Contents

Chapter 1

Introduction

1.1 Thesis Context

1.2 Thesis Framework

Fig.1.1 shows the process schema of this thesis. It proposes an ideal way for image segmentation. Given an input image, at the end of the process we'll have an output image; the segmented image. For this goal, we are going two follow two diferent strategies; first one by using specular free images and second one a robustness strategy by using Illumination source Chromaticity (ISC) estimation. The specular free strategy is an ultra-fast method for image segmentation applayed in real time. The ISC is applied also in real time (20 f.p.s) and is robust against illumination changes. Both strategies have a step of preprocessing before the segmentation. Fig.1.1 shows the diagram flow between process, where each process is idenfied by a number too.

1.3 Thesis Outline

This thesis is otlined following the diagram flux drown on Fig.1.1.

Background chapter provides the basic physical and mathematical background. This chapter discuss about illumination, surface reflectance, color, color spaces (focusing on the RGB color space), reflection models (focusing on the dichromatic reflecction model (DRM), besides we'll see a spheric interpretation both of RGB and DRM corresponding to the process 1 .

Following chapters are performed by contributions which refer to process of the diagram flux.

Illumination Correction charpter is compound by four contributions. First one is a ISC estimation which correspond to the process 2 and 3. After that this charper has three works more. An evolutionary method for specular component

reduction and two methods for diffuse and specular components separation. First one, is grounded in a Bayesian approcah and second one is a geometrical approach which take profit of the pixels distribution within RGB under the DRM point of view. Last two contributions correspond with the fourth process on the flowchart.

Specular Free Images chapter deals of this kind of images. How applications can take profit of these images for image segmentation. In this chapter we are going to propose two new class of specular free images with its application in robotic context. Both contributions correspond with the process 8 and 9 of the flowchart.

Gradients chapter correspond with the process 5 of the flowchart. This chapter present an RGB color gradient following color constancy preservation.

Segmentation chapter correspond with the process 6 and 7 of the flowchart. This chapter has the main aplication of this thesis the segmentation and it has two contributions. On the one hand A Hybrid Color Distance for Image Segmentation. On the other hand a robust color Watershed transformation and image segmentation defined on RGB spherical coordinates.

1.4 Publications Along this Thesis

1. Z. Echegoyen, I. Villaverde, R. Moreno, M. Graña, and A. d'Anjou. Linked multi-component mobile robots: Modeling, simulation and control. Robotics and Autonomous Systems, 58(12):12921305, December 2010.

2. M. Grana, R. Moreno, and F. X. Albizuri. Convex coordinates based on lattice independent sets as pattern features. In Proc. IEEE Int Fuzzy Systems Conf, pages 225230, 2006.

3. Manuel Grana, Miguel A Veganzones, and Ramon Moreno. A remote mycological assistant. 4th IEEE Workshop on Intelligent Data Acquisition and Advanced Computing Systems Technology and Applications IDAACS, pages 408412, 2007.

4. Ramón Moreno Ivan Villaverde, Zelmar Echegoyen. Control visual para multi-robots, implementación de un aprueba de concepto de un srmc enlazado. In Actas de las III jornadas de Inteligencia Computacional. Servicio editorial UPV/EHU, 2009. ISBN: 978-84-9860-320-0.

5. Vassilis G. Kaburlasos and Gerhard X. Ritter, editors. Computational Intelligence Based on Lattice Theory, volume 67. Springer Berlin Heidelberg, Berlin, Heidelberg, 2007.

6. R. Moreno. Reectance analysis i. In A. Savio R. Moreno, editor, Actas de las II jornadas de Inteligencia Computacional, pages 258269. Servicio editorial UPV/EHU, 2008. ISBN 978-84-296-2316-1.

7. R. Moreno. Estimación de la cromacidad de la luz. In R. Moreno, editor, Actas de las III Jornadas de Inteligencia Computacional, pages 159170. Servicio editorial UPV/EHU, 2009. ISBN: 978-84-9860-320-0.

8. R. Moreno, M. Graña anda, and A. d'Anjou. An image color gradient preserving color constancy. In Fuzzy Systems (FUZZ), 2010 IEEE International Conference on, pages 1 5, july 2010.

9. R. Moreno, M. Graña, and E. Zulueta. Rgb colour gradient following colour constancy preservation. Electronics Letters, 46(13):908910, 2010.

10. Ramón Moreno. Inteligencia ambiental, sistemas ubícuos y visión por computador en telefonía móvil. estado del arte. In M.A. Veganzones R. Moreno, editor, Actas de las I jornadas de Inteligencia Computacional, pages 417 431. Servicio editorial UPV/EHU, 2007. ISBN 978-84-9860-019-3.

11. Ramón Moreno, Manuel Graña, and Alicia d'Anjou. Evolutive parametric approach for specular correction in the dichromatic reection model. In Hybrid Articial Intelligence Systems, volume 5271 of Lecture Notes in Computer Science, pages 665672. Springer Berlin / Heidelberg, 2008.

12. Ramón Moreno, Manuel Graña, and Alicia d'Anjou. A color transformation for robust detection of color landmarks in robotic contexts. In Trends in Practical Applications of Agents and Multiagent Systems, volume 71 of Advances in Soft Computing, pages 665672. Springer Berlin / Heidelberg, 2010.

13. Ramón Moreno, Manuel Graña, and Alicia d'Anjou. A geometrical method of diuse and specular image components separation. In Joan Cabestany, Ignacio Rojas, and Gonzalo Joya, editors, Advances in Computational Intelligence, volume 6692 of Lecture Notes in Computer Science, pages 8389. Springer Berlin / Heidelberg, 2011.

14. Ramón Moreno, Manuel Graña, and Alicia d'Anjou. A hybrid color distance for image segmentation. In Emilio Corchado, Marek Kurzynski, and MichalWozniak, editors, Hybrid Articial Intelligent Systems, volume 6679 of Lecture Notes in Computer Science, pages 447454. Springer Berlin / Heidelberg, 2011.

15. Ramón Moreno, Manuel Graña, and Alicia d'Anjou. Illumination source chromaticity estimation based on spherical coordinates in rgb. Electronics Letters, 47(1):2830, 2011.

16. Ramón Moreno, José López-Guede, and Alicia d'Anjou. Hybrid color space transformation to visualize color constancy. In Hybrid Articial Intelligence Systems, volume 6077 of Lecture Notes in Computer Science, pages 241 247. Springer Berlin / Heidelberg, 2010.

17. Dominik M. Ramík Kurosh Madani Ramón Moreno, Manuel Graña. Image segmentation by spherical coordinates. In Proceedings of the 11th International Conference on Pattern Recognition and Information Processing, pages 112115. Belarusian State University of Informatics and Radioelectrics, MAY 2011.

18. Manuel Graña Ramón Moreno, Alicia d'Anjou, and Carmen Hernandez. Bayesian reectance component separation. In Knowledge-Based and Intelligent Information and Engineering Systems, volume 5712 of Lecture Notes in Computer Science, pages 846852. Springer Berlin / Heidelberg, 2009.

19. R.Moreno. Visión por computador y clasicación aplicadas a la micología. In M. A. Veganzones R. Moreno, editor, Actas de las I jornadas de Inteligencia Computacional, pages 208224. Servicio editorial UPV/EHU, 2007. ISBN 978-84-9860-019-3.

20. Ivan Villaverde, Zelmar Echegoyen, Ramón Moreno, and Manuel Graña. Experiments on robotic multi-agent system for hose deployment and transportation. In Trends in Practical Applications of Agents and Multiagent Systems, volume 71 of Advances in Soft Computing, pages 573580. Springer Berlin / Heidelberg, 2010.

21. Ramon Moreno Lucile Rossi Kurosh Madani Manuel Graña. Véronique Amarger, Dominik M. Ramík. Wildland res' outlines extraction a spherical coordinates framed rgb color space dichromatic reection model based image segmentation approach. In Proceedings of the 11th International Conference on Pattern Recognition and Information Processing, pages 451454. Belarusian State University of Informatics and Radioelectrics, May 2011.

1.5 Symbols

Symbol	Description
θ	Zenith angle
ϕ	Azimuth angle
Ψ	Chromaticity
Π_Ψ	Chromatic plane, Maxwell plane
Λ	Diffuse chromaticity
Λ^{sf}	Diffuse specular free chromaticity
Γ	Specular chromaticity
I^{sf}	Specular free image
I^{est}	Illumination source chromaticity stimation
I^{norm}	Normalized image respect to the illumination
L_d	Diffuse line
L_s	Specular line
L_w	Achromatic line
Π_{dc}	Dichromatic plane

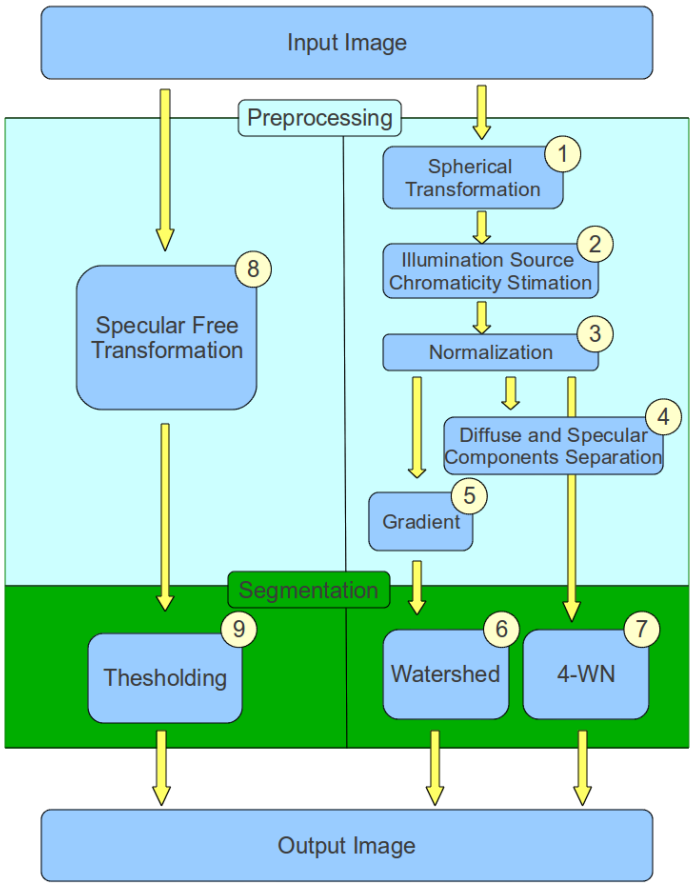


Figure 1.1: Main schema of the thesis

Chapter 2

Background

This chapter provides the basic mathematical and physical background concerning to this thesis. Sec.2.1 refers to illumination, the main factor on computer vision. On it depends the surface reflectance explained on Sec.2.2. After that we explain in Sec.2.3 what are color and color spaces, explaining in detail the RGB color space. Sub-sec.2.3.3 explains the "chromaticity" a key concept in this thesis. In Sec.2.4 we introduce a spheric interpretation of the RGB color space, where we provide a new equivalent definition of "chromaticity". After that in Sec. 2.5 we explain the sense of "color constancy" both for human vision and computer vision. Sec.2.6 discusses about models of reflectance, focusing on the reflection model where is grounded this thesis; the Dichromatic Reflection Model (DRM) in Sub-sec.2.6.1. Finally Sec.2.8 summarizes this chapter with the main ideas.

2.1 Illumination

Illumination is the main topic in vision systems, of it depends the full vision system. It is easy to understand which a scene without illumination turn useless any vision system, both animal and artificial. If the illumination is into the infrared spectrum, only few animals could see some things, and only artificial systems with sensors designed for this light wavelength interval are able to detect some things. In the same way with ultraviolet light, with red light or any single color light. The reference for vision systems is the human vision system (HVS), and the typical kind of illumination is the white illumination provided by the Sun. But this illumination is not a constant, natural illumination is different morning and evening, summer and spring, sunny days or days with clouds. Artificial illumination has strong changes in artificial environments; shops, pubs, city ways, town streets, industrial environments and then.

Literature is plenty of works which estimate illumination properties when this one belong to the visible spectrum. Hara [1] proposes a method for illumination source position and reflectance estimation from a single view without the

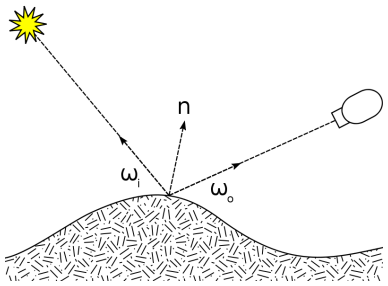


Figure 2.1: Surface reflectance by BRDF

distant illumination assumption. Sato [2] introduces a method for recovering an illumination distribution of a scene from image brightness inside shadows cast by an object of known shape in the scene. Other works are focused only in the estimation of the chromaticity of the illumination [3, 4, 5, 6].

2.2 Surface reflectance

There are many concepts related with the "reflectance", and some times exist ambiguity between them if it is studied from different sciences. There are different sciences which study the reflectance fenomenom; optics, physics, radiometry, astronomy, and recently computer sciences for visualization. We are interested from the point of view of computer vision.

When researching digital image processing, we can not ignore the physic effect of the reflectance, because it has all "a priory knowledge". Unfortunately the reflectance is different depending of the illumination and of the surface properties, and therefore this "a priory knowledge" generally is unknown. In computer vision, we refer to the reflectance phenomenon as the "surface reflectance".

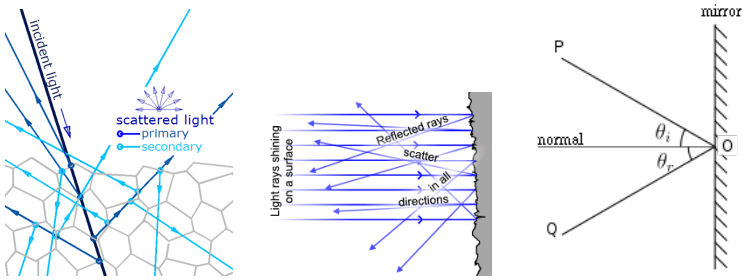
Strictly speaking, the measurement of light is a field in itself know as radiometry. A brief and clear introduction to radiometry is exposed by Forsyth and Ponce [7]. Fig. 2.1[1] shows the basic case of reflectance where a beam of light strikes a surface with an angle ω_1 respect to the normal angle n , and the outgoing beam leaves it with the same magnitude but at different direction ω_2. Here, we can formulate some interesting questions; the outgoing beam has the same energy and wavelength that the incoming beam? and the angle between ω_1 and n is the same that the angle between n and ω_2? The answer for both questions is no, it depends of the surface properties.

Generally some assumption have done respect to the surfaces for simplification; radiance leaving a point (irradiance) is due to the radiance arriving to this point, all light leaving at a wavelength is due to the incoming light to this wavelength (surfaces only can absorb some light at some wavelength) , and fluorescent surfaces are ignored.

[1] http://en.wikipedia.org/wiki/File:BRDF_Diagram.svg

Surfaces are classified in two main classes; specular surfaces and diffuse or lambertian surfaces. The first one, specular surfaces (after Latin word 'speculum', a mirror) has a behavior like a mirror; all incoming radiance leaves the surface without changes (ideally). The second one, diffuse surfaces or lambertian surfaces (after Lamber) have the property to absorb some of the incoming radiance, and the outgoing beam follows the Lambert cosine law [8]. In a diffuse surface we can detect the surface color and textures.

Fig.2.2 helps to understand these kind of refectances. Sketch 2.2(a)[2] shows a light beam arriving the surface and interacting within the matter. In this way, some wavelengths of light beam are absorbed by the matter. After that light beam leaves surface in a undermined direction. Sketch 2.2(b)[3] shows the behavior of a lot of light beams striking a surface, then the general reflection can be modeled by the Lambert cosin law. At last Sketch 2.2(c) [4] shows the behavior of a specular surface. In this case, the incoming light beam leaves the surface following the symmetrical geometry respect the normal angle of the strike point.



(a) (b) (c)

Figure 2.2: Diffuse and specular reflections

In nature most of the surfaces have a mixture of both reflectances due to the diversity of the surface composition, e.g. human skin could be matte, however grease and sweat can originate shines, and so on in the common surfaces. Usually surfaces has not only a component, the opposite is too must abundant in the nature. Vegetables has a wax layer and light can go trough them. Rocks and stones have different material composition with different reflectance properties.

Such an approximation of the surface reflectance, for computer vision and visualization are used *reflectance models* which try to explain the light behavior in different surfaces.

[2] http://en.wikipedia.org/wiki/File:Diffuse_reflection.gif
[3] http://en.wikipedia.org/wiki/File:Reflection_angles.svg
[4] http://en.wikipedia.org/wiki/File:Diffuse_reflection.PNG

2.3 Color and Color Spaces

Color is an important descriptor for object recognition[9], however is not a very accurate concept, because color is a human (and mamarian) ability and its knowingness is under current research [10]. Color is the visual perceptual property corresponding in humans to the categories called red, green, blue and others. Color derives from the spectrum of light interacting in the eye with the spectral sensitivities of the light receptors. Color categories and physical specifications of color are also associated with objects, materials, light sources, etc., based on their physical properties such as light absorption, reflection, or emission spectra. The perception of color stems from the varying spectral sensitivity of different types of cone cells in the retina to different parts of the spectrum, colors may be defined and quantified by the degree to which they stimulate these cells. These physical or physiological quantifications of color, however, do not fully explain the psychophysical perception of color appearance.

By defining a color space, colors can be identified numerically by their coordinates. Mainly there are two kind of color spaces. First one, the set of color spaces created by addition of primary colors (Red, Green and Blue) those spaces are named "additive color spaces" because all colors are represent as a sum of dimensional units, those colors are used for visualization using light, lamps, screens and so on. This set is compound by RGB, sRGB, RGBa, HSV,HSI, HSL, CIE l*a*b, CIE L*u*v color spaces and then. Second set of color spaces is performed by the addition of secondary color (Cyan, Magenta , Yellow and Black) those are named "subtractive color spaces" because they absorb the respective wavelength. These spaces are used in paints, inks and printing systems. The most famous is CMYK . There are also some specific color spaces for specific devices like NTCS for TV sets.

2.3.1 RGB

The first part of this thesis is grounded on the RGB color space [11]. Fig.2.3[5] shows the RGB cube in the range [0-255]. In general, RGB is defined on the Natural (discrete) range [0-255] because $2^8 = 256$ and first computers work with 8 bits providing enough quality for the human vision on screens. We are going to work in the Real (continuous) range [0-1] for a better precision. In this color space the three orthogonal lines represent the primary colors (Red, Green and Blue) which give the name to this space R-G-B. The other three orthogonal lines have the secondary colors (Cyan, Magenta and Yellow). These six lines begin on the origin [0,0,0] and finish on its respective extrema. The achromatic line begins on the black origin [0,0,0] and finishes on the white corner.

2.3.2 Color

Colorimetrically, color is compound by two concepts; Chromaticity and Intensity. Chromaticity has all chromatic information regardless to the intensity

[5] http://upload.wikimedia.org/wikipedia/commons/0/03/RGB_farbwuerfel.jpg

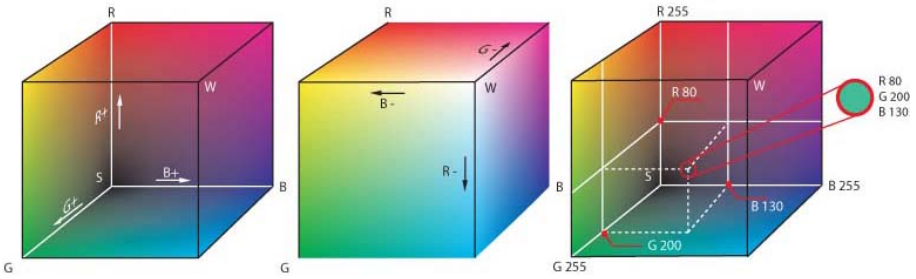


Figure 2.3: RGB color space in the range [0-255].

$$Color \begin{cases} Chromaticity & \begin{cases} Hue \\ Saturation \end{cases} \\ Intensity \end{cases}$$

Figure 2.4: Decomposition of color concepts

whereas intensity means the energy, the amount of electron by unit surface. Hence chromaticity is compound by Hue and Saturation. First one has the difference between color (perceived wavelength class) it detects the difference between red and blue (e.g). Second one represents the relative mixture with white. These ideas are drawn on Fig.2.4. If we interpret color on RGB like vectors, previous ideas can be explained as follow: Intensity is the vector length, Saturation is the distance with the achromatic line, and Hue is the rotational distance over the achromatic line beginning from the Red corner.

2.3.3 Chromaticity

Chromaticity is an objective specification of the quality of a color regardless of its luminance. In HSx family chromaticity is defined as the couple (H, S). In RGB Chromaticity is known as the normalized RGB; $r + g + b = 1$ and by dimensional reduction usually us used the couple (r, g) as the chromaticity where $b = 1 - r + g$. When normalizing an image by applying aforegoing equation, all color pixels are proyected into the Maxwell triangle [12]. This triangle is performed by the plane defined by the points $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$. We will name this plane chromatic plane Π_Ψ. Fig.2.5 shows these ideas.

Next section explain too, how to get the chromatic information of a RGB color through spherical coordinates representation.

2.4 Spherical coordinates

An image pixel's color corresponds to a point in the RGB color space $c = \{R_c, G_c, B_c\}$ The vector going from the origin up to this point shown in Fig.

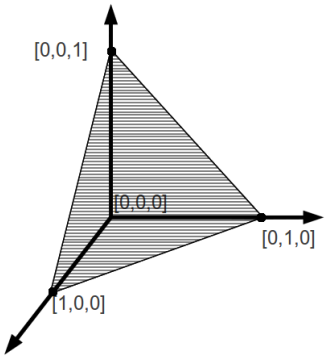


Figure 2.5: Maxwell triangle

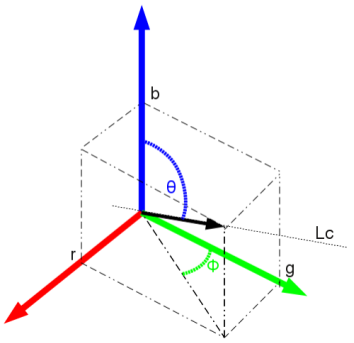


Figure 2.6: The vector corresponding to a color point in the RGB space

2.6 can be represented using spherical coordinates $c = \{\theta_c, \phi_c, l_c\}$, where θ is zenithal angle, ϕ is the azimuthal angle and l is the vector's magnitude.

In the RGB color space, chromaticity Ψ_c of a color point is represented by its normalized coordinates $r_c = \frac{R_c}{R_c+G_c+B_c}$, $g_c = \frac{G_c}{R_c+G_c+B_c}$, $b_c = \frac{B_c}{R_c+G_c+B_c}$, such that $r_c + g_c + b_c = 1$. That is, chromaticity corresponds to the projection on the chromatic plane Π_Ψ, defined by the collection of vertices of the RGB cube $\{(1,0,0),(0,1,0),(0,0,1)\}$, along the line defined as $L_c = \{y = k \cdot \Psi_c; k \in \mathbb{R}\}$. In other words, all the points in line L_c have the same chromaticity Ψ_c, which is a 2D representation equivalent to one provided by the zenithal and azimuthal angle components of the spherical coordinate representation of the a color point. Given an image $\mathbf{I}(x) = \left\{(R,G,B)_x ; x \in \mathbb{N}^2\right\}$, where x refers to the pixel coordinates in the image grid domain, we denote the corresponding spherical representation as $\mathbf{P}(x) = \left\{(\phi,\theta,l)_x; x \in \mathbb{N}^2\right\}$, which allows us to use $(\phi,\theta)_x$ as the chromaticity representation of the pixel's color.

2.5 Color Constancy

Color Constancy (CC) is the mental ability to identify chromatically homogeneous surfaces under illumination changes [13, 14]. This mental ability is still an open neupsicological research topic [10]. The CC property is inversely proportional to the color discontinuity represented by the chromatic edges (CE). Foster [15] speaks about CC, Relational Color Constancy and CE as different and complementary aspects under the neurological and retinal activity. CC is the perceptual mechanism which provides humans with color vision which is relatively independent of the spectral content of the illumination of a scene. It is the ability of a vision system to diminish or, in the ideal case, remove the effect of the illumination, and therefore "see" the true physical scene as the invariant to illumination changes.

Measurements on human subjects lead to the conclusion that retinal processing is not enough to extract chromatic features and chromatic based structural image information. Some works demonstrate that CC analysis is done in the visual cortex, in the areas V4 and V4A [13]. Assuming the analogy with the human vision biology, artificial vision systems need no trivial processing to ensure CC results on the processing real images. Dark scenes are critical for CC, because dark image regions are usually very noisy, that is, the signal to noise ratio is very high due to the low magnitude of the visual signal. In these regions, the ubiquitous thermodynamical noise has an amplified effect that distorts region and edge detection ensuring CC conditions. Our approach obtains remarkable good results in these critical regions.

