_{C_k}^j(\vec{s_i}; \theta_{C_k}^j) - m_k^i)^2, \qquad j = 1, \cdots, N_L \qquad (4.41)$$

where $\hat{\theta}^j = (\hat{\theta}_{C_1}^j, ..., \hat{\theta}_{C_{n_c}}^j)$ is the estimation of the parameters of the model for

the chromatic categories on the jth chromaticity plane, $\theta_{C_k}^j$ is the set of parameters of the category C_k for the jth chromaticity plane, n_c is the number of chromatic categories, n_{cp} is the number of samples of the chromaticity plane, $\mu_{C_k}^j$ is the membership function of the colour category C_k for the jth chromaticity plane, and m_k^i is the membership value of the ith sample of the learning set to the kth category.

The previous minimisation is subject to the unity-sum constraint:

$$\sum_{k=1}^{11} \mu_{C_k}^j(\vec{s}; \theta_{C_k}^j) = 1, \qquad \forall \vec{s} = (I, c_1, c_2) \quad | \quad I_{j-1} < I \le I_j \tag{4.42}$$

which is imposed to the fitting process through two assumptions. The first one is related to the membership transition from chromatic categories to achromatic categories:

Assumption 1: All the chromatic categories in a chromaticity plane share the same Elliptical-Sigmoid function which models the membership transition to the achromatic categories. This means that all the chromatic categories share the set of estimated parameters for ES:

$$\theta_{ES_{C_p}}^j = \theta_{ES_{C_q}}^j \quad and \quad \vec{t}_{C_p}^j = \vec{t}_{C_q}^j, \qquad \forall p, q \in \{1, \ldots, n_c\} \tag{4.43}$$

where n_c is the number of chromatic categories.

The second assumption refers to the membership transition between adjacent chromatic categories:

Assumption 2: Each pair of neighbouring categories, C_p and C_q share the parameters of slope and angle of the Double-Sigmoid function which define their boundary:

$$\beta_y^{C_p} = \beta_x^{C_q} \qquad and \qquad \alpha_y^{C_p} = \alpha_x^{C_q} - \left(\frac{\pi}{2}\right) \tag{4.44}$$

where the superscripts indicate the category to which correspond the parameters.

These assumptions considerably reduce the number of parameters to be estimated. Hence, for each chromaticity plane, we must estimate 2 parameters for the translation, $\vec{t} = (t_x, t_y)$, 4 for the ES function, $\theta_{ES} = (e_x, e_y, \phi, \beta_e)$, and a maximum of $2 \times n_c$ for the DS functions, since the other two parameters of $\theta_{DS} = (\alpha_x, \alpha_y, \beta_x, \beta_y)$ can be obtained from the neighbouring category (equation (4.44)).

Hence, following the two previous assumptions, the parameters of the chromatic categories at each chromaticity plane, $\hat{\theta}_{C_k}^j = (\hat{\vec{t}}^j, \hat{\theta}_{DS_{C_k}}^j, \hat{\theta}_{ES}^j)$, with $k = 1, \ldots, 8$, are estimated in two steps:

1. According assumption 1 we estimate the parameters of a unique ES function, $\hat{\vec{t}}^j$ and $\hat{\theta}_{ES}^j$, for each chromaticity plane by minimising:

$$(\hat{\vec{t}}^j, \hat{\theta}_{ES}^j) = \arg\min_{\vec{t}^j, \theta_{ES}^j} \frac{1}{n_{cp}} \sum_{i=1}^{n_{cp}} (ES(\vec{s_i}; \vec{t}^j, \theta_{ES}^j) - \sum_{k=9}^{11} m_k^i)^2 \tag{4.45}$$

where n_{cp} is the number of samples from the learning set in the *jth* chromaticity plane, and m_k^i is the membership to the *kth* category of the *ith* sample for values of k between 9 and 11, which correspond to the achromatic categories according to equation (1.21).

2. Considering assumption 2 allows estimating the rest of the parameters, $\hat{\theta}^j_{DS_{C_k}}$, of each colour category by minimising the following expression for each pair of neigbouring categories, C_p and C_q:

$$(\hat{\theta}^j_{DS_{C_p}}, \hat{\theta}^j_{DS_{C_q}}) = \arg \min_{\theta^j_{DS_{C_p}}, \theta^j_{DS_{C_q}}} \sum_{i=1}^{n_{cp}} \left((\mu^j_{C_p}(\vec{s_i}; \theta^j_{C_p}) - m^i_p)^2 + (\mu^j_{C_q}(\vec{s_i}; \theta^j_{C_q}) - m^i_q)^2 \right)$$

(4.46)

where $\theta^j_{C_k} = (\hat{\vec{t}^j}, \theta^j_{DS_{C_k}}, \hat{\theta}^j_{ES})$.

Once all the parameters of the chromatic categories have been estimated for all the chromaticity planes, the parameters to differentiate between the three achromatic categories, $\hat{\theta}_A = (\hat{\theta}_{C_9}, \hat{\theta}_{C_{10}}, \hat{\theta}_{C_{11}})$ are estimated by minimising the expression:

$$\hat{\theta}_A = \arg \min_{\theta_A} \sum_{i=1}^{n_s} \sum_{k=9}^{11} (\mu_{C_k}(\vec{s_i}; \theta_{C_k}) - m^i_k)^2$$

(4.47)

where n_s is the number of samples in the learning set and the values of k correspond to the three achromatic categories, that is, $C_9 = Black$, $C_{10} = Grey$, and $C_{11} = White$ (see equation (1.21)).

All the minimisations to estimate the parameters are performed by using the simplex search method proposed in [84]. After the fitting process, we obtain the parameters that completely define our colour-naming model and that are presented and discussed in the next section.

4.5.3 Results

The essential result of this work is the set of parameters of the colour-naming model that are summarized in table 4.15. Subscripts x and y, used in the model formulation, are changed to a and b respectively in order to make parameter interpretation easier, since parameters in this work have been estimated for the CIELab space.

The fitting error of the model to the learning dataset was of $MAE_{fit} = 0.0168$. Table 4.16 shows the distribution of this total error through the different categories.

The MAE_{fit} measure was also computed for a test dataset of 3149 samples. To build the test dataset, the Munsell space was sampled at Hues 1.25, 3.75, 6.25 and 8.75, Values from 2.5 to 9.5 at steps of 1 unit and Chromas from 1 to the maximum available with a step of 2 units. As in the case of the learning set, the memberships of the test set that were considered the ground truth, were computed with Seaborn's algorithm. The corresponding CIELab values to apply our parametric functions were computed with the Munsell Conversion software. The value of MAE_{fit} obtained was 0.0218 which confirms the accuracy of the fitting that allows the model to provide membership values with very low error even for samples that were not used in the fitting process.

To compute the $MAE_{unitsum}$ measure, we have sampled each one of the six chromaticity planes with values from -80 to 80 at steps of 0.5 units on both a and b axis, which means a total number of points of $n_p = 153600$. The value obtained of

Achromatic axis

Black-Grey boundary	$t_b = 28.28$	$\beta_b = -0.71$
Grey-White boundary	$t_w = 79.65$	$\beta_w = -0.31$

Chromaticity plane 1

	$e_a = 5.89$		$\beta_e = 9.84$
$t_a = 0.42$			
$t_b = 0.25$	$e_b = 7.47$		$\phi = 2.32$
	α_a	α_b	β_a	β_b

	α_a	α_b	β_a	β_b
Red	-2.24	-56.55	0.90	1.72
Brown	33.45	14.56	1.72	0.84
Green	104.56	134.59	0.84	1.95
Blue	224.59	-147.15	1.95	1.01
Purple	-57.15	-92.24	1.01	0.90

Chromaticity plane 2

$t_a = 0.23$	$e_a = 6.46$		$\beta_e = 6.03$
$t_b = 0.66$	$e_b = 7.87$		$\phi = 17.59$

	α_a	α_b	β_a	β_b
Red	2.21	-48.81	0.52	5.00
Brown	41.19	6.87	5.00	0.69
Green	96.87	120.46	0.69	0.96
Blue	210.46	-148.48	0.96	0.92
Purple	-58.48	-105.72	0.92	1.10
Pink	-15.72	-87.79	1.10	0.52

Chromaticity plane 3

$t_a = -0.12$	$e_a = 5.38$		$\beta_e = 6.81$
$t_b = 0.52$	$e_b = 6.98$		$\phi = 19.58$

	α_a	α_b	β_a	β_b
Red	13.57	-45.55	1.00	0.57
Orange	44.45	-28.76	0.57	0.52
Brown	61.24	6.65	0.52	0.84
Green	96.65	109.38	0.84	0.60
Blue	199.38	-148.24	0.60	0.80
Purple	-58.24	-112.63	0.80	0.62
Pink	-22.63	-76.43	0.62	1.00

Chromaticity plane 4

$t_a = -0.47$	$e_a = 5.99$		$\beta_e = 7.76$
$t_b = 1.02$	$e_b = 7.51$		$\phi = 23.92$

	α_a	α_b	β_a	β_b
Red	26.70	-56.88	0.91	0.76
Orange	33.12	-9.90	0.76	0.48
Yellow	80.10	5.63	0.48	0.73
Green	95.63	108.14	0.73	0.64
Blue	198.14	-148.59	0.64	0.76
Purple	-58.59	-123.68	0.76	5.00
Pink	-33.68	-63.30	5.00	0.91

Chromaticity plane 5

$t_a = -0.57$	$e_a = 5.37$		$\beta_e = 100.00$
$t_b = 1.16$	$e_b = 6.90$		$\phi = 24.75$

	α_a	α_b	β_a	β_b
Orange	25.75	-15.85	2.00	0.84
Yellow	74.15	12.27	0.84	0.86
Green	102.27	98.57	0.86	0.74
Blue	188.57	-150.83	0.74	0.47
Purple	-60.83	-122.55	0.47	1.74
Pink	-32.55	-64.25	1.74	2.00

Chromaticity plane 6

$t_a = -1.26$	$e_a = 6.04$		$\beta_e = 100.00$
$t_b = 1.81$	$e_b = 7.39$		$\phi = -1.19$

	α_a	α_b	β_a	β_b
Orange	25.74	-17.56	1.03	0.79
Yellow	72.44	16.24	0.79	0.96
Green	106.24	100.05	0.96	0.90
Blue	190.05	-149.43	0.90	0.60
Purple	-59.43	-122.37	0.60	1.93
Pink	-32.37	-64.26	1.93	1.03

Table 4.15: Parameters of the Triple-Sigmoid with Elliptical centre model (angles are expressed in degrees to make the results interpretation easier).

	Red	Orange	Brown	Yellow	Green	Blue
MAE_{fit}	1.40	1.58	1.02	1.48	2.28	2.62
	Purple	Pink	Black	Grey	White	**Global**
MAE_{fit}	2.21	2.22	0.43	2.00	1.23	**1.68**

Table 4.16: Results in terms of the MAE_{fit} $(\times 10^{-2})$ of the Triple-Sigmoid with Elliptical centre functions for each one of the 11 colour categories considered. In the last column, the global measure of error of the model.

$MAE_{unitsum} = 6.41e - 04$ indicates that the model provides a great fulfilment of that constraint making the model consistent with the proposed fuzzy framework. Table 4.17 shows these values compared to the obtained by the previous proposed models.

Model	MAE_fit	MAE_unitsum
MGM	0.0393	0.8206
SGM	0.0281	0.3612
DSM	0.0304	0.1219
TSCM	0.0298	0.1135
TSEM	**0.0168**	**6.41e-04**

Table 4.17: Results obtained by the five models proposed up to now (Multivariate Gaussian (MGM), Sigmoid-Gaussian (SGM), Double-Sigmoid (DSM), Triple-Sigmoid with Circular centre (TSCM) and Triple-Sigmoid with Elliptical centre (TSEM)).

Hence, for any point of the CIELab space we can compute the membership to all the categories and, at each chromaticity plane, these values can be plotted to generate a membership map. In Figures 4.22 and 4.23 we show the membership maps of the six chromaticity planes considered. In figure 4.22 each map is plotted in the centre of its lightness level in the CIELab space. Figure 4.23 shows each map separately with the membership surfaces labelled with their corresponding colour term.

As we did before, the model has also been evaluated in terms of the categorization of the Munsell space. To evaluate the plausibility of the model with psychophysical data, we compare our categorization to the results reported in two works of reference: the study of Berlin and Kay [27] and the experiments of Sturges and Whitfield [135]. Figure 4.24 shows the boundaries found by Berlin and Kay in their work, superimposed on our categorization. Samples inside these boundaries assigned with a different name by our model are marked with a cross. As can be seen, there are a total of 17 samples out of 210 inside Berlin and Kay boundaries with a different name. The errors are concentrated on certain boundaries, namely, *Green-Blue*, *Blue-Purple*, *Purple-Pink* and *Purple-Red*.

The comparison to Sturges and Whitfield results is presented in figure 4.25. In Sturges and Whitfield experiment the samples labelled with the same name by all the subjects defined the consensus areas for each category. Amongst these samples, the fastest named sample for each category was its focus. These areas are superimposed over our categorization to show that all the consensus and focal samples from Sturges and Whitfield experiment are assigned the same name by our model.

The analysis done to our Triple-Sigmoid-Elliptical model (TSEM) was also performed on some previous categorizations. These are obtained by Lammens Gaussian model (LGM) [85], an English speaker presented by MacLaury (MES) in [96], Seaborn's Fuzzy Colour Category Map (FCCM) [124], and our previous best approach, the Triple-Sigmoid with Circular centre model (TSCM). The results are summarized in table 4.18 where it can be seen that the results of our TSE model equal the previous best of Seaborn's non-parametric model (FCCM), but adding the advantages of having a parametric model that have been previously discussed in section 4.1. Notice that although the learning process of both models was based on data derived from
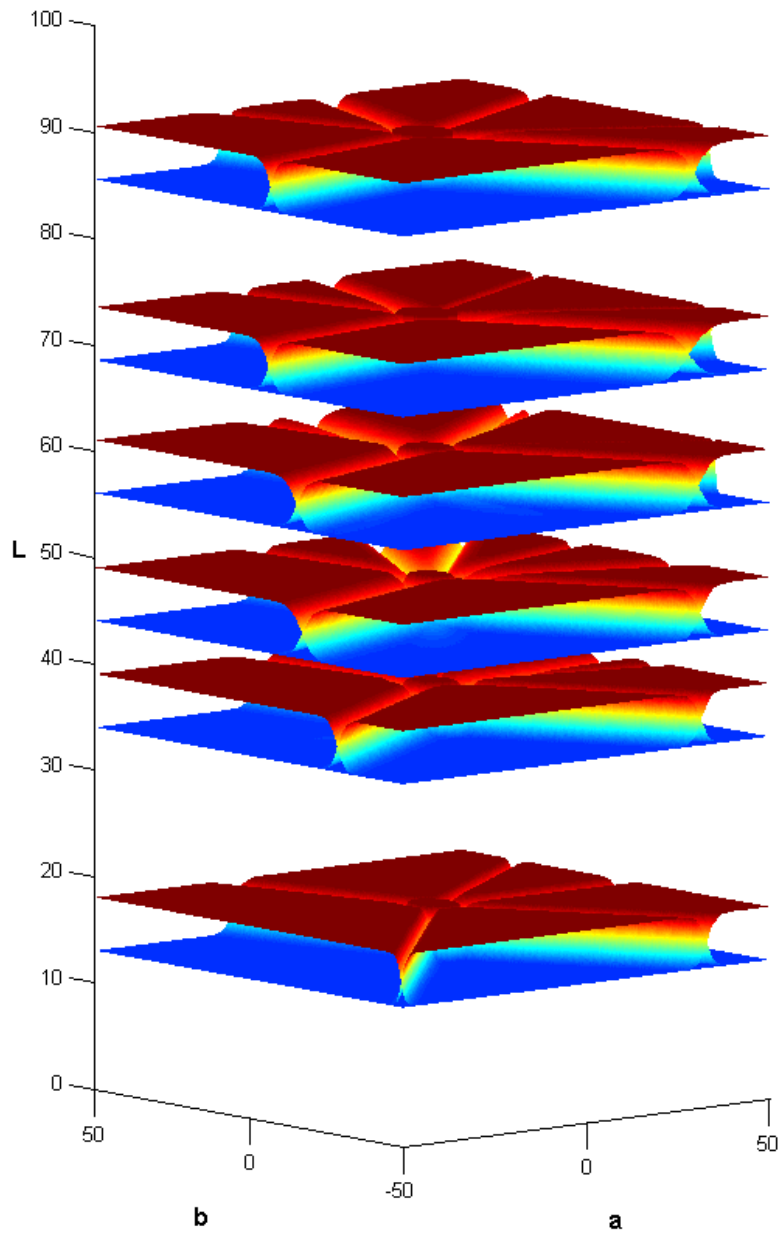
Figure 4.22: Membership maps for the six chromaticity planes of the TSE model. The planes are plot in the centre of each lightness level.