In computer vision CC refers to the control of the illumination. Mainly, there are two ways to face it. On the one hand by automatic image transform like Spitzer [16] who proposes an algorithm which is based on retinal mechanisms of adaptation (gain control): 'local' and 'remote'. Geusebroek [17] proposes a physics-based method, valid for Lambertian reflectance where considering spatial and spectral derivatives of the image formation model, object reflectance properties are derived independent of the spectral energy distribution of the illuminant. On the other hand, a lot of methods estimate the illuminant chromaticity and after that normalize the image respect to the illumination, transforming the illumination to pure white. Yoon [18] uses the $r - g$ chromaticity and look for the chromatic lines, R. Tan [4] proposed a method by using the inverse intensity space and the Hough transform. Usually these methods are grounded on DRM [19]. Other approaches try to obtain segmentation procedures which are inherently robust to illumination effects [20, 21].

2.6 Models of Reflectance

A model of reflectance could be defined as "a mathematical model which try to explains the surface reflection of a specific kind of materials under a specific kind of illumination". It means that now a days we haven't got an universal reflection model that explains the reflection of all surfaces. Therefore there are a lot of

reflectance models who try to explain the reflectance phenomenon. Mainly we can group them in some sets:

- The used for image processing (photo retouching and color restoring, they are inspired in human vision)
- The used for visualization (games, virtual reality and synthetic images)
- The used for computer vision (we can extract true information from images)

The first set, is compound by reflectance models which are not based on physics or optics. These models are inspired in the human (or mammalian) perception, therefore they are based in the more recently ophthalmological and neural discoveries, focused on retinal behavior and visual cortex understanding. The widely used of this set is *Retinex* [22, 23] which is inspired in the human vision. It is inspired on the rods and cones retinal cells. Mainly it is used for color restoration, but it present some shortcomings [24]. The retinex algorithm is too sensitive to changes in the color of nearby objects to serve as an adequate of human color constancy. More recently, new models have been proposed in this line, like the *Neuromorphic* model introduced by Hong and Grossberg [25] which shows more sophisticated than retinex.

The second set is compound by all models used for visualization, that is for synthetic digital images. The widely used is the *Bidirectional Reflection Distribution Function BRDF* [26] which express the ratio between the irradiance and surface radiance depending of the point of view and direction of the illumination. As an improvement of this model Torrance & Sparrow [27] which add a pseudospheric globe on the specular component. A betterment of both models have been proposed by Ragheb and Edwin [28] who can simulate as wax (or glass) layer over surfaces. Besides of this models *Bidirectional Texture Functions BTF* are used to simulate textures [29, 30]. With these models (and similar models) objects can be drawn showing diffuse and specular reflections. However these models are not enough for to simulate all surfaces, e.g. human vision system is very exact beholding the human skin, indeed no one of the previous models are valid for simulate human faces. For this goal new models have been proposed. *Bidirectional Surface Scattering Distribution Function BSSRDF* [31] can simulate light behavior through epidermis. A well parametrized BSSRDF function have been introduced by Jensen & Donner[32] which can simulate all human skin kinds.

The third set is compound by all reflectance models who can be applied in computer vision. These models are grounded in physics measurements, hence we can extract some true conclusions. The first and widely used is *BRDF* [33, 34]. By observing a homogeneous surface, we can measure the BRDF reflection parameters using a scatterometer [35] even in the infrared [36]. Other important reflection model in this set is the *Dichromatic Reflection Model* (DRM). It was introduced by Shafer [37] and by difference with BRDF this one has better meaning within linear color spaces. It has been widely used to separate diffuse

and specular component, and for estimation of the illuminat chromaticity. An evolution of this model is the DRM under bi-illuminat conditions [38]. This model try to help to understand illumination changes on chromatic edges.

This thesis is grounded in DRM, then we are going to explain it in detail.

2.6.1 DRM

The Dichromatic Reflection Model (DRM) was introduced by Shafer [37]. It explains the perceived color intensity $I \in \mathbb{R}^3$ of each pixel in the image as the addition of two components, one diffuse component $D \in \mathbb{R}^3$and a specular component $S \in \mathbb{R}^3$. The diffuse component refers to the chromatic properties of the observed surface, while the specular component refers to the illumination color. Surface reflections are pixels with a high specular component. Therefore, if an image does not have a bright area were the specular component is strong relative to adjacent image regions, it is not possible to make Illumination Source Chromaticity (ISC) estimation by any means. The mathematical expression of the model, when we have only one surface color in the scene, is as follows:

$$\mathbf{I}(x) = m_d(x)\mathbf{D} + m_s(x)\mathbf{S}, \tag{2.1}$$

where m_d and m_s are weighting values for the diffuse and specular components.

Equivalently, Eq.2.1 can be expressed in spherical coordinates as:

$$\mathbf{I}(x) = (\theta_\mathbf{D}, \phi_\mathbf{D}, l_\mathbf{D}(x)) + (\theta_\mathbf{S}, \phi_\mathbf{S}, l_\mathbf{S}(x)),$$

where $\mathbf{\Lambda} = (\theta_\mathbf{D}, \phi_\mathbf{D})$ is the diffuse chromaticity,

$$l_\mathbf{D}(x) = \sqrt{(m_d(x)\mathbf{D}_R)^2 + (m_d(x)\mathbf{D}_G)^2 + (m_d(x)\mathbf{D}_B)^2},$$

and $\mathbf{\Gamma} = (\theta_\mathbf{S}, \phi_\mathbf{S})$ is the specular chromaticity and

$$l_\mathbf{S}(x) = \sqrt{(m_s(x)\mathbf{S}_R)^2 + (m_s(x)\mathbf{S}_G)^2 + (m_s(x)\mathbf{S}_B)^2}.$$

For a scene with several surface colors, the DRM equation must assume that the diffuse component may vary spatially, while the specular component is constant across the image domain:

$$\mathbf{I}(x) = m_d(x)\mathbf{D}(x) + m_s(x)\mathbf{S}, \tag{2.2}$$

that in spherical coordinates is expressed as:

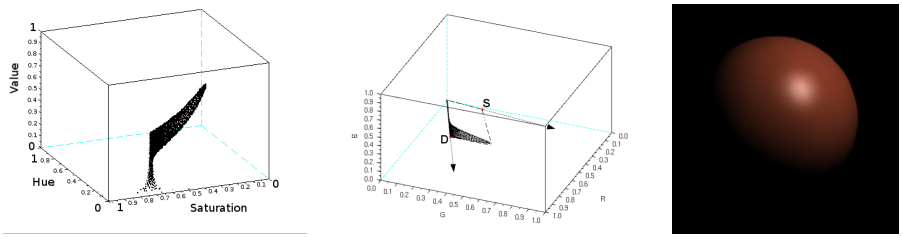
$$\mathbf{I}(x) = (\theta_\mathbf{D}(\mathbf{x}), \phi_\mathbf{D}(\mathbf{x}), l_\mathbf{D}(x)) + (\theta_\mathbf{S}, \phi_\mathbf{S}, l_\mathbf{S}(x)),$$

where $\mathbf{\Lambda}(x) = (\theta_\mathbf{D}(x), \phi_\mathbf{D}(x))$, and

$$l_\mathbf{D}(x) = \sqrt{(m_d(x)\mathbf{D}_R(x))^2 + (m_d(x)\mathbf{D}_G(x))^2 + (m_d(x)\mathbf{D}_B(x))^2},$$

and $\mathbf{\Gamma} = (\theta_\mathbf{S}, \phi_\mathbf{S})$ and

$$l_\mathbf{S}(x) = \sqrt{(m_s(x)\mathbf{S}_R(x))^2 + (m_s(x)\mathbf{S}_G(x))^2 + (m_s(x)\mathbf{S}_B(x))^2}.$$



(a) (b) (c)

Figure 2.8: Distribution of the ball image in the HSV color space (a) and in the RGB color space (b)

The chromaticity of the specular component $\theta_{\mathbf{S}}, \phi_{\mathbf{S}}$ is space invariant, meaning that the ISC is constant all over the scene. This is the most common situation in practice, where we have one colored illumination source irradiating over a scene with objects of different colors.

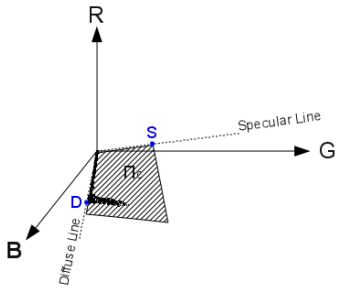


Figure 2.7: Dichromatic reflection model

RGB has some advantages versus HSx family when working with DRM. First, the DRM is defined as a vectorial sum in an euclidean space. This linearity exist in RGB however do not exist in other spaces like in the HSx color family. In Fig. 2.8(a) we can see the distribution of the pixels in the HSV color space. On the one hand pixels with low 'Value' are very separated each others, when 'Value' is increasing they are shaping in a line (the chromatic line). On the other hand specular pixels are in a curved shape like a horn. Comparing with the Fig.2.8(b) the linearity is lost. Hence is difficult to express DRM in HSx parameters. Fig.2.8(c) shows original image.

A digital image taken with a camera is defined on DRM as:

$$I(x) = w_d(x) \int_{\Omega} S(\lambda, x) E(\lambda) q(\lambda) d\lambda + w_s(x) \int_{\Omega} E(\lambda) q(\lambda) d\lambda \tag{2.3}$$

$$I(x) = w_d(x) B + w_s(x) G \tag{2.4}$$

where:
$I = \{I_r, I_g, I_b\}$ is the color of an image pixel obtained through a camera sensor.
$x = \{x, y\}$ are the two dimensional coordinates of the pixel in the image.
$q = \{q_r, q_g, q_b\}$ is the three element vector of sensor sensitivity.
$w_d(x)$ and $w_s(x)$ are the weighting factors for diffuse and specular components, respectively. They depend on the geometric structure at location x.
$S(\lambda, x)$ is the diffuse spectral reflectance.
$E(\lambda)$is the illumination spectral power distribution function, it is independent of the spatial location x when we assume an uniform illumination color.
The integration is done over the visible light spectrum Ω.

Eq.2.4 is equivalent to Eq.2.1 therefore the spheric transformation can be done in the same way.

2.7 Hyperspectral?

2.8 Chapter conclusion

In this chapter we have discussed about the main topics on the physical, cognitive and mathematical aspects respect to surface reflectance. Illumination is the main aspect, on it depends it. This reflectance is modeled in images by colors of whatever color space. We have explained in detail the RGB color space in which is grounded the first part of this thesis. Color is compound by different elements where chromaticity is a key concept explained accurately in RGB, both in euclidean and in spherical representation. Chromaticity is the main image feature which we'll use for CC in segmentation process. For a physical understanding of the surface phenomenon there are a lot of models of reflectance, where we are focused on DRM and its meaning on RGB color through a spherical interpretation. We advocate that the spherical interpretation of RGB is the best way to take profit of the DRM.

Part I

Contributions on RGB

Chapter 3

Illumination Correction

This chapter is outlined in the following manner: Sec.3.1 presents a review of the state of the art respect to the current advances on illumination correction, after that we will show the contributions in this area. Sec.3.2 corresponds with a method for Illumination source chromaticity estimation [3], this section corresponds with the process 2 and 3 of this thesis flowchart . The following sections are three manners to separate the diffuse and specular components each one with its advantages, all them corresponds with the 4 process of this thesis. First one, in Sec.3.3 we propose a way to reduce the shines using an evolutionary algorithm[39], Sec.3.4.2 shows a Bayesian approach [40] and Sec.3.5 describes a geometrical method [41] for separation of diffuse and specular component. At the end of this chapter we will give the chapter conclusions.

Note that the ISC contribution works with Spherical coordinates whereas the other contributions works with Euclidean coordinates.

3.1 State of the Art

Illumination information is contained within the specular component of the image, hence a key step is the detection and separation of the diffuse and specular components. Besides spots and shines are considered as a kind of noise or interferences in the image. Therefore the detection of the specular component is key in image processing and it covers two goals. On the one hand the specular component of an image contains all illumination information, it depends also of the geometrical properties of the image, then this information is used too for 'shape from shanding'[42, 43, 44]. On the other hand, once we know the specular component of an image, we can remove it from the original image obtaining then a diffuse image which has only the true chromatic surface properties.

The estimation of the illumination source chromaticity (ISC) [45] is a necessary step for color image normalization which is a critical step for constant color perception either in biological human perception [13, 10] or in the design of robust artificial vision systems [46]. Color normalization to a reference ISC,

usually white, allows the robust estimation of reflectance components, and subsequent segmentation of the image. Most ISC estimation algorithms [47, 45] work on the normalized RGB color space ($r + g + b = 1$).

There are works on the specular reflection reduction by using multi-flash mechanism [48] . The main drawback of this method is that it requires many images taken with flash devices. Other advance technics use only an image giving a reflectance estimation [1]. There are several works which are focused in the illumination chromaticity estimation (ISC) [5, 6, 4]. These methods are grounded in the DRM and only need one image. Kuk-jin method takes profit of the DRM and uses chromatic lines estimation by local ratios. Roby Tan method[4, 49] uses the inverse intensity space and by the Hough transform can estimate the illumination chromaticity. Ebner [6] estimates the illuminat chromaticity by segmentation and filtering looking for dichromatic lines by performing a linear regression on the x- and y-coordinates in CIE XYZ chromaticity space. All these methods which estimate the illuminant chromaticity are well-know as 'color constacy' methods, because once the illuminat chromaticity is estimated, original image can be normalized respect to the illumination, hence output image is independent of the illumination. The normalized image change the original illuminat to a pure white illumination.In this thesis we will propose a method for ISC estimation based on spheric coordinates [3] in Sec.3.2.

There are some works which separate the diffuse and specular components of the image [20, 41, 50, 51, 52, 53]. Mallick [20] presents a photometric stereo method for non-diffuse materials that does not require an explicit reflectance model or reference object. By computing a data-dependent rotation of RGB color space, the specular reflection effects can be separated from the much simpler, diffuse (approximately Lambertian) reflection effects for surfaces that can be modeled with dichromatic reflectance. Hui-Liang [50] proposes a method to separate reflections in a color image based on the error analysis of chromaticity and appropriate selection of body color for each pixel. By solving the least-squares problem of the dichromatic reflection model, reflection separation is implemented on a single pixel level, without requiring image segmentation and even local interactions between neighboring pixels. R. Tan method [51] is based solely on colors, particularly chromaticity, without requiring any geometrical information. One of the basic ideas is to iteratively compare the intensity logarithmic differentiation of an input image and its specular-free image. Umeyama [52] shows a method where diffuse and specular components of surface reflection can be separated as two independent components applying Independent Component Analysis to the images observed through a polarizer of different orientations. Kuk-jin method [54] proposes a specular-free two-band image that is a specularity-invariant color image representation and then reflection components separation is achieved by comparing local ratios at each pixel and making those ratios equal in an iterative framework.

Our contributions on this topic; separation of diffuse and specular components are carried out from three points of view. First one we propose an evolutive approach by using Legendre polynomials for specular reduction [39] in Sec.3.3. Second one we show a bayesian method supported by Random Markov Fields

[40]in Sec.3.4. Finally we propose a geometrical method grounded on DRM and on the spheric interpretation of the RGB color space [3] in Sec.3.5.

3.2 Illumination Source Chromaticity Estimation

Spherical coordinates in the RGB color space provide direct chromatic information. In this section we propose a novel method that uses this information for ISC estimation. We test our method on synthetic images whose ISC is known. This allows a quantitative comparison with a state of the art algorithm. This algorithm improves over the competing algorithm.

In previous works [55] we have already noted that the zenith θ and azimuth ϕ coordinates of a point in RGB space give the chromatic information of the corresponding color. Moreover, we were able to derive efficient chromatic gradients in color normalized images. Here we will use this information in a novel procedure to obtain an ISC estimation. This method does not need any previous image segmentation and its computational complexity is linear in the number of pixels, therefore it is suitable for real time.

3.2.1 Method

The proposed ISC estimation method is illustrated in Fig.3.1. The objects in the Fig.3.1(a) are labeled as (1), (2) and (3). These object labels are used in Fig.3.1(b),(c) and (d) for easy tracking the results of each step of the algorithm in each of the figures. The ISC information lies in the spherical coordinates of the pure specular pixels, according to the DRM of equation (2). The process' first step is the detection of the specular pixels in the image.

This detection is performed as follows: first we compute the specular free image [56], second, we compute the pixelwise difference between the specular free and original image intensities, third, we detect specular pixels setting a threshold on this difference image. The specular regions identified in Fig.3.1(a) are shown in Fig.3.1(b) as corresponding blobs.

The second step is to compute the spherical coordinates in RGB space of the specular pixels. We construct a 2D representation using the Zenith and Azimuth angle values. In this plane, the specular regions are represented by elongated shapes due to the transition between the pure ISC to the surface color that occurs in those regions. Fig.3.1(c) shows the Zenith-Azimuth representation corresponding to the specular regions in Fig.3.1(b). For each connected specular region we independently compute a linear regression. Fig.3.1(d) shows the regression lines computed for the region corresponding regions in Fig.3.1(c). The intersection between those lines corresponds to the estimation of the ISC's spherical coordinates $\theta_{\mathbf{S}}, \phi_{\mathbf{S}}$.

To obtain the corresponding ISC's normalized RGB coordinates, Ψ_{ISC} we compute the intersection of the line determined by $\theta_{\mathbf{S}}, \phi_{\mathbf{S}}$ with the chromatic plane Π_{Ψ}. If we want to perform the image color normalization to the pure white illuminant, it suffices to compute the differences $\theta_{\mathbf{S}} = \theta_{\mathbf{S}} - \frac{\pi}{4}$ and $\phi_{\mathbf{S}} = \phi_{\mathbf{S}} - \frac{\pi}{4}$. Applying this correction to the spherical coordinates of the pixels in the image

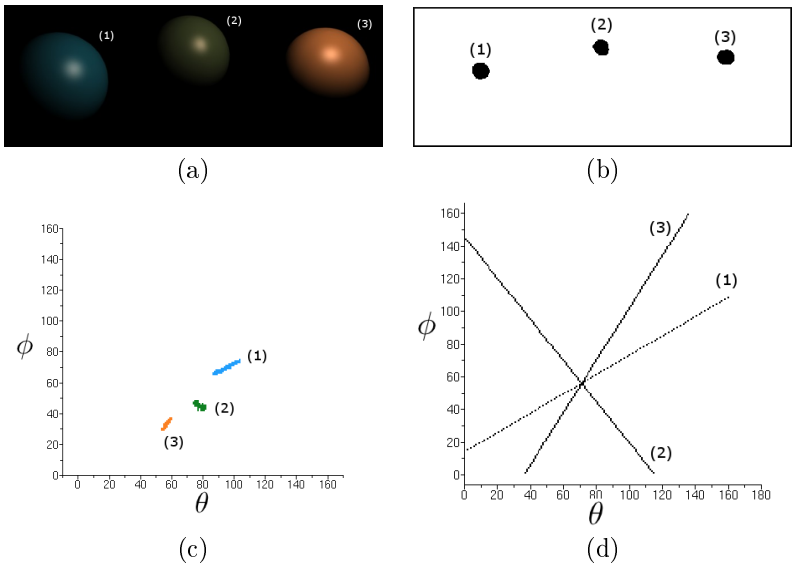


Figure 3.1: (a) Original Synthetic Image, (b) Specular regions detected in the original image, (c) distribution of spherical coordinates of specular pixels, (d) linear regressions of each specular region and ISC detected as their intersection.

we obtain the image colors under pure white ISC. An Scilab implementation of the algorithm is available at http://www.ehu.es/ccwintco/index.php/GIC-source-code-free-libre.

In summary, the proposed method's steps are the following: (1) Compute the specular free image, (2) perform detection of specular regions, extracting their pixels, (3) plot them in the zenith-azimuth plane, (4) compute the linear interpolation for each region, (5) estimate the intersection of the interpolation lines.

3.2.2 Experimental Results

We apply the proposed approach and Tan's method [4] to synthetic images generated under a known ISC. Therefore, we can compute the squared error of the estimations given by the algorithms to achieve a quantitative comparison of both approaches. Fig.3.1(a) shows an instance of the synthetic test images.

The experimental images have been generated as follows:

1. We selected three ISC values (left column in Table 3.1) as the means of the Gaussian distribution of ISC, with standard deviation $\sigma = 005$.

2. We generated 30 samples of these ISC Gaussian distributions

3. We generated the synthetic images using a common reflectance image

	Proposed Method		R. Tan Method	
Ψ_{ISC}	$\widetilde{\Psi}_{ISC}$	$\bar{e}$	$\widetilde{\Psi}_{ISC}$	$\bar{e}$
$r = .28$	$r = .2684$	.010790	$r = .2664$	.052410
$g = .32$	$g = .3102$		$g = .3403$	
$b = .40$	$b = .4212$		$b = .3733$	
$r = .32$	$r = .3301$	.006410	$r = .3232$	.052410
$g = .40$	$g = .4061$		$g = .4118$	
$b = .38$	$b = .2637$		$b = .2627$	
$r = .40$	$r = .3824$	.005964	$r = .4062$	.008008
$g = .32$	$g = .3246$		$g = .3321$	
$b = 28$	$b = .2929$		$b = .2595$	

Table 3.1: Experimental ISC, estimated vales of the normalized RGB and the estimation errors.

4. We apply both our approach and the competing method to estimate the ISC from the synthetic images.

5. Finally, we computed the estimation error for each image and give separate mean estimation errors for each separate Gaussian distribution. The error is computed as the angle between the true and the estimated ISC divided by the maximum possible error $\frac{\pi}{2}$. Table 3.1 contains the average ISC estimation $\widetilde{\Psi}_{ISC}$ and the mean error $\bar{e}$ for each method. It can be appreciated that our approach improves always over Tan's method.

3.2.3 Work Conclusions

We present a method for Illumination Source Chromaticity (ISC) estimation that is accurate and its complexity is linear in the number of pixels, therefore it is suitable for real time. It works on the spherical coordinates of the specular pixels of the image, which is a small fraction of the whole image. It does not lose luminosity information, because it is preserved in the magnitude component of the spherical representation which is not affected by color normalization. It can be applied in real time. We have shown in computational experiments that it improves over state of the art competing algorithms.

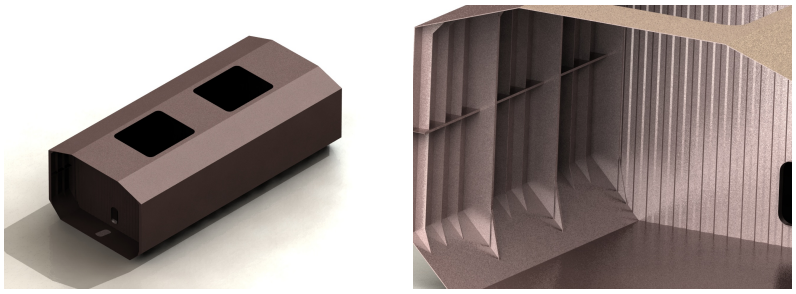


Figure 3.2: Synthetic images of a ship hold. The images used in the experiments were taken from the upper aperture of the hold.

3.3 Evolutive Parametric Approach for Specular Correction

Assuming the dichromatic image model we propose a global reduction of specularity effects by means of parametric illumination gradient images obtained by fitting 2D Legendre polynomials to the specular component of the images. Fitting is done applying an Evolution Strategy. The method could be applied to static robotic monitoring in teams of robots, where the illumination gradient image could be computed once and applied to successive frames until the illumination conditions change drastically. The method could be useful for the detection of image regions with different chromatic properties.

The framework of this contribution is the work on the design of multirobot systems for highly unstructured environments, such as shipyards. There a potentially critical role is that of observer or monitoring, that is a member of the team located in a position where it can monitor all the environment and serve this information to the remaining members of the team. This robot will be static once it has reached the surveillance position, so that images will be relatively static also. That means that illumination conditions will vary slowly and it is possible to perform illumination correction under real time constraints. The Fig.3.2 shows two views of a ship hold that illustrate the point of view of the images used in the experiments reported below. Naturally, the kind of images obtained contain a lot of speculaties, because of the presence of water and metal surfaces. Moreover, the illumination will change continuously (slowly) due to changes in natural illumination.

Legendre polynomials have applied successfully to intensity inhomogeneity correction in MRI [57], as a parametric model of the inhomogeneity field that can be estimated by an energy minimization method like the Evolution Strategies (ES). Our approach mimics that one, applying it to the dichromatic model. We obtain a specularity bias that can be used to normalize the images easing further segmentation and detection processes. Due to the slow change in natural illumination, we expect that the estimated bias would valid for several frames, reducing the time constraint for real life application.

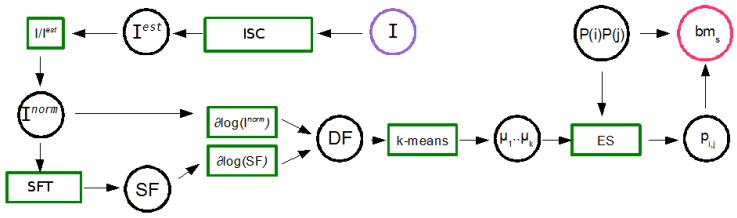


Figure 3.3: Flow diagram of the process of estimation of bm_s.