CHROMATICITY PLANE 1

CHROMATICITY PLANE 2

(a)

(b)

CHROMATICITY PLANE 3

CHROMATICITY PLANE 4

(c)

(d)

CHROMATICITY PLANE 5

CHROMATICITY PLANE 6

(e)

(f)
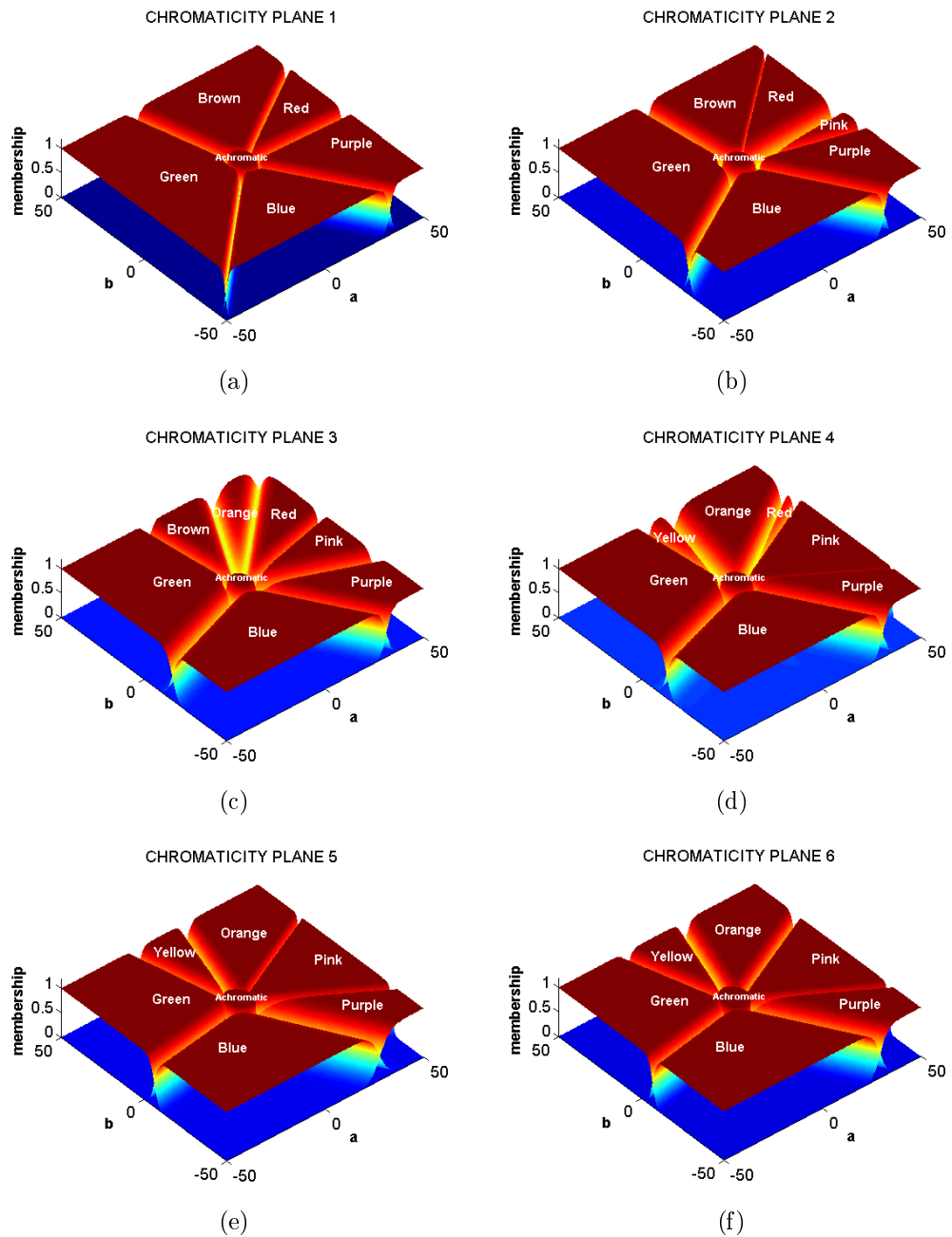
Figure 4.23: Membership maps for the six chromaticity planes of the TSE model plotted separately.
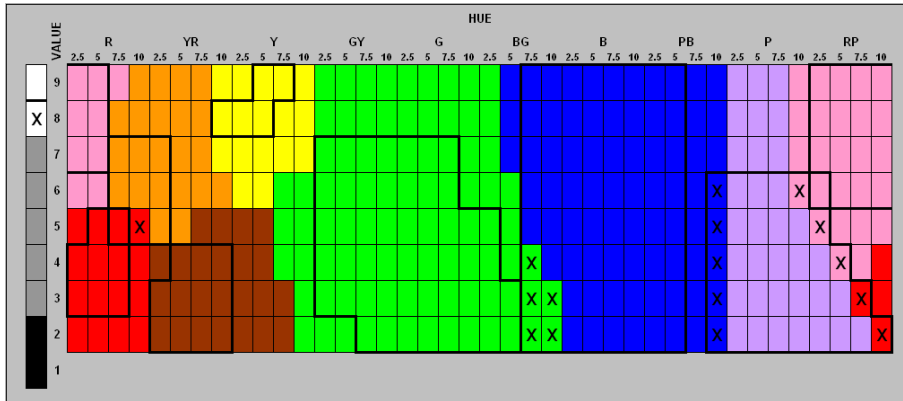
Figure 4.24: Comparison between the TSE model's Munsell categorization and Berlin and Kay boundaries. Samples named differently by our model are marked with a cross.
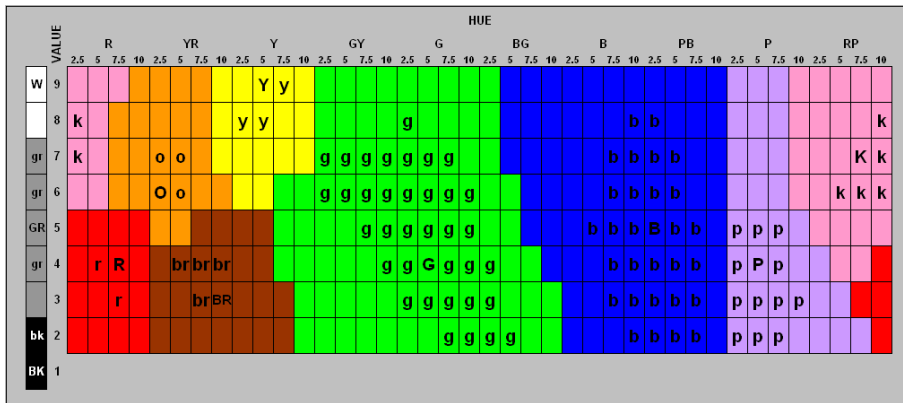
Figure 4.25: Consensus areas and focus from Sturges and Whitfield experiment superimposed on the TSE model's categorization of the Munsell array. Each letter corresponds to a category: r-*Red*, o-*Orange*, br-*Brown*, y-*Yellow*, g-*Green*, b-*Blue*, p-*Purple*, k-*Pink*, bk-*Black*, gr-*Grey*, w-*White*. Samples marked with capital letters are the focus.

Sturges results, they are the most consistent with Berlin and Kay's experiments and are also better than the results of the English speaker's categorization.

Model	Berlin and Kay data			Sturges and Whitfield data		
	Coincidences	Errors	% Errors	Coincidences	Errors	% Errors
LGM	161	49	23.33	92	19	17.12
MES	182	28	13.33	107	4	3.60
TSCM	185	25	11.90	108	3	2.70
FCCM	193	17	8.10	111	0	0.00
TSEM	**193**	**17**	**8.10**	**111**	**0**	**0.00**

Table 4.18: Comparison of different Munsell categorizations to the results from colour-naming experiments of Berlin and Kay [27], and Sturges and Whitfield [135].

4.6 Discussion
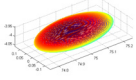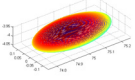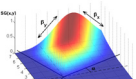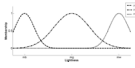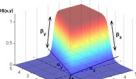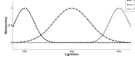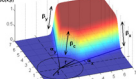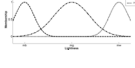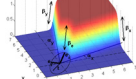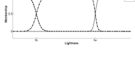
In this chapter we have proposed a parametric fuzzy model for colour naming based on different membership functions. The use of a parametric model introduces several advantages with respect to the previous non-parametric approaches. These advantages that have been discussed in section 4.1, include a reduction of the implementation costs in terms of memory and computation time, a compact data representation, simplicity for model analysis since each parameter has a meaning in terms of the characterization of the data and, consequently, the model can be easily updated by just tuning some of the parameters.

The model has been conceived for any colour space with two chromatic dimensions and a lightness dimension, but in the present work the parameters have been estimated for the CIELab space. After presenting a preliminary Gaussian model we have proposed four different membership functions based on the product of sigmoid functions. From the initial Sigmoid-Gaussian function, the model has been evolved until defining the final model based on the Triple-Sigmoid with Elliptical centre (TSE) function. Each proposed function has been analysed and in all the cases the next proposal has been defined to solve the drawbacks detected in the preceding function. Hence, the results obtained have progressively been improved until defining the final model based on the TSE which obtains very good results in terms of the measures defined.

Another aspect that must be evaluated when defining parametric models is the number of parameters that define the model. Usually, a function defining a complex form needs more parameters than one defining a simpler shape to be analytically defined. However, this is not always the case. Although the TSE function has 10 parameters, the constraints imposed on the estimation process allows to define each chromatic category by 2 parameters plus 6 parameters more that are the same for all the categories of a chromaticity plane. Furthermore, the simplification on the model for achromatic categories has implied a reduction of the total number of parameters for the global model.

Althouth the constraints imposed in the estimation process of the TSE parameters also imply a reduction of the final number of parameters of the model, they were defined to assure the fulfilment of the fuzzy framework where the memberships sum for any point must be one. The result of such estimation process is the set of parameters that define a model which achieves low fitting error for both the learning and test datasets and also fulfils the unity-sum constraint. The evaluation of the model when compared to previous results from the colour-naming experiments of Berlin and Kay, and Sturges and Whitfield, demonstrates that the TSE model is plausible with these psychophysical data. Table 4.19 summarizes the functions that have been proposed in this chapter and the results obtained with each one.

As a final result of the model, the memberships to the 11 basic colour categories can be obtained for any point in the CIELab space to provide a colour-naming descriptor with meaningful information about how humans name colours. The model has many applications on different computer vision tasks, such as image annotation, image indexing and segmentation amongst others, where the inclusion of this high level information might improve their performance. The proposed representation of colour information could also be used as a more perceptual measure of similarity in colour spaces, instead of the frequently used Euclidean distance.

Function for Chromatics and Model Name	Functions for Achromatics	Parameters of chrom. function	Parameters (Total)	MAE_fit	MAE_unitsum
Multivariate Gaussian (MG)	3D-Gaussians	9	99	0.0393	0.8206
Sigmoid-Gaussian (SG)	1D-Gaussians	7	168	0.0281	0.3612
Double-Sigmoid (DS)	1D-Gaussians	6	147	0.0304	0.1219
Triple-Sigmoid with Circular centre (TSC)	1D-Gaussians	8	189	0.0298	0.1135
Triple-Sigmoid with Elliptical centre (TSE)	1D-Sigmoids	10	114	0.0168	6.41e-04

Table 4.19: Summary of the models defined in this chapter and results obtained.

Chapter 5

Results and Discussion

Once the colour-naming model has been defined, the next step is to apply it to real computer vision problems. As we have said previously in this thesis, the proposed model has been designed on ideal conditions. Unfortunately, this will not usually be the case in real applications. Hence, some previous considerations have to be made. In this chapter we first briefly review the applications that the use of colour names can have on computer vision. Next we detail the steps that must be performed before applying the colour-naming model to real problems and analyse the viability of the model for real problems. Finally, we propose some colour-naming descriptors for images and the TSE model is tested on a real application of automatic annotation for image retrieval.

5.1 Colour Naming in Computer Vision Applications

Although colour names are a high-level information that can be very useful for some computer applications, they have not been widely used up to now. The essential goal of colour naming in computer vision is to reduce the semantic gap between the low-level features that can be extracted from an image and the high-level semantics of humans. Colour naming has been used in some image retrieval works. Mojsilović et al. [104] have used colour names jointly with other high-level features on a retrieval system that allows making queries on a semantic-friendly language. The Fuzzy Colour Category Map of Seaborn et al. [124] has been used as similarity measure for CBIR on the Pisaro system [123]. Gagaudakis and Rosin [57] also combined several features, amongst them colour labels, for CBIR.

There are other simpler approaches based on defining a division of different colour spaces such as HSV [93, 94], HSI [146, 147] or the Munsell space [13]. Another proposed approach is based on the definition of a set of centroids and pixel clustering by the nearest-neighbour rule. In these algorithms centroids have been defined on the ISCC-NBS colour system [109] or the RGB space [119].

Although colour naming has normally been applied for general CBIR, colour names have also been used on reduced domain applications such as indexing of flower images [40] or paintings retrieval [131].

99

Finally, in [18, 19, 17, 16, 141, 142] colour naming was applied to clothes description on a surveillance application. In [76] colour-naming information is used to detect targets on a tracking application and in [102] colour naming is applied to image segmentation.

5.2   Colour Naming in Practice

The use of the proposed colour-naming model for real computer vision problems requires to consider some issues that were not considered in the definition of the model. Firstly, the model has been fitted on a standard colour space such as CIELab while images are normally represented on a device-dependent RGB space. Secondly, the learning data was obtained in calibrated experiments under controlled conditions, that is, without neither illumination changes nor context influences.

However, the working conditions in computer vision applications are normally quite different. Firstly, the illumination rarely will be constant. Hence, the same surface will have different RGB values on different images depending on the illuminant under which the image is acquired. Secondly, in most images, the interactions between the objects in the scene will make the colour induction mechanisms of the human visual system to act. Thus, the same surface can be differently perceived on different images depending on the image content. Finally, many applications must deal with images from multiple sources which represent colour in its own device-dependent RGB space. In most cases, a gamma correction to account for non-linearity of display devices will also be applied to images. Hence, a transform from the device-dependent colour spaces to the standard CIELab where the model is defined will be needed.

Hence, before applying the proposed colour-naming method, the input image should be preprocessed. First, a calibration step will be performed to represent the image on the sRGB space (standard RGB [7]). Next, processes to account for colour constancy and colour induction issues will be applied. At this point, it will be possible to transform the image to the CIELab space and apply the colour-naming method which will provide the colour-naming information needed for a specific computer vision application. Figure 5.1 shows a module diagram of all these steps.

5.2.1   Colour Space Transforms

Since the colour-naming model has been defined on the CIELab space and images will be normally represented on a device-dependent RGB space, a transform between both spaces must be performed to be able to apply the model. As it was presented in section 1.2, CIELab is defined as a non-linear transform from XYZ values, and XYZ is a linear transform from CIE RGB. As acquisition devices have different response functions, the transform from devices' RGB to XYZ will depend on each device. Furthermore, a gamma transform to take into account the non-linearities of display devices is normally applied to images.

Ideally, the image should be represented on a standard RGB space [137], e.g. the default sRGB colour space [7], for which the transform to XYZ is already described. In the case of the sRGB standard this transform is defined in two steps, as follows:
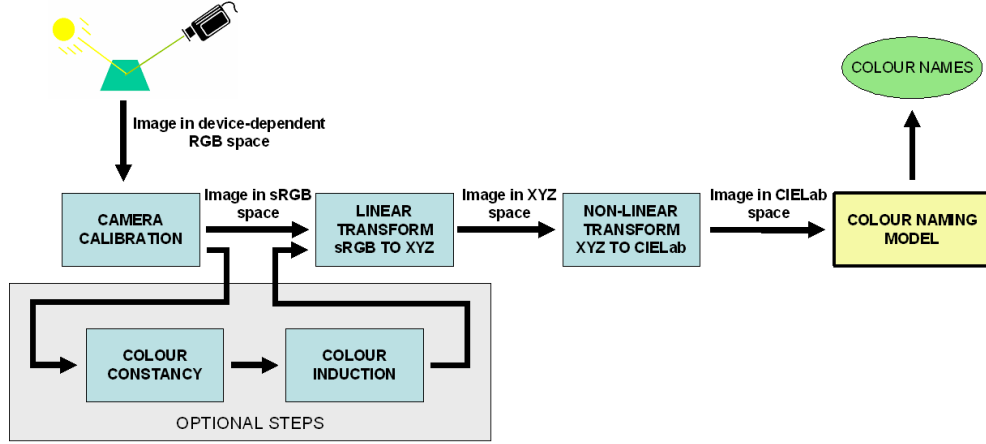
Figure 5.1: Module diagram of the process steps to apply the colour-naming model to real problems.