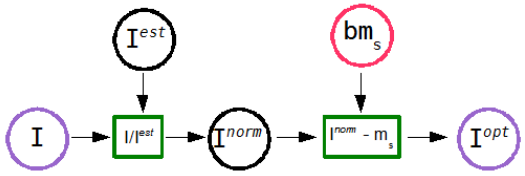


Figure 3.4: Illumination correction using the specular field bm_s.

3.3.1 Description of the approach

Fig.3.3 shows a flow diagram describing the process followed to obtain an specularity field. The starting point is the captured image I that need to be normalized in chromaticity. We use ISC estimation to obtain the illuminant chromaticity estimation I^{est} and the normalized image is given by the ratio $I^{norm} = \frac{I}{I^{est}}$. It is also needed to obtain in parallel an Specular-Free image [58] (SF), we use the Specular Free Two Band method [54]. Then we obtain the derivatives of the logarithm of SF and I^{norm}, to obtain the diffuse pixels. From them we select the most representative k classes $\{\mu_1, ..., \mu_k\}$, corresponding to chromatic regions in the image. Finally, the ES estimates the parameters of the 2D Legendre polynomials that gives the estimation of bm_s. Fig.3.4 shows the correction of the image removing the specular field.

3.3.2 Normalization

In some methods before separating specular and diffuse components it is necessary to carry out a process of normalization, because they need that the specular component must be pure white. This process requires the value of I^{est} which can be obtained by some methods [5, 59, 6, 4, 47]. It is computed as $I^{norm} = \frac{I}{I^{est}}$. After normalization, the image change its specular component Γ for pure white $[1, 1, 1]$ that is $I^{norm}(x) = m'_d(x)\Lambda'(x) + m_s(x)[1, 1, 1]$

$$I^{est} \backsimeq \Gamma \Rightarrow \frac{\Gamma}{I^{est}} \simeq [1, 1, 1] \Rightarrow$$

$$I^{norm}(x) = \frac{m_d(x)\Lambda(x) + m_s(x)\Gamma}{I^{est}} = \frac{m_d(x)\Lambda(x)}{I^{est}} + \frac{m_s(x)\Gamma}{I^{est}} \Rightarrow$$

$$\frac{m_d(x)\Lambda(x)}{I^{est}} + m_s(x)\frac{\Gamma}{I^{est}} = \frac{m_d(x)\Lambda(x)}{I^{est}} + m_s(x)[1,1,1]$$

where

$$\frac{m_d(x)\Lambda(x)}{I^{est}} = m'_d(x)\Lambda'(x)$$

then

$$I^{norm}(x) = m'_d(x)\Lambda'(x) + m_s(x)[1,1,1]$$

when working in normalized RGB

$$I^{norm}(x) = m'_d(x)\Lambda'(x) + m_s(x)/3$$

3.3.3 Specular-Free Image

The specular free image is critical for the detection of the reflectance component, we used the Specular-Free Two-Band method proposed by [5]. The process is simple: it subtracts to each pixel its minimum band. The image is geometrically identical to the original: $I^{sf}(x) = m'_d(x)\Lambda^{sf}(x)$. An specular free image, is a geometrical transformation of a image such, the returned image hasn't got specular component, it means $m_s = 0$.

3.3.4 Intensity Logarithmic Differentiation

This technique allows the detection of pure diffuse pixels $m_s = 0$, and, as a consequence, the specular pixels. Pure diffuse pixels allows us estimate the chromaticity of the surface. Assuming uniform colour pixels (Λ becomes independent from x) applying the logarithm and spatial differentiation

$$\frac{\partial}{\partial x}\log(I^{norm}(x)) = \frac{\partial}{\partial x}\log\left(m'_d(x)\Lambda'(x) + \frac{m_s(x)}{3}\right)$$

For diffuse pixels $m_s = 0$, then $\frac{\partial}{\partial x}\log(I^{norm}(x)) = \frac{\partial}{\partial x}\log\left(m'_d(x)\Lambda'(x)\right)$. Assuming which two neighboring pixels has the same diffuse chromaticity, Λ' don't depend of the spatial coordinate x, then

$$\frac{\partial}{\partial x}\log(I^{norm}(x)) = \frac{\partial}{\partial x}\log\left(m'_d(x)\Lambda'\right) = \frac{\partial}{\partial x}\left(\log\left(m'_d(x)\right) + \log\left(\Lambda'\right)\right) = \frac{\partial}{\partial x}\log\left(m'_d(x)\right)$$

For the Specular-Free image we have that

$$\frac{\partial}{\partial x}\log\left(I^{sf}(x)\right) = \frac{\partial}{\partial x}\log\left(m'_d(x)\Lambda^{sf}(x)\right) = \frac{\partial}{\partial x}\left(\log\left(m'_d(x)\Lambda^{sf}(x)\right)\right)$$

$$\frac{\partial}{\partial x}\log\left(I^{sf}(x)\right) = \frac{\partial}{\partial x}\left(\log\left(m'_d(x) + \Lambda^{sf}(x)\right)\right)$$

Assuming which two neighboring pixels has the same diffuse chromaticity, Λ^{sf} don't depend of the spatial coordinate x, then

$$\frac{\partial}{\partial x}\log\left(I^{sf}(x)\right) = \frac{\partial}{\partial x}\left(\log\left(m'_d(x) + \Lambda^{sf}\right)\right) = \frac{\partial}{\partial x}\log\left(m'_d(x)\right)$$

Therefore, the test for a diffuse pixel is that

$$\Delta(x) = \frac{\partial}{\partial x}\log(I^{norm}(x)) - \frac{\partial}{\partial x}\log\left(I^{sf}(x)\right)$$

If $\Delta(x) = 0$ then x is a diffuse pixel.

3.3.5 Evolutionary Strategy

We are adapting the ideas about intensity inhomogeneity correction in MRI [57] to our problem. The first step is to propose an energy function whose minimization would solve our problem. This energy function comes from the dichromatic model

$$E_{tot} = \sum_{x \in I^{norm}} \left(I^{norm}(x) - bm_s(x,p) - \mu_k(x)\right)^2 \tag{3.1}$$

where the specular field is given by

$$bm_s(x,p) = \sum_{i=0}^{l}\sum_{j=0}^{l} p_{i,j}P(i)P(j)$$

and the $P(i)P(j)$ denote products of 1D Legendre polynomials.

The parameters of the minimization are the Legendre polynomials linear coefficients. The energy includes some class representatives μ_k of the image reflectance given by the representatives obtained form the diffuse pixels by some clustering process (i.e. k-means). Image pixels must be classified according to the classes before performing the bias search.

To search for the optimal parameters of the Legendre 2D field, we use an -ES. Each individual in the ES is a matrix of coefficients of the Legendre polynomials. They are mutated by random Gaussian values (still we have not implemented any adaptive search method modifying their variances). The starting point is population of 50 individuals, which are the seeds for the process. We select the best 20 in an elitist selection process. They are the parents for the next generation.

If we introduce the dichromatic model into the energy function of Eq.3.1, we have

$$E(x) = (m'_d(x)\Lambda'(x) + m'_s(x) - bm_s(x,p) - \mu k)^2$$

so that

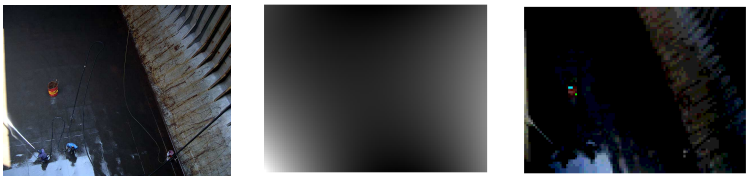


Figure 3.5: From left to right: Original image, the estimated specular bias composed of polynomials of degree up to 2, and the corrected image obtained removing the specular bias

$$E(x) = (m'_d(x)\Lambda'(x) - \mu_k)^2$$

and the energy is proportional to the diffuse component of the image:

$$E(x) \approxeq (m'_d(x))^2$$

3.3.6 Experimental Results

Experiments have been performed using Scilab and the SIP toolbox. The ES always converge to good solutions.

Fig.3.10 shows the original full image of the ship hold being watered for cleaning. It also shows the recovered image when the higher degree of the Legendre polynomials composing the bias is 2. Fig.3.6 shows the bias and the recovered image when the higher polynomial degree is 3. These figures illustrate how increasing the model order allowing higher degree polynomials the estimated bias tends to fit also the variations in reflectance. Lower order models are desired to obtain more robust estimations of the illumination. Fig.3.7 and 3.8 show the effect of the algorithm on the region containing the images of the human operators.

The resulting corrected image is, of course, more dark than the original image, but it retains all the geometric information, which can be observed computing the spatial gradient of the images, which we are not including for lack of space. The effect on the hold floor is that we get a constant intensity (color) image of it. The almost specular region on the lower left corner is greatly enhanced, making it more similar to the remaining floor surface. One of the goals of this work is to obtain robust segmentations of objects lying in the surface of the ship hold. It can be appreciated in Fig.3.7 and 3.8 that the human operators are easily segmented despite the increased darkness of the image.

3.3.7 Work Conclusions

This work shows an image correction approach based on the modeling of the specular component of the image as bias field obtained as the linear composition of 2D Legendre polynomials. Model fitting is done by a $(\lambda - \mu) - ES$ on the

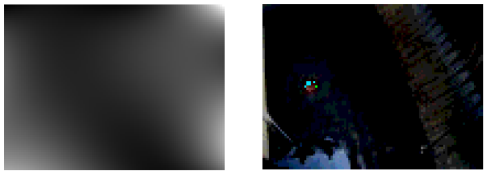


Figure 3.6: From left to right: The estimated specular bias composed of polynomials of degree up to 3, and the corrected image obtained removing the specular bias

Figure 3.7: Detail from figure 4 images. From left to right: original image, estimated specular bias, corrected image obtained removing the specular bias.

Figure 3.8: Detail from figure 5 images. From left to right: original image, estimated specular bias, corrected image obtained removing the specular bias.

space of the linear coefficients of the Legendre polynomials. The approach is based on the dichromatic reflectance model. The effect is that we can remove strong specularity effects from difficult scenes, such as the ship hold treated as example, allowing more robust segmentation of objects in the image.

3.4 Bayesian Reflectance Component Separation

We work on a Bayesian approach to the estimation of the specular component of a color image, based on DRM. The separation of diffuse and specular components is important for color image segmentation, to allow the segmentation algorithms to work on the best estimation of the reflectance of the scene. In this work we postulate a prior and likelihood energies that model the reflectance estimation process. Minimization of the posterior energy gives the desired reflectance estimation. The approach includes the illumination color normalization and the computation of a specular free image to test the pure diffuse reflection hypothesis.

Works on reflectance map estimation usually need to impose some assumptions like the knowledge of a color segmentation of the image, the detection of color region boundaries or color discontinuities, or the knowledge of the decomposition into linear basis functions of the surface color. The approach presented here does not impose any such assumption and does not need previous segmentations of the image. We follow a Bayesian approach [60] to model the desired result as constraints implemented in an *a priori* distribution. We postulate the *a priori* distribution based on the idea developed in [58] that the derivatives of the logarithmic images of both diffuse image and specular free must be equal in order to have pure diffuse pixels.

3.4.1 Separation Method

We will base our Bayesian model in the derivative of the logarithm of the normalized image respect to ISC and the specular free transform[4, 58]. As in the aforegoing section, the pure diffuse pixels can be characterized by the following relation:

$$\triangle(x) = dlog(\mathbf{I}'(x)) - dlog(\mathbf{I}^{sf}(x)) = 0, \tag{3.2}$$

where $dlog(\mathbf{I}^{sf}(x)) = \frac{\partial}{\partial x} log(\mathbf{I}^{sf}(x))$ and $dlog(\mathbf{I}'(x)) = \frac{\partial}{\partial x} log(\mathbf{I}'(x))$, the logarithm is computed pixel wise, and the spatial derivative can be computed in several ways, for instance in [**?**] it is computed on the scalar value image given by the summation of the three channels. It can be easily verified that $dlog(\mathbf{I}'(x)) = \frac{\partial}{\partial x} log(m'_d(x)) = dlog(\mathbf{I}^{sf}(x))$ for pure diffuse pixels if the diffuse chromaticity of neighboring pixels is the same. That means that the method works well inside homogeneous color regions, and needs the estimation of color region boundaries. When $\triangle(x) > 0$ in Eq. 3.2 and the pixel is not at a color boundary and a pure specular pixel, then it has some specular component that can be removed to get the diffuse reflectance component. The method proposed in [58] follows from an heuristic observation about the distribution of pixels in the maximum chromaticity versus (normalized illumination color) intensity space. Non diffuse pixels are decreased in intensity iteratively to search for the pure diffuse pixel value.

3.4.2 Bayesian Modeling

Given an image f and a desired unknown response of a computational process d, Bayesian reasoning gives, as the estimate of d, the image which maximizes the *A Posteriori* distribution $P(d|f) \propto e^{-U(d|f)}$, where the *A Posteriori* energy can be decomposed in to the *A Priori* $U(d)$ and Likelihood (Conditional) $U(f|d)$ energies $U(d|f) = U(f|d) + U(d)$. The Maximum A Posteriori (MAP) estimate is equivalent minimize the posterior energy function

$$d^* = \arg\min_d U(d|f) \tag{3.3}$$

The Likelihood energy $U(f|d)$ measures the cost caused by the discrepancy between the input image f and the solution d. The A Priori energy $U(d)$ is a model of the desired solution, usually built as a Random Markov Field (RMF), so that the A Priori energy can be built up as the summation of the local energies at the pixels, which are expressed as summations over the set of cliques including the pixel, weighted by the local potential parameter. A Priori energy usually incorporates any desired constraint, such as smoothness, into the model.

We will assume a Gaussian Likelihood distribution plus a Chromaticity preservation constraint, therefore the Likelihood energy will have the following expression:

$$U(d|f) = \sum_{i=1}^{m} \frac{(f_i - d_i)^2}{2\sigma^2} + \sum_{i=1}^{m} \left(\Psi_i^f - \Psi_i^d\right)^2,$$

where f_i and d_i are the RGB pixel values a the i-th pixel position for the observed and desired image, respectively. Also, Ψ_i^f and Ψ_i^d denote the chromaticity pixels of the observed and desired image, respectively.

The A Priori energy is built up from two components. The first one is the Chromaticity continuity:

$$U_\Psi(d) = \sum_{i=1}^{m} \sum_{j \in N_i} \sum_{c \in \{r,g,b\}} \left(\Psi_{i,c}^d - \Psi_{j,c}^d\right)^2.$$

The second modeling the estimation of the derivatives in Eq. **??** as the cliques of the RMF. That is, we assume that the local energy at pixel d_i is defined as

$$U_\triangle(d_i) = \left(dlog(d_i) - dlog(d_i^{sf})\right)^2,$$

where d_i^{sf} is the i-th pixel of the specular free image, computed as described above, and $dlog(.)$ in means the local estimation of the derivative, which is approximated as follows:

$$dlog(d_i) = \frac{1}{\#N} \sum_{j \in N_i} \log(\frac{I(x_j)}{I(x_i)}),$$

where N_i is the local neighborhood of pixel d_i, and $\#N$ is its cardinality. After some manipulations, the local derivative component of the A Priori energy is derived as:

$$U_{\triangle}(d_i) = \left(\sum_{j \in N_i} \sum_{c \in \{r,g,b\}} \log \frac{d_{j,c} d^{sf}_{i,c}}{d_{i,c} d^{sf}_{j,c}} \right)^2 .$$

This local energy is equivalent to the Kuk-Jin ratio criterion [54]. The derivative component of the A Priori energy is, therefore, the addition of these local energies:

$$U_{\triangle}(d) = \sum_{i=1}^{m} U(d_i),$$

and the A Priori energy is given by the addition $U(d) = U_{\triangle}(d) + U_{\Psi}(d)$.

3.4.3 Experimental results

In this section we report some experimental results applying the Bayesian approach described above. The starting value for the energy minimization process is set to $f = d(0) = \mathbf{I}'$. Each iteration step of the energy minimization involves the computation of the specular free image $d^{sf}(t)$ of the current hypothesis $d(t)$ of the optimal estimation d^*. Instead of using a Monte Carlo minimization technique [60], such as Simulated Annealing, we have employed a simple heuristic to determine the new hypothesis $d(t+1)$, consisting in the reduction of the intensity of the pixels preserving their chromaticity components relative ratios. Although simple, this strategy does in fact produce a minimization of the energy function, as can be appreciated in Fig.3.9, where we plot an instance of the energy function evolution. We have tested our approach on some images already tested by some authors in the literature i.e. [49, 58] among others. Fig.3.10 shows the result over a well known test image with two colors and two light sources. Our algorithm does not include any modeling of the underlying color regions in the scene, such as in [49], so it can be appreciated that the almost pure specular pixels can not be corrected, because there almost no chromatic information left in them. To improve our approach we will be including a color map field in the model, to be able to assign those pixels the most likely color. The Fig.3.11 shows a complex geometry image. Our estimation of the diffuse reflectance component recovers the underlying geometry, with some blurring effects.

3.4.4 Conclusions

We have presented a Bayesian approach to the problem of reflection component separation. As in previous works, our approach works with only one image [58] and does not need any additional assumption, such as models of the colors in scene o previous color segmentations of the image. We compute the specular free image, which can be done on the fly for each hypothesis. We have tested the approach applying a simple heuristic to provide new hypothesis from the

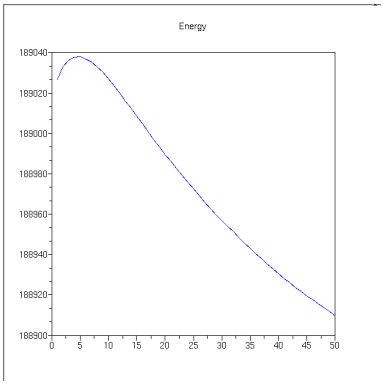


Figure 3.9: Evolution of the energy function in an instance run of the algorithm

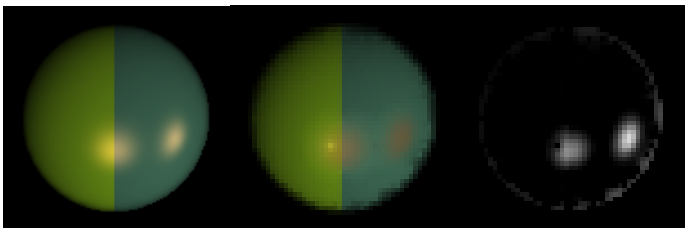


Figure 3.10: From left to righ, the original image, the estimated diffuse reflection component, and the estimated specular component

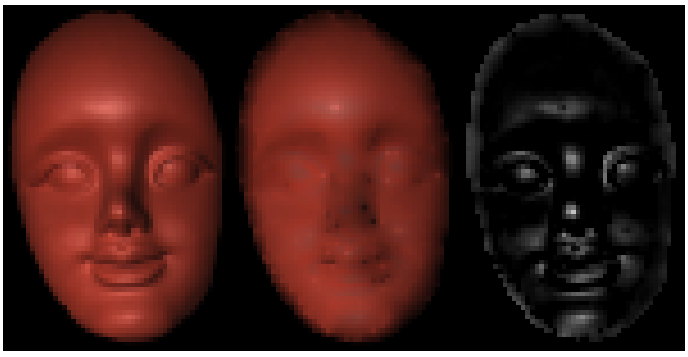


Figure 3.11: From left to right, the original image, the estimated diffuse reflection component, and the estimated specular component

previous iteration, with quite encouraging results. From the experiments we detect the need to incorporate a color map field in the *A Priori* model, so that the color of almost purely specular pixels can be recovered more easily. The problem of diverse color illumination sources will be dealt with in further works.

3.5 A geometrical method of diffuse and specular image components separation

The approach presented here is based on observed properties of the distribution of pixel colors in the RGB cube according to DRM. We estimate the lines in the RGB cube corresponding to the diffuse and specular chromaticities. Then the specular component is easily removed by projection on the diffuse chromaticity line. The specular component is computed by a straightforward difference. The proposed algorithm does not need any additional information besides the image under study.

The diffuse component estimation is useful for color based artificial vision processes, while the specular component contains surface topological information and is required for the estimation of reflectance maps. Recent solutions [61, 42] require uniform illumination and the identification of constant color regions, working on synthetic "clean" images. Color constancy analysis is a requirement for such algorithms. Our approach does not need such analysis. The process is as follows: first we estimate the chromatic lines, then we perform a dychromatization process, we estimate the diffuse image component, then we compute the specular image component. We show some computational results on well known benchmark images.

3.5.1 Pixels distribution on RGB color space

Following DRM we can classify image pixels into: Diffuse pixels: showing the observed surface color, with an almost null specular component. Specular pixels: whose specular component is much bigger than the diffuse component.

Placement of diffuse and specular pixels is qualitative different in the RGB cube. Let us focus on the proximity of pixels to the black-white cube diagonal, defined as $L_w : r, g, b = P + su; \forall s \in \mathbb{R}$ where $P = [0, 0, 0]$ and $u = [1, 1, 1]$. Given a uniform color region, without any specular component, its representation in the RGB cube would be a line, the diffuse chromaticity line for this region. However, due to noise, it appears as an elongated point cloud.

Given a uniform color region, with high specular component, from the DRM point of view, it must appear like a line parallel to line L_w or approaching it. Again, due to noise, an elongated pint cloud appears. Specular image regions have RGB representations far from the color space origin.

Finally, a uniform color region (color constancy) with some non negligible specular component must show a V shape. The point cloud beginning in the coordinate origin and go away from line L_w contain the diffuse points, while the ones close to it are the specular ones. Using this knowledge, we can penalize the specular component and magnify the diffuse component.

3.5.2 General description of the method

We assume that the observed surface is decomposable into patches of homogeneous chromatic characteristics. The proposed method has the following phases:

1. Chromatic line estimation: estimate the diffuse line L_d and the specular line L_s .
2. Dichromatization: We compute the parameters of the chromatic plane Π_{dc} in the RGB cube, and we project all the pixel colors into this plane. This step involves some additive noise removal.
3. Component separation: We compute the pure diffuse image component and the specular image component.

3.5.3 Chromatic line estimation

In figure 3.12 we have a plot of the pixels in the image of figure 3.13 in the three-dimensional RGB cube. Let us denote them $\{I_i; i = 1, ..., M\}$. We can easily appreciate the two main directions in the data. The most clear is the one corresponding to the diffuse line L_d which rises from the coordinate system origin. The second, less defined, appearing at the end of the diffuse elongation, is the specular direction identified by the specular line L_s.

To estimate the diffuse line, we start selecting the less bright pixels in the image region corresponding to the surface, which will have the greatest diffuse component. We plot them in the RGB cube and we estimate the best linear regression on the RGB data. In fact, we perform a Principal Component Analysis [**?**] (PCA) which give us the direction of the chromatic line $\overrightarrow{u}$. Therefore the diffuse chromatic line is defined as $L_d : (r, g, b) = P + s\overrightarrow{u}; \forall s \in \mathbb{R}$. Analogously, we select the brightest pixels, obtaining a mean point Q in the RGB cube and the largest eigenvector $\overrightarrow{v}$ for the specular color, therefore the specular chromaticity line is expressed as follows $L_s : (r, g, b) = Q + t\overrightarrow{v}; \forall t \in \mathbb{R}$.

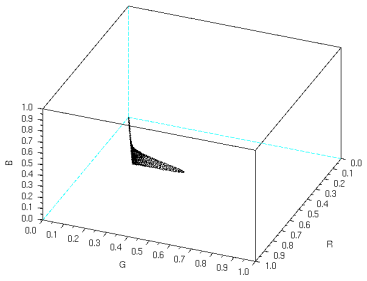


Figure 3.12: Synthetic image plotted in the three-dimensional RGB space

3.5.4 Image dichromatic regularization

Once we know the chromatic lines, we build the dichromatic plane Π_{dc} in $\mathbb{R}^3$ which is the best planar approximation to the color distribution in RGB. It can

be expressed as follows: $\Pi_{dc} : (r, g, b) = P + s\overrightarrow{u} + t\overrightarrow{v}; \forall s, t \in \mathbb{R}$, and the normal vector is $\overrightarrow{N} : \overrightarrow{u} \times \overrightarrow{v}$, where $\times$ denotes the conventional vector product. To remove noise and regularize the image colors we project the pixel's colors into this dichromatic plane Π_{dc} . For each image point color in the RGB cube I_i we compute the line $L_i : (r, g, b) = I_i + k\vec{N}; \forall k \in \mathbb{R}$, which is orthogonal to the dichromatic plane Π_{dc}, and to regularize I_i we compute its projection I_i^c as the intersection of L_i with Π_{dc}.

3.5.5 Component separation

Recalling the DRM definition $I(x) = m_d(x)D + m_s(x)S$ our goal is to bring the pixels to the chromatic line, that is $\forall x : m_s(x) = 0$. We proceed as follows: for each regularized image point I_i^c lying in the plane Π_{dc} we draw the line $L_i : (r, g, b) = I_i^c + t\overrightarrow{v}; \forall t \in \mathbb{R}$ where $\overrightarrow{v}$ is the specular line vector director. The pixel diffuse component corresponds to the intersection point I_i^d of this line with the diffuse line $L_d : (r, g, b) = P + s\overrightarrow{u}; \forall s \in \mathbb{R}$ and it exists because they lie in the same plane Π_{dc} and they are not parallel lines. We have obtained $I^d(x) = m_d(x)D$ so that $\forall x, : m(x) = 0$, and the resulting image $I^d(x)$ is purely diffuse, without specular components. Obtaining the specular image component is then trivial if we recall the DRM definition: $I^s(x) = I(x) - I^d(x) = I(x) - m_d(x)D = m_s(x)S$.

3.5.6 Experimental results

The experimental demonstration of our approach is shown in figures 3.13 and3.14. The first is a synthetic image (using Blender), and the second is a natural image. Both are monochromatic. Original image is the leftmost image in both figures. Following our approach we obtain the diffuse and specular images, shown at the center and rightmost images, respectively, in both figures. Both original images can be downloaded from http://www.ehu.es/ccwintco/index.php/Images. The natural image has been used as benchmark by several researchers [58, 49]. The visual results are comparable or better than the state of the art results in [58, 49]. As we know the original surface color ($r = 0.790$, $g = 0.347$ and $b = 0.221$) in the synthetic image, we can compute an estimation of the error committed by our estimation of the diffuse image. If we denote Q the original color, the error is the distance of this point to the diffuse line, computed as $d(Q, L_d) = ||\vec{PQ} - \perp(\vec{PQ}, \vec{u})||$, where $\perp(\vec{a}, \vec{b})$ denotes the projection operator. In the images shown in figure 3.13 the error committed is 0.0116. Variations in the error are due to the diffuse region pixel selection.

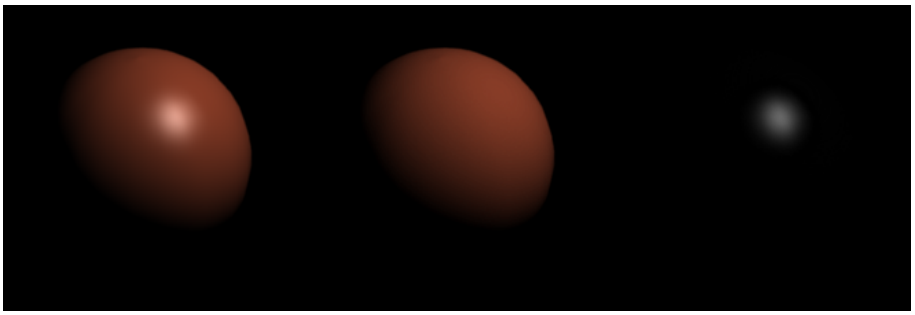


Figure 3.13: Synthetic image, diffuse image and specular image

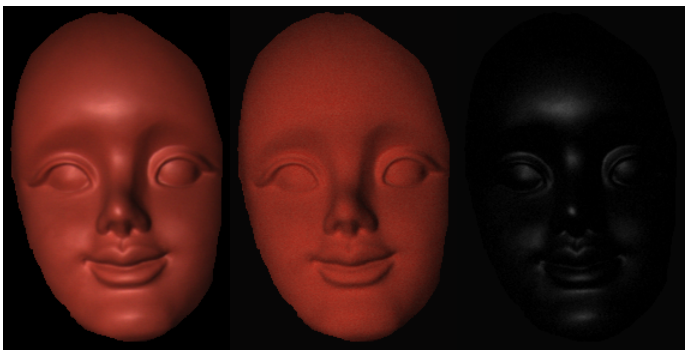


Figure 3.14: Natural image, diffuse image and specular image

3.5.7 Work Conclusions

As conclusion of this work, we have described an image component separation for mono-color images which is very effective, fast and robust. It has been developed from the DRM and is well theoretically grounded despite its simplicity. It consists in the estimation of the diffuse and specular lines as the principal components of diffuse and specular point clouds, respectively, selected from the image by hand. Contrary to other approaches [48, 52] our approach does not need specific hardware devices, and only needs one image. Our approach's complexity time is linear in the image size $O(M)$, while others [54, 58] are quadratic $O(M^2)$. Our approach does not need a Specular Free image, it provides almost simultaneously both image components. On going work is addressing the extension of this approach to images containing several surface colors, i.e. $I(x) = m_d(x)D + m_s(x)S$, and to images with illumination sources of different colors, i.e. $I(x) = m_d(x)D + m_s(x)S(x)$.

3.5.8 Chapter conclusions

This chapter shows the importance of illumination correction in computer vision systems. First we have propose a method for ISC estimation that is accurate and its complexity is linear in the number of pixels, therefore it is suitable for real time. It is the first step in our segmentation model that let us normalize the images respect to the illumination chromaticity obtaining robustness respect to illumination changes. After that we have proposed three technics for illumination correction. First one, an evolutive strategy for specular component reduction with chromaticity preservation, which let us to obtain good results in segmentation. By other hand we have propose two manners to separate the specular and the diffuse components from a single image. A Bayesian method which does not need any additional assumption, such as models of the colors in scene o previous color segmentations of the image. It converges linearly by intensity reduction in the ISC line. Finally we have propose a geometrical method for separate the diffuse and specular components. This method is faster that other ones and provides excellent results. A further work is to extend this model to multi-chromatic surfaces.

Chapter 4

Specular Free Images

In this chapter cover the process 8 and 9 of the thesis flowchart , where process 8 correspond with the Specular Free image and the process 9 its application for image segmentation, and it is outlined as follow: Sec.4.1 discuss about the state of the art, Sec.4.2 present the first contribution and Sec.4.3 presents the second contribution. The chapter finish in Sec.4.4 with the chapter conclusions.

4.1 State of the Art

There are few works which discusses about specular free images (SF). These were introduced by Roby Tan [49] and Kuk-jin Yoon [54] more or less at the same time. In the first case, the basic idea is to make the saturation constant for all surface colors while retaining their hue values. In other words, if these images preserve their hue and saturation, then preserve the chromaticity, therefore they change only the intensity respect to the original image. In the case of Kuk-Jin technique, the transformation project all RGB points within the nearest plane using the achromatic direction. In both cases, specular free transformations are geometrical and injective transformations within a color space (RGB) where the output image, the specular free image has not specular component.

4.1.0.1 SF2 and SF3

We are going to present two new kind of specular free images and its applications. First one, Sec.4.2 [62] introduces a SF transformation (SF2) based on the pseudo-norm of the saturation. This work is applied to landmark detection in robotic context, illumination correction (as a reduction of specular component) and in robot detection in real time. Second one, Sec.4.3 shows other SF transformation (SF3) which is based in an angular definition of the saturation following DRM behavior in RGB. SF3 applies the intensity depending of the pixel chromaticity. This SF image is applied on natural and synthetic images. Both SF transformations uses the HSV color space for a better understanding.

4.2 A color transformation for robust detection of color landmarks in robotic contexts

This work presents a robust color transformation which has been applied successfully to natural scenes allowing the fast and precise segmentation of regions corresponding to color landmarks under uncontrolled lightning. The process is grounded in the DRM and the properties of the RGB space.

Robust and fast detection of color regions is one of the typical artificial vision problems. Given our color perception, color clustering is not an appropriate approach most of the times. Among the various color spaces, the HSV and CIE L*a*b are the ones closest to human perception.

The need to detect color regions stems from its conventional use in signaling: red for danger, blue and green for informative, yellow for danger advice. Also Red, Green and Blue are the basic colors in the RGB space unit cube. All the remaining colors are represented as linear combinations of these colors. In robotic contexts, working on artificial environments, we must benefit from this information source by the robust detection of signaling symbols drawn in the basic colors.

A critical problem is removing the reflections in the image, which interfere with the observed surface. The two goals of the color image processes are identified as: efficient color detection and reflection removal.

In the following, Sec. 4.2.1 presents a brief explanation of the DRM model and its justification in the RGB space. We present our method in Sec. 4.2.2. We present some experimental results in Sec. 4.2.4. We give our conclusions and further work lines in Sec. 4.3.3.1

4.2.1 Some properties of the RGB cube and the DRM

Human chromatic perception is the result of biological evolution along millions of years. The mental interpretation of colors is subject to subjective aspects: philosophical, cultural and evolution. We can say that the human beings have developed individual color perception traits. However we have a consensus on the basic color interpretation which is represented in the color space used for their representation. The HSV color space is one that matches the human perception better than the RGB space. The pair Hue-Saturation defines the chromatic space, while V is the light intensity. The most used color space is RGB. From the computational point of view, and the artificial vision one, the RGB space has the following interesting properties:

1. It is the default color representation space for all the machines, from perception (Bayer's mosaic) up to the monitor visualization.

2. The vertices of the unit RGB cube represent the primary colors (red, blued, green), the secondary colors (yellow, cyan, magenta) and the black and white colors. The ones most used in signalization.

3. The reflections or brights are characterized in the RGB cube for its proximity to the black-white diagonal.

Algorithm 4.1 Code for SF2

```
//I is a RGB image
// IR is the transformed image
Function IR = SF2(I)
New_Intensity = (max(I,3) - min(I,3));
Imghsv = rgb2hsv(I);
Imghsv(:,:,3) = New_Intensity;
IR = rgb2hsv(Imghsv);
Endfunction
```

4. DRM has been defined in the RGB space.

As we explain in the background chapter, one of the main properties of DRM on RGB is that chromatically homogeneous and diffuse regions are collinear within RGB as we can see in Fig. . Therefore, we prefer to work with RGB than with other color spaces.

4.2.2 Method

Being interested in pure color regions, we expect their color representation in the RGB cube far from the achromatic line. On the other hand, we want to penalize specular regions, those close to achromatic line and far from the coordinate system origin.

A main feature of achromatic line is that the three components of its points are equal $r = g = b; \forall r, g, b \in [0, 1]$. For pixels close to this region, we $r \approxeq g \approxeq b; \forall r, g, b \in [0, 1]$. As the pixels fall away from this line, the differences among their components are greater. We use this difference as the intensity of the processed image. As we want to preserve the chromatic information, only the intensity is modified, boosting the diffuse pixels and nullifying the specular pixels. The new intensity of the pixels is computed as difference between the maximum and minimum of their RGB components:

$$Intensity(x) = max_{\{r,g,b\}}(x) - min_{\{r,g,b\}}(x)$$

This intensity replaces the V component in the HSV representation, thus preserving the chromatic content of the pixel. We show in Algorithm 4.1 an implementation for SciLab.

4.2.3 Application

The SF2 image, the one obtained after the described transformation, is characterized by the absence of reflections, substituted by dark spots. Also the diffuse regions are boosted in the image. With an straightforward analysis we can find all the diffuse regions.

Figure 4.1: Natural images

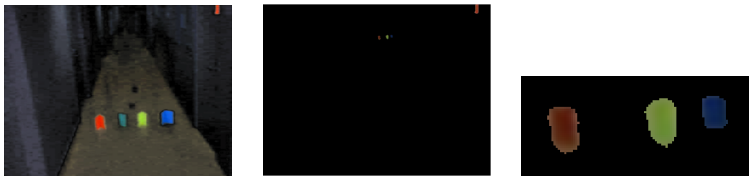


Figure 4.2: SF2 images

4.2.4 Experiments

We have performed experiments in three different contexts: first the detection of markers in real scenes, other with synthetic images, and the last about the detection of robots in real time. All the results can be viewed in the following web address: http://www.ehu.es/ccwintco/index.php/SMC

4.2.4.1 Mark detection

The definition of the experiment is as follows:

1. Context:

a) Place: a lab corridor, with artificial illumination of diverse intensity and uniform color.

b) Markers are DIN A4 sheets of different colors: red, cyan, yellow and blue.

c) Standard web cam Phillips SPC 900NC/00

2. Experiment:

From each image (recorded in a MPEG file) we find the SF2 images, and there we find the markers.

In Fig.4.1 we have three images from the described scenario. The ones on the left are the closest ones to the camera, the ones on the right are the farthest ones. Notice variations in illumination along the corridor. In Fig.4.2 we show the SF2 images as follows: left corresponding to the middle one in Fig.4.1 , middle after the analysis of the intensity and to the right a zoom of the previous one, showing that one mark is missing. In table 4.1 we show the detections performed on each mark, where 'x' means good detection and '+' incomplete detection.

Milestone	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Distance	2.6	4	6	8.4	10.8	12.8	14.5	17.3	20.7	26.02	31.7	36	41.9	46	50
Label 1	x	x	x	x	x	x	x	x	x	x		x	+	x	
Label 2	x	x	x	x	x	x	+								
Label 3	x	x	x	x	x	x	x	x	x	x	x	x	+	x	
Label 4	x	x	x	x	x	x	x	x	x	x	+				

Table 4.1: Measurements

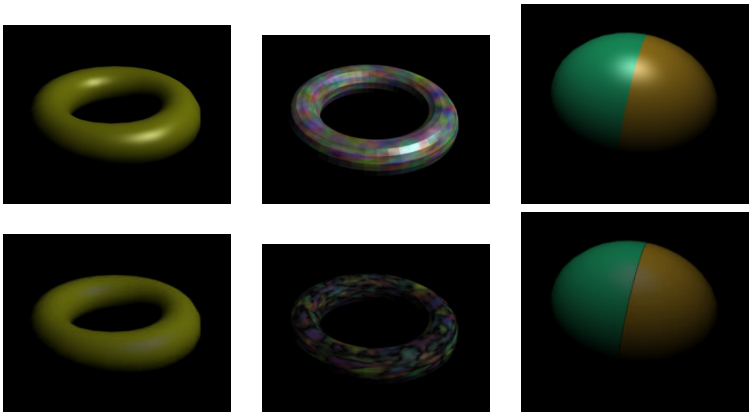


Figure 4.3: Synthetic images

4.2.4.2 Synthetic images

The above color transformation has been applied to natural and synthetic images. Synthetic images have the advantage that we know with precision the color and geometry of the surface, as well as the illumination color. Fig.4.6 shows some of these images, in the top row we place the original image and on the bottom the computed SF2 images. First image is a monochromatic image, with a green surface. The second is a Voronoi tessellated surface painted with random colors. Last image is a bi-chromatic oval. We observe that SF2 images remove completely all the reflections, canceling the specular component. In the Voronoi tessellated ring surface, besides canceling brights spots, colors have been enhanced.

The SF2 method has been ideated for robotic contexts. In Fig.4.8 we show results on three natural images. The two first ones are customary marks in the previously described experiment, and the last one is used by other researchers in the literature of specular correction. The first two scenes show the magnification of the markers in the image. In the last case we see that the bright spots are cleanly removed, respecting original color.

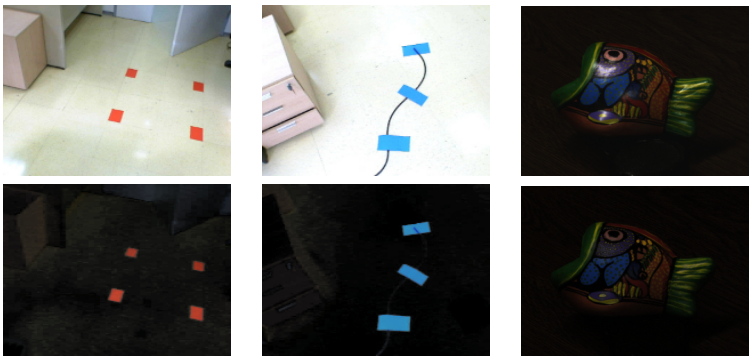


Figure 4.4: Natural images

Figure 4.5: Robots detection

4.2.4.3 Real robot detection

The last experiment is the detection of small robots (SR1) in a real scene and real time. The robots are yellow color against a yelowist background, making visual detection tricky. The floor is very bright with many bright spots front above illumination. Besides, robot's upper part contains the printed board and some fixing for the cable being carried. The robots have lots of shadows, thus only a small part of the robot can be clearly detected as pure yellow. Fig.4.5 contains three images: first the capture from the scene, second its SF2 image, third the SF2 image intensity analysis to detect the robots. The web address http://www.ehu.es/ccwintco/index.php/SMC contains the original video. We must point out that illumination is not constant, there are doors, windows, etc.

4.2.5 Work Conclusions

The work presented here proposes a method for color detection in images, characterized by:

1. Being fast and efficient.
2. Removes the specular component.
3. Magnifies color, preserving scene chromaticity, modifying only the intensity.
4. Can work in real time.

Other methods for the removal of the specular component are based on iterative methods that render them unsuitable for real time processing.

4.3 Hybrid Color Space Transformation to Visualize Color Constancy

Color constancy and chromatic edge detection are fundamental problems in artificial vision. Here we present a way to provide a visualization of color constancy that works well even in dark scenes where such humans and computer vision algorithms have hard problems due to the noise.

This work presents a hybrid and non linear transformation of the RGB image based on the assignment of the chromatic angle of the pixel (computed in the RGB space) as the luminosity value in the HSV space. The image is preprocessed to remove the specular component. The chromatic angle was defined on the basis of DRM, having thus a physical interpretation supporting it. In the HSV color space the intensity is represented in the V value, changing it does not change the pixel chromatic information. Thus, to visualize CC we assign constant intensity to the pixels having common chromatic features, by assigning the chromatic angle as the V value in HSV space.

4.3.1 Chromaticity and Chromatic distance

In the HSV color space, chromaticity is identified with the pair (H, S), and the V variable represents the luminosity or light intensity. Plotting on the RGB space a collection of color points that have constant (H, S) components and variable intensity I component, we have observed that chromaticity in the RGB space is geometrically characterized by a straight line crossing the RGB space's origin, determined by the ϕ and θ angles of the spherical coordinates of the points over this chromaticity line. The plot of the pixels in a chromatically uniform image region appear as straight line in the RGB space. We denote L_d this *diffuse line*. If the image has surface reflection bright spots, the plot of the pixels in these highly specular regions appear as another line L_s intersecting L_d. This idea is correctly explained on [ref]

For diffuse pixels (those with a small specular weight $m_s(x)$, of the image expresed on the DRM Eq.[ref]) the zenith ϕ and azimuthal θ angles are almost constant, while they are changing for specular pixels, and dramatically changing among diffuse pixels belonging to different color regions. Therefore, the angle between the vectors representing two neighboring pixels $\mathbf{I}(x_p)$ and $\mathbf{I}(x_q)$, denoted $\angle(I_p, I_q)$, reflects the chromatic variation among them. For two pixels in the same chromatic regions, this angle must be $\angle(I_p, I_q) = 0$ because they will be collinear in RGB space.

The angle between I_p, I_q is calculated with the equation:

$$\angle(I_p, I_q) = \arccos\left(\frac{\mathbf{I}(x_p)^T \mathbf{I}(x_q)}{\sqrt{\|\mathbf{I}(x_p)\|^2 + \|\mathbf{I}(x_q)\|^2}}\right). \tag{4.1}$$

4.3.2 An Approach for Regular Region Intensity

The basic idea of our approach is to assign a constant luminosity to the pixels inside an homogeneous chromatic region. To do that we must combine manipulations over the two color space representations of the pixels, the HSV and RGB. The process is highly non linear and it is composed of the following steps:

1. Isolate the diffuse component removing specular components ($m_s = 0$): we are interested only in the diffuse component because it is the representation of the true surface color. We use the Roby T. method presented in [58] to perform the diffuse and specular component separation.
2. Transform the diffuse RGB image into the HSV color space.
3. Compute for each pixel in the image the chromaticity angle as the angle between the gray diagonal line in the RGB space, going from the black space origin to the pure white corner, and the chromaticity line of the pixel.
4. Assume the normalized chromaticity angle as the new luminosity value in the HSV space pixel representation.

In an homogeneous chromatic region, all pixels fall on the same diffuse line $L_d : (r, g, b) = \mathbf{O} + s\mathbf{\Psi}; \forall s \in \mathbb{R}^+$ where $\mathbf{O} = [0, 0, 0]$ and $\mathbf{\Psi} = [\Psi_r, \Psi_g, \Psi_b]$ is the region chromaticity expressed in Euclidean coordinates. The chromatic reference is the achromatic (pure white) line L_{pw} which is defined as $L_{pw} : (r, g, b) = c + s\mathbf{u}; \forall s \in \mathbb{R}^+$ where $\mathbf{O} = [0, 0, 0]$ and $\mathbf{u} = [1, 1, 1]$. Therefore, if all pixels is a region belong to the same chromatic line, the angle between each pixel and the line L_{pw} must be the same, and the result of this angular measurement is a constant for whole region. Our strategy is to normalize this measure in his domain of definition (the RGB cube) and assume it as the constant luminosity value V. This method is expressed with the equation:

$$V^{new}(x) = \frac{\angle(\mathbf{I}(x), \mathbf{u})}{\arccos(\vartheta)} \tag{4.2}$$

where the denominator $\arccos(\vartheta)$ is the normalization constant corresponding to the maximum angle between the extreme chromatic lines of the RGB space (red, green or blue axes) and the pure white line. Algorithm 4.2, shows a Matlab/Scilab implementation of the method, where ϑ takes the value $\frac{1}{3}$ and $\arccos(\vartheta) = 0.9553166$.

4.3.3 Experimental Results

We present the results from three computational experiments. The first one using a synthetic image and the remaining using natural images. Fig.4.6 displays the first experimental results. Fig.4.6a is the original image. Fig.4.6b is

Algorithm 4.2 Regular Region Intensity

```
function IR = SF3(I)
  Idiff = imDiffuse(I); // look for the diffuse component
  new_intensity = angle(Idiff, [1 1 1]); // return a matrix of chromatic angles
  Ihsv = rgb2hsv(Idiff);
  Ihsv(:,:,3) = new_intensity; // assign the normalized angles as image intensity
  IR = hsv2rgb(Ihsv);
endfunction
```

the diffuse image obtained applying the method in [49]. Fig.4.6(c) is the result applying our proposed method in Fig.4.6(a). Fig.4.6(d) display the result applying the method in Fig.4.6(b). It can be appreciated that our method is able to identify the main chromatic regions even without component separation (Fig.4.6c), with some artifact due to the bright reflections. After removal of these reflections, the method has a very clean identification of the chromatic regions.

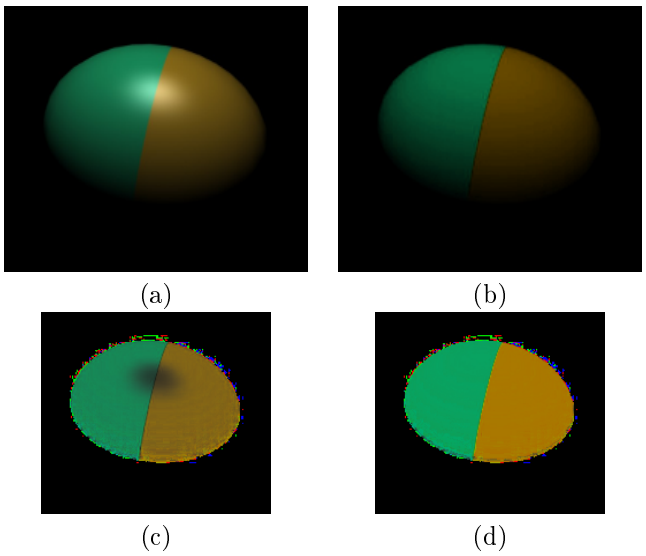


Figure 4.6: Synthetic image results (a) original image, (b) diffuse component of the image, (c) our method on image (a), our method on image (b).

For the next experiments we use natural images that have been used by other researchers previously. The Fig. 4.7 and 4.1 show the experimental results. In both cases the sub-figure (a) has the original image, sub-figure (b) shows the diffuse image, sub-figure (c) displays the results applying our proposed method to the original image (a), sub-figure (d) show the results applying our method in

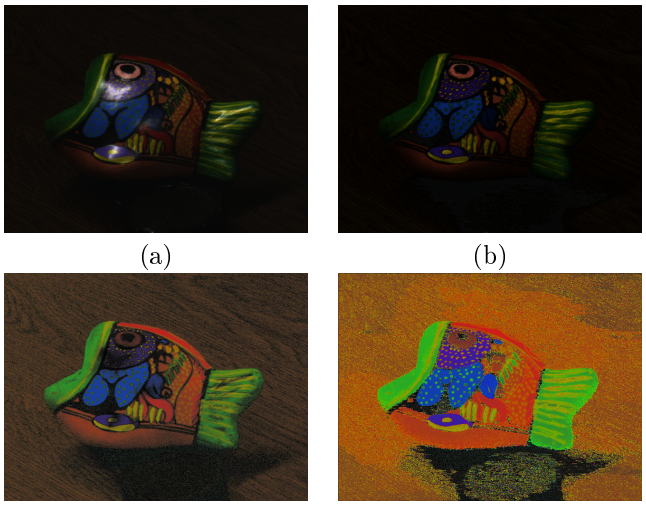
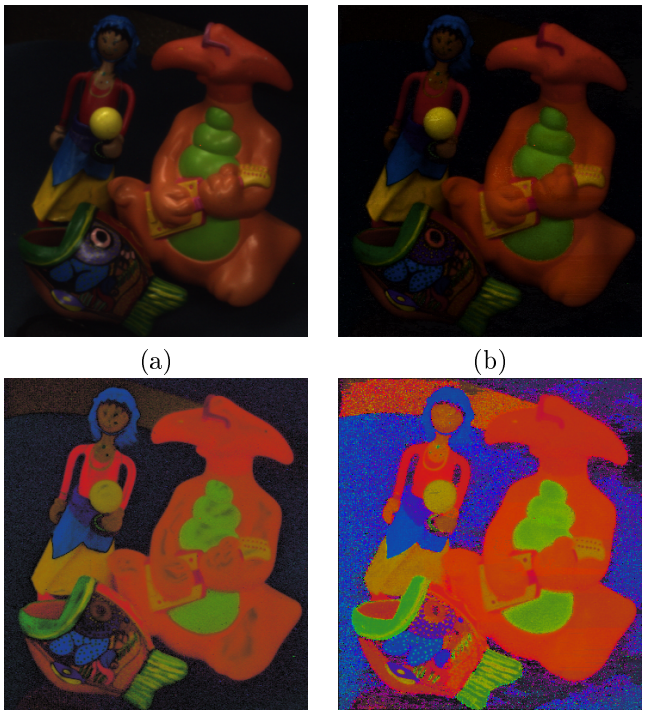


Figure 4.7: Natural image results, (a) original image, (b) diffuse component of the image, (c) our method on image (a), our method on image (b).

the diffuse image (b). In both experiments we can see a similar effect of applying specular correction. The images (c) obtained without component separation, show a better chromatic preservation, although with some degradation in the regions corresponding to the specular brights. The images obtained after diffuse component identification [49] are less sensitive to specular effects, however they show some chromatic region oversegmentation. It is important to note that no clustering process has been performed to obtain these images.