1. Obtaining linear RGB values by removing the gamma transform:

   If R,G,B≤0.04045

$$
\begin{aligned}
R' &= R/12.92 \\
G' &= G/12.92 \\
B' &= B/12.92
\end{aligned}
\tag{5.1}
$$

   If R,G,B>0.04045

$$
\begin{aligned}
R' &= \left(\frac{R + 0.055}{1.055} \right)^{2.4} \\
G' &= \left(\frac{G + 0.055}{1.055} \right)^{2.4} \\
B' &= \left(\frac{B + 0.055}{1.055} \right)^{2.4}
\end{aligned}
\tag{5.2}
$$

   where R, G, and B are the values of the image normalized to $[0, 1]$.

2. A linear transform to the XYZ space multiplying the linear (R', G', B') triplet by a 3×3 matrix:

$$
\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} 0.4124 & 0.3576 & 0.1805 \\ 0.2126 & 0.7152 & 0.0722 \\ 0.0193 & 0.1192 & 0.9505 \end{pmatrix} \begin{pmatrix} R' \\ G' \\ B' \end{pmatrix}
\tag{5.3}
$$

The non-linear transform from XYZ to CIELab is done by applying the standard equations [157] that were explained in section 1.2.2.

Unfortunately, sRGB is not widely used yet and in most cases images will not be represented in this standard space. Similarly as in [53] we propose to perform a pre-processing step to transform the non-sRGB input image to a new one representing the same image as having been taken by an sRGB camera in order to apply the standard transform to XYZ and then to CIELab space. To do this, several situations are possible:

Non-sRGB sensor responses are known

If the input images are not sRGB but the sensitivity curves of the sensor used are known it is possible to compute the camera responses for a set of known colour reflectances (e.g. a Macbeth Colour Checker). Then it is possible to compute a 3×3 matrix, M, that transforms RGB values of the non-sRGB sensor to sRGB values:

$$\vec{p_i'}^T = M \cdot \vec{p_i}^T \qquad (5.4)$$

where $\vec{p_i}$ is a pixel from the non-sRGB image and $\vec{p_i'}$ are the corresponding sRGB values.

Least-squares regression is the simplest way to compute such a matrix. However, other methods have been proposed [99, 72].

Another possibility is to directly map device-dependent RGB values to XYZ tristimulus values. This transform can be computed in the same way as in the previous case. Several methods have been proposed [153, 47, 48].

Unknown sensor but control on acquisition is possible

When the response curves of the camera used to acquire the images are unknown it will not be possible to use sensor information to compute the device-dependent RGB values of the calibration set. However, if it is possible to acquire images of the calibration set then it will be possible to apply the same methods described in the previous case.

It will not always be possible to acquire images of the reflectances used as calibration set. However, if at least a white reference is available it will be possible to apply a white-balance transform [150] as:

$$\vec{p_i'}^T = D \cdot \vec{p_i}^T \qquad (5.5)$$

where D is a diagonal matrix with R_{sRGB}/R_{dev}, G_{sRGB}/G_{dev}, and B_{sRGB}/B_{dev} in the diagonal. $(R_{sRGB}, G_{sRGB}, B_{sRGB})$ are the sRGB values of the white reference and $(R_{dev}, G_{dev}, B_{dev})$ are the device-dependent values of the white reference acquired with the non-sRGB sensor.

In these conditions, another possibility is to fit the model's parameters on the CIELab space derived from the device-dependent RGB space. Obviously the error of the transform will propagate to the CIELab values used on the learning step, but as the CIELab values of samples to be named will also include the same error the problem will not be critical. However, this solution implies to acquire the set of samples used as learning-set on the parameter estimation step.

Unknown sensor and no control on acquisition

If the acquisition conditions are unknown, i.e. uncalibrated conditions, and it is not possible to acquire a set of calibration samples, some assumptions must be done.

One possibility is to apply a colour constancy method that makes assumptions on the content of the image. Hence, Grey-world method [35] assumes the mean of the image is grey, White-Patch Retinex [34] assumes the maximum values on the image correspond to a white, gamut mapping methods [46, 52] assume the image has a representative set of colours and the method based on nameability of Tous [140] assumes that most colours in an image have a definite name.

Applying one of these approaches can reduce the error of assuming the sensor is sRGB and applying the standard transform from sRGB to XYZ to the non-sRGB values of the image.

In section 5.2.4 we consider the different possible working conditions and evaluate the performance that can be expected on some of them. Different conditions are evaluated for three different commercial cameras and the maximum performance of the model for each camera and condition is evaluated.

5.2.2 Colour Constancy

As we saw in section 1.4.1 the same surface acquired under different illuminants can be represented by very different RGB values. The visual system has mechanisms that allow humans to disregard the effects of the illuminant and thus perceive colours avoiding the illuminant influence. Hence, if we aim to obtain the same results as the human visual system when the colour-naming model is applied on a real application, the problem of colour constancy must be taken into account.

Previous works on computational models for colour naming have considered this problem before applying the model on real images. Lammens [85] applied a modified Von Kries rule [157] to solve the colour constancy problem. In this approach, the value of a pixel, p_i is modified as:

$$\vec{p_i'} = \frac{\vec{p_i} - \vec{b}}{\vec{w} - \vec{b}} \tag{5.6}$$

where $\vec{p_i'}$ is the corrected pixel, $\vec{b}$ and $\vec{w}$ are the black and white representatives of the image.

This method can be seen as a scaling and a reorientation of the grey axis of the image to align it to the theoretical grey axis, i.e. the axis from (0,0,0) to (1,1,1). Mojsilović [103] used the same idea but with some modifications. In this case, the image was previously median filtered to remove noise and the selection of black and white representatives was done by choosing the nearest pixels to black and white prototypes of her model, where distances were computed with the colour-namimg metric defined.

In both cases the results are quite good although comparatives of colour constancy algorithms [56, 71] have shown that the best results are obtained by gamut mapping approaches [46, 52] and the colour by correlation method [50].

In our case, to consider the colour constancy problem in the model tests we will apply the white-patch Retinex method proposed in [56] that was based in the Retinex version of [34]. This method, that has been chosen by its simplicity, assumes that the image has a white surface which corresponds to the maximum response on each of the channels of the image. This implementation of the Retinex algorithm finds the maximum response in each of the three channels of the image, R_{max}, G_{max}, and B_{max}, and applies a diagonal transform to each pixel, $\vec{p_i}$ in the form:

$$\vec{p_i}'^T = \begin{pmatrix} \frac{R_{white}}{R_{max}} & 0 & 0 \\ 0 & \frac{G_{white}}{G_{max}} & 0 \\ 0 & 0 & \frac{B_{white}}{B_{max}} \end{pmatrix} \cdot \vec{p_i}^T \qquad (5.7)$$

where $\vec{p_i}'$ is the pixel under a reference illuminant, and R_{white}, G_{white}, and B_{white} are the RGB values of a white surface under the reference illuminant. In our case, a perfect white with values $(R_{white}, G_{white}, B_{white}) = (1, 1, 1)$ is assumed.

Figure 5.2 shows an example with synthetic images of a Macbeth Colour Checker where the application of the colour constancy method improves the naming results. In the figure, images 5.2(a) and 5.2(b) show the Macbeth Checker acquired under a daylight and a blue illuminant respectively. The result of applying the white-patch Retinex algorithm to image 5.2(b) is shown in figure 5.2(c). As can be seen, the blue cast due to the illuminant has been removed to obtain an image very similar to the image under daylight (image 5.2(a)). Images 5.2(d)-5.2(f) show the labels assigned by the colour-naming method. As it can be observed, colour patches in the image under the blue illuminant are very badly named, but the problem is solved if the method is applied to the image obtained after applying the colour constancy method (image 5.2(c)).

5.2.3   Colour Induction

In section 1.4.2 we explained the induction effects that can appear in an image. As we explained there, the same RGB value on an image can be perceived as being of a different colour depending on the image content and the relations between the colours in the image. As happened with the colour constancy problem, if we want to assign colour names in a perceptual way, colour induction effects in images must be considered previously to apply the colour-naming model.

In most previous works on computational colour naming, colour induction effects have not been considered. However, one approach to this problem was included in Mojsilović model [103], where the assimilation effect of human perception is modelled as an adaptive low-pass filtering operation. The process starts with the LBG quantization [92] of the image followed by a computation of the local contrast of each pixel which is used to detect edges in the image. The density of edges in the neighborhood of each pixel is used to classify pixels in six classes: uniform, noise, contour, texture edge, coarse texture, and fine texture. Finally, an averaging operation by means of a convolution with a Gaussian kernel is done. The amount of averaging depends on the type of region, being higher for textured regions and lower for uniform regions. Contour and transition pixels are not smoothed.
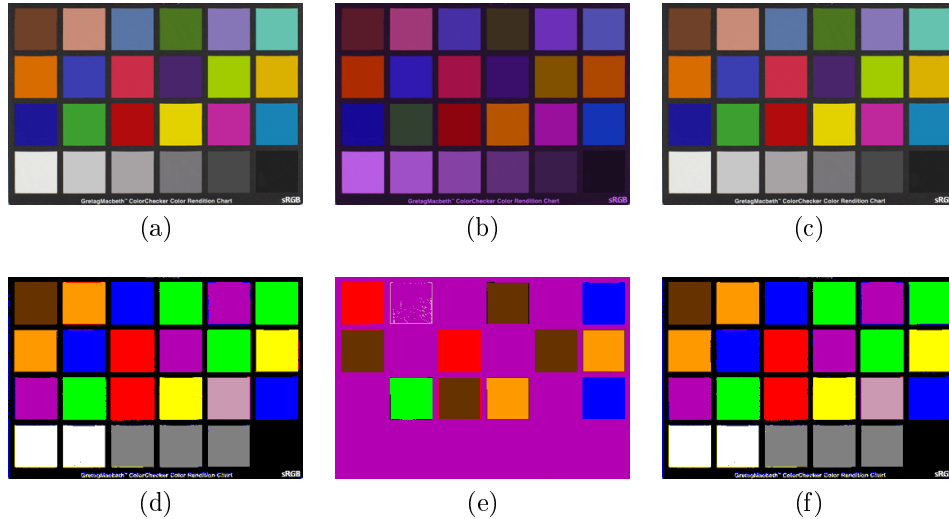
Figure 5.2: Example of the need to apply a colour constancy method before applying the colour-naming model. (a) Image under daylight. (b) Image under a blue illuminant. (c) Image acquired under a blue illuminant after applying the white-patch Retinex colour constancy method. (d)-(f) Colour names assigned by the colour-naming model to images (a)-(c) respectively.

On the other hand, Seaborn's Fuzzy Colour Category Map (FCCM) is also defined for ideal conditions on the Munsell space and does not consider neither colour constancy nor colour induction effects.

In our case, to consider the induction effects we will apply the method of Otazu and Vanrell [112, 114, 113]. This method is based on a multiresolution decomposition of the original image to build a perceived image, that is, an image where colour representation has been modified accordingly with the colour changes that are produced in the human visual system. In this model, colour changes are caused by the following image properties:

- the relationship between an stimulus spatial frequency and the frequency of its surround

- the relationship between the orientation of the stimulus and that of its surround

- the chromatic contrast variation of the stimulus surround

Depending on these properties, the colour of an stimulus can be sharpened with respect to its surround (i.e. a colour contrast effect), or can be blurred with respect to its surround (i.e. a colour assimilation effect).

To reproduce all these effects on an image, it is decomposed with a multiresolution wavelet model and afterwards it is recovered with a modulation introduced by a perceptual function that causes all these effects. Hence given an original image, I, we can build its perceived image, $I_{perceived}$, by the following formulation:

$$I_{perceived} = \sum_{i=1}^{n} CSF(i) \cdot (w_i^h + w_i^v + w_i^d) + c_r \qquad (5.8)$$

where, n is the number of frequency planes in which each image channel is decomposed, w_i^k represents the coefficient decomposition at the plane of frequency i and orientation k (h:horizontal, v:vertical, d:diagonal), and c_r is the residue term. The Contrast Sensitivity Function, $CSF(i)$, causes the induction effects on the decomposition, by combining the information on each pixel and its surround on this frequency plane.

Figures 5.3 and 5.4 show two examples on synthetic images of the usefulness of applying the colour induction model previously to the colour-naming method. In figure 5.3 we show an example of an image with colour contrast effects. Without applying the colour induction model (figures 5.3(a) and 5.3(c)), the same RGB values from the rectangular regions in the sides of the images are assigned the same colour name by the colour-naming model, although they are perceived as being of different colour. The application of the colour induction model provides images in which the RGB value of each pixel better resembles the perceived colour by a human observer. After the application of the induction model (image 5.3(b)), the colour-naming model is able to assign the same colour names as a human observer would do (image 5.3(d)).

(a)   (b)

PURPLE   PURPLE   BLUE   PURPLE
(c)   (d)
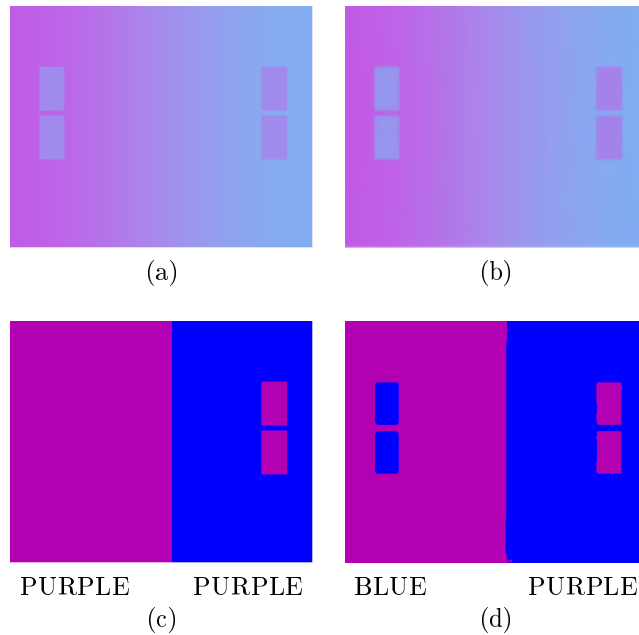
Figure 5.3: Example of the need to take colour contrast into account. (a) Original image where contrast effects are present. (b) Image after applying the colour induction model. (c) and (d) Results of the colour-naming model applied to (a) and (b) respectively. Colour labels under images (c) and (d) are the names assigned to the rectangular areas on the sides of the images.

In figure 5.4 we show an example of an image with colour assimilation effects. As happened in the previous example, the small rectangular regions in image 5.4(a) have the same RGB values. The colour naming model assigns the label 'Blue' to all these regions (image 5.4(c)) although the regions on the right are perceived as green. After applying the colour induction model (image 5.4(b)), the colour-naming model assigns the labels that correspond to the colours actually perceived (image 5.4(d)).

(a)   (b)

BLUE   BLUE   BLUE   GREEN

(c)   (d)
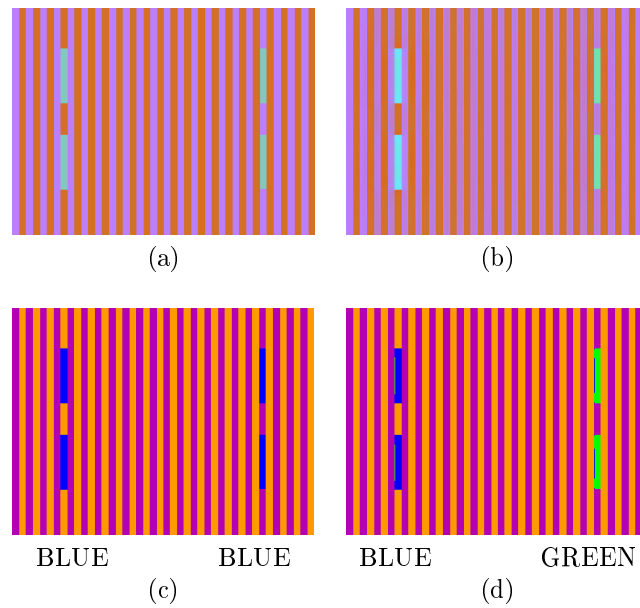
Figure 5.4: Example of the need to take colour assimilation into account. (a) Original image where assimilation effects are present. (b) Image after applying the colour induction model. (c) and (d) Results of the colour-naming model applied to (a) and (b) respectively. Colour labels under images (c) and (d) are the names assigned to the rectangular areas on the sides of the images.

5.2.4 Analysis of the Colour-Naming Model Under Uncalibrated Conditions

As we explained in the previous sections, there are three issues, namely space transforms, illumination variability and induction effects, that must be considered previously to apply the colour-naming model. The last two can be overcome, at least to a certain extent, by the existing methods cited for each case. However, transforms from device-dependent RGB spaces to CIELab will depend on the knowledge we have about the sensor used to acquire the images.