(a) (b)

(c) (d)

Figure 4.8: Natural images, (a) original image, (b) diffuse component of the image, (c) our method on image (a), our method on image (b).

4.3.3.1 Work Conclusions

This work presents a color transformation that enables good visualization of Color Constancies in the image, changing only the image luminosity and preserving its chromaticity. The result is a new image with strong contrast between chromatic homogeneous regions, and good visualization of these regions as uniform regions in the image. This method performs very well in dark regions, which are critical for most CC methods and image segmentation based on color clustering processes. The method could be the basis for such a process, applying the clustering process to the chromaticity angle.

We have found that specular correction of the image improves the results on highly specular regions of the image, however our approach performs well also on images that have not been preprocessed.

4.4 Chapter Conclusions

SF images are a resourceful solution to overcome the noise aberration induced by shines. Those straightforward transformations provides good results despite its easiness. In both cases those transformations are based on the image saturation and hue, hence in its chromaticity. The main difference between both methods are the definition of the saturation; whereas first one trusts in the well-know pseudo-norm of the saturation, second one improves it with an angular definition of the saturation. This one has a physical support on the DRM, providing CC properties.

SF image is a new topic in computer vision or image processing which can help to solve task related with to avoid specular effects on images. We advocate its application and further improvements.

Chapter 5

Gradients

This chapter correspond with the process 5 of the thesis flowchart and is outlined as follow: Sec.5.1 discuss about the state of the art, Sec.5.2 presents the chapter contribution and Sec.5.3 concludes with the contribution conclusion and chapter conclusions.

5.1 State of the Art

Edge detection is a key step in most computer vision applications, on it depends the successful of the application because it is one of the first steps, and the fiability of following steps depend on it. The detection of edges provides a way to find shapes in images. The knowledge about object boundaries is enough to describe its shape, and hence it is equivalent to the detection of regions with constant values. Edge is understood as a line whose neighbor pixels have a significant variation each others respect to its gray level, that is, respect to the intensity. Speaking about edges we refers about significant differences in a neighborhood due to the scene features and not due to the noise. Edges are lines whereas isolated points with strong differences in a neighborhood are referred as edgels.

Images edges are studied respect to the gray image or image intensity. The limited resolution of the imaging systems blurs the edges, the nonlinear sensor and the electronic thermal noise disturbs on the task of find edges.

The principal method for edge detection is based on the Gradient Vector. It is the application of the first derivative in a bidimensional space. This approach gives the edge magnitude and direction, but generally in computer vision, the direction is ignored (regarding to edge detection). Another well-known method for edge detection is the Laplacian operator. It is a second-order derivative operator, it is the addition of the derivation in all directions, and it is an isotropic edge detector. A good use of these tools is using the magnitude of the gradient vector as edge detector and using the sign of the Laplacian to determine if the edge point belong to the dark side of the border or to the light side. There are

some methods based on the Laplacian of the Gaussian these ones have biological justification in accordance with retinal cells structure.

It is significantly more difficult to analyze edges en multichannel images (color images, multi-spectral and hyper-spectral images). The first idea detecting edges on his kind of images is applying the gradient vector on each channel and sum them. However in this case, the main drawback comes from that each channel has different information of the image, then in a point the gradient could have different directions at each channel and the sum of all channels could be 0. To overcome this shortcoming, can be used $S = J^T J$ where J is the Jacobian matrix and S is a diagonal matrix. If only one member of S is non-zero, then is is called a perfect edge and it gives de direction of the variation. However, usually this diagonal matrix has more members bigger than zero, as a consequence of the noise. The trace of S measures the edge strength. A good review of edges is shown in [63] where these topics are more deeply explained.

In the literature we can find some well-known methods based on he first derivative like Sobel [64] and Prewitt [65] that are performed by convolution masks. Canny [66] has the property to avoid edgels, therefore all line edges are connected. Canny is an algorithm which uses the direction of the gradient vector to find the nearest edge point, then the output is compound by convex regions, whereas Sobel and Prewitt approached follows the mathematical definition of gradient but these ones not use the gradient direction. Sobel and Prewitt need a threshold to decide whether a point is a edge or not.

Roberts operator [67] was one of the first contributions on edge detection, it marks only the edge points without their direction. This operator is too fast and has good results when working with binary images. Kirsch masks [68] also known as 'compass' because they are defined by using a simple mask and rotating in the eight main directions (North, Northwest, West, Southwest, South, Southeast, East and Northeast). Robinson masks [69] follow the same compass idea but in this case by using one of the Sobel masks. These aforegoing methods are performed in the eight neighborhood but they could be extended easily to a bigger neighborhood like 5x5, 7x7, and then.

Now a days, edge detection keeps as an active research topic. Watershed technique [70, 71, 72] has not physical support, it has a topographic inspiration. We'll see a deep insight on watershed in the next chapter. These topological technics have an important contribution in image segmentation, where for example Ségonne[73] presents a novel framework to exert topology control over a level set evolution. It is an active contour which works in a three dimensional space and have excellent results in medical image segmentation. Respect to the evolution of the traditional gradients, McIlhagga [74] discuss about Canny's work. He improves it solving two problems. First, he provides a more accurate localization criterion, and second, the width of the optimal detector is limited by considering the effect of the neighboring edges in the image.

5.1.1 Chromatic Gradients

A common feature of all gradients explained in previous section, is the lack of chromatic information. That is, all these gradients works within gray scale. We advocate color is an important descriptor and we must take profit of all information; intensity and chromaticity too. In fact, all gradients which works only with intensity can not detect a lot of chromatic edges, because for each color to intensity transformation exist infinite colors with identical intensity. Firstly we focus in chromatic gradient because chromaticity is independent of the intensity and therefore of the surface geometry having invariance respect to the illumination. This is one of the desired goals for optimal image segmentation.

Next section presents an image color gradient which preserves color constancy [75].

5.2 An Image Color Gradient preserving Color Constancy

This work presents a color gradient with good color constancy preservation properties. The approach does not need a priori information or changes in color space. It is based on the angular distance between pixel color representations in the RGB space. It is naturally invariant to intensity magnitude, implying high robustness against bright spots produced be specular reflections and dark regions of low intensity.

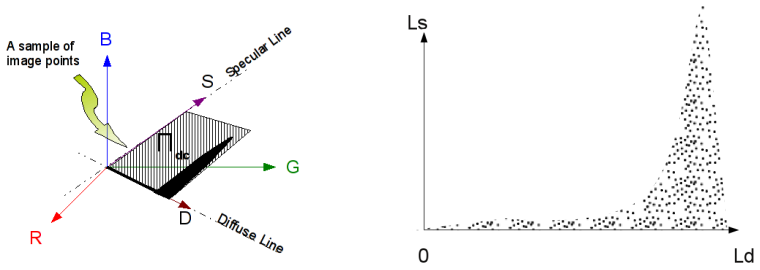
Color constancy (CC) is fundamental problem in artificial vision [18, 46, 76], and it has been the subject of neuropsicological research [13], it can be very influential in Color Clustering processes [47, 77, 78]. In the artificial vision framework, CC assumes some color space, the illumination chromaticity estimation [51, 18] and the separation of diffuse and specular image components [58, 54, 50].

In this work, we assume a physical interpretation of the image reflectance and its behavior in the RGB space. We use polar coordinates to specify points in the RGB space, because we will be interested in the zenithal ϕ and azimuthal angles θ, because the characterize the chromatic component of the RGB point. We are looking for color image edge detection under a CC constraint and founded on DRM.

5.2.1 Dichromatic Reflection Model

From the DRM we can deduce some interesting features of the distribution of the pixels in the RGB cube. In Fig. 5.1 we illustrate the main expected effects for a single color image (disregarding the black background) with a bright spot due to the illumination source. According to DRM we need to know only two colors: D corresponding to the observed surface and S corresponding to the illumination source. Drawing a line in the RGB cube passing over these colors and the RGB origin (black), we obtain two chromatic lines L_d and L_s, respectively. These two lines define a chromatic plane in RGB illustrated as the stripped region in Fig. 5.1a. All the image pixels must fall in this plane, discounting additive Gaussian noise perturbations, according to DRM equation **??** for image colors D and S. Looking to the image pixel distribution inside the chromatic plane, we obtain the plot in Fig. 5.1b, whose axes are the chromatic lines L_d and L_s. We have that non-specular pixels fall close to the diffuse line L_d, while specular pixels go away from the origin and the diffuse line parallel to the specular line L_s. There is an intensity threshold for the pixels having a significative specular component ($m_s(x) >> 0$). This threshold is the albedo of the material in the scene. For intensities greater than the albedo, pixels fall away from the L_d diffuse line along the direction of L_s.

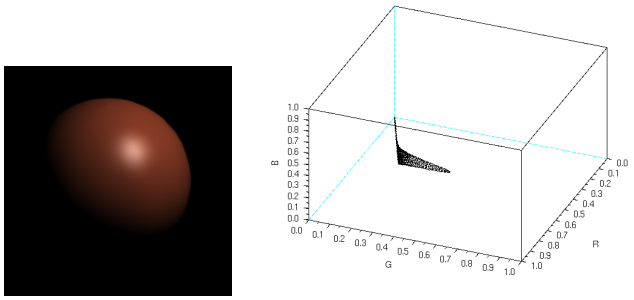
Fig.5.2 shows the pixel distribution for a synthetic image. The RGB cube plot in 5.2(b) shows the pixel RGB color distribution of the image 5.2(a). These images confirm our previous discussion, for a case of a single color object in the image. When there are more than one color in the image, we can expect



(a) (b)

Figure 5.1: Expected distribution of the pixels in the RGB cube according to DRM for a single color image.

several diffuse lines, so that the we can base our image segmentation on this observation. All these lines cross the RGB origin, therefore the pixel polar coordinates of diffuse pixels contain much information relative to underlying reflectance regions.



(a) (b)

Figure 5.2: Distribution of pixels in the RGB space

For an scene with several surface colors, the DRM equation assumes that the diffuse component may vary spatially: $I(x) = m_d(x)D(x) + m_s(x)S$. However, the specular component is space invariant in both cases, because the illumination is constant for all the scene. Finally, assuming several illumination colors we have the most general DRM $I(x) = m_d(x)D(x) + m_s(x)S(x)$ where the surface and illumination chromaticity are space variant.

5.2.2 Color Constancy in the RGB Space

The CC is the mental ability to identify chromatically homogeneous surfaces under illumination changes. This mental ability is still an open neupsicological research topic [10]. The CC property is inversely proportional to the color discontinuity represented by the color edges (CE). In essence, given a chromatic

image gradient, low intensity gradient magnitude corresponds to CC and high magnitude to CE. In the HSI and HSV color spaces, chromaticity is identified with the pair (H, S) and the I or V variable represents the intensity. We have observed that chromaticity in the RGB space is characterized by a straight line crossing the RGB space's origin, determined by the ϕ and θ angles of the polar coordinates of the points over the line, by plotting on the RGB space a collection of color points that have constant HS components and variable intensity I component. The plot of the pixels in a chromatically uniform image region appear as straight line in the RGB space. We denote L_d this *diffuse line*. If the image has surface reflection bright spots, the plot of the pixels in these regions appear as another line L_s intersecting L_d.

For diffuse pixels (those with a small specular weight $m_s(x)$) the zenithal ϕ and azimuthal θ angles are almost constant, while they are changing for specular pixels, and dramatically changing among diffuse pixels belonging to different color regions. Therefore, the angle between the vectors representing two neighboring pixels I_p and I_q, denoted $\angle(I_p, I_q)$, reflects the chromatic variation. For two pixels in the same chromatic regions, this angle is $\angle(I_p, I_q) = 0$ because they will be collinear in RGB space.

5.2.3 Gradient Operators

The notion of CC is closely related to the response to the gradient operators [17]. Regions of constant color must have low gradient response, while color edges must have a strong gradient response. To set the stage for our chromatic gradient proposition, we must recall the definition of the image gradient

$$G[I(i,j)] = \begin{bmatrix} G_i \\ G_j \end{bmatrix} = \begin{bmatrix} \frac{\partial}{\partial i} I(i,j) \\ \frac{\partial}{\partial j} I(i,j) \end{bmatrix}, \tag{5.1}$$

where $f(i,j)$ is the image function at pixel (i,j). For edge detection, the usual convention is to examine the gradient magnitude:

$$G(I) = |G_i| + |G_j|. \tag{5.2}$$

For color images, the basic approach to perform edge detection is to drop all color information, computing the intensity $Intensity = (Red + Green + Blue)/3$ (sometimes computed as $Intensity = .2989 * Red + .587 * Green + .114 * Blue$), and then convolve the intensity image with a pair of high-pass convolution kernels to obtain the gradient components and gradient magnitude. The most popular edge detectors are the Sobel and the Prewitt detectors, illustrated in Fig.5.3 because we will build our own operators following a similar pattern structure. To take into account color information, the easiest approach is to apply the gradient operators to each color band image and to combine the results afterwards: $G(I) = [G(I_r) + G(I_g) + G(I_b)]/3$. Fig.5.4 illustrates these ideas. It can be appreciate how the gradient magnitude amplifies noise on one hand when we combine the color band gradient magnitudes, and how the color edge is not detected by the edge operator applied to the intensity image, because the two

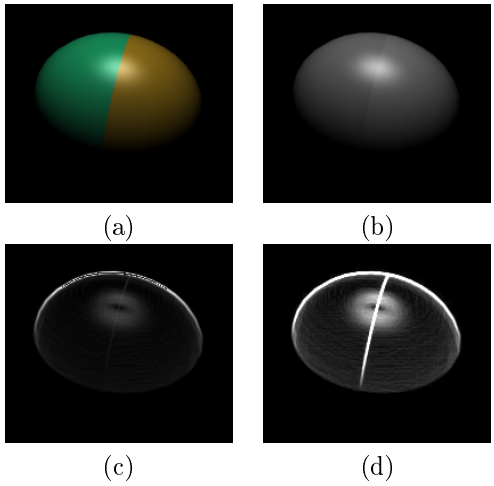
$$\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$$

(a)

$$\begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

(b)

Figure 5.3: Convolution kernels for the (a) Sobel and (b) Prewitt edge detection operators.



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

Figure 5.4: (a) Original synthetic RGB image, (b) Intensity image, (c) Gradient magnitude computed on the intensity image, (d) gradient magnitude combining the gradient magnitudes of each color band

color regions have quite near intensity values. The edge magnitude computed by the straightforward approaches is also misled by the specular surface reflections, which highlighted as can be appreciated in Fig.5.4(d).

5.2.4 Proposed Method

We first discuss how do we build a distance between color pixel values which preserves chromatic coherence and, thus, color consistency. Then we formulate the gradient operators which are consistent with this color distance.

5.2.5 A chromatic coherent RGB pixels distance

First, we convert the RGB Cartesian coordinates of each pixel to spherical coordinates, with the black color as the RGB space origin. Let us denote the Cartesian coordinate image as $I = \left\{(r,g,b)_p ; p \in \mathbb{N}^2\right\}$ and the spherical coordinate as $P = \left\{(\phi,\theta,l)_p ; p \in \mathbb{N}^2\right\}$, where p denotes the pixel position. In this second expression, we discard the l because it does not contain chromatic information. For a pair of image pixels p and q, the color distance between them is defined as:

$$\angle(P_p, P_q) = \sqrt{(\theta_q - \theta_p)^2 + (\phi_q - \phi_p)^2}, \tag{5.3}$$

that is, the color distance corresponds to the euclidean distance of the Azimuth and Zenith angles of the pixel's RGB color polar representation. This distance is not influenced by the intensity and, thus, will be robust against specular surface reflections.

5.2.6 Chromatic coherent gradient operators

We will formulate a pair of Prewitt-like gradient convolution operations on the basis of the above distant. Note that the $\angle(P_p, P_q)$ distance is always positive. Note also that the process is non linear, so we can not express it by convolution kernels. The row convolution is defined as

$$CG_R\left(P\left(i,j\right)\right) = \sum_{r=-1}^{1} \angle\left(P\left(i-r, j+1\right), P\left(i-r, j-1\right)\right),$$

and the column convolution is defined as

$$CG_C\left(P\left(i,j\right)\right) = \sum_{c=-1}^{1} \angle\left(P\left(i+1, j-c\right), P\left(i-1, j-c\right)\right),$$

so that the color distance between pixels substitutes the intensity subtraction of the Prewitt linear operator. The color gradient image is computed as:

$$CG(P) = CG_R\left(P\right) + CG_C\left(P\right) \tag{5.4}$$

5.2.7 Experimental Results

To demonstrate the efficiency of our proposed approach, we will show three experimental results. Two of the experiments are done on synthetic images whose ground truth is know.

Fig.5.5 contains two synthetic images Fig.5.5(a) and 5.5(b) which are chromatically identical. The image in Fig. 5.5(a) hs constant intensity inside each color region, while the image in Fig.5.5(b) contains a central square with lower intensity (0.8), preserving the chromatic content of Fig. 5.5(a). Applying the Prewitt operator to each color band of Fig.5.5(b) we obtain the detection shown in Fig.5.5(c), while applying our color edge detection of Eq.5.4 we obtain the

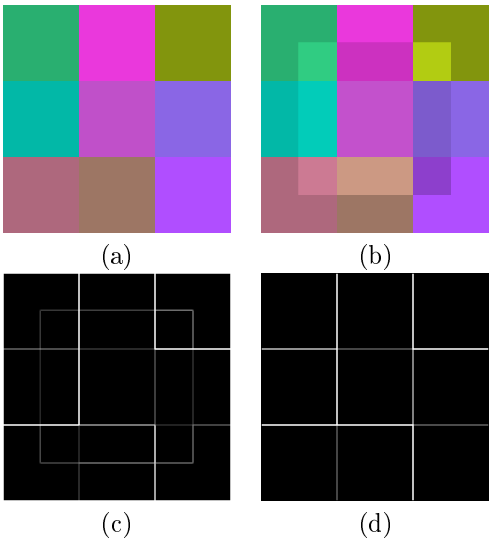


Figure 5.5: Results of the color edge detection on a synthetic image with nine uniform chromatic regions and a variation of intensity. (a) Original color distribution, (b) lower intensity central square, (c) Prewitt detection on RGB bands, (c) our approach in equation (5.4).

detection in Fig.5.5(d). It is clear that our approach has superior CC properties and an improved intensity invariant detection of color edges.

The second computational experiment was performed on the image shown in Fig.5.4(a). This image has a strong specular reflection region, and two color regions with a black background. We have tested a Sobel like and a Prewitt like variation of the basic schema of Eq.5.4. The Fig.5.6 gives the results of the RGB band combined detection and our approach. It can be appreciated that our approach discovers the edge even in very dark areas, it is also robust against specular reflections, which the linear operators do confound with color edges. The color edge between the two regions is better detected in both cases by our approach.

Final results are given on a natural image, shown in Fig.5.7. This image contains many color regions, with specular reflections, shadows and light effects. Fig.5.8 shows the results of the linear operators based on the Sobel and Prewitt masks. Besides the lower response of the Prewitt operator, it can be appreciated the high sensitivity to specular reflections and low color constancy. All bright spots are interpreted as color edges. In the Fig.5.9 we show the results of our approach under two variations of the neighborhood considered. The 4 neighborhood follows the same pattern of Eq.5.4 but over a reduced set of neighboring pixels. Again our approach is very robust against specular reflectance. Bright spots do not appear to be detected. Dark regions of the image are equalized in their results relative to brighter regions. A very significative result is the

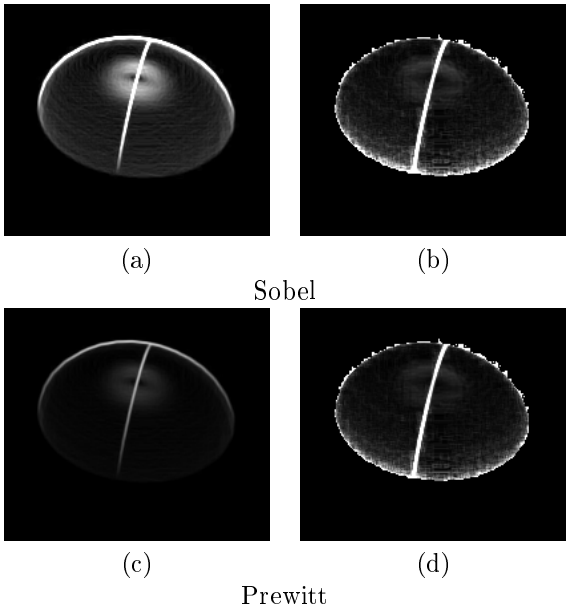


Figure 5.6: Color edge on the synthetic image of Fig.5.4(a) with two color regions. (a) The Sobel operator over the RGB bands with specular component, (b) our approach in a Sobel-like structure, (c) the Prewitt linear operator, (d) our approach in a Prewitt like structure.

Figure 5.7: Natural image

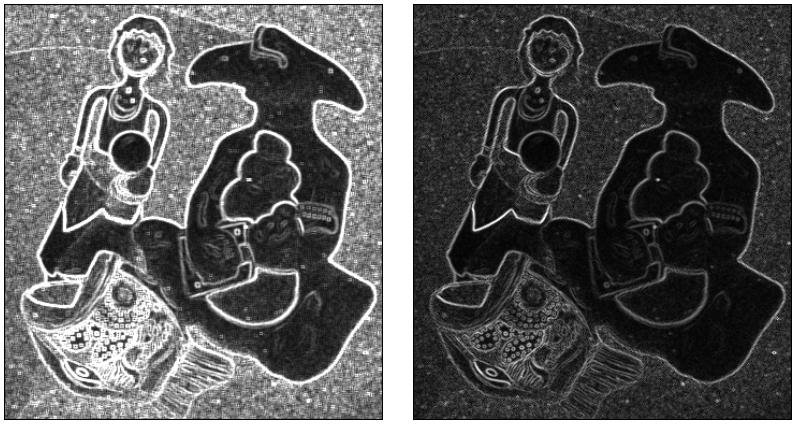
detection of color edges even in the almost black background. A drawback that appears in our approach is the high spurious detection in the black background. This is due to the high angular variations induced by noise. It could be avoided by a simple intensity thresholding.

5.3 Work and Chapter Conclusions

This chapter has only one contribution where we have presented an innovative chromatic gradient computation, which is chromatically coherent, preserves the Color Constancy and gives good detection of Color Edges. The method is grounded in the DRM which is a widely accepted image model for reflectance analysis. The method is intensity invariant, and, thus, is robust against the bright spots of specular reflections. It does not imply or need color segmentation, on the contrary can provide good color region separation with little assumptions. It works on the RGB space, which the most common color processing space.

(a) (b)

Figure 5.8: Results of the linear operators on the natural image (a) Sobel detector, (b) Prewitt detector



(a) (b)

Figure 5.9: Results of our approach on the natural image (a) taking 8 neighbors, (b) taking 4 neighbors

Chapter 6

Segmentation

This chapter corresponds with the process 6 and 7 of the thesis flowchart and it has two contributions; Sec.6.4 presents a segmentation method which follows a region growing strategy, whereas in Sec.6.5 presents a segmentation method based on watershed. Previously to these contributions, Sec.6.1 defines segmentation, Sec.6.2 present a state of the art, Sec.6.3 talks about the main segmentation problems and we'll finish this chapter with the Sec.6.6 where we summarize this chapter.

6.1 Image Segmentation

Image segmentation is one of the foremost topics in image processing and in computer vision. It is the first step in a large number of computer vision systems. The whole system performance may be dependent on the segmentation results. Unfortunately previous gradient detectors are not enough for image segmentation due to not exist an universal edge detector based on gradients. Popular instances of segmentation methods for color images proposed in the literature are based on watershed transform [79, 80, 72, 71, 70], and on clustering procedures [81, 82, 78, 83].