In this section we analyse the performance we can expect from our colour-naming model when it is applied to computer vision applications under some of the different working conditions explained in section 5.2.1. To this purpose we use three commercial cameras with known response curves to make the simulations: Sony DXC930 [12],

Kodak DCS420[4], and Nikon D70 [1]. In order to isolate the effect of the space transform from the effect of illuminant changes on the colour-naming performance, we have white balanced the cameras used in the analysis, that is, cameras have its white response normalised to $(1, 1, 1)$.

The three working conditions considered are:

Condition 1: Fitting the model on CIELab derived from camera RGB's
We assume we are able to obtain (synthetically or acquiring images of the samples) the RGB values for the camera of the samples used in the learning step of the colour-naming model. The CIELab values are derived from these RGB's. Note that the model fitted in this way will only be useful for that camera and that all the CIELab values will have a certain error.

Condition 2: Calibration step with a Macbeth Colour Checker
We assume we are able to obtain the RGB values for the camera of the Macbeth Colour Checker and we are able to compute the calibration matrix to transform RGB values of the camera to sRGB values. Hence, an approximation of the real CIELab values of the image can be obtained after that calibration step by applying standard equations to transform from sRGB to CIELab.

Condition 3: Fully uncalibrated conditions
In this case it will not be possible to perform any calibration step and RGB values of the camera are assumed to be sRGB and transformed to CIELab by applying standard equations without any calibration step, with the error this assumption involves.

To test the performance of the colour-naming model on the different cameras and under the different working conditions, the model is tested on the data set used to define the model (see section 4.5.2). Since we have the reflectances of these samples, the device-dependent initial RGB values can be computed for the simulation. These device-dependent RGB values are transformed to CIELab following the steps defined for each condition. Then, for each sample, memberships to the 11 colour categories are computed and a colour name is assigned by applying the maximum criteria (see equation (1.20)). Memberships and names obtained under the ideal conditions in which the model was defined are considered the ground truth to which the results from different cameras are compared.

For each camera and each working condition, two measures are computed. First, we compute the percentage of samples correctly named comparing with the ground-truth. Second, the Mean Absolute Error, MAE_{fit} (see equation 4.3), between the memberships obtained on the CIELab space computed from the camera responses and the ground truth. Note that although both error measures are related, they evaluate different aspects of the problem. The MAE_{fit} value will give a precise measure of the error in terms of membership to the different colour categories. The percentage of correct naming will give us a measure of the impact of the previous error in the final result of the colour-naming model. In many cases, a certain error in terms of membership will not imply a change in the name assigned by the model. In these cases, the model will work well although the MAE_{fit} is high. However, in samples

close to a boundary a small change in the membership values can imply a change in the name assigned. The results obtained are summarized on table 5.1.

	Condition 1		Condition 2		Condition 3	
Camera	Naming (%)	**MAE$_{\text{fit}}$**	Naming (%)	**MAE$_{\text{fit}}$**	Naming (%)	**MAE$_{\text{fit}}$**
Nikon D70	95.09	0.0129	87.60	0.0215	71.06	0.0512
Sony DXC930	90.70	0.0153	80.10	0.0364	71.83	0.0520
Kodak DCS420	95.61	0.0105	80.62	0.0351	62.53	0.0822

Table 5.1: Results obtained on the analysis of the model under different working conditions.

As can be seen on the table, the model is fully applicable for those cases when control on acquisition is possible. In this case, the results obtained are of $90 - 95\%$ of the samples labelled with the same colour name as the model and the value of the error (MAE_{fit}) is low. These values are the maximum performance that the model obtains on these devices.

When we are able to make a colour calibration previously to the application of the colour naming, the performance decrease depending on the camera. The best results are obtained with Nikon D70 camera with 87.60% of the samples correctly named. In this condition, the results will be conditioned by the accuracy of the calibration step.

In the third case, without calibration, the results that we can expect are considerably lower than before with performance falling to values around 71% or 62% depending on the camera used. Note that this is the worst possible case when RGB values of the non-sRGB camera are transformed to CIELab as if they were sRGB.

From these results several conclusions can be extracted. The colour-naming model can obtain good results in computer vision when it is applied to images for which we know the response curves of the camera used, we have control on image acquisition or, at least, it is possible to do a colour calibration step previously to the colour naming application. The performance of colour naming on real images will be highly conditioned to the camera used to acquire the images. For those cameras with response curves similar to the sRGB ones, a good calibration will be easier to compute.

Finally, on those applications where conditions are not controlled and no colour calibration is possible, the performance of colour naming will decrease considerable as it could be expected. However, in real images, it will be normally possible to apply a colour constancy method that can reduce the error of working on fully uncalibrated conditions. In this experiment this fact is not included because samples were considered individually and computations were all done synthetically. However, the important result of this analysis is that it provides a measure of the error order that can be expected on each of the working conditions considered.

5.3 Global Image Colour-naming Descriptors

As we explained in Chapter 1, the direct application of a colour-naming model in computer vision is image annotation for image retrieval. Given a colour sample, $\vec{p}$,

the colour naming model provides us with a membership value to each of the 11 colour categories. However, if we aim to use the model to describe the content in an image, global colour-naming descriptors for the whole image are needed.

In the last years some efforts to define global colour descriptors for images have been done. The standard for multimedia content description MPEG7 [8] has defined four descriptors for colour. These descriptors are:

- Dominant colour

- Colour distribution

- Global spatial distribution (colour layout)

- Local spatial distribution

In this section, we will propose three global colour-naming descriptors of image content that will be used in the next section to apply the model to automatic image annotation. With these descriptors we do not aim to propose an alternative to those of MPEG7, but to show the usefulness and potentiality of colour-naming information for describing image colour, and particularly for automatic image annotation.

As we proposed in section 1.5, the colour-naming information in form of memberships that the model provides can be represented by a simple colour descriptor, $CD(\vec{p})$, which is a vector of 11 components where each component is the membership to one of the 11 colour categories (see equation (1.19)). The correspondence between components and categories is ruled by equation (1.21).

Given an image, I, we can compute the colour-naming descriptor, $CD(\vec{p_i})$, for each pixel, $\vec{p_i}$, with $0 < i < N - 1$, where N is the number of pixels in the image. Hence, for a given image we have a set of N colour descriptors (one for each pixel) that provide a considerable amount of useful information to define global descriptors of the image content in terms of colour names. Amongst all the possible, we propose three simple descriptors that will be used on the application we will present later in this chapter. The descriptors proposed are:

- **Colour Naming Histogram (CNH)**
  The CNH descriptor, $\overrightarrow{CNH}(I)$, is a 11-dimensional vector where each component, $CNH_k(I)$, is the relative frequency of pixels in the image with its maximal membership corresponding to category C_k:

  $$CNH_k(I) = \frac{np_k}{N} \qquad (5.9)$$

  where np_k is the number of pixels with maximal membership corresponding to category C_k and N is the number of pixels in the image.

  In this descriptor, each pixel only contributes to the bin corresponding to the category for which the pixel has its maximal membership and the second and subsequent maximal membership values are disregarded. The goal of this descriptor is to capture the predominant colours in the different image regions of the image.

- **Mean Membership Vector (MMV)**
  The MMV descriptor, $\overrightarrow{MMV}(I)$, is a 11-dimensional vector where each component, $MMV_k(I)$, is the mean of the membership to category C_k of all the pixels in the image, that is:

$$MMV_k(I) = \sum_{i=0}^{N-1} \frac{CD_k(p_i)}{N}, \qquad k = 1, \ldots, 11 \tag{5.10}$$

  where $\vec{p_i}$ is the ith pixel of the image and N is the number of pixels in the image.

  In this descriptor, each pixel contributes to the descriptor with the memberships to all the categories, no matter whether they are high or low. Hence, with this descriptor it will be possible to obtain a more perceptual measure of the global image content since no colour-naming information is disregarded.

- **Predominant Colour (PC)**
  The predominant colour of an image is obtained by applying the decision function defined in equation (1.20) which chooses the maximum from a 11-dimensional colour-naming vector. In our case, the colour-naming vector can be any of the two image descriptors previously defined. Therefore, the PC descriptor is defined as:

$$PC(I) = t_{k_{\max}} \quad | \quad k_{\max} = \arg\max_{k=1,\ldots,11}\{CD_k(I)\} \tag{5.11}$$

  where t_k is the linguistic term to name the colour category C_k, and $CD(I)$ is an 11-dimensional vector containing colour-naming information from image I.

  This measure can be generalized to the M predominant colours by selecting the list of M maximal values on the colour-naming vector.

Obviously, although the three descriptors have been proposed as global descriptors, they can be used locally in any given region. Hence, given an image it can be divided in several parts and descriptors for each part could be computed to easily derive the colour layout in the image. However, this is out of the scope of this work and in next section, descriptors are used globally.

5.4 Automatic Image Annotation for Image Retrieval

In this section we present the application of the colour-naming model on a real problem of automatic image annotation for image retrieval. As we saw in Chapter 1, two kind of techniques have been used in image retrieval. On one hand, the first systems indexed and retrieved images based on a set of textual annotations that described the image content. On the other hand, Content Based Image Retrieval (CBIR) systems have computed low-level features from images and these values have been used to index and establish similarity between them.

The growth of multimedia contents in the last decade, caused that annotating by hand entire image databases became unfeasible. In this context, CBIR has been given more attention in the last years due to the lack of robust automatic textual

annotators. However, textual descriptors present important advantages over CBIR. Firstly, queries in CBIR normally require providing the system with an image similar to the one the user wants to retrieve. Hence, if a similar image to the desired one is not available, results in CBIR can be unsatisfactory. Secondly, a query based on natural language is easier to define than a query based on low-level features since it allows to express more exactly what is desired by the user. This is the case of the real application in which our colour-naming model will be tested.

5.4.1   Experiments

Age Fotostock is a photographic agency that manages image reproduction rights for professional use in the corporate, editorial, advertising and design fields. The company has a huge image database where each image is annotated by human operators with a set of textual labels describing information related to the image. The information described includes very different topics such as colour, texture, objects, emotions, acquisition conditions, place and others. In the system, there is not any ontology defined and human operators use their own vocabulary to describe images. Searches on the system are done through string matching between the user's query and the annotations of the whole database. Figure 5.5 shows an example of the annotations attached to an image.

Air	Fojtik	Museum	Stone
Black	Folk	Nature	Stony
Blue sky	Grass	Nobody	Summer
Buildings	Green	Open	Sunny
Country	Horizontal	Outside	Three
Countryside	House	Red	Village
Daytime	Houses	Roof	White
Europe	Iceland	Skogar	Wood
Exterior	Island	South	Wooden

Figure 5.5: Example of the annotations attached to an image of Age Fotostock database.

The company is interested in automating the process of image annotation (or at least, a part of it) and, in this context, an automatic system of colour naming would be very useful. In this section we present a first approach to this problem with experiments on a reduced subset of the image database. Hence, the model will be

tested on a real computer vision application where the working conditions will be far from the ideal conditions under which the model was defined.

In the experiments, the annotations made by the operators will be used as the ground truth and the annotations provided by the model will be compared to this ground truth to evaluate performance of the model.

Before the presentation of the experiments, some considerations about this problem must be done:

- The images provided are in 'jpg' format and have low resolution.

- The database includes images from very different sources.

- The content of images includes a wide range of objects, situations and conditions.

- Information about acquisition conditions of the images is not available.

- Most images are not completely annotated in terms of colour, that is, there are images that do not have attached labels for all the colours in the image.

- Some colour labels attached to images correspond to very small areas of a colour that is not predominant in the image.

- There are erroneous annotations, that is, labels corresponding to colours that are not in the image.

Therefore, the model will be tested on fully uncalibrated conditions. Furthermore, the presence of errors on the annotations will complicate the evaluation of the model since, as we mentioned before, these annotations will be used as ground truth to compare the results of our model.

To annotate images, the descriptors defined in the previous sections will be used. The two global descriptors defined, the Colour Naming Histogram (CNH) and the Mean Membership Vector (MMV) will be compared in the experiments. These descriptors will be used as input for the Predominant Colour (PC) descriptor to obtain the labels attached to each image. The PC descriptor has been extended to the 5 predominant colour with the constraint that the returned labels must have a value of the global descriptor (either CNH or MMV), higher than a threshold that in our case has been set to 0.05, that is, the colour name represents more than 5% of the pixels in the image.

At this point we must emphasize that the goal of these experiments are not to propose an annotation method (actually the method used is very simple), but to test the possibilities that the colour-naming model has of being applied to real problems in computer vision.

The performance of the model on the experiments will be measured by the percentage of the annotations from the human operators that are also provided by the colour-naming model. It would be interesting to compute the inverse measure, that is, the percentage of model labels that have also been provided by human annotators, but the fact that images are not completely well annotated makes it unfeasible to compute such measure.

Experiment 1

For this experiment, we retrieved 10 images from the database for each colour category to build a test set of 110 images with their corresponding annotations. The images selected from each query were the first 10 in the results retrieved from the company database. From all the annotations, only those corresponding to the 11 colour categories were maintained. The rest (non-colour annotations and non-basic colour annotations) were disregarded. Some images have associated more than one colour annotation. The total number of colour annotations for the whole set is of 173 labels. The distribution of the annotations across colour categories can be seen in table 5.2. As a query for each colour name is made, there are at least 10 annotations for each colour category. In figure 5.6 some examples of the images used for the experiment are shown.

	Red	**Orange**	**Brown**	**Yellow**	**Green**	**Blue**
Annotations	18	11	12	18	16	20
	Purple	**Pink**	**Black**	**Grey**	**White**	**Total**
Annotations	13	10	16	13	26	173

Table 5.2: Distribution of colour annotations in the set of images used in experiment 1.
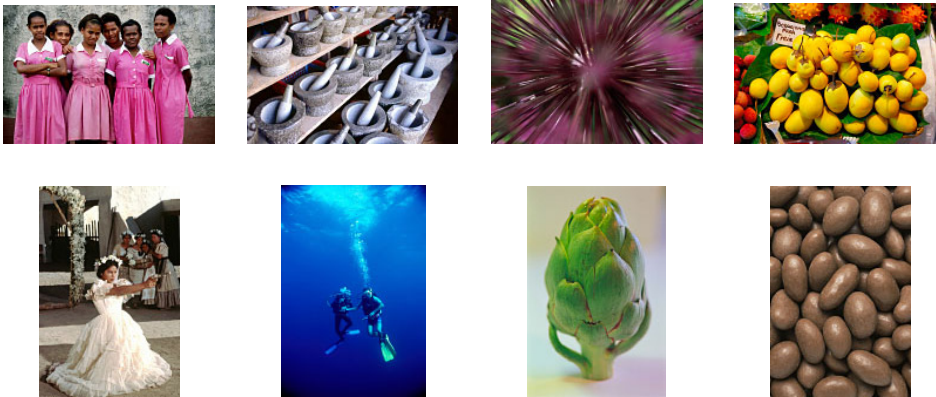
Figure 5.6: Examples of images in the set used for experiment 1.

The three global image descriptors defined in section 5.3 are used in this experiment. The results obtained with the Colour Naming Histogram (CNH) and the Mean Membership Vector (MMV) are compared. As we explained above, the selection of colour names to annotate images is done by applying the Predominant Colour (PC) descriptor to give a maximum of 5 colour terms. Images are assumed to be in sRGB and no calibration step is done. The algorithms to consider colour constancy and colour induction explained in sections 5.2.2 and 5.2.3 respectively are also tested for both descriptors. Results are summarized in table 5.3.

	Descriptor	
	CNH	**MMV**
None	80.92	81.50
Colour Constancy	82.08	82.08
Colour Induction	83.82	83.24
Col. Constancy and Induction	84.39	83.82

Table 5.3: Results for experiment 1. Percentages of database annotations also returned by the model. The descriptors defined, Colour Naming Histogram (CNH) and Mean Membership Vector (MMV), are tested after applying different preprocessing steps.

As can be seen in table 5.3 the results for the different combinations of descriptors and preprocessing algorithms are around 80-84%. The results obtained for the two proposed descriptors are very similar, although the best performance is obtained with the Colour Naming Histogram when colour constancy and colour induction algorithms are applied previously to the colour-naming model.

The use of both the colour constancy and the colour induction algorithms improves the results obtained when the model is applied without preprocessing. The best results are obtained when both are used jointly. In this case, an improvement of 3% is achieved with respect to the application of the model without preprocessing.

The annotations from the database that were not correctly assigned by the colour-naming model have been separated by colours to analyse how errors are distributed in the different categories. Table 5.4 shows the percentage of errors for each category and for each method tested. The fact of having only 173 annotations makes that, in some cases, small variations in the number of correct or wrong annotations imply high changes in the percentages. Despite this fact, the table allows us to have a measure of how well performs the model for each colour category. Thus, in the table it can be seen that Red, Purple, Brown and White are the categories for which the method has more errors.

In figure 5.7 some results are presented. For each image, the annotations in the database and the ones provided by the model are presented. Pixels of the original images have been labelled with a colour representing the colour category for which it has maximum membership.

The analysis of the results showed that some of the errors of the model corresponded in fact to incorrect annotations in the original database. Figure 5.8 shows some of these problematic cases.

In some cases, the colour annotations that did not correspond to colours in the image were mistakes of the human annotators (e.g. 'White' in Figure 5.8(a)). In other cases, the colour terms did not describe colours in the image, but were related to the image content. This is the case of annotations such as 'red wine' or 'red sea' (see 5.8(b)). Since searches in the company database are done by simple string matching, these kind of images are also returned when a query with a colour name is done. Another group of problematic annotations are those describing small areas of the image that are important for the meaning of the image but are not predominant in

	CNH				MMV			
	None	CC	CI	CC+CI	None	CC	CI	CC+CI
Red	33.33	33.33	27.78	27.78	33.33	33.33	27.78	27.78
Orange	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Brown	25.00	16.67	25.00	33.33	16.67	8.33	33.33	33.33
Yellow	11.11	11.11	11.11	11.11	11.11	11.11	11.11	11.11
Green	12.50	12.50	12.50	12.50	12.50	12.50	12.50	12.50
Blue	10.00	10.00	10.00	5.00	10.00	10.00	10.00	10.00
Purple	38.46	30.77	53.85	38.46	30.77	30.77	46.15	38.46
Pink	20.00	20.00	0.00	10.00	20.00	20.00	0.00	0.00
Black	6.25	6.25	0.00	6.25	12.50	12.50	6.25	6.25
Grey	7.69	7.69	7.69	7.69	7.69	7.69	7.69	7.69
White	34.62	34.62	23.08	19.23	34.62	34.62	23.08	23.08

Table 5.4: Percentage of errors in the model annotations for experiment 1 separated by colours. Abbreviations used: CNH - Colour Naming Histogram, MMV - Mean Membership Vector, CC - Colour Constancy, CI - Colour Induction

White	White	Green	Green
Orange	Orange	Blue	Blue
Green	Green	Red	Red
Yellow	Yellow		
Red	Brown		

Red	Blue	Blue	Blue
Black	Black	Orange	Green
Yellow	Green	Red	White
White	White		Grey
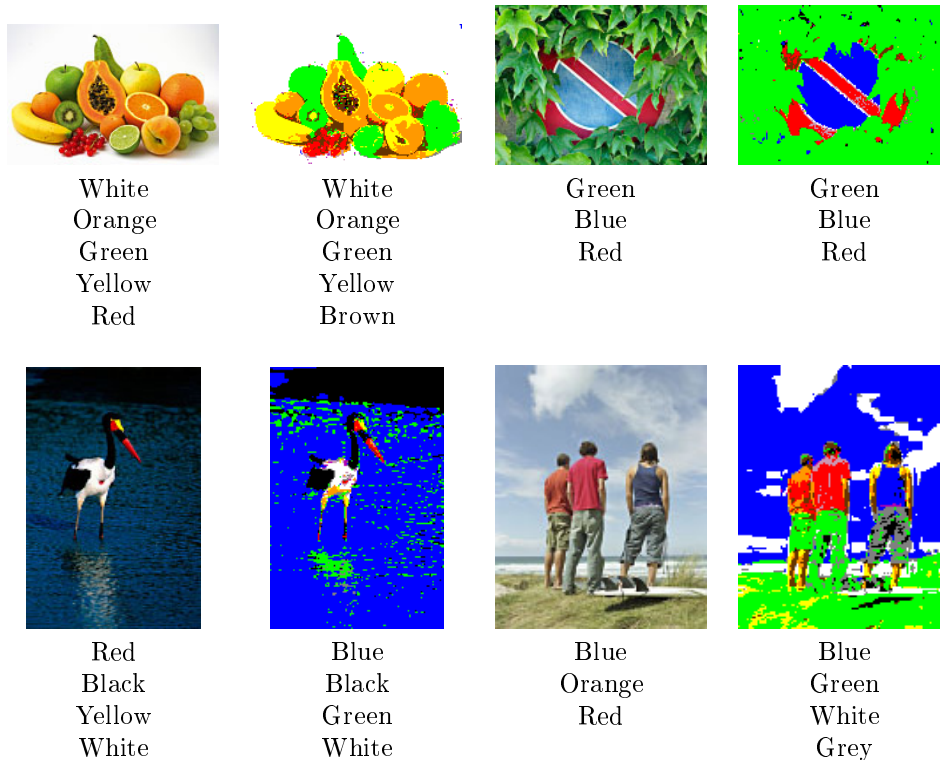
Figure 5.7: Examples of annotations provided by the model. Images on the right of each original image show the colour that has been assigned to each pixel by the model. Under each image, the annotations associated are presented.

the image. This is the case of images with annotations such as 'blue eyes' (see 5.8(c) and 5.8(d)).

Blue	Black	Blue	Red
Yellow	Purple	Green	Blue
Orange	Red	White	Green
White	Green	Red	
(a)	(b)	(c)	(d)

Figure 5.8: Examples of images with problematic annotations.

The main conclusion of this experiment is that the annotations available on the database of Age Fotostock are not adequate to be considered the ground truth to validate the model, although results obtained by the TSE model are near to 85%. At least 9 out of the 173 annotations (5% approx.) on the test set are clearly incorrect and correspond to colours that are not present in the images to which these annotations are associated. The second conclusion is that the global colour-naming descriptors defined are too simple to be able to annotate small regions that are important in the image. To solve this problem, more complex descriptors should be considered in order to be able to correctly annotate these regions.

Experiment 2

In this experiment, the 110 images used in experiment 1 were re-annotated by two human observers in order to obtain a set of annotations better than the set of the database. The two subjects in this experiment were asked to label each image with the colour names of the regions that were important in the image and that described most of its content. For each image, the intersection between the two sets of annotations from the subjects was considered as the final annotations set of the image. As a result of this experiment, a total of 241 colour annotations were obtained. Table 5.5 shows the new distribution of annotations for the different categories. Figure 5.9 shows some examples of images in which the annotations obtained in this experiment were different from the ones in the company database.

Once a more complete set of annotations for the 110 images was available, the percentage of image annotations that were also provided by the model was computed for the two proposed descriptors and for the preprocessing algorithms as it was done in experiment 1. The results obtained are summarized in table 5.6.

As it can be seen in the table, all the results have improved between 4% and 7%, which agrees with one of the conclusions of experiment 1 that about 5% of the annotations in the company's database are wrong. As in experiment 1, the inclusion of

	Red	Orange	Brown	Yellow	Green	Blue
Annotations	20	12	18	24	27	27
	Purple	Pink	Black	Grey	White	Total
Annotations	13	14	24	17	45	241

Table 5.5: Distribution of colour annotations in the set of images used in the second experiment.
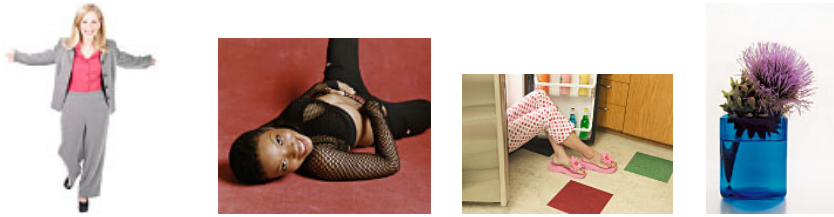
ORIGINAL ANNOTATIONS FROM THE DATABASE

Grey	Black	White	Purple
Red		Pink	

NEW ANNOTATIONS FOR EXPERIMENT 2

Grey	Black	White	Purple
Pink	Red	Pink	Blue
White	Brown	Red	White
		Green	

Figure 5.9: Examples of images that were assigned different annotations from the ones in the company's database.

	Descriptor	
	CNH	MMV
None	87.97	87.97
Colour Constancy	89.63	88.80
Colour Induction	87.14	87.14
Col. Constancy and Induction	90.04	89.21

Table 5.6: Results for experiment 2. Percentages of database annotations also returned by the model. The descriptors defined, Colour Naming Histogram (CNH) and Mean Membership Vector (MMV), are tested after applying different preprocessing steps.

preprocessing steps taking into account perceptual issues improve the results in most of the cases. In this experiment, when only the colour induction process is applied before the colour-naming model, results slightly decrease from the ones obtained applying only the colour-naming model. However, when it is combined with the colour constancy algorithm the results are the best ones for both descriptors. As the colour constancy algorithm is applied before it corrects some effects in the images that make the colour induction to fail.

As we did in experiment 1, the percentage of errors for the different categories is showed in table 5.7. As can be seen in the table, the distribution of errors is similar to the one from experiment 1 (see table 5.4) with small variations for most categories. However, two of the categories that worst performed in experiment 1, Red and White, show a considerable improvement. The two categories that obtain higher error percentages are Brown and Purple, but it can be observed that these categories considerable decrease its performance when the colour induction is applied. The same problem appears for Green, but the error values are not so high as in the other two categories.

	CNH				MMV			
	None	**CC**	**CI**	**CC+CI**	**None**	**CC**	**CI**	**CC+CI**
Red	15.00	15.00	15.00	10.00	15.00	15.00	15.00	15.00
Orange	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Brown	16.67	11.11	27.78	33.33	16.67	11.11	33.33	33.33
Yellow	4.17	4.17	4.17	8.33	4.17	4.17	4.17	4.17
Green	3.70	3.70	14.81	14.81	3.70	3.70	14.81	14.81
Blue	3.70	3.70	3.70	3.70	3.70	3.70	3.70	3.70
Purple	23.08	15.38	46.15	30.77	15.38	15.38	38.46	30.77
Pink	28.57	28.57	14.29	21.43	28.57	28.57	14.29	14.29
Black	4.17	4.17	0.00	4.17	8.33	8.33	4.17	4.17
Grey	11.76	5.88	11.76	11.76	11.76	11.76	11.76	11.76
White	22.22	20.00	15.56	13.33	22.22	20.00	13.33	13.33

Table 5.7: Percentage of errors in the model annotations for experiment 2 separated by colours. Abbreviations used: CNH - Colour Naming Histogram, MMV - Mean Membership Vector, CC - Colour Constancy, CI - Colour Induction

Finally, the analysis of the images for which the colour-naming method provided wrong annotations, showed that some of the errors were caused by the presence of shadows in the image. Figure 5.10 presents examples in which shadows cause wrong annotations of image areas. Some algorithms to solve this problem have been proposed [49, 55, 51] and it would be interesting to include a previous step to apply one of these algorithms before applying the colour-naming model.

5.5 Discussion

In this chapter we have considered the issues related to the application of the TSE colour-naming model to real problems in the computer vision field. To take into
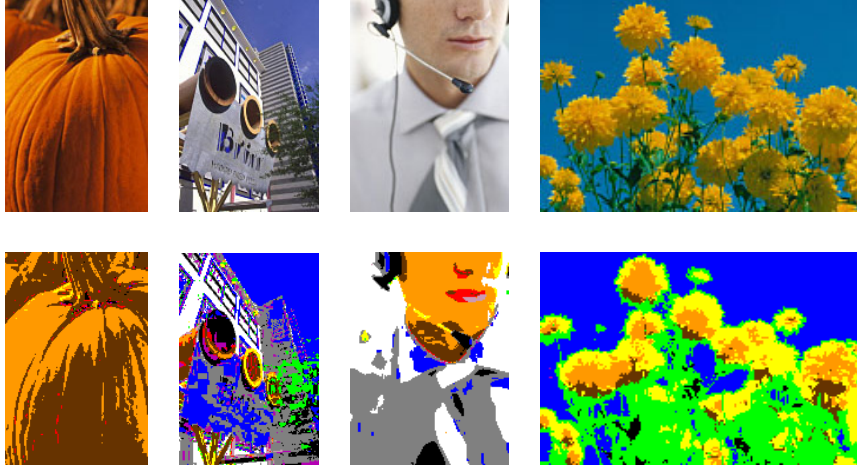
Figure 5.10: Examples of wrong annotations caused by shadows in images.

account these issues, we have defined a modular scheme following the proposal in
[148], with the preprocessing steps that must be performed before applying the colour-
naming model.

The fact that the model was defined using data obtained in psychophysical ex-
periments where colour samples are presented in very controlled conditions, implies
that some perceptual issues that will appear in real conditions must be considered.
As we explained in Chapter 1 the model was defined assuming that it would work on
perceived images where the effects of perceptual adaptation to the illuminant and to
the surround had been previously removed.

Another aspect that has been considered is the fact that the colour-naming model
was defined for the CIELab space while images will normally be acquired by an RGB
device and, therefore, a space transform must be done. Ideally, the RGB device should
provide standard RGB (sRGB) for which the transform to XYZ space is known.
However, in most of the cases images are represented on a device-dependent RGB
space and a previous calibration step is needed. Depending on the knowledge and
control on the acquisition conditions this calibration step can be done in different
ways.

In many cases, the acquisition conditions will be unknown and the colour-naming
model will work under uncalibrated conditions. This problem has been analysed in
an experiment simulating the different possible working conditions and for different
RGB sensors. The results of the experiment show that under calibrated conditions
the model can obtain good results working with different sensors. Obviously, under
completely uncalibrated conditions the performance of the model will decrease and it
will depend on the application whether the colour-naming results are acceptable or
not.

The model has been tested on a real problem of automatic annotation for image
retrieval. The goal of this test is to have a measure of the model performance on a

real computer vision application. With this goal, we have defined the preprocessing steps for images before the application of the colour-naming model. We assume that images are represented in sRGB and therefore no calibration step is done. To consider the colour constancy problem the White-patch Retinex algorithm [34] is applied and colour induction effects are solved with the model of Otazu and Vanrell [113]. The solutions adopted for the problem of non-sRGB images and for colour constancy are not the best option and obviously, the selection of approaches with better results would probably improve the perceived image and therefore, the performance of the colour-naming model.

Three colour-naming descriptors for images have been proposed and analysed on the experiments to test the model on the automatic annotation task. The experiments have been done on a set of 110 images. Although it would be desirable to validate the model on a larger set of images the experiments allow extracting some conclusions. Firstly, the two colour-naming descriptors used to derive the final colour names (Colour Naming Histogram (CNH) and Mean Membership Vector (MMV)) obtain similar results and none of them shows significant better results than the other. Secondly, an additional previous step removing shadows from images is needed since in some cases the presence of shadowed regions leads to wrong name assignments. Finally, two colour categories, Brown and Purple, obtain poor results in the experiment for all the configurations evaluated.

In the experiments, the model achieves up to 90% of correct annotations. However, the most interesting of the results is the fact that it obtains good results although some of the preprocessing steps were not the best choices and the colour-naming descriptors used are quite simple. These results are promising and point out future research lines and applications. Hence, for example, if all the pixels of an image are assigned with a membership value to the 11 colour categories, the information of the image can be represented in a 11-dimensional space where each dimension is associated to one of the 11 categories. This representation can be considered more perceptual since each dimension will be related to one of the colour categories that humans perceive. This representation could facilitate some computer vision tasks such as segmentation. Figures 5.11 and 5.12 show two examples of this representation in 11 planes. In both cases, the important regions of the image are almost perfectly segmented in the channel corresponding to its colour.
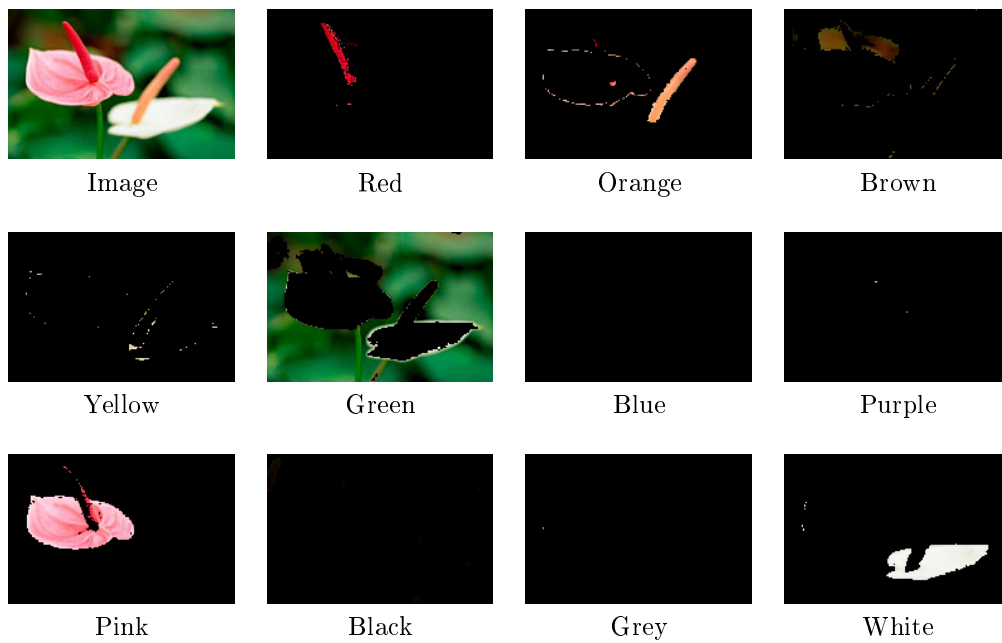
Figure 5.11: Example of an image represented on a 11-dimensional space where each dimension corresponds to a colour category (Example 1).