Image segmentation can be defined as a process which defines a partition of the image domain F according to some pixel property [84, 85]: If $P(x)$ is some homogeneity predicate defined on groups of connected pixels, then the image segmentation is a partition of the set F into connected regions $(S_1, S_2, ..., S_n)$ such that $\bigcup_{i=1}^{n} S_i = F$ with $\forall i \neq j, S_i \cap S_j = \varnothing$, and each pixel in a region fulfills the same property, i.e. $\forall x, y \in S_i; P(x) = P(y)$.

In other words, image segmentation is the process by which the input image is divided in regions following some criterion, and the union of all these regions gives the input image.

6.2 State of the Art

The segmentation task can be carried out by two different ways; first one, by clustering technics within the color space (or gray level) then, some sets could be disjoint into the image domain. The other way is having in account the pixel neighborhood, in this case the algorithm works in the image domain. The first one is referred as *pixel based* segmentation and the second one is refereed as *region based* segmentation.

In color images, *pixel based* segmentation have been widely performed by the c-means algorithm [83, 86]. There are three major difficulties in this way: (1) determining the optimal number of clusters to be created, (2) choosing the initial cluster centroids, and (3) handling data characterized by variabilities in cluster shape, cluster density, and the number of points in different clusters. Usually, the two first ones problems are solved by user specification whereas the third one is application dependent. There are some improvements. Dae-Won [86] proposes an automatic initialization of the fuzzy c-means for color clustering (1 and 2 points) however, the third point is defined by the user, it is the color space where the author advocate by the CIElab. More recently, the same idea has been applied on hypersptectral images [83], where pixel feature is defined by its hyperspectral signature. In both cases, the shortcoming is the absolute dependence with the color space, spectral signature for hyperspectral images or the strategy for dimensional reduction in the second case. In fact, only by changing the color space or the method for dimensionality reduction, the results are going to be different. This drawback comes from the distance used in c-means, usually the euclidean distance. In this way, Oussalah[87] defines the distance in terms of divergence distance, which builds a bridge to the notion of probabilistic distance. A good idea for image segmentation, is to have in account a pixel distance and the neighborhood, like in [88] where for clustering uses c-means and for the neighborhood Random Markov Fields.

Mathematical morphology (MM) belong to the *region based* segmentation methods and it is growing up regard to image segmentation. It is the application of lattice theory [89, 90, 91]. First introduced as a shape based tool for binary images, MM has become a powerful non-linear image analysis technique with operators for image segmentation, image filtering and extraction of features in binary or gray-scale images. The definition of morphological operators is based in a totally ordered complete lattice structure: each element in a set is ordered without uncertain, in other words, for whatever two elements in a lattice, we can determine if they are equals or one is bigger than the other one. Here, one more time, despite it works with the pixel neighborhood, we have the drawback of the *pixel based* segmentation methods; the difficult task to determine a correct distance and the dependence of the color space. Like in the aforementioned clustering strategies, changing the color space or the method for dimensionality reduction, the results are going to be different. Hanbury [92, 93, 94] presents a work of MM for the CIElab and HSL color spaces, where at last he present IHSL color space. Angulo [95] applies this technics extracting information from cartographic images.

Watershed algorithm is defined within the MM methods, therefore it is a *region based* segmentation method. It was introduced by Beucher [70] three decades ago. This algorithm has a topological inspiration, it takes a gray-scale image as a topological surface and after a flooding process after a flooding process, each region is a catchment basin. The widely used gray-scale image is the gradient image, therefore, watershed algorithm find the regions based in a edge image. This original algorithm has done some improvements. For example, this algorithm is based on the gradient image, therefore depending of the used gradient we are going to obtain different results. In this way, Wang [96] proposes a multiescale gradient for watershed, he uses morphological gradient and try to avoid the over-segmentation detecting isolated minimal local. In fact, this one is the main drawback of the watershed transformation, the over-segmentation. To avoid over-segmentation, a solution comes by applying a "region merging" process. There are some interesting solutions too in the literature, like Dagher [97] who presents Waterballoons. Waterballoons is a mixture of "balloon snakes" and watershed; on each minimal of the gradient image an active contour is initialized, now, the flooding process is replaced by a growing up of the snakes (like balloons). This strategy helps to avoid isolated little regions, like points, which can suppose that they are a noise consequence. Other improvements of the standard watershed are focused on the speed-up. An original idea is to apply a raining simulation instead of the flooding process [79]. This strategy helps to find in a faster manner the edges. The key idea to sort the computational cost is by optimizing the algorithm in the neighborhood looking for shorter paths [80, 71].

For whatever used algorithm, one of the more difficult task is the evaluation of the method. To design a good measure for segmentation quality is a known hard problem. Each person has different idea of segmentation and different applications may function better using different segmentations, while the criteria of a good segmentation are often application-dependent. A good review of measures for the validation of segmentation process is done by H. Zhang [98] where a interesting feature is the lack of measures for color images and multi or hyperspectral images.

6.3 Main Segmentation Problems

There are four main sources of problems in image segmentation: illumination, noise, edge ambiguity and computational cost. In this section we discuss those topics, offering some ideas about how our segmentation algorithm addresses these problems.

Illumination in real environments: both real and artificial illumination sources introduce some important problems in image understanding. Because of shines or shadows, both depending on the illuminant position, a surface of a scene can produce a lot of different image perceptions. In digital image processing, it is very usual to assume a uniform chromatic illumination because it makes easier to deal with the problems derived from the illumination. Our approach works

on a chromatic representation derived from a spherical coordinate interpretation of the RGB color space which is rather insensitive to achromatic variations in the image pixel colors. That means that we avoid shines introducing falsely detected regions. Dark image regions, such as object shadows, are characterized by the fact that small pixel color perturbations introduce strong chromatic shifts. Therefore, the chromatic representation is useless in these regions. Imitating the human vision system (HVS), we shift our computation of pixel color distance to a grayscale representation in such dark regions.

Noise: there are two main sources of image noise, introduced, on one hand, by thermal noise of the camera and robot motion, and, on the other hand, by lossy compression algorithms. The chromatic representation is rather insensitive to additive noise in high intensity regions. However, it is very sensitive to it in dark image regions. The proposed compound hybrid distance is robust in all cases.

Edge ambiguity: The proposed segmentation algorithm is region-oriented, therefore edge detection is obtained indirectly as the boundaries between detected regions. An edge appears where two neighboring pixels have different properties. Therefore, any edge detector is always based on the definition of a distance between pixel properties. If this distance is bigger than a given threshold, we can declare that there is an edge between the pixels, otherwise they belong to the same region. A chromatic representation helps to avoid spurious edge detection due to shines. Usually shines modify color in a neighborhood slowly whereas in true boundaries between surfaces the chromatic information changes dramatically.

Computing time: Most image segmentation methods are not designed with real time applications in mind, therefore they have high computational costs. However, for some applications like robotics, computing time is critical because these applications have to run in real time. Our algorithm performs only one-pass over the image, processing each pixel only once. The order of pixel processing is row-wise therefore we use the 4-WN neighborhood for pixel processing. The algorithm implementation that we have published does obtains real-time performance on off-the-shelf personal computers for small images.

In this chapter we present two contributions; first one works in a 4-NW neighborhood in order to obtain a good speed-up whereas the second one follows watershed strategy. Both contributions works with spherical coordinates taking profit of colorimetric properties of this approach and avoid strict dependence of the euclidean RGB representation. Finally both segmentations has a human validation of the results.

6.4 Image Segmentation on the Spherical Coordinate Representation of the RGB Color Space

This work presents an image segmentation algorithm working on the spherical interpretation of the RGB color space. The algorithm uses a hybrid chromatic distance inspired in the human vision system (HVS) that shifts its emphasis from the chromatic to the grayscale distance depending on the pixel's luminance value. For dark areas, the chromatic distance is too much sensitive and the gray scale distance is used instead. Color constancy properties of this segmentation can be easily deduced from the dichromatic reflection model (DRM). The segmentation doesn't need preprocessing steps, such as illuminant source color estimation. The approach is strongly robust regarding shines and dark spots, and it is amenable to work in real time on a robotic platform. We give results on benchmark databases and robot camera images. A public implementation is made available for independent test of the algorithm image segmentation results. Following the segmentation definition done in Sec.6.1, we assign one region label to each and all image pixels, where each label corresponds to a connected region characterized by a chromaticity vector, which is computed along with the segmentation. Therefore, two separated regions with the same chromatic representation will have two different labels. The algorithm's output are a bi-dimensional matrix of integer labels and a bidimensional matrix of chromatic vectors corresponding to the identified image regions.

This work introduces a new hybrid distance to measure the similarity between pixel colors, which is used in a one-pass pixel region labeling algorithm. This hybrid distance is a mixture of an intensity difference and a chromatic distance based on the spherical representation of the RGB color space, inspired in the sensitivity of the HVS. Its definition allows to parameterize the algorithm's noise tolerance, and to tune it for optimal color edge detection. Furthermore, it is easy to see that the chromatic component of this distance has some inherent color constancy, analyzing its behavior under the dichromatic reflection model (DRM). The labeling algorithm uses only the four north-west (4-NW) topological neighbors of the current pixel, because it process each pixel onces and does not perform relaxation processes. According to [85] "the image segmentation problem is basically one of psycho-physical perception, and therefore not susceptible to a purely analytical solution". The parameters of the mixture of the hybrid distance allow a fine tuning of the algorithm to the characteristics of the image being segmented.

To allow for users to test our algorithm we make available[1] a C# implementation, using Emgu [2] running on windows platforms.

[1] http://www.ehu.es/ccwintco/index.php/Hybrid_Image_Segmentation. In the sources, the method name corresponding with this paper is "fastSegmentation2"

[2] http://www.emgu.com/wiki/index.php/Main_Page

6.4.1 The Chromatic Distance

Image segmentation detect an edge between two neighboring pixels when they have different color properties [96]. The simplest edge detectors are based on spatial gradients of the image intensity, i.e. the Sobel or Prewitt convolution kernels. These approaches ignore the chromatic information leading to poor color edge detection. To improve this approach, the convolution masks can be extended to color representations using hybrid distances [55]. The core of all edge detection methods is the definition of an appropriate distance between pixel colors.

According to the DRM model presented in section **??**, the diffuse component is expressed by the angular components (θ, ϕ), which are almost constant in regions with homogeneous chromatic properties, and it is independent of the luminosity component, therefore independent of the illumination assuming an uniform chromatic illumination. If we use only the diffuse component for edge detection, black-white borders may be undetected because both colors belong to the same achromatic line, and hence they have the same chromaticity. On the other hand it is important to avoid the effect of shines, because they do not correspond to a true surface. Fig.6.1 illustrates the problem of false image regions in the image due to shines that do not correspond to real objects in the scene.

Figure 6.1: Image with shines

Fortunately, some differences exist between edge features due to shines and due to boundaries between diffuse color surfaces. When two surface regions of different chromatic properties are adjacent, this chromatic difference is clear and detectable measuring the difference on the zenithal and azimuthal angles of the spherical representation of pixel colors. In the shines, the chromaticity of pixels changes smoothly, offering the opportunity to filter the false detection setting a threshold on the chromatic distance based on the diffuse color component . For the human vision the main difficulty is to detect a color accurately in dark regions. In fact, chromaticity in dark regions is very unstable because it amplifies small color perturbations due to noise. According to the HVS, in regions with poor illumination it is more appropriate to use the luminance component. In the human eye's retina we have two kind of photoreceptor cells;

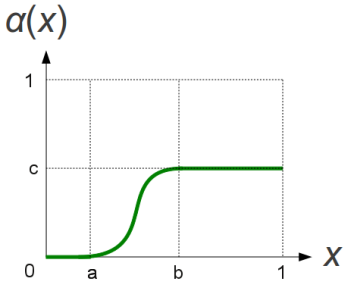


Figure 6.2: Chromatic activation function $\alpha(x)$

rods and cones. The first one is an luminance detector and the other one is a chromatic detector. Both need different energy for his activation. Rods need few energy for its activation, for this reason under poor illumination human vision becomes grayscale. Cones needs more energy, for this reason color are detected only with a good illumination.

Fig.6.2 shows the activation function $\alpha(x)$ of the chromatic distance component of the proposed hybrid distance. For luminance values below a, the chromatic component of the distance is inactive, for intensity values in the interval $[a, b]$, we smoothly change the contribution of the chromatic component of the hybrid distance from zero to its maximum $c \leq 1$ according to a sinusoidal function. Finally, for intensity values above b its contribution is always c. The three parameters a, b, c are in the range $[0, 1]$.

The function $\alpha(x)$ depends of the image intensity. The below Eq.6.1 is the mathematical expression of $\alpha(x)$:

$$\alpha(x) = \begin{cases} 0 & x \leq a \\ \frac{c}{2} + \frac{c}{2}\cos\left(\frac{(x-a)\cdot\pi}{b-a} + \pi\right) & a < x < b \\ c & x \geq b \end{cases} \tag{6.1}$$

We can define a hybrid distance between the color of two pixels p, q as follows:

$$d_H(p,q) = \left(1 - \alpha\left(\frac{l_p + l_q}{2}\right)\right) \cdot d_I(p,q) + \alpha\left(\frac{l_p + l_q}{2}\right) \cdot d_C(p,q) \tag{6.2}$$

where l_p, l_q are the intensity l in spherical coordinates of the pixels p and q respectively. d_I is a grayscale intensity distance computed as $d_I(p,q) = |l_p - l_q|$ and d_C is a chromatic distance computed as

$$d_C(p,q) = \sqrt{(\theta_q - \theta_p)^2 + (\phi_q - \phi_p)^2}. \tag{6.3}$$

The foregoing Eq.6.2 follows a HVS inspiration, where the first term express the behavior of the rod retinal cells which are sensitive to intensity. They need few energy for its activation. On the other hand, the second term express the

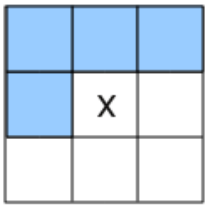


Figure 6.3: 4-WN Neighbors of pixel site x.

behavior of cone retinal cells which can detect chromaticity. They need more energy for its activation. Alfa function depends of the image intensity and the intensity of two pixels is the mean of both.

6.4.2 Proposed Segmentation Method

The segmentation method proposed in this section combines the spherical interpretation of the RGB space and the aforegoing hybrid distance expressed in the Eq.6.2. Edge detection can be accomplished using a threshold on the distance between pixels. In order to decrease the computing time, we use a 4-WN neighborhood as illustrated in Fig.6.3. This structure allows to obtain acceptable results processing only once each pixel. This method travels over the image by rows, hence along the computing the 4-WN neighbors are always labeled. Our segmentation method is explained by the following algorithm .

6.4.3 Algorithm

This algorithm returns a bi-dimensional integer matrix where each label is an integer. While performing the computation, we also need a structure that relates each label with a chromaticity and the amount of pixels labeled with it. That is necessary because each time that we assign a new pixel to a label we must actualize the chromaticity of this label, the chromaticity of a label is the mean chromaticity of all pixels labeled with it.

The most important parameter for this algorithm is the threshold δ. The granularity and the noise tolerance depends on it. For small threshold values we will obtain a lot of regions, and, conversely, with a high value we obtain big coarser regions. On the other hand, the parameters a, b, c of Eq.6.2 allow to tune the hybrid distance . If $b = 0$ and $c = 1$ it becomes a pure chromatic distance. If $a = 1$ it becomes a pure intensity distance. In any other cases it is a hybrid distance.

The algorithm 6.1 gives the details of our method. In this algorithm $L(x)$ denotes the region label of pixel x, $L_4(x)$ denotes the set of labels of the 4-WN neighbors of pixel x, that can be expressed as $L_4(x) = \bigcup_{x' \in N_4(x)} L(x')$, where $N_4(x)$ the 4-WN neighborhood of pixel x illustrated in Fig.6.3. The

algorithm may be applied to any color image $\Omega(x)$ represented in RGB spherical coordinates. It needs the specification of the distance $d_H(x,y)$ that gives a measure of the similarity between pixel colors $\Omega(x)$ and $\Omega(y)$. We build a map $\Psi_\ell = (\theta_\ell, \phi_\ell)$ assigning to region labeled ℓ a chromatic value. We denote R_ℓ the current set of pixels labeled in region ℓ, $R_\ell = \{x \,\text{s.t.}\, L(x) = \ell\}$, the number of pixels in a region is its cardinality $|R_\ell|$. Function newlabel creates a new region label. Function merge $(\ell_1, \ldots, \ell_n)$ creates the union of the regions $R_{\ell_1} \cup \ldots \cup R_{\ell_n}$ relabeling pixels accordingly with label ℓ, finally, merge returns the new label. The chromatic distance of Eq. (6.3) is extended to the chromatic representation of two regions of labels ℓ_1 and ℓ_2 as follows:

$$d_C(\ell_1, \ell_2) = \sqrt{(\theta_{\ell_1} - \theta_{\ell_2})^2 + (\phi_{\ell_1} - \phi_{\ell_2})^2}. \tag{6.4}$$

Algorithm 6.1 Image Segmentation Algorithm

Input: $\Omega(x)$ the color image in spherical coordinates
Threshold δ, distance parameters values a, b, c
-Initialize the pixel labels $\forall x; L(x) = \emptyset$
-The first region is composed of the first pixel $x_0 = (0,0)$: $L(x_0) =$ newlabel. The region chromaticity is that of the first pixel $\Psi_{L(x_0)} = \Psi(x_0)$.
for each x **do**
  if $L_4(x) = \{\ell\}$ /there is only one region label in $N_4(x)$/
    $d \leftarrow \min\{d_H(x,y) \,|\, y \in N_4(x)\}$
    if $d < \delta$ /some neighbor's color is similar enough/
      $L(x) = \ell$; /assign region label/
      $\Psi_\ell = \frac{1}{|R_\ell|} \sum_{y \in R_\ell} \Psi(y)$ /update region label chromaticity representation/
    else
      $L(x) =$ newlabel. /create a new region label/
      $\Psi_{L(x)} = \Psi(x)$. /update region label chromaticity/
  else
    $D \leftarrow \{d_C(L(y), L(z)) \,|\, y, z \in N_4(x) \;\&\; L(y) \neq L(z)\}$
    if $\min(D) < \delta$
      $L(x) = \text{merge}(L_4(x))$ /create region fusion label /
      $\Psi_{L(x)} = \frac{1}{|R_{L(x)}|} \sum_{y \in R_{L(x)}} \Psi(y)$ /update region chromaticity rep./
    else /regions can not be merged/
      $d \leftarrow \min\{d_H(x,y) \,|\, y \in N_4(x)\}$
      if $d < \delta$ /assign to region with the lower chromatic distance/
        $L(x) = L(y)$ s.t. $d_H(x,y) = d$;
        $\Psi_{L(x)} = \frac{1}{|R_{L(x)}|} \sum_{y \in R_{L(x)}} \Psi(y)$ /update region chromaticity representation/
      else /current pixel can not be assigned to existing regions/
        $L(x) =$ newlabel. /create a new region label/
        $\Psi_{L(x)} = \Psi(x)$. /update region label chromaticity/
end for

6.4.3.1 Experimental results

In this section we present some computational results of the algorithm described in section 6.4.3. The algorithm parameter values were: $\delta = 0.02$, $a = 0.2$, $b = 0.4$ and $c = 0.5$. With these values for the dark regions (intensity less than .2) we use only the intensity distance, whereas for pixels with an intensity greater than .4 we use a hybrid distance where intensity and chromaticity have the same importance. In order to validate our approach, we have tested the proposed algorithm firstly with the well-known Berkeley database [99], and, secondly, with images obtained from the camera of a real robot Nao, Aldebaran Robotics, Paris, because we have developed the algorithm with robotic applications in mind. The main features of the images taken by the robot on-board camera are, firstly, the low signal to noise ratio, due to the poor quality of this cameras and, secondly, the appearance of many shines in the images due to illumination sources of the real environment where the robot is working. For a visual assessment of the results, the output images are pictured using the mean color image of each region.

6.4.3.2 Results on the Berkeley database

Fig.6.4 shows the segmentation results on some images from the Berkeley image database [99]. The first row show original pictures as provided with the database. The second row shows segmentation results using the chromatic distance of Eq.6.3. We present results using the hybrid distance of Eq.6.2 in the third row. Comparison of images in those rows show the improvement obtained using the hybrid distance, obtaining more natural segmentations, specially in shadowy regions, like the tree or the sky in the right-most image. Smooth regions in the images are identified as homogeneous regions despite small color fluctuations and brightness. The fourth row shows the edges between regions identified in the images of the third row. Finally, the last row shows the human segmentation provided with the Berkeley data-set. Comparison between these edge images must take into account that the human edge delineation is an idealization of the actual image, drawing regions whose identification involve semantic processing of the image. Nevertheless, our approach captures most of the salient image partitions. Small detail edges appear in regions with randomized textures such as the tree, forest or rocky soil. Notice that in the chromatic images of the third row this over-segmentation is less apparent, suggesting that post-processing the chromatic segmentation image those textures could be removed if desired. Smooth regions, like the river in the left-most image or the wave in the surf image are identified quite closely to the hand made edge delineations, minimizing spurious detection due to shines.

6.4.3.3 Results on NAO's camera images

The images obtained from the Nao robot's camera are characterized by strong illumination effects and high noise ratios due to the poor camera quality and

Figure 6.4: Segmentation of Berkeley data-set images using the chromatic distance of Eq.6.3 (second row) and the proposed hybrid distance (third row) of Eq.6.3

robot motion. In Fig.7 we show the segmentation results for some original images. The left column shows the original images. The middle column shows the color region segmentations, and the right column shows the edges between identified regions. The most salient feature of the segmentation results obtained on these images is that most of the shines on the floor are not identified as distinct image regions and the object shadows are ignored as well. All the segmented regions correspond to actual objects in the scene. For robotic applications, this robustness may be critical for task accomplishment. Although it was not the main goal of our work at this point, to obtain real time performance (of the order of 50 milliseconds per image) we downsampled the images to 80×60 pixels. This experiment has been carried out in a laptop with a processor Intel Core i3 M330 with 4GB of memory. The code has been written in C#.

6.4.4 Conclusions

In this work we have presented a novel image region segmentation method, with the following features: (a) It is a one-pass method which can achieve real time performance for small images. (b) It doesn't need an image preprocessing for edge detection contrary to other approaches. (c) It is grounded in the dichromatic reflection model and, therefore, it has a physical modeling support. (d) It uses a hybrid distance inspired in the HVS whose parameters can be tuned to optimize segmentation for different image conditions. (e) It avoids spurious region detection due to shines produced by illumination, detecting color edges with a physical interpretation as boundaries of surfaces with different reflectances.

Even thought the returned label matrix is the desired image segmentation information, we maintain that the chromatic information is also important for further processes. A region label is associated with a chromaticity value which must be close to the true chromatic characterization of the imaged surface. This chromatic information is a photometric invariant, which will be used in future works to implement new algorithms for visual saliency and unsupervised learning of objects.

Regarding real-time performance, current experiments on the NAO robot off-loading the image segmentation to an auxiliary workstation give real-time responses for small image frames (20 frame/second).

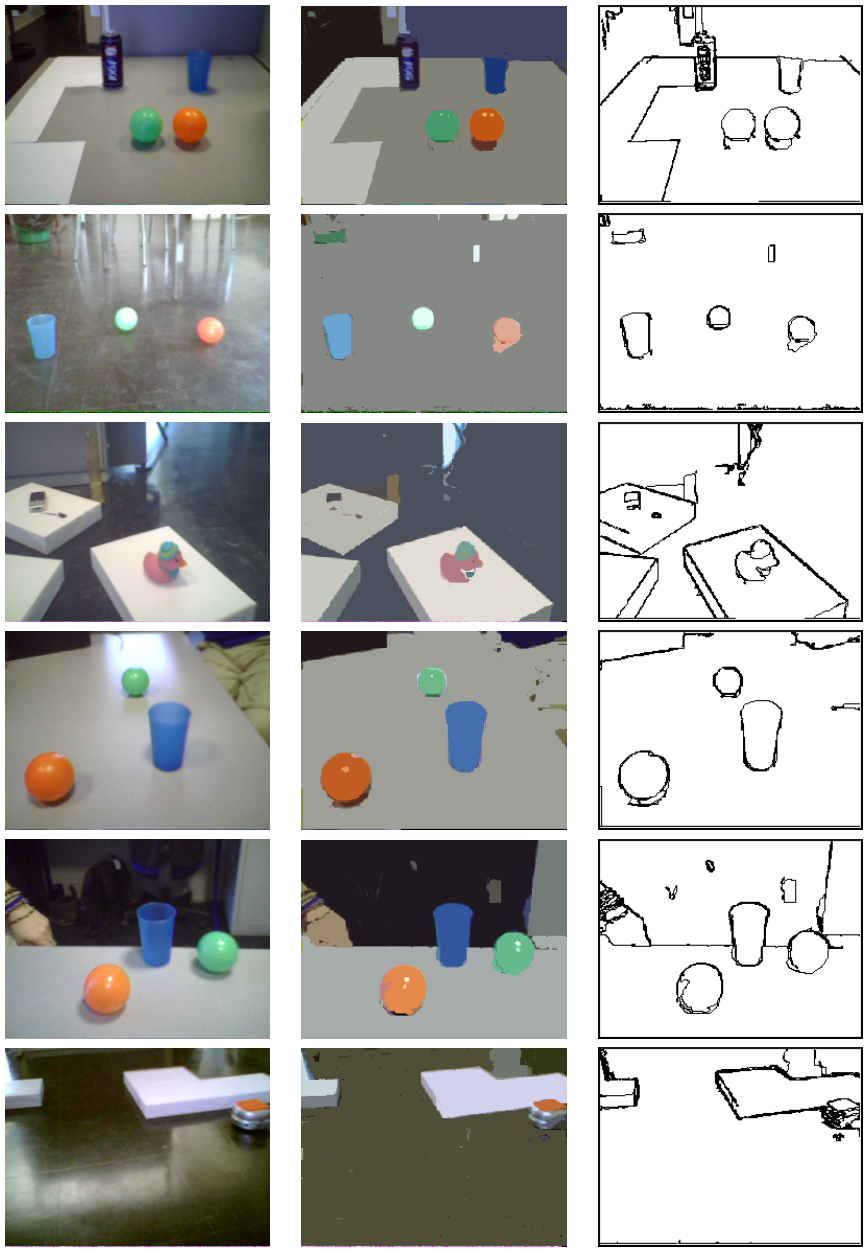


Figure 6.5: Segmentation of the images captured by the Nao Robot camera, using the hybrid distance of Eq.6.2.