Image	Red	Orange	Brown
Yellow	Green	Blue	Purple
Pink	Black	Grey	White
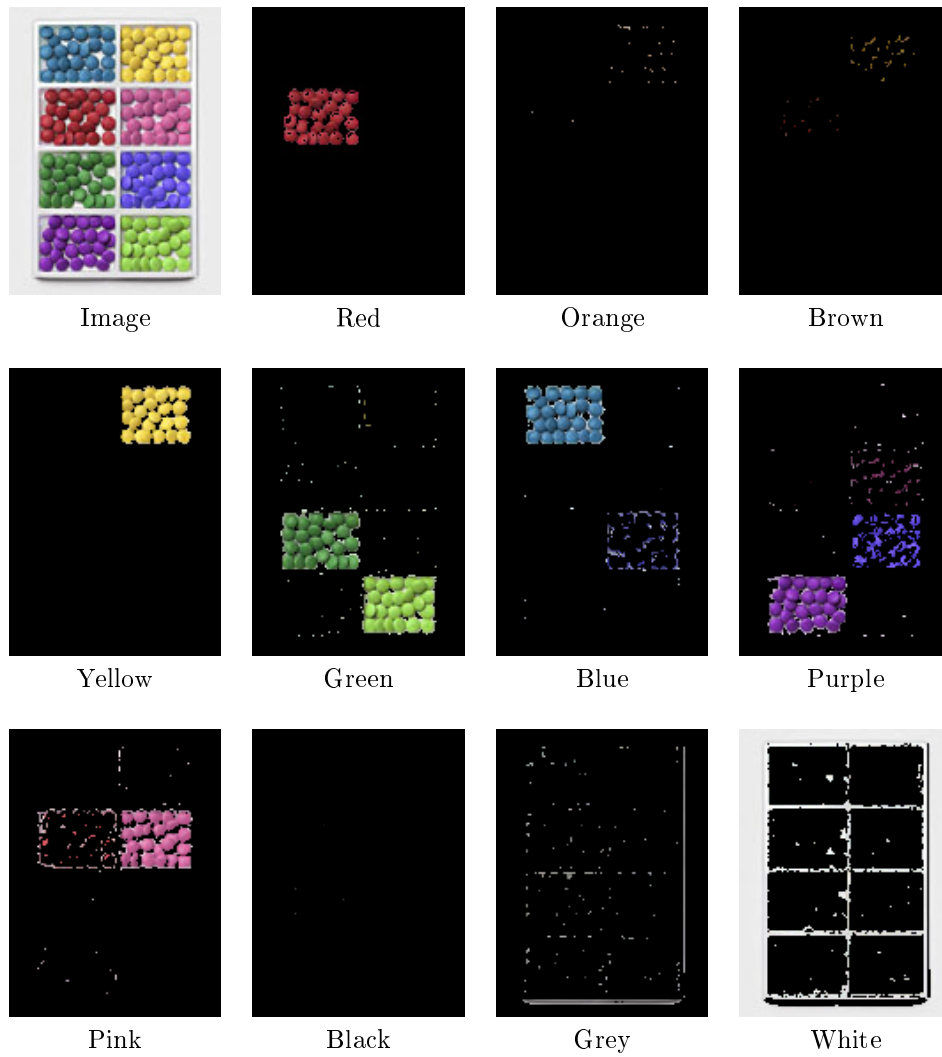
Figure 5.12: Example of an image represented on a 11-dimensional space where each dimension corresponds to a colour category (Example 2).

Chapter 6

Conclusions and further work

The main goal of this thesis was to propose a colour-naming model to reduce the existing semantic gap between the low-level information that can be obtained from images and the high-level semantics that human beings use. Hence, the main contribution of this work is the definition of a parametric colour-naming model for digital images. However, this thesis makes some additional contributions and opens several research lines for future work. In this chapter we present some conclusions of the work done and summarize the contributions of this thesis. Finally, some applications of the defined model and directions for future work are proposed.

6.1   Conclusions

At the beginning of this thesis, we stated that the main goal of this work was to automate the colour-naming task in the frame of computer vision. The most important result of this work is the definition of a parametric model to automate the colour-naming task. The model has been posed on a fuzzy framework where each colour category is modelled by a membership function. The categories considered are the eleven basic colour categories defined by Berlin and Kay [27].

In order to obtain a colour-naming model providing the same colour names a human being would assign, a colour-naming experiment was developed. The goal of this experiment was to obtain a set of fuzzy colour-naming judgements that could be used as learning set in a later fitting process to estimate the parameters of the models for the CIELab space. The results of the experiment provided valuable information to define a set of properties that membership functions of the model should fulfil.

After a preliminary approach using multivariate Gaussians as membership functions, the subsequent proposals aimed to fulfil the properties defined from the observation of the experiment results. The functions proposed are based on the well-known one-dimensional sigmoid function which has been extended to two dimensions and in

which rotations and translations have been included. Our first proposal, the Sigmoid-Gaussian function (SG), combined two oriented 2D-sigmoids with a Gaussian function. The effect of the Gaussian multiplying the product of the two sigmoids caused that this model could not be well fitted to some of the categories, such as *Green*, which have a wide area of the colour space with membership 1. This fact implied high values of the fitting error, MAE_{fit} (see equation (4.3)), for some of the categories and a high value of the unity-sum measure, $MAE_{unitsum}$ (see equation (4.6)).

To solve this problem, individual rotation angles were included on each of the two sigmoids and the Gaussian function was removed to define the Double-Sigmoid function (DS). This model considerably improved the unity-sum constraint fulfilment (reflected on an important reduction of the $MAE_{unitsum}$ value) but presented problems to correctly model the transition from chromatic categories to achromatic ones.

The Triple-Sigmoid with Circular centre function (TSC) included a Circular-Sigmoid to improve the fitting of the model to the chromatic-achromatic boundary. The results obtained slightly improved the ones from the previous model, but the values of the error measures were still too high.

To obtain the final Triple-Sigmoid with Elliptical centre function (TSE) some important modifications were included in the previous model:

- The Circular-Sigmoid was replaced by an Elliptical-Sigmoid in order to improve the fitting to the chromatic-achromatic boundary.

- The 1D-Gaussian functions that allowed differentiating between the three achromatic categories (Black, Grey and White) were replaced by 1D-sigmoid functions.

- The learning set used up to then was replaced by a wider set of samples where membership values were derived from data of the psychophysical experiment of Sturges and Whitfield [135] by applying the Fuzzy Colour Category Map of Seaborn et al. [124].

- The fitting process was modified to include some constraints on the parameters of the functions in order to achieve that neighbouring categories had in common the parameters that defined their boundary.

The analysis of the results of the TSE model shows a good fitting to the learning set and almost a perfect fulfilment of the unity-sum constraint. The categorization of the Munsell space has been used to evaluate the model in terms of the names assigned to a set of samples and the results of the TSE model equals previous non-parametric best.

The final result of all this work is the table of parameters that define the membership functions to the 11 basic colour categories. These parameters allow obtaining the membership values to the eleven basic categories for any point of the colour space.

Once the main contribution of the thesis has been presented, the other contributions of this thesis are summarized in the following paragraphs:

- **Definition of properties for membership functions**

  The study of the results of the psychophysical experiment has allowed us to define the set of necessary properties that any membership function should fulfil (see section 4.3). The proposed TSE model fulfils the four properties, although other functions fulfilling these properties could be found. The set of properties that have been defined in this thesis can be a start point for future definitions of different membership functions to model colour-naming memberships.

- **Definition of a methodology for fuzzy colour-naming experiments**

  We have proposed a new methodology for colour-naming experiments. After a review of the literature on psychophysical experimentation for colour naming, we have defined a different methodology to obtain an adequate data set for the final purpose of modelling the colour-naming task.

  The results of the experiment have been validated by computing some usual statistics in colour naming and comparing them to the same statistics from previous experiments. With this analysis we have proved the equivalence of the results obtained with our fuzzy methodology to the results from previous experiments.

- **Providing psychophysical data for fuzzy colour-naming research**

  The colour-naming experiment has provided us with a set of psychophysical fuzzy colour-naming data . The most valuable of this data set is that it provides fuzzy membership values to the eleven basic colour categories and therefore it is suitable to be used as learning set or test set in a modelling process of the colour-naming task.

  This dataset has been made available online at [3] to the scientific community in order to be used to model or validate future colour-naming models.

- **Providing psychophysical data for colour constancy research**

  The colour-naming experiment developed was repeated with four additional illuminants. The final result is a set of human colour-naming judgements under five different illuminants. These data can be very useful for future research on the relationship between colour naming and colour constancy.

- **Definition of a tool for psychophysical analysis**

  The proposed model can be an useful tool for the analysis of psychophysical colour-naming experiments. Up to this moment, colour-naming discussion in psychophysics has been focused on areas or simpler statistics such as centroids.

  With the TSE model, evaluation of results can be done in terms of the model parameters instead of in terms of statistics that depend on the set of samples used. This will allow correct comparison between experiments done with different samples sets.

- **Proposal of a modular scheme to apply the colour-naming model**

Since the colour-naming model has been defined under ideal conditions, perceptual issues such as colour constancy and colour induction effects must be addressed before applying the model to real problems. We have analysed these topics and have proposed a modular scheme where colour constancy and colour induction algorithms are applied to obtain a perceived image, i.e. an image where influences of the illuminant and the surround have been removed. For the moment, we have proposed simple existing methods to take into account these problems before applying the colour-naming model, but more complex methods could be applied in order to improve the results.

- **Analysis of the model application to computer vision problems under uncalibrated conditions**

  In order to evaluate the performance that the model can achieve under different working conditions we have made simulations with three different cameras. This analysis has been focused on evaluating the performance that can be expected from the model working under uncalibrated conditions. The results show that the model can provide good results on calibrated conditions but the performance will decrease under uncalibrated conditions.

- **Proposal of colour-naming descriptors for images**

  We have proposed three descriptors to describe images in terms of colour naming. Although the descriptors defined are quite simple, they have performed well on the experiments done on a real application. The descriptors that have been proposed are a global measure of the colour content in the image and can provide useful information for several applications.

6.2   Future Work

The colour-naming model developed in this thesis is a further step on the automation of the colour-naming task and the reduction of the semantic gap in computer vision. Furthermore, this work opens several research directions for future work which are presented in the following paragraphs:

- **Psychophysical validation**

  It would be interesting to test the model with more psychophysical data obtained with the proposed methodology for fuzzy colour-naming experiments. As the first data set and the definition of the colour-naming model have determined the areas of the colour space where confusion is higher, validation could be focused in these areas which normally correspond to boundaries between colour categories. Some preliminary experiments have already been done [115].

- **Improving the TSE model**

  The proposed model can be improved in several points. Firstly, the division of the colour space in different lightness levels should be removed. Observation

of the membership maps of the TSE model (figure 4.23) allows us to detect some tendencies in the displacement of the boundaries between colour categories across lightness levels. Hence, the parameters of the membership functions could be interpolated along the levels defined in the current model to obtain the parameters of the membership functions for any given value of lightness.

Secondly, it would be also interesting to relax or even eliminate the first assumption done in the fitting process to allow that the membership transition from chromatic categories to the achromatic center was different for each category.

- **Extension of the vocabulary**

  Although the 11 basic colour terms considered in the model are enough for many applications, the vocabulary could be easily extended. The TSE model has shown good efficiency and versatility to be fitted to membership values. Hence, the inclusion of a new category in the model would only imply to estimate its parameters and re-estimate the boundary parameters of the neighbouring categories.

  Furthermore, the vocabulary could also be extended by using the fuzzy information provided by the model in different ways:

  - Compound nouns could be used for samples with a membership of 0.5 to two categories (e.g. samples with memberships 0.5 to *Green* and 0.5 to *Blue* could be named as *Blue-Green*)

  - Modifiers such as the '-*ish*' suffix could be used on samples with a high membership to a category and up to a certain membership to another (e.g. samples with memberships 0.7 to *Green* and 0.3 to *Blue* could be named as *bluish Green*).

  - Luminance information could also be used to obtain some modifiers such as 'light' or 'dark'.

  - New colour terms such as 'beige' or 'turquoise' by defining the membership pattern that characterize these colours.

  Finally, it would be interesting the inclusion of a category for skin colour which is important to describe content of images that content human beings.

- **Application of the model to computer vision problems**

  The proposed model has been tested on a real computer vision application of image annotation. The model has obtained good results on a set of 110 images that had been previously annotated by human operators of the company *Age Fotostock*. However, a validation with a larger set of images should be done.

  The test on the annotation application allowed us to detect that the model was not robust to the presence of shadows in the images. Hence, a preprocessing step accounting for shadows should be added in order to remove them from the perceived image on which the colour-naming model is applied.

  Apart from automatic image annotation, the model can be applied to other computer vision problems such as the following:

– **Segmentation**

The use of the information provided by the colour-naming descriptors defined in Chapter 5 can be useful to obtain more perceptual segmentations than current methods based on pixel distributions. The representation of images in terms of the 11 memberships can be a first step to segment regions as it has been shown in section 5.5 (see examples of figures 5.11 and 5.12).

– **Content based image retrieval**

Colour has been widely used as a feature for content based image retrieval. Similarity between images has been usually done by using the histogram of the image [138]. Although histograms are a good measure of the colour content of the image, it does not have any information in terms of how this content is perceived by humans. Hence, the pixels of a region can be included in different bins of the histogram because the pixels values present variability due to the geometry and illumination conditions, although a human observer perceives them as being of the same colour. The use of a histogram based on the colour names, similar to the proposed Colour Naming Histogram (CNH) and the Mean Membership Vector (MMV) could be useful to compare colour content of images in terms of the perceived colour names.

– **Tracking of regions** Colour-naming information could also be used for tracking of regions in image sequences. As in the previous applications, this task has been previously based on pixel distributions [9] but without taking into account how these pixels are perceived in terms of colour names. Tracking of regions based on colour naming would also allow better interaction with human users. Hence, for example, in a surveillance application it would be possible to define actions in natural language such as "follow the red car" or "find someone wearing a blue dress".

• **Colour naming and colour constancy**

It would be interesting to study the relationship between colour naming and colour constancy to evaluate up to what degree colour naming is affected by changes in the illuminant.

The data set obtained under different illuminants could be very useful for this purpose. As the inside of the booth where the experiment was developed was white we can assume that colour constancy mechanisms of the human visual system acted during the experiment. Hence, hypothetic differences in the colour-naming responses obtained under the different illuminants would indicate that human colour constancy not always is able to deal with them. Since some of the illuminants used are very extreme it would be interesting to study the limitations of human colour constancy.

Bibliography

[1] http://scien.stanford.edu/class/psych221/projects/05/joanmoh/index2.html. Last accessed on March 27, 2007.

[2] http://spectral.joensuu.fi/databases/index.html. Last accessed on March 27, 2007.

[3] http://www.cvc.uab.cat/color_naming. Last accessed on March 27, 2007.

[4] http://www.graphics.cornell.edu/online/measurements/sensors/index.html. Last accessed on March 27, 2007.

[5] http://www.gretagmacbeth.com. Last accessed on March 27, 2007.

[6] Munsell Book of Color - Matte Finish Collection. Munsell Color Company, Baltimore,MD, 1976.

[7] IEC 61966-2-1 / FDIS: Multimedia systems and equipment - Colour measurement and management - Part 2-1: Colour management - Default RGB colour space - sRGB, 1999.

[8] ISO/IEC 15938-3 / Information Technology - Multimedia content description interface - Part 3: Visual, 2002.

[9] D. Alexander. *Statistical Modelling of Colour Data and Model Selection for Region Tracking*. PhD thesis, Department of Computer Science, University College London, 1997.

[10] K. Barnard, V. Cardei, and B. Funt. A comparison of computational color constancy algorithms. I: Methodology and experiments with synthesized data. *IEEE Transactions on Image Processing*, 11(9):972–984, 2002.

[11] K. Barnard, L. Martin, A. Coath, and B. Funt. A comparison of computational color constancy algorithms. II experiments with image data. *IEEE Transactions on Image Processing*, 11(9):985–996, 2002.

[12] K. Barnard, L. Martin, B. Funt, and A. Coath. A data set for color research. *Color Research and Application*, 27(3):147–151, 2002.

[13] M. Belkhatir, P. Mulhem, and Y. Chiaramella. A full-text framework for the image retrieval signal/semantic integration. In K. V. Andersen, J. Debenham, and R. Wagner, editors, *Lecture Notes in Computer Science*, volume 3588, pages 113–123. Springer-Verlag, 2005.

[14] T. Belpaeme. Reaching coherent color categories through communication. In B. Kröse, M. De Rijke, G. Schreiber, and M. van Someren, editors, *Proceedings of the 13th Belgium-Netherlands Conference on Artificial Intelligence*, pages 41–48, 2001.

[15] T. Belpaeme. Simulating the formation of color categories. In B. Nebel, editor, *Proceedings of the International Joint Conference on Artificial Intelligence*, pages 393–398, 2001.

[16] R. Benavente. Dealing with colour variability: Application to a colour naming task. Master's thesis, Universitat Autònoma de Barcelona, Bellaterra, 2000.

[17] R. Benavente, R. Baldrich, M.C. Olivé, and M. Vanrell. Colour naming considering the colour variability problem. *Computación y Sistemas. Revista Iberoamericana de computación*, 4(1):30–43, 2000.

[18] R. Benavente, M.C. Olivé, M. Vanrell, and R. Baldrich. Colour perception: A simple method for colour naming. In *Proceedings of the 2nd Congrés Català d'Intel·ligència Artificial (CCIA'1999)*, pages 340–347, 1999.

[19] R. Benavente, G. Sànchez, R. Baldrich, M. Vanrell, and J. Lladós. Normalized colour segmentation for human appearance description. In *Proceedings of the 15th International Conference on Pattern Recognition (ICPR'2000)*, pages 637–641, September 2000.

[20] R. Benavente, F. Tous, R. Baldrich, and M. Vanrell. Statistical modelling of a colour naming space. In *Proceedings of the 1st European Conference on Color in Graphics, Imaging, and Vision (CGIV'2002)*, pages 406–411, Poitiers (France), April 2002.

[21] R. Benavente and M. Vanrell. A colour naming experiment. Technical Report 56, Computer Vision Center, Barcelona, September 2001.

[22] R. Benavente and M. Vanrell. Fuzzy colour naming based on sigmoid membership functions. In *Proceedings of the 2nd European Conference on Colour in Graphics, Imaging, and Vision (CGIV'2004)*, pages 135–139, Aachen (Germany), 2004.

[23] R. Benavente, M. Vanrell, and R. Baldrich. Estimation of fuzzy sets for computational colour categorization. *Color Research and Application*, 29(5):342–353, October 2004.

[24] R. Benavente, M. Vanrell, and R. Baldrich. A data set for fuzzy colour naming. *Color Research and Application*, 31(1):48–56, 2006.

[25] R. Benavente, M. Vanrell, and R. Baldrich. Parametric fuzzy sets for automatic colour naming. *Pattern Recognition*, (Submitted), 2007.

[26] T. Berk, L. Brownston, and A. Kaufman. A new color-naming system for graphics languages. *IEEE Computer Graphics and Applications*, 2(3):37–44, 1982.

[27] B. Berlin and P. Kay. *Basic Color Terms: Their Universality and Evolution*. University of California Press, Berkeley, CA, 1969.

[28] C.M. Bishop. *Pattern Recognition and Machine Learning*. Spinger, 2006.

[29] R.M. Boynton, L. Fargo, and B.L. Collins. Categorical color rendering of four common light sources. *Color Research and Application*, 15(4):222–230, 1990.

[30] R.M. Boynton, R.E. MacLaury, and K. Uchikawa. Centroids of color categories compared by two methods. *Color Research and Application*, 14(1):6–15, 1989.

[31] R.M. Boynton and C.X. Olson. Locating basic colors in the OSA space. *Color Research and Application*, 12(2):94–105, 1987.

[32] R.M. Boynton and C.X. Olson. Salience of chromatic basic color terms confirmed by three measures. *Vision Research*, 30(9):1311–1317, 1990.

[33] R.M. Boynton and K.F. Purl. Categorical colour perception under low-pressure sodium lighting with small amounts of added incandescent illumination. *Lighting Research and Technology*, 21(1):23–27, 1989.

[34] D.H. Brainard and B.A. Wandell. Analysis of the retinex theory of color vision. *Journal of the Optical Society of America*, 3(10):1651–1661, 1986.

[35] G. Buchsbaum. A spatial processor model for object colour perception. *Journal of the Franklin Institute*, 310:1–26, 1980.

[36] V. Cardei, B.V. Funt, and K. Barnard. Modeling color constancy with neural networks. In *Proceedings of the International Conference on Vision Recognition, Action: Neural Models of Mind and Machine*, Boston, 1997.

[37] V.C. Cardei. *A neural network approach to colour constancy*. PhD thesis, Simon Fraser University, 2000.

[38] T.F. Coleman and Y. Li. On the convergence of reflective newton methods for large-scale nonlinear minimization subject to bounds. *Mathematical Programming*, 67(2):189–224, 1994.

[39] T.F. Coleman and Y. Li. An interior, trust region approach for nonlinear minimization subject to bounds. *SIAM Journal of Optimiztion*, 6:418–445, 1996.

[40] M. Das, R. Manmatha, and E.M. Riseman. Indexing flowers by color names using domain knowledge-driven segmentation. In *Proceedings of the 4th IEEE Workshop on Applications of Computer Vision*, pages 94–99, 1998.

[41] R.L. De Valois, I. Abramov, and G.H. Jacobs. Analysis of response patterns of LGN cells. *Journal of the Optical Society of America*, 56(7):966–977, 1966.

[42] R.L. De Valois and K.K. De Valois. Neural coding of color. In E.D. Carterette and M.P. Friedman, editors, *Handbook od Perception*, volume 5, pages 117–166. Academic Press, 1975.

[43] R.L. De Valois and K.K. De Valois. A multi-stage color model. *Vision Research*, 33(8):1053–1065, 1993.

[44] R.O. Duda, P.E. Hart, and David G. Stork. *Pattern Classification*. Wiley-Interscience, second edition, 2001.

[45] A. Ebenhoh and H. Hemminger. Scaling of color sensation by magnitude estimation: A contribution to opponent-colors theory. *Biological Cybernetics*, 39(3):227–237, 1981.

[46] G.D. Finlayson. Color in perspective. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(10):1034–1038, 1996.

[47] G.D. Finlayson and M.S. Drew. Constrained least-squares regression in color spaces. *Journal of Electronic Imaging*, 6(4):484–493, 1997.

[48] G.D. Finlayson and M.S. Drew. White-point preserving color correction. In *Proceedings of the 5th Color Imaging Conference (CIC'97)*, pages 258–261, 1997.

[49] G.D. Finlayson, S.D. Hordley, and M.S. Drew. Removing shadows from images. In *Proceedings of the 7th European Conference on Computer Vision (ECCV'2002)*, pages 823–836, 2002.

[50] G.D. Finlayson, S.D. Hordley, and P.M. Hubel. Color by correlation: A simple, unifying framework for color constancy. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(11):1209–1221, 2001.

[51] G.D. Finlayson, S.D. Hordley, C. Lu, and M.S. Drew. On the removal of shadows from images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(1):59–68, 2006.

[52] G.D. Finlayson, S.D. Hordley, and I. Tastl. Gamut constrained illuminant estimation. *International Journal of Computer Vision*, 67(1):93–109, 2006.

[53] G.D. Finlayson, C. Lu, and M.S. Drew. Invariant image improvement by sRGB colour space sharpening. In *Proceedings of the 10th Congress of the International Colour Association (AIC'2005)*, 2005.

[54] D.A. Forsyth. A novel algorithm for color constancy. *International Journal of Computer Vision*, 5(1):5–36, 1990.

[55] C. Fredembach and G.D. Finlayson. Hamiltonian path based shadow removal. In *Proceedings of the 16th British Machine Vision Conference (BMVC'2005)*, 2005.

[56] B.V. Funt, K. Barnard, and L. Martin. Is machine colour constancy good enough? In *Proceedings of the 5th European Conference on Computer Vision (ECCV'1998)*, pages 445–459, 1998.

[57] G. Gagaudakis and P.L. Rosin. Incorporating shape into histograms for CBIR. *Pattern Recognition*, 35(1):81–91, 2002.

[58] K.R. Gegenfurtner. Color in the cortex revisited. *Nature neuroscience*, 4(4):339–340, 2001.

[59] K.R. Gegenfurtner and D.C. Kiper. Color vision. *Annual Review of Neuroscience*, 26:181–206, 2003.

[60] K.R. Gegenfurtner and L.T. Sharpe. *Color Vision. From Genes to Perception.* Cambridge University Press, 1999.

[61] J.M. Geusebroek, R. van den Boomgaard, A.W.M. Smeulders, and A. Dev. Color and scale: The spatial structure of color images. In *Proceedings of the 6th European Conference on Computer Vision*, pages 331–341, 2000.

[62] T. Gevers and A.W.M. Smeulders. Piktoseek: A content-based image search engine for the world wide web. In *Proceedings of VISUAL'97*, 1997.

[63] H. Grassman. Zur theorie der farbenmischung. *Poggendorf's Annalen Physik Chemie*, 89:69–84, 1853.

[64] W.C. Graustein. *Homogeneous Cartesian Coordinates. Linear Dependence of Points and Lines.*, chapter 3, pages 29–49. Introduction to Higher Geometry. Macmillan, New York, 1930.

[65] S. Guest and D. Van Laar. The structure of colour naming space. *Vision Research*, 40:723–734, 2000.

[66] S. Guest and D. Van Laar. The effect of name category and discriminability on the search characteristics of colour sets. *Perception*, 31:445–461, 2002.

[67] N. Hadjikhani, A.K. Liu, A.M. Dale, and P. Cavanagh. Retinotopy and color sensitivity in human visual cortical area V8. *Nature neuroscience*, 1(3):235–241, 1998.

[68] A. Hard and L. Sivik. NCS Natural Color System: a Swedish standard for color notation. *Color Research and Application*, 6(3):129–138, 1981.

[69] E. Hering. *Zur Lehre von Lichtsinne.* Carl Gerolds Sohn, Wien, 1878. (English translation: Outlines of a theory of the light sense. Translated by Hurvich, L.M. and Jameson, D., Cambridge, MA (USA), Harvard University Press).

[70] T. Hofmann. Probabilistic latent semantic indexing. In *Proceedings of the ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 50–57, 1999.

[71] S.D. Hordley and Finlayson. G.D. Re-evaluating colour constancy algorithm performance. *Journal of the Optical Society of America, A*, 23(5):1008–1020, 2006.

[72] P.M. Hubel, J. Holm, G.D. Finlayson, and M.S. Drew. Matrix calculations for digital photography. In *Proceedings of the 5th Color Imaging Conference: Color Science, Systems, and Applications*, pages 105–111. Society for Imaging Science and Technology, IS&T, 1997.

[73] L.M. Hurvich and D. Jameson. An opponent-process theory of color vision. *Psychological Review*, 64(6):384–404, 1957.

[74] K Jameson and R. D'Andrade. It's not really red, green, yellow, blue: An inquiry into color space. In C.L. Hardin and L. Maffi, editors, *Color categories in thought and language*, pages 295–319. Cambridge University Press, 1997.

[75] K.A. Jameson and N. Alvarado. Differences in color naming and color salience in vietnamese and english. *Color Research and Application*, 28(2):113–138, 2003.

[76] P. KaewTrakulPong and R. Bowden. A real time adaptive visual surveillance system for tracking low-resolution colour targets in dynamically changing scenes. *Image and Vision Computing*, 21(10):913–929, 2003.

[77] M. Karaman, L. Goldmann, and T. Sikora. A new segmentation approach using Gaussian color model and temporal information. In J.G. Apostolopoulos and A. Said, editors, *Proceedings of SPIE - Visual Communications and Image Processing*, volume 6077, 2006.

[78] P. Kay and C.K. McDaniel. The linguistic significance of the meanings of basic color terms. *Language*, 3(54):610–646, 1978.

[79] K.L. Kelly and D.B. Judd. The ISCC-NBS method of designating colors and a dictionary of color names. *National Bureau of Standards (USA)*, 553, 1955.

[80] K.L. Kelly and D.B. Judd. Color universal color language and dictionary of names. *National Bureau of Standards (USA)*, Spez. publ. 440, 1976.

[81] G.J. Klir and B. Yuan. *Fuzzy sets and fuzzy logic: theory and applications.* Pearson Education POD, Prentice Hall, Upper Saddle River, 1995.

[82] A. König and C. Dieterici. Die grundempfindungen und ihre intensitäts-vertheilung im spektrum. *Sitzungsberichte Akad. Wiss Berlin*, 2:805–829, 1886.

[83] J. Krauskopf, D.R. Williams, and D.W. Heeley. Cardinal directions of color space. *Vision Research*, 22(9):1123–1131, 1982.

[84] J.C. Lagarias, J.A. Reeds, M.H. Wright, and P.E. Wright. Convergence properties of the Nelder-Mead simplex method in low dimensions. *SIAM Journal of Optimization*, 9(1):112–147, 1998.

[85] J.M. Lammens. *A Computational Model of Color Perception and Color Naming.* PhD thesis, University of New York, 1994.

[86] E.H. Land. The retinex theory of colour vision. *Scientific American*, 237(6):108–129, 1977.

[87] P. Lennie. Single units and visual cortical organization. *Perception*, 27(8):889–935, 1998.

[88] M.D. Levine. *Vision in Man and Machine*. McGraw-Hill, 1985.

[89] H. Lin, M.R. Luo, L.W. MacDonald, and A.W.S. Tarrant. A cross-cultural colour-naming study. Part I: Using an unconstrained method. *Color Research and Application*, 26(1):40–60, 2001.

[90] H. Lin, M.R. Luo, L.W. MacDonald, and A.W.S. Tarrant. A cross-cultural colour-naming study. Part II - Using a constrained method. *Color Research and Application*, 26(3):193–208, 2001.

[91] H. Lin, M.R. Luo, L.W. MacDonald, and A.W.S. Tarrant. A cross-cultural colour-naming study. Part III - A colour-naming model. *Color Research and Application*, 26(4):270–277, 2001.

[92] Y. Linde, A. Buzo, and R.M. Gray. An algorithm for vector quantization design. *IEEE Transactions on Communications*, 28(1):84–95, 1980.

[93] Y. Liu, D. Zhang, G. Lu, and W. Ma. Region-based image retrieval with perceptual colors. In K. Aizawa, Y. Nakamura, and S. Satoh, editors, *Lecture Notes in Computer Science (Pacific-Rim Multimedia Conference)*, volume 3332, pages 931–938. Springer-Verlag, 2004.

[94] Y. Liu, D. Zhang, G. Lu, and W. Ma. Region-based image retrieval with high-level semantic color names. In *Proceedings of the 11th International Multimedia Modelling Conference*, pages 180–187, 2005.

[95] Y. Liu, D. Zhang, G. Lu, and W.Y. Ma. A survey of content-based image retrieval with high-level semantics. *Pattern Recognition*, 40(1):262–282, 2007.

[96] R.E. MacLaury. From brightness to hue: An explanatory model of color-category evolution. *Current Anthropology*, 33:137–186, 1992.

[97] A. Maerz and M.R. Paul. *A Dictionary of Color*. McGraw-Hill, 1st edition, 1930.

[98] L.T. Maloney and B.A. Wandell. Color constancy: a method for recovering surface spectral reflectance. *Journal of the Optical Society of America*, 3(1):29–33, January 1986.

[99] E. Marszalec and M. Pietikäinen. On-line color camera calibration. In *Proceedings of the 12th International Conference on Pattern Recognition (ICPR '94)*, volume 1, pages 232–237, 1994.

[100] D.J. McKeefry and S. Zeki. The position and topography of the human colour centre as revealed by functional magnetic resonance imaging. *Brain*, 120:2229–2242, 1997.

[101] S.J. McKenna, Y. Raja, and S. Gong. Tracking colour objects using adaptive mixture models. *Image and Vision Computing*, 17(3-4):225–231, 1999.

[102] G Menegaz, A. Le Troter, J. Sequeira, and J.M. Boi. A discrete model for color naming. *EURASIP Journal on Advances in Signal Processing*, 2007:Article ID 29125, 10 pages, 2007.

[103] A. Mojsilović. A computational model for color naming and describing color composition of images. *IEEE Transactions on Image Processing*, 14(5):690–699, 2005.

[104] A. Mojsilović, J. Gomes, and B. Rogowitz. Semantic-friendly indexing and quering of images based on the extraction of the objective semantic cues. *International Journal of Computer Vision*, 56(1-2):79–107, 2004.

[105] N. Moroney. Unconstrained web-based color naming experiment. In Reiner Eschbach and Gabriel G. Marcu, editors, *Color Imaging VIII: Processing, Hardcopy, and Applications. Proceedings of the SPIE*, volume 5008, pages 36–46. SPIE, 2003.