6.5 A robust color Watershed transformation and image segmentation defined on RGB Spherical Coordinates

The representation of the RGB color space points in spherical coordinates allows to retain the chromatic components of image pixel colors, pulling aparteasily the intensity component. This representation allows the definition of a chromatic distance and a hybrid gradient with good properties of perceptual color constancy. In this work we present a watershed based image segmentation method using using this hybrid gradient. Over-segmentation is solved by applying a region merging strategy based on the chromatic distance defined on the spherical coordinate representation. We show the robustness and performance of the approach on well known test images and the Berkeley benchmarking image database.

Color images have additional information over grayscale images that may allow the development of robust segmentation processes. There have been works using alternative color spaces with better separation of the chromatic components like HSI, HSL, HSV, Lab [100, 93] to obtain perceptually correct image segmentation. However, chromaticity's illumination can blur and distort color patterns.

Color constancy is closely related to the response of the gradient operators [17]. Regions of constant color must have low gradient response, while color edges must have a strong gradient response. Image segmentation methods based on spatial gradients need a correct definition of the spatial color gradient and unambiguous contour definition. In fact, formulation of watershed segmentation methods in color images is still an open research issue. A straightforward but inexact approach is the independent application of the watershed segmentation on image channel [101]. This approach loses chromatic information, and has difficulties merging the subsequent independent segmentations into one.

In this work we will use the RGB spherical coordinates representation to archive color constancy properties of our image segmentation approach [55, 75, 56]. We define a chromatic distance on this representation. The robustness and color constancy of the approach is grounded in the dichromatic reflection model (DRM) [37]. We propose a chromatic gradient operator suitable for the definition of a watershed transformation on color images and a robust region merging for meaningful color image segmentation. The baseline chromatic gradient operator [55, 56] suffers from noise in the dark areas of the image. We propose in this work a hybrid gradient operator overcoming this problem and we use it to build a watershed transformation on color images. To achieve a natural segmentation, we perform region merging on the basis of our proposed chromatic distance over the chromatic characterization of the watershed regions. We give a general schema that combines watershed flooding with region merging in a single process. Finally, we specify our proposal as an instance of the aforementioned general schema.

6.5.1 Gradient operators

Here, we remember the **chromatic** gradient definition presented on the previous chapter.

The row pseudo-convolution is defined as

$$CG_R\left(P\left(i,j\right)\right) = \sum_{r=-1}^{1} d_C\left(P\left(i-r,j+1\right), P\left(i-r,j-1\right)\right),$$

and the column pseudo-convolution is defined as

$$CG_C\left(P\left(i,j\right)\right) = \sum_{c=-1}^{1} d_C\left(P\left(i+1,j-c\right), P\left(i-1,j-c\right)\right),$$

where d_C is the chromatic distance of Eq. 6.3. So, the color distance between pixels substitutes the intensity subtraction of the Prewitt linear operator. The color gradient image is computed as:

$$CG(x) = CG_R\left(x\right) + CG_C\left(x\right) \tag{6.5}$$

And the **intensity** gradient definition is:

The row pseudo-convolution is defined as

$$G_R\left(J\left(i,j\right)\right) = \sum_{r=-1}^{1} \|J\left(i-r,j+1\right) - J\left(i-r,j-1\right)\|,$$

and the column pseudo-convolution is defined as

$$G_C\left(J\left(i,j\right)\right) = \sum_{c=-1}^{1} \|J\left(i+1,j-c\right) - J\left(i-1,j-c\right)\|,$$

where J is the intensity image, then the intensity gradient image is computed as:

$$G(x) = G_R\left(x\right) + G_C\left(x\right) \tag{6.6}$$

6.5.2 Hybrid Gradient

Empirical experiments show that the aforegoing chromatic gradient is very susceptible to image noise. The angular distance of Eq.6.3 is more sensitive to noise for pixel colors lying close to the origin in RGB space. This is due to the fact that the angular distance between two points at a given euclidean distance grows as the points are closer to the origin. Small perturbations as measured by the euclidean distance are mapped into big angular differences. The background noise which has little effect in lighted regions is amplified in the dark regions.

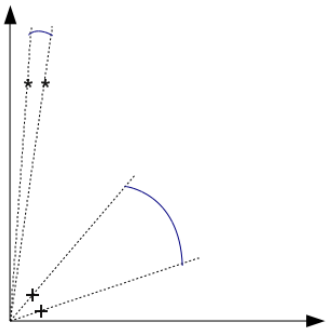


Figure 6.6: Effect of distance to the origin in the angular distance for pairs points at the same euclidean distance.

Inspired in HVS, we propose a hybrid gradient which is an intensity gradient when the illumination is poor, and a chromaticity gradient in better illuminated image regions. For intensity values below a threshold a it is an intensity gradient, for values above another threshold b it is a chromatic gradient, and for values between both it is a mixture of the two kinds of gradients whose mixing coefficient is sinusoidal function of the image intensity. This idea is expressed mathematically as a convex combination of the two gradient operators:

$$HG(x) = \beta(x)\, G(x) + \overline{\beta}(x)\, CG(x) \tag{6.7}$$

where x is the pixel location, $G(x)$ is the intensity gradient magnitude of Eq.6.6, $CG(x)$ is the chromatic gradient of Eq.6.5 and $\overline{\beta}(x) = 1 - \beta(x)$, hence $\overline{\beta}(x) + \beta(x) = 1$, $\beta(x)$ is normalized to the range $[0, 1]$ which corresponds with Eq.6.1 of the previous work.

$$\beta(x) = \begin{cases} 1 & I(x) < a \\ \frac{1}{2} + \frac{\cos(\frac{x-a}{b-a}\pi)}{2} & a \leq I(x) < b \\ 0 & b \leq I(x) \end{cases}, \tag{6.8}$$

where $I(x)$ is the pixel intensity.

Note that by difference with the α mixing function presented in the previous work, this β function has not the c parameter, therefore this hybrid function is a chromatic gradient or a intensity gradient except in the range [a,b]. Fig.6.7 shows the activation of the intensity gradient depending of the intensity.

This hybrid gradient does not suffer from noise sensitivity in dark regions of the image, the effect of bright spots is reduced because it is chromatically consistent in bright image regions, and it detects chromatic edges.

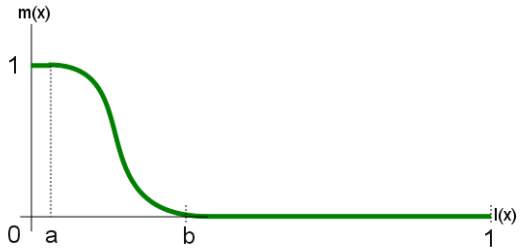
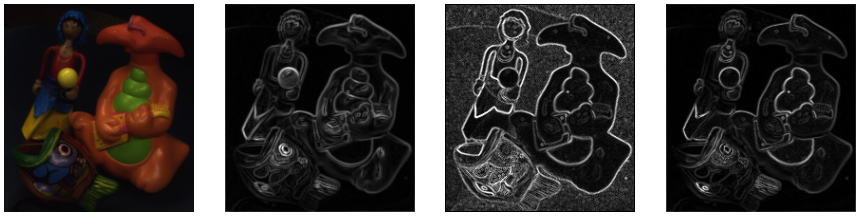


Figure 6.7: Hybrid gradient convex combination factor as a function of the image intensity.



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

Figure 6.8: Response of different gradient operators. (a) original image (b) intensity gradient, (c) chromatic gradient, (d) hybrid gradient.

Fig.6.8 shows the response of different gradient operators on the same test image. Fig. 6.8(a) presents the original image. Fig. 6.8(b) contains the response of the intensity gradient, it shows false border detection in some diffuse regions, i.e. the yellow ball, green thorax. It also shows false edge detection in bright spots. Fig.6.8(c) shows the response of the chromatic gradient operator. It does not give false edge detection inside diffuse regions. It does not give false edge detection in bright spot areas. However, it is very sensitive to noise in the dark regions, showing false edge detections due to small random variations. Fig.6.8(d) presents the response of the hybrid gradient which has the good detection properties of the chromatic gradient operator and it is not sensitive to noise in the dark image regions.

6.5.3 Watershed

Watershed transformation is a powerful mathematical morphology technique for image segmentation. It was introduced in image analysis by Beucher and Lantuejoul [70], and subsequently a lot of algorithm variations and applications have been proposed [101, 72, 97].

The watershed transform considers a bi-dimensional image as a topographic relief map. The value of a pixel is interpreted as its elevation. The watershed lines divide the image into catchment basins, so that each basin is associated

with one local minimum in the topographic relief map. The watershed transformation works on the spatial gradient magnitude function of the image. The crest lines in the gradient magnitude image correspond to the edges of image objects. Therefore, the watershed transformation partitions the image into meaningful regions according to the gradient crest lines.

The baseline watershed transformation computation method is as follows. The image pixel sites with local minimum gradient are selected as the sources of their respective catchment basins. A flooding process fills each catchment basin from its respective source. When a catchment basin is full, the contour points which are in touch with a neighbor catchment basin are the dam points. The process is finished when all gradient surface is covered. The closed lines defined by the dam points give us the watershed transformation, and, implicitly, the image partition. Usually, this partition is very fine, therefore a further step of region merging is needed to obtain partitions closer to the natural segmentation of the image. Region merging needs the specification of when and why two neighboring catchment basins are merged into one region. In other words, region merging criterion defines which watershed lines are going to be removed. Watershed regions are image regions with homogeneous properties. We look for homogeneous chromatic regions on the basis of the aforegoing hybrid chromatic gradient.

6.5.4 General schema of a watershed method

The general schema of the watershed method perform a flooding process which performs a region growing based on the ordered examination of the level sets of the gradient image. In fact, an ordered succession of thresholds are applied to produce the progression of the flooding. The image is examined iteratively n times, each iteration step the threshold is raised and pixels of the gradient image below the new threshold are examined to be labeled with a corresponding region. Initially each region will contain the source of its catchment basin when the flooding level reaches it. Each flooded region is also characterized by a chromaticity value, that corresponds to the source pixel chromaticity. This chromaticity value is used to perform region merging simultaneously with the flooding process. A pixel whose neighboring pixels belong to different regions is a watershed pixel. When a watershed pixel is detected, the adjacent regions may be merged into one if the chromatic distance between the region chromatic values is below a chromatic threshold. The merged region chromatic value is the average of that of the merged regions. The final labeling of the image regions is performed taking into account the equivalences established by the merging process. Watershed pixels whose adjacent regions do not merge into one are labeled as region boundary pixels and retain their chromaticity.

The algorithm 6.2 gives the details of our method. In this algorithm $L(x)$ denotes the region label of pixel x, $L_8(x)$ denotes the set of labels of the 8 neighbors of pixel x, that can be expressed as $L(x) = \bigcup_{x' \in N_8(x)} L(x')$, where $N_8(x)$ the 8-th neighborhood of pixel x. The algorithm may be applied to any color image $\Omega(x)$ and gradient magnitude image $\Phi(x)$. The algorithm needs

the specification of a chromatic distance $\Delta(\Omega(x), \Omega(y))$ that gives a measure of the similarity between pixel colors $\Omega(x)$ and $\Omega(y)$. To label the regions we keep a counter R, and we build a map Ψ_R assigning to each region label a chromatic value. While the flooding process performs region growing, the region chromatic value is updated to the average chromaticity of the pixels in the region. Each region R has a corresponding chromatic value Ψ_R which can be used for visualization.

6.5.5 The proposed approach

Our color image segmentation process proposal can be precisely specified by Algorithm 6.2 applied on the zenithal and azimuthal angles of the color representation $\mathbf{P}(x)$ of Eq., the gradient magnitude image computed by the hybrid gradient $HG(x)$ of Eq.6.7, using the chromatic distance of Eq.6.5

6.5.6 Experimental results

The watershed-merge Algorithm 6.2 is parametrized by:

- The number of iterations n, which determines the resolution of the flooding process going over the gradient magnitude image level sets.
- The gradient operator used to compute the gradient magnitude image, which can be either the intensity gradient $G(x)$ of Eq. 6.6 or the hybrid gradient $HG(x)$ of Eq.6.7.
- The color representation of the image. Assuming the RGB space, it can be either the Cartesian representation $\mathbf{I}(x)$ or the zenithal and azimuthal angles of the Spherical representation $\mathbf{P}(x)$. This selection determines the selection of the chromatic distance.
- The Chromatic distance, which can be either the Euclidean distance in the RGB Cartesian space, or the chromatic distance of Eq.6.3.
- The Chromatic distance threshold δ, which determines the chromatic resolution of the region merging process.

This section reports results of two experiments, the first one compares our proposal of section 6.5.5 with other instances of the algorithm, whereas in the second one we will to provide a more extensive qualitative validation our method using the well know Berkeley benchmark image collection[99] which provides hand-draw artistic shape boundaries.

6.5.6.1 Behavior

In this section we will use a well known benchmark image [58] to compare our proposed segmentation process with variations of Algorithm 6.2 obtained with other parameter settings. The dark regions are critical to the perceptually

Algorithm 6.2 General scheme of watershed and region merge for color image segmentation.

Set number of iterations n, the chromatic distance threshold δ, initialize the pixel labels $\forall x; L(x) = \emptyset$, the region label counter $R = 0$,
$\Omega(x)$ is the color image, $\Phi(x)$ is the gradient magnitude image.

1. Calculate $\Phi_{\min} = \min_x \{\Phi(x)\}$ and $\Phi_{\max} = \max_x \{\Phi(x)\}$. Calculate the step at each interaction $s = (\Phi_{\max} - \Phi_{\min})/n$
   Initialize $t = \Phi_{\min}$;

2. Iterate n times, setting

   (a) Calculate threshold $t = t + s$.

   (b) Consider $X'(t) = \{x' | \Phi(x') < t\}$, for each $x \in X'(t)$ perform:

      i. If $L(x) = \emptyset$ the pixel is unprocessed, then one of the following cases apply

         A. If $L_8(x) = \emptyset$
            - Assign new label $R \leftarrow R + 1$; $L(x) = R$.
            - Assign the region chromatic value $\Psi_R = \Psi(\Omega(x))$.

         B. If $|L_8(x)| = 1$
            - $L(x) = L_8(x)$
            - Update $\Psi_{L(x)}$ using $\Psi(\Omega(x))$.

         C. If $|L_8(x)| > 1$ there are at least two adjacent regions, x is a gradient watershed pixel. Consider all pairs of adjacent regions of labels r_1 and r_2
            - If $\triangle(\Psi_{r_1}, \Psi_{r_2}) \leq \delta$ then we can merge both regions into one of label r^*. Compute $\Psi_{r^*} = \left(\frac{|r_1|\Psi_{r_1} + |r_2|\Psi_{r_2}}{|r_1| + |r_2|}\right)$. We keep record of the detected equivalence. Update Ψ_{r^*} using $\Psi(\Omega(x))$. $L(x) = r^*$.
            - If $\triangle(\Psi_{r_1}, \Psi_{r_2}) > \delta$ the pixel x is a region boundary pixel with a special label $L(x) = b$.

3. From the recorded label equivalences compute the final region labels, and assign definitive labels.

4. Each region R has a corresponding chromatic value Ψ_R which can be used for visualization

correct gradient computation, while the bright spots may induce false edge detection. The algorithm does not compute any specular free image to remove this latter problem.

The operational parameter setting are $n = 100$ and $\delta = 0.1$. In Fig.6.9 we show the segmentation results on this image for all combinations of the remaining Algorithm 6.2 parameter settings. The column of images labeled "Gradient" has the gradient magnitude images. From top to bottom, Fig.6.9(a), 6.9(e), 6.9(i) show, respectively the result of the intensity gradient, the chromatic gradient of equation Eq.5.4, and the hybrid gradient of Eq.6.7. The column of images labeled "Watershed" correspond to the image region partition performing only to the flooding process, without any region merging, on the corresponding gradient magnitude images. It can be appreciated that the hybrid gradient watershed removes most of the dark microregions originated by the chromatic gradient. There are, however, some regions with different colors in this rough dark region which are not fully identified by the intensity gradient watershed of Fig.6.9 (b) and are better detected by the hybrid gradient watershed in Fig.6.9(j). The two image columns with the heading "segmentation" show the results of the region merging from the corresponding gradient watershed in the same row. The left column shows the results of using of the Euclidean distance on the RGB Cartesian coordinates. The right segmentation column show the results of the using the chromatic distance of Eq.**??**. If we want to ascertain the effect of the color representation and the chromatic distance we must compare the rightmost columns in Fig.6.9. We find that the general effect is that the chromatic distance on polar coordinates is better identifying the subtle color regions in the darkest areas of the image, it detects better the shape of the objects, has better color constancy properties, and it is much less sensitive to bright spots or shining areas. Comparing the gradient operators attending to the final segmentation we observe that the hybrid gradient is better than the others in removing noise from the dark regions and maintain the object integrity. Overall the best result is obtained with our proposal as shown in Fig.6.9(l), where we can easily identify the subtle regions in the upper dark area, the shadow of the lowermost object, and we can clearly identify object with the same color unaffected by shading and bright spots.

6.5.6.2 Validation on the Berkeley images

In the Fig.6.10 we can show the experimental results using the Berkeley DB [99]. The first and fourth rows shows the original images, the second and fifth shows our respective outputs, whereas the third and sixth rows shows the human segmentation reference. As we can see our method gives always homogeneous regions, and the segmentation output is close to the human segmentation. Some facts that we find comparing our segmentation with the hand-drawn segmentation:

- Large chromatically smooth regions are well segmented by our approach despite variations in intensity, e.g. the face skin of the portraited man, the river, the road in the road race image.

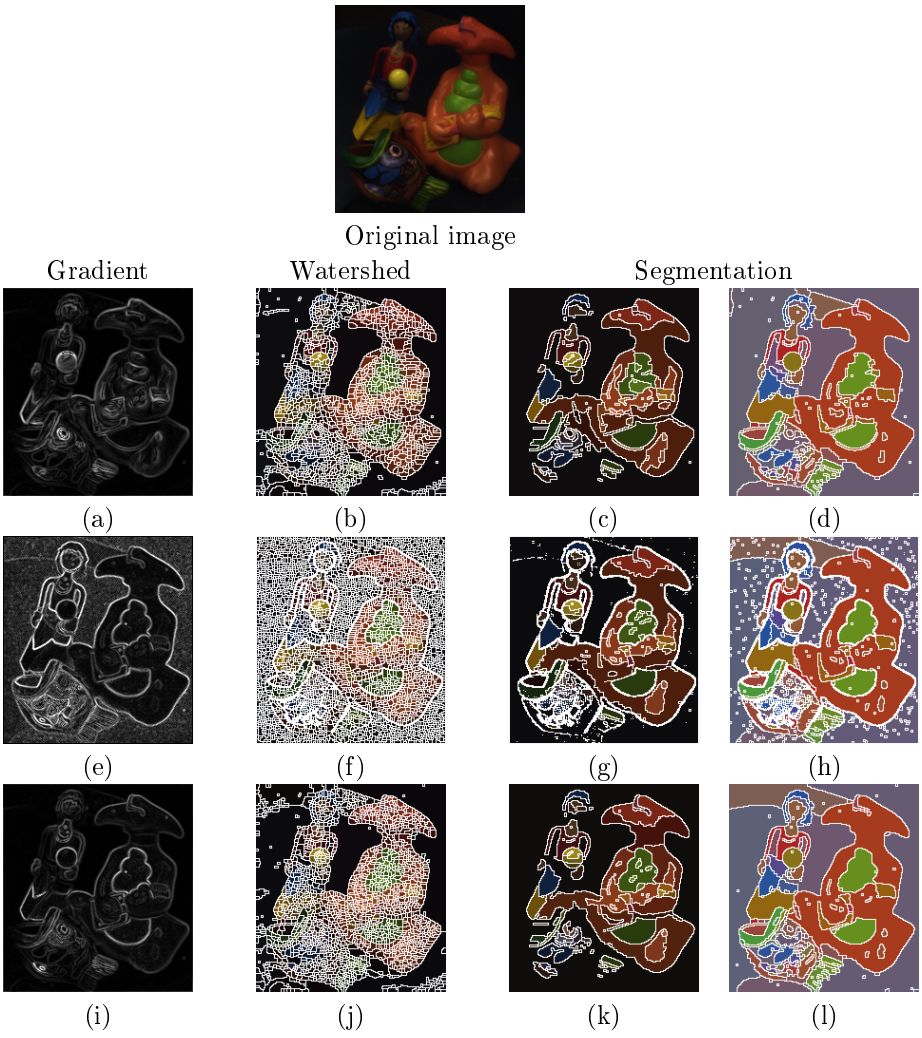


Figure 6.9: Image segmentation results with different parametrizations of Algorithm 6.2.

- Some subtle chromaticity variations are detected and segmented, like the reflections in the water of the jungle river image.
- The algorithm does not use any spatial information to segment textured objects. However it can cope with some textured spatial intensity variations of chromatically constant regions, outlining the corresponding object, i.e. the clouds in the flying plane image, the yellow skirt in the jungle river image.
- The hand-draw contours obviate some regions of the image that the artist may have found irrelevant, i.e. the clouds in the sky in some images, the texture details of some bushes. Some of these regions can not be segmented as a unit unless some spatial texture information is used, like the bushes in the jungle river image, or the skyscraper windows.

6.5.7 Work Conclusions

This work introduces a watershed and region merging segmentation algorithm based on the zenithal and azimuthal angles of the spherical representation of colors in the RGB space. We have shown that this representation is equivalent to the chromaticity representation of the color. Considering the DRM we find that most of the diffuse reflectance is preserved. Moreover the color representation and the chromatic distance defined on it possesses has color constancy properties. These definitions allows the construction of a robust hybrid chromatic spatial gradient that we use to realize a robust chromatic watershed segmentation. This gradient operator has good color edge detection in lightened areas and does not suffer from the noise in the dark areas. The watershed is complemented by a region merging based on the defined chromatic distance. We give a general schema of the algorithm performing both watershed and region merging. Our proposal can be stated by this algorithm fixing the color representation, gradient operator, and region merging distance. We compare our approach with other algorithms obtained with different setting of the general schema, obtaining the best qualitative segmentation. The results on the Berkeley database images find excellent approximations to the provided hand-drawn segmentations, without using spatial or semantic information.