[106] W. Niblack, R. Barber, W. Equitz, M. Flickner, E. Glasman, D. Petkovic, P. Yanker, C. Faloutsos, and G. Taubin. The qbic project: Querying images by content using color, texture, and shape. In *Proceedings SPIE Storage and Retrieval for Image and Video Databases*, pages 173–181, 1993.

[107] D. Nickerson. OSA Color scale samples: A unique set. *Color Research and Application*, 6(1):7–33, 1981.

[108] D. Nickerson and S.M. Newhall. Central notations for ISCC-NBS color names. *Journal of the Optical Society of America*, 31:587–591, 1941.

[109] T. Nurminen, J. Kivinen, and P. Oittinen. Color naming and computational prediction from natural images. *Graphic Arts in Finland*, 35(2):13–28, 2006.

[110] V.E. Ogle and M. Stonebraker. Chabot: Retrieval from a relational database of images. *IEEE Computer*, 28(9):40–48, 1995.

[111] K. Okajima, A.R. Robertson, and G.H. Fielder. A quantitative network model for color categorization. *Color Research and Application*, 27(4):225–232, 2002.

[112] X. Otazu and M. Vanrell. Building perceived colour images. In *Proceedings of the 2nd European Conference on Colour in Graphics, Imaging, and Vision (CGIV'2004)*, pages 140–145, April 2004.

[113] X. Otazu and M. Vanrell. Decimated multiresolution framework with surround-induction function to unify assimilation and contrast mechanisms. Technical Report 99, Computer Vision Center, Bellaterra, 2005.

[114] X. Otazu and M. Vanrell. Representation of textured images. *Journal of Imaging Science and Technology*, 49(3):262–271, 2005.

[115] C.A. Parraga, R. Benavente, M. Vanrell, and R. Baldrich. Modelling colour naming space with fuzzy sets. In *30th European Conference on Visual Perception (ECVP'2007)*, 2007.

[116] A. Pentland, R.W. Picard, and S. Sclaroff. Photobook: Content-based manipulation of image databases. *International Journal of Computer Vision*, 18(3):3–14, 1996.

[117] M. Petrou, M. Mirmehdi, and M. Coors. Perceptual smoothing and segmentation of colour textures. In *Proceedings of the 5th European Conference on Computer Vision (ECCV'1998)*, pages 623–639, 1998.

[118] T. Regier, P. Kay, and R.S. Cook. Universal foci and varying boundaries in linguistic color categories. In B.G. Gara, L. Barsalou, and M. Bucciarelli, editors, *Proceedings of the 27th Meeting of the Cognitive Science Society*, pages 1827–1832, 2005.

[119] J. Ren, Y. Shen, and L. Guo. A novel image retrieval based on representative colors. In *Proceedings of the Image and Vision Computing New Zealand*, pages 102–107, 2003.

[120] D. Roberson, I. Davies, and J. Davidoff. Color categories are not universal: Replications and new evidence from a stone-age culture. *Journal of Experimental Psychology: General*, 129(3):369–398, 2000.

[121] S.J. Schein, R.T. Marrocco, and F.M. de Monasterio. Is there a high concentration of color-selective cells in area V4 of monkey visual cortex? *Journal of Neurophysiology*, 47(2):193–213, 1982.

[122] J.A. Schirillo. Tutorial on the importance of color in language and culture. *Color Research and Application*, 26(3):179–192, 2001.

[123] M. Seaborn, L. Hepplewhite, and J. Stonham. Pisaro: Perceptual colour and texture queries using stackable mosaics. In *Proceedings of the IEEE International Conference on Multimedia Computing and Systems*, volume 1, pages 171–176, 1999.

[124] N. Seaborn, L. Hepplewhite, and J. Stonham. Fuzzy colour category map for the measurement of colour similarity and dissimilarity. *Pattern Recognition*, 38(1):165–177, 2005.

[125] S.A. Shafer. Using color to separate reflection components. *Color Research and Application*, 10(4):210–218, 1985.

[126] A.W.M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain. Content-based image retrieval at the end of the early years. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(12):1349–1380, 2000.

[127] V.C. Smith and J. Pokorny. Spectral sensitivity of the foveal cone photopigments between 400 and 500nm. *Vision Research*, 15(2):161–171, 1975.

[128] R. Snowden, P. Thompson, and Tom Troscianko. *Basic Vision: An introduction to visual perception.* Oxford University Press, 2006.

[129] S.G. Solomon and P. Lennie. The machinery of colour vision. *Nature Reviews Neuroscience*, 8(4):276–286, 2007.

[130] J.M. Speigle and D.H. Brainard. Is color constancy task independent? In K. Braun and R. Eschbach, editors, *Recent progress in color science*, pages 328–332. The Society for Imaging Science and Technology, 1997.

[131] P.L. Stanchev, D. Green, and B. Dimitrov. High level color similarity retrieval. *International Journal Information Theories and Applications*, 10(3):363–369, 2003.

[132] A. Stockman and L.T. Sharpe. The spectral sensitivities of the middle- and long-wavelength-sensitive cones derived from measurements in observers of known genotype. *Vision Research*, 40(13):1711–1737, 2000.

[133] A. Stockman, L.T. Sharpe, and C. Fach. The spectral sensitivity of the human short-wavelength sensitive cones derived from thresholds and color matches. *Vision Research*, 39(17):2901–2927, 1999.

[134] M. Störring, T. Kočka, H.J. Andersen, and E. Granum. Tracking regions of human skin through illumination changes. *Pattern Recognition Letters*, 24(11):1715–1723, 2003.

[135] J. Sturges and T.W.A. Whitfield. Locating basic colors in the Munsell space. *Color Research and Application*, 20(6):364–376, 1995.

[136] J. Sturges and T.W.A. Whitfield. Salient features of Munsell color space as a function of monolexemic naming and response latencies. *Vision Research*, 37(3):307–313, 1997.

[137] S. Süsstrunk, R. Buckley, and S. Swen. Standard RGB color spaces. In *Proceedings on the 7th Color Imaging Conference: Color Science, Systems, and Applications*, pages 127–134, 1999.

[138] M.J. Swain and D.H. Ballard. Color indexing. *International Journal of Computer Vision*, 7(1):11–32, 1991.

[139] S. Tominaga. A color-naming method for computer color vision. In *Proceedings on the IEEE International Conference on Cybernetics and Society*, pages 573–577, 1985.

[140] F. Tous. *Computational framework for the white point interpretation based on colour matching.* PhD thesis, Universitat Autònoma de Barcelona, 2006.

[141] F. Tous, A. Borràs, R. Benavente, R. Baldrich, M. Vanrell, and J. Lladós. Textual descriptors for browsing people by visual appearance. In M.T. Escrig, F. Toledo, and E. Golobardes, editors, *Lecture Notes in Artificial Intelligence*, volume 2504, pages 419–429. Springer-Verlag, 2002.

[142] F. Tous, A. Borràs, R. Benavente, R. Baldrich, M. Vanrell, and J. Lladós. Textual descriptors for browsing people by visual appearance. In *Proceedings of the 5th Congrés Català d'Intel·ligència Artificial (CCIA '2002)*, pages 80–86, 2002.

[143] J.M. Troost and C.M. De Weert. Naming versus matching in color constancy. *Perception and psychophysics*, 50(6):591–602, 1991.

[144] R. Uchikawa and R.M. Boynton. Categorical color perception of Japanese observers: comparison with that of Americans. *Vision Research*, 27(10):1825–1833, 1987.

[145] J. van de Weijer, C. Schmid, and J. Verbeek. Learning color names from real-world images. In *Proceedings of the 2007 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2007.

[146] E.L. van den Broek, P.M.F. Kisters, and L.G. Vuurpijl. Content-based image retrieval benchmarking: Utilizing color categories and color distributions. *Journal of Imaging Science and Technology*, 49(3):293–301, 2005.

[147] E.L. van den Broek, E.M. van Rikxoort, and T.E. Schouten. Human-centered object-based image retrieval. In S. Singh et al., editor, *Lecture Notes in Computer Science (Advances in Pattern Recognition)*, volume 3687, pages 492–501. Springer-Verlag, 2005.

[148] M. Vanrell, R. Baldrich, X. Otazu, R. Benavente, A. Salvatella, F. Tous, and S. Alvarez. A five-stage model for computational colour-texture perception. In *First Iberian Congress on Perception*, Barcelona, 2005.

[149] M. Vanrell, R. Baldrich, A. Salvatella, R. Benavente, and F. Tous. Induction operators for a computational colour texture representation. *Computer Vision and Image Understanding*, 94(1):92–114, 2004.

[150] J.A.S. Viggiano. Comparison of the accuracy of different white balancing options as quantified by their color constancy. In M.M. Blouke, N. Sampat, and R.J. Motta, editors, *Proceedings of the SPIE - Sensors and Camera Systems for Scientific, Industrial, and Digital Photography Applications V*, volume 5301, pages 323–333, 2004.

[151] H. von Helmholtz. *Handbuch der Physiologischen Optik*. Voss, 1866.

[152] J.J. Vos and P.L. Walraven. On the derivation of the foveal receptor primaries. *Vision Research*, 11(8):799–818, 1971.

[153] B. Wandell and J.E. Farrell. Water into wine: Converting scanner RGB to tristimulus XYZ. In R.J. Motta and H.A. Berberian, editors, *Proceedings of the SPIE - Device-Independent Color Imaging and Imaging Systems Integration*, volume 1909, pages 92–101, 1993.

[154] Z. Wang, M.R. Luo, B. Kang, H. Choh, and C. Kim. An algorithm for categorising colours into universal colour names. In *Proceedings of the 3rd European Conference on Colour in Graphics, Imaging, and Vision*, pages 426–430. Society for Imaging Science and Technology (IS&T), 2006.

[155] B.R. Wooten. *The effects of simultaneous and successive chromatic contrast on spectral hue.* PhD thesis, Brown University, 1970.

[156] J.A. Worthey and M.H. Brill. Heuristic analysis of Von Kries color constancy. *Journal of the Optical Society of America*, 3(10):1708–1712, 1986.

[157] G. Wyszecki and W.S. Stiles. *Color Science: concepts and methods, quantitative data and formulae.* John Wiley & Sons, New York, second edition, 1982.

[158] H Yaguchi. Color categories in various color spaces. In *Proceedings of the 9th Color Imaging Conference*, pages 6–8. Society for Imaging Science and Technology, IS&T, 2001.

[159] S.N. Yendrikhovskij. A computational model of colour categorization. *Color Research and Application*, 26:S235–S238, 2001.

[160] S.N. Yendrikhovskij. Computing color categories from statistics of natural images. *Journal of Imaging Science and Technology*, 45(5):409–417, 2001.

[161] T. Young. On theory of light and colours. *Philosophical transactions of the Royal Society of London*, 92:12–48, 1802.

[162] S. Zeki. Color coding in rhesus monkey prestriate cortex. *Brain research*, 53(2):422–427, 1973.

[163] S. Zeki, D.J. McKeefry, A. Bartels, and R.S.J. Frackowiak. Has a new color area been discovered? *Nature neuroscience*, 1(5):335, 1998.

[164] S.Y. Zhu, M.R. Luo, and G.H. Cui. New experimental data for investigating uniform colour spaces. In *Proceedings of the SPIE - 9th Congress of the International Colour Association*, volume 4421, pages 626–629, 2002.

Publications

Publications of this thesis

JCR papers

- R. Benavente, M. Vanrell and R. Baldrich. Parametric Fuzzy Sets for Automatic Colour Naming. *Pattern Recognition*, Submitted 2007.

- R. Benavente, M. Vanrell and R. Baldrich. A Data Set for Fuzzy Colour Naming. *Color Research and Application*, 31(1):48-56, 2006.

- R. Benavente, M. Vanrell and R. Baldrich. Estimation of fuzzy sets for computational colour categorization. *Color Research and Application*, 29(5):342-353, 2004.

MSc Thesis

- R. Benavente. *Dealing with colour variability: Application to a colour naming task*. Master Thesis. Universitat Autònoma de Barcelona, Bellaterra, July 2000.

Conferences and other publications

- C.A. Parraga, R. Benavente, M.Vanrell, and R. Baldrich. Modelling Colour Naming Space with fuzzy sets. *30th European Conference on Visual Perception*, Arezzo (Italy), August, 2007.

- R. Benavente and M. Vanrell. Automatic colour naming based on fuzzy sets. J.Lladós (ed.), *Computer Vision: Progress of Research and Development. Proceedings of the 1st CVC Internal Workshop*, 1:42-47, Bellaterra, October, 2006.

- R. Benavente and M. Vanrell. Fuzzy colour naming based on sigmoid membership functions. *Proceedings of the Second European Conference on Color in Graphics, Imaging and Vision (CGIV'04)*, 1:135-139, Aachen(Germany), April 5-8, 2004.

- R. Benavente, F. Tous, R. Baldrich and M. Vanrell. Statistical modelling of a colour naming space. *Proceedings of the First European Conference on Color in Graphics, Imaging and Vision (CGIV'02)*, 1:406-411, Poitiers(France), April 2-5, 2002.

- R. Benavente and M. Vanrell. *A color naming experiment.* CVC technical report, number 56, Computer Vision Center, Bellaterra, September, 2001.

- R. Benavente, R. Baldrich, M.C. Olivé and M. Vanrell. Colour naming considering the colour variability problem. *Computación y Sistemas. Revista Iberoamericana de Computación*, 4(1):30-43, 2000.

- R. Benavente, M.C. Olivé, M. Vanrell and R. Baldrich. Colour perception: A simple method for colour naming. *Proceedings of the Segon Congrés Català d'Intel·ligència Artificial (CCIA'99)*, 1:340-347, Girona, October 25-27, 1999.

Publications related to this thesis

JCR papers

- M. Vanrell, R. Baldrich, A. Salvatella, R. Benavente and F. Tous. Induction operators for a computational colour texture representation. *Computer Vision and Image Understanding*, 94(1-3):92-114, 2004.

Conferences and other publications

- F. Tous, A. Borràs, R. Benavente, R. Baldrich, M. Vanrell and J. Lladós. Textual descriptors for browsing people by visual appearance. *Lecture Notes in Artificial Intelligence*, Springer-Verlag, 2504:419-429, October, 2002.

- M. Vanrell, R. Baldrich, X. Otazu, R. Benavente, A. Salvatella, F. Tous and S. Alvarez. A five-stage model for computational colour-texture perception. *First Iberian Congress on Perception*, Barcelona, July 6-8, 2005.

- R. Baldrich, M. Vanrell, A. Salvatella and R. Benavente. Color enhancement based on perceptual sharpening. *Proceedings of the IEEE International Conference on Image Processing (ICIP'03)*, 3:401-404, Barcelona, September 14-17, 2003.

- F. Tous, A. Borràs, R. Benavente, R. Baldrich, M. Vanrell and J. Lladós. Textual descriptors for browsing people by visual appearance. *Proccedings of the Cinquè Congrés Català d'Intel·ligència Artificial (CCIA'02)*, 1:80-86, Castelló de la Plana, October 24-25, 2002.

- R. Benavente, G. Sànchez, R. Baldrich, M. Vanrell and J. Lladós. Normalized colour segmentation for human appearance description. *Proceedings of the 15th International Conference on Pattern Recognition (ICPR'2000)*, 3:637-641, Barcelona, September 3-8, 2000.

Publications of other works in computer vision

JCR papers

- D. Ponsa, R. Benavente, F. Lumbreras, J. Martínez and X. Roca. Quality control of safety belts by machine inspection for real-time production. *Optical Engineering*, 42(4):1114-1120, 2003.

Conferences and other publications

- F. Lumbreras, X. Roca, D. Ponsa, R. Benavente, J. Martínez, S. Sànchez, C. Antens and J.J. Villanueva. Visual inspection of safety belts. *Proceedings of the Fifth Conference on Quality Control by Artificial Vision (QCAV'01)*, 2:526-531, Le Creusot-Burgundy(France), May 21-23, 2001.

- A. Pujol, J. Vitrià, P. Radeva, X. Binefa, R. Benavente, E. Valveny, C. Von Land. Real time pharmaceutical product recognition using color and shape indexing. *Proceedings of the Second Workshop on European Scientific and Industrial Collaboration on Promoting Advanced Technologies in Manufacturing (WESIC'99)*, 1:275-282, Newport(UK), September 1-3, 1999.

- J. Vitrià, P. Radeva, X. Binefa, A. Pujol, E. Valveny, R. Benavente and C. Von Land. *Real time recognition of pharmaceutical products by subspace methods.* CVC technical report, number 35, Computer Vision Center, Bellaterra, October, 1999.

- A. Martínez and R. Benavente. *The AR face database.* CVC technical report, number 24, Computer Vision Center, Bellaterra, June, 1998.