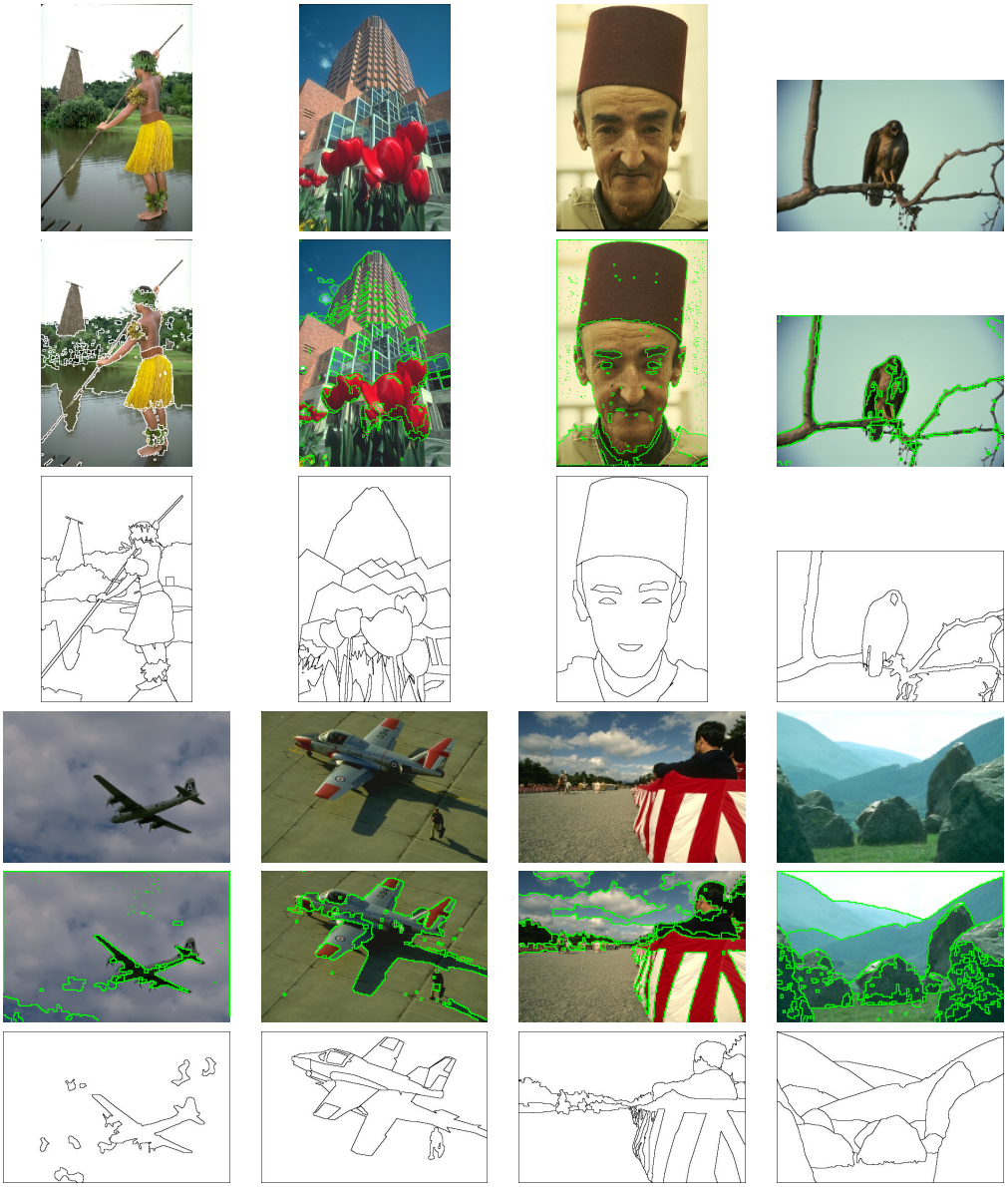


Figure 6.10: Segmentation results on some of the Berkeley images. Second and fourth rows show the results of our approach. Third and last row show the hand-drawn shapes.

6.6 Chapter Conclusions

In this chapter we are discussed about image segmentation, making a brief survey of some strategies for image segmentation. After that, we have presented two innovative image segmentation methods, first one in Sec.6.4 and second one in Sec.6.5. Both methods has somethings in common:

- they are grounded on DRM having a physical support
- they use the spheric approach introduced too in this thesis
- they have a good behavior avoiding shines and shadows effect
- they belong to *region based* segmentation methods
- they are fast and can be used for real time applications.

By other hand they have some differences:

- they use different hybridization equations α and β respectively
- first method follow a strict order by rows whereas second one follows an algorithm which depending of the image the seeds are going to be located at different positions.
- first method works in the 4-NW neighborhood whereas the second one works on the 8th neighborhood
- due to the neighborhood first method looks faster than second one.

Part II

Contributions on Hyperspectral Images

6.7 Background

6.8 Hyperspheres

Chapter 7

Gradients

7.1 State of the Art

7.2 A chromatic Gradient

7.3 A Hybrid Gradient for n-Dimensional Images

Chapter 8

Segmentation

Part III

Glosary & References

Chapter 9

Glosary

Albedo of an object is the extent to which it diffusely reflects light from the sun. It is therefore a more specific form of the term reflectivity. Albedo is defined as the ratio of diffusely reflected to incident electromagnetic radiation. It is a unit less measure indicative of a surface's or body's diffuse reflectivity. The word is derived from Latin albedo "whiteness", in turn from albus "white". The range of possible values is from 0 (dark) to 1 (bright).

Azimuth is the horizontal angular distance from a reference direction, usually the northern point of the horizon, to the point where a vertical circle through a celestial body intersects the horizon, usually measured clockwise. Sometimes the southern point is used as the reference direction, and the measurement is made clockwise through 360 grades.

Chromaticity is an objective specification of the quality of a color irrespective of its luminance, that is, as determined by its colorfulness (or saturation, chroma, intensity, or excitation purity) and hue.

Colorfulness, chroma, and saturation are related concepts referring to the intensity of a specific color. More technically, colorfulness is the perceived difference between the color of some stimulus and gray, chroma is the colorfulness of a stimulus relative to the brightness of a stimulus that appears white under similar viewing conditions, and saturation is the colorfulness of a stimulus relative to its own brightness. Though this general concept is intuitive, terms such as chroma, saturation, purity, and intensity are often used without great precision, and even when well-defined depend greatly on the specific color model in use.

Diffuse reflection is the reflection of light from an uneven or granular surface such that an incident ray is seemingly reflected at a number of angles. It is the complement to specular reflection. If a surface is completely non specular, the reflected light will be evenly spread over the hemisphere

surrounding the surface (2π steradians). The most familiar example of the distinction between specular and diffuse reflection would be matte and glossy paints as used in home painting. Matte paints have a higher proportion of diffuse reflection, while gloss paints have a greater part of specular reflection.

Fresnel equations, deduced by Augustin-Jean Fresnel, describe the behavior of light when moving between media of differing refractive indices. The reflection of light that the equations predict is known as Fresnel reflection.

Hue is one of the three main attributes of perceived color, in addition to lightness and chroma (or colorfulness). Hue is also one of the three dimensions in some color spaces along with saturation, and brightness (also known as lightness or value). Hue is that aspect of a color described with names such as "red", "yellow", etc.

Insolation (Incident solar radiation) is a measure of solar radiation energy received on a given surface area in a given time. It is commonly expressed as average irradiance in watts per square meter (W/m^2) or kilowatt-hours per square meter per day ($\frac{kW*h}{m^2*day}$) , or in the case of photovoltaic it is commonly measured as $kWh/kWp*y$ (kilowatt hours per year per kilowatt peak rating). Sometimes, as in the text below, a long-term average intensity of incoming solar radiation will be given in units such as watts per square meter (W/m2 or W*m-2) and called insolation, with the duration (such as daily, annual, or historical) stated or only implied.

Irradiance, radiant emittance, and radiant exitance are radiometry terms for the power of electromagnetic radiation at a surface, per unit area. "Irradiance" is used when the electromagnetic radiation is incident on the surface. "Radiant exitance" or "radiant emittance" is used when the radiation is emerging from the surface. The SI units for all of these quantities are watts per square meter ($W*m^{-2}$).

Lambert's cosine law in optics says that the radiant intensity observed from a "Lambertian" surface is directly proportional to the cosine of the angle θ between the observer's line of sight and the surface normal. The law is also known as the cosine emission law or Lambert's emission law. It is named after Johann Heinrich Lambert, from his Photometria, published in 1760.

Lambertian reflectance, light falling on it is scattered such that the apparent brightness of the surface to an observer is the same regardless of the observer's angle of view. More technically, the surface luminance is isotropic. For example, unfinished wood exhibits roughly Lambertian reflectance, but wood finished with a glossy coat of polyurethane does not (depending on the viewing angle, specular highlights may appear at different locations on the surface). Not all rough surfaces are perfect Lambertian reflectors,

but this is often a good approximation when the characteristics of the surface are unknown. Lambertian reflectance is named after Johann Heinrich Lambert.

Luminance is a photometric measure of the density of luminous intensity in a given direction. It describes the amount of light that passes through or is emitted from a particular area, and falls within a given solid angle. The SI unit for luminance is candela per square met re (cd/m2). The CGS unit of luminance is the stilb, which is equal to one candela per square centimeter or 10 kcd/m2.

Radiance and spectral radiance are radiometric measures that describe the amount of light that passes through or is emitted from a particular area, and falls within a given solid angle in a specified direction. They are used to characterize both emission from diffuse sources and reflection from diffuse surfaces. The SI unit of radiance is watts per steradian per square meter ($W * sr^{-1} * m^{-2}$).

Reflectivity is the fraction of incident radiation reflected by a surface. In full generality it must be treated as a directional property that is a function of the reflected direction, the incident direction, and the incident wavelength. However it is also commonly averaged over the reflected hemisphere to give the hemispherical spectral reflectivity: $\rho(\lambda) = \frac{G_{refl}(\lambda)}{G_{incid}}$ where Grefl(λ) and Gincid(λ) are the reflected and incident spectral (per wavelength) intensity, respectively.

Reflection is the change in direction of a wave front at an interface between two different media so that the wave front returns into the medium from which it originated. Common examples include the reflection of light, sound and water waves.

Snell's law (also known as Descartes' law or the law of diffraction), is a formula used to describe the relationship between the angles of incidence and refraction, when referring to light or other waves, passing through a boundary between two different isotropic media, such as water and glass. The law says that the ratio of the sines of the angles of incidence and of refraction is a constant that depends on the media.

Specular reflection is the perfect, mirror-like reflection of light (or sometimes other kinds of wave) from a surface, in which light from a single incoming direction (a ray) is reflected into a single outgoing direction. Such behavior is described by the law of reflection, which states that the direction of incoming light (the incident ray), and the direction of outgoing light reflected (the reflected ray) make the same angle with respect to the surface normal, thus the angle of incidence equals the angle of reflection; this is commonly stated as $\theta i = \theta r$.

Zenith is the direction pointing directly above a particular location (perpendicular, orthogonal). Since the concept of being above is itself somewhat vague, scientists define the zenith in more rigorous terms. Specifically, in astronomy, geophysics and related sciences (e.g., meteorology), the zenith at a given point is the local vertical direction pointing away from direction of the force of gravity at that location.

Bibliography

[1] K. Hara, K. Nishino, and K. Ikeuchi, "Light source position and reflectance estimation from a single view without the distant illumination assumption.," *IEEE Trans Pattern Anal Mach Intell*, vol. 27, pp. 493–505, Apr 2005. 10, 23

[2] I. Sato and K. Ikeuchi, "Illumination from shadows," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 25, pp. 290–300, 2003. 11

[3] R. Moreno, M. Graña, and A. d'Anjou, "Illumination source chromaticity estimation based on spherical coordinates in rgb," *Electronics Letters*, vol. 47, no. 1, pp. 28–30, 2011. 11, 22, 23, 24

[4] T. Tan, K. Nishino, and K. Ikeuchi, "Illumination chromaticity estimation using inverse-intensity chromaticity space," in *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on*, vol. 1, pp. I–673–I–680vol.1, 18-20 June 2003. 11, 16, 23, 26, 29, 35

[5] Y.-J. C. Kuk-Jin Yoon and I. S. Kweon, "Illuminant chromaticity estimation using dichromatic slope and dichromatic line space," in *Korea-Japan Joint Workshop on Frontiers of Computer Vision*, pp. 219–224, FCV, 2005. 11, 23, 29, 30

[6] c. H. Marc Ebner, "On determining the color of the illuminant using the dichromatic reflection model," in *Pattern Recognition* (S. B. . Heidelberg, ed.), vol. 3663/2005, pp. 1–8, 2005. 11, 23, 29

[7] D. A. Forsyth and J. Ponce, *Computer Vision: A Modern Approach*. Prentice Hall, US ed ed., Aug. 2002. 11

[8] J. Lambert, *Photometria sive de mensure de gratibus luminis, colorum umbrae*. Eberhard Klett, 1760. 12

[9] T. Gevers and A. W. M. Smeulders, "Color-based object recognition," *Pattern Recognition*, vol. 32, pp. 453–464, Mar. 1999. 13

[10] P. Skorupski and L. Chittka, "Is colour cognitive?," *Optics & Laser Technology*, vol. In Press, Corrected Proof. 13, 16, 22, 62

[11] D. Pascale, "A review of rgb color spaces," 2003. http://www.babelcolor.com/download/Af. 13

[12] J. C. Maxwel, "Experiments on colour, as perceived by the eye, with remarks on colour-blindness," *TRANSACTIONS OF THE ROYAL SOCIETY OF EDlNBURGH*, vol. 21, pp. 255–299, 1885. 14

[13] J. L. Barbur and K. Spang, "Colour constancy and conscious perception of changes of illuminant," *Neuropsychologia*, vol. 46, no. 3, pp. 853–863, 2008. 16, 22, 61

[14] K. Barnard, *Practical colour constancy*. PhD thesis, Burnaby, BC, Canada, Canada, 1999. AAINQ51839. 16

[15] D. H. Foster, S. M. C. Nascimento, B. J. Craven, K. J. Linnell, F. W. Cornelissen, and E. Brenner, "Four issues concerning colour constancy and relational colour constancy," *Vision Research*, vol. 37, pp. 1341–1345, Apr. 1997. 16

[16] H. Spitzer and S. Semo, "Color constancy: a biological model and its application for still and video images," *Pattern Recognition*, vol. 35, pp. 1645–1659, Aug. 2002. 16

[17] J. Geusebroek, R. van den Boomgaard, A. W. M. Smeulders, and T. Gevers, "Color constancy from physical principles," *Pattern Recognition Letters*, vol. 24, pp. 1653–1662, July 2003. 16, 63, 83

[18] K.-J. Yoon, Y. J. Chofi, and I.-S. Kweon, "Dichromatic-based color constancy using dichromatic slope and dichromatic line space," in *Image Processing, 2005. ICIP 2005. IEEE International Conference on*, vol. 3, pp. III–960–3, 11-14 Sept. 2005. 16, 61

[19] G. D. Finlayson and G. Schaefer, "Solving for colour constancy using a constrained dichromatic reflection model," *INTERNATIONAL JOURNAL OF COMPUTER VISION*, vol. 42, pp. 127—144, 2002. 16

[20] S. Mallick, T. Zickler, D. Kriegman, and P. Belhumeur, "Beyond lambert: reconstructing specular surfaces using color," in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, vol. 2, pp. 619–626 vol. 2, 2005. 16, 23

[21] T. Zickler, S. Mallick, D. Kriegman, and P. Belhumeur, "Color subspaces as photometric invariants," in *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, vol. 2, pp. 2000–2010, 2006. 16

[22] E. LAND, "Retinex theory of color vision," *JOURNAL OF THE OPTICAL SOCIETY OF AMERICA*, vol. 57, pp. 1428–&, 1967. 17

[23] R. Zia-ur, W. G. A, and J. D. J, "A comparison of the multiscale retinex with other image enhancement techniques," 1997. 17

[24] D. H. Brainard and B. A. Wandell, "Analysis of the retinex theory of color vision," *Journal of the Optical Society of America A*, vol. 3, pp. 1651–1661, Oct. 1986. 17

[25] S. Hong and S. Grossberg, "A neuromorphic model for achromatic and chromatic surface representation of natural images," *Neural Networks*, vol. 17, pp. 787–808, 2004. 17

[26] F. E. Nicodemus, "Directional reflectance and emissivity of an opaque surface," *Applied Optics*, vol. 4, pp. 767–773, July 1965. 17

[27] K. E. Torrance and E. M. Sparrow, "Theory for off-specular reflection from roughened surfaces," in *Radiometry*, pp. 32–41, Jones and Bartlett Publishers, Inc., 1967. 17

[28] H. Ragheb and E. Hancock, "A light scattering model for layered dielectrics with rough surface boundaries," *International Journal of Computer Vision*, vol. 79, pp. 179–207, 2008. 17

[29] K. Zhou, P. Du, L. Wang, Y. Matsushita, J. Shi, B. Guo, and H. Shum, "Decorating surfaces with bidirectional texture functions," *Visualization and Computer Graphics, IEEE Transactions on*, vol. 11, pp. 519–528, 2005. 17

[30] G. Hu, S. Ong, Y. Chen, and A. Nee, "Reflectance modeling for a textured object under uncontrolled illumination from high dynamic range maps," *Computers & Graphics*, vol. 31, pp. 262–270, Apr. 2007. 17

[31] H. W. Jensen, S. R. Marschner, M. Levoy, and P. Hanrahan, "A practical model for subsurface light transport," in *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, pp. 511–518, ACM, 2001. 17

[32] C. Donner and H. W. Jensen, "A Spectral BSSRDF for Shading Human Skin," in *Rendering Techniques 2006: 17th Eurographics Workshop on Rendering*, pp. 409–418, June 2006. 17

[33] H. W. Gary, *A BRDF Database Employing the Beard-Maxwell Reflection Model*. 17

[34] R. L. Cook and K. E. Torrance, "A reflectance model for computer graphics," *ACM Transactions on Graphics (TOG)*, vol. 1, p. 7–24, Jan. 1982. ACM ID: 357293. 17

[35] B. HAPKE, "Bidirectional reflectance spectroscopy .1. theory," *JOURNAL OF GEOPHYSICAL RESEARCH*, vol. 86, pp. 3039–3054, 1981. 17

[36] F. Weiwei and W. qingnong, "A scatterometer for measuring the polarized bidirectional reflectance distribution function of painted surfaces in the infrared," *Infrared Physics & Technology*, vol. In Press, Corrected Proof, 2008. 17

[37] S. A. Shafer, "Using color to separate reflection components," *color research and aplications*, vol. 10, pp. 43–51, april 1984. 17, 18, 83

[38] B. Maxwell, R. Friedhoff, and C. Smith, "A bi-illuminant dichromatic reflection model for understanding images," in *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pp. 1 –8, june 2008. 18

[39] R. Moreno, M. Graña, and A. d'Anjou, "Evolutive parametric approach for specular correction in the dichromatic reflection model," in *Hybrid Artificial Intelligence Systems*, vol. 5271 of *Lecture Notes in Computer Science*, pp. 665–672, Springer Berlin / Heidelberg, 2008. 22, 23

[40] M. G. Ramón Moreno, A. d'Anjou, and C. Hernandez, "Bayesian reflectance component separation," in *Knowledge-Based and Intelligent Information and Engineering Systems*, vol. 5712 of *Lecture Notes in Computer Science*, pp. 846–852, Springer Berlin / Heidelberg, 2009. 22, 24

[41] R. Moreno, M. Graña, and A. d'Anjou, "A geometrical method of diffuse and specular image components separation," in *Advances in Computational Intelligence* (J. Cabestany, I. Rojas, and G. Joya, eds.), vol. 6692 of *Lecture Notes in Computer Science*, pp. 83–89, Springer Berlin / Heidelberg, 2011. 22, 23

[42] J. Lellmann, J. Balzer, A. Rieder, and J. Beyerer, "Shape from specular reflection and optical flow," *International Journal of Computer Vision*, vol. 80, pp. 226–241, Nov. 2008. 22, 40

[43] J. Seyama and T. Sato, "Shape from shading: estimation of reflectance map," *Vision Research*, vol. 38, pp. 3805–3815, Dec. 1998. 22

[44] K. Yoon, E. Prados, and P. Sturm, "Joint estimation of shape and reflectance using multiple images with known illumination conditions," *International Journal of Computer Vision*, vol. 86, no. 2, pp. 192–210, 2010. 22

[45] J. Toro, "Dichromatic illumination estimation without pre-segmentation," *Pattern Recognition Letters*, vol. 29, pp. 871–877, May 2008. 22, 23

[46] A. Gijsenij, T. Gevers, and J. van de Weijer, "Generalized gamut mapping using image derivative structures for color constancy," *International Journal of Computer Vision*, vol. 86, no. 2, pp. 127–139, 2010. 22, 61

[47] R. T. Tan, K. Nishino, and K. Ikeuchi, "Color constancy through inverse-intensity chromaticity space.," *J Opt Soc Am A Opt Image Sci Vis*, vol. 21, pp. 321–334, Mar 2004. 23, 29, 61

[48] R. Feris, R. Raskar, K.-H. Tan, and M. Turk, "Specular reflection reduction with multi-flash imaging," in *Computer Graphics and Image Processing, 2004. Proceedings. 17th Brazilian Symposium on*, pp. 316–321, 17-20 Oct. 2004. 23, 43

[49] R. Tan and K. Ikeuchi, "Reflection components decomposition of textured surfaces using linear basis functions," in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, vol. 1, pp. 125–131vol.1, 20-25 June 2005. 23, 37, 42, 45, 54, 55

[50] H. Shen, H. Zhang, S. Shao, and J. H. Xin, "Chromaticity-based separation of reflection components in a single image," *Pattern Recognition*, vol. 41, pp. 2461–2469, Aug. 2008. 23, 61

[51] R. Tan and K. Ikeuchi, "Separating reflection components of textured surfaces using a single image," in *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on*, pp. 870–877 vol.2, 2003. 23, 61

[52] S. Umeyama and G. Godin, "Separation of diffuse and specular components of surface reflection by use of polarization and statistical analysis of images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 5, pp. 639–647, 2004. 23, 43

[53] K. Yoon and I. Kweon, "Voting-based separation of diffuse and specular pixels," *Electronics Letters*, vol. 40, no. 20, pp. 1260–1261, 2004. 23

[54] K.-J. Yoon, Y. Choi, and I. S. Kweon, "Fast separation of reflection components using a specularity-invariant image representation," in *Image Processing, 2006 IEEE International Conference on*, pp. 973–976, 8-11 Oct. 2006. 23, 29, 37, 43, 45, 61

[55] R. Moreno, M. Graña, and E. Zulueta, "Rgb colour gradient following colour constancy preservation," *Electronics Letters*, vol. 46, no. 13, pp. 908–910, 2010. 25, 75, 83

[56] R. Moreno, J. López-Guede, and A. d'Anjou, "Hybrid color space transformation to visualize color constancy," in *Hybrid Artificial Intelligence Systems*, vol. 6077 of *Lecture Notes in Computer Science*, pp. 241–247, Springer Berlin / Heidelberg, 2010. 25, 83

[57] M. Styner, C. Brechbuhler, G. Szckely, and G. Gerig, "Parametric estimate of intensity inhomogeneities applied to mri," *Medical Imaging, IEEE Transactions on*, vol. 19, pp. 153–165, 2000. 28, 31

[58] R. T. Tan, K. Nishino, and K. Ikeuchi, "Separating reflection components based on chromaticity and noise analysis.," *IEEE Trans Pattern Anal Mach Intell*, vol. 26, pp. 1373–1379, Oct 2004. 29, 35, 37, 42, 43, 53, 61, 88

[59] T. M. Lehmann and C. Palm, "color line search for illuminant estimation in real-word scenes," *Journal of the Optical Society of America A*, vol. 18, pp. 2679–2691, 2001. 29

[60] G. Winkler, *Image analysis, random fields and dynamic Monte Carlo methods*. Springer Verlag, 1995. 35, 37

[61] S. Li, Y. Manabe, and K. Chihara, "Accurately estimating reflectance parameters for color and gloss reproduction," *Computer Vision and Image Understanding*, vol. 113, pp. 308–316, Feb. 2009. 40

[62] R. Moreno, M. Graña, and A. d'Anjou, "A color transformation for robust detection of color landmarks in robotic contexts," in *Trends in Practical Applications of Agents and Multiagent Systems*, vol. 71 of *Advances in Soft Computing*, pp. 665–672, Springer Berlin / Heidelberg, 2010. 45

[63] B. Jahne, *Practical Handbook on Image Processing for Scientific and Technical Applications, Second Edition*. CRC Press, 2 ed., Mar. 2004. 59

[64] J. Kittler, "On the accuracy of the sobel edge detector," *Image and Vision Computing*, vol. 1, no. 1, pp. 37 – 42, 1983. 59

[65] I. Abdou and W. Pratt, "Quantitative design and evaluation of enhancement/thresholding edge detectors," *Proceedings of the IEEE*, vol. 67, pp. 753 – 763, may 1979. 59

[66] J. Canny, "A computational approach to edge detection," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. PAMI-8, pp. 679 –698, nov. 1986. 59

[67] L. G. Roberts, "MACHINE PERCEPTION OF THREE-DIMENSIONAL SOLIDS„" tech. rep., May 1963. 59

[68] R. A. Kirsch, "Computer determination of the constituent structure of biological images," *Computers and Biomedical Research*, vol. 4, no. 3, pp. 315 – 328, 1971. 59

[69] G. S. Robinson, "Edge detection by compass gradient masks," *Computer Graphics and Image Processing*, vol. 6, no. 5, pp. 492 – 501, 1977. 59

[70] S. Beucher and C. Lantuejoul, "Use of Watersheds in Contour Detection," in *International Workshop on Image Processing: Real-time Edge and Motion Detection/Estimation, Rennes, France.*, Sept. 1979. 59, 70, 72, 86

[71] P. De Smet, "Optimized high speed pixel sorting and its application in watershed based image segmentation," *Pattern Recogn.*, vol. 43, pp. 2359–2366, July 2010. 59, 70, 72

[72] M. Elwaseif and L. Slater, "Quantifying tomb geometries in resistivity images using watershed algorithms," *Journal of Archaeological Science*, vol. 37, no. 7, pp. 1424 – 1436, 2010. 59, 70, 86

[73] F. Segonne, "Active contours under topology Control—Genus preserving level sets," *International Journal of Computer Vision*, vol. 79, pp. 107–117, 2008. 59

[74] W. Mcilhagga, "The canny edge detector revisited," *Int. J. Comput. Vision*, vol. 91, pp. 251–261, February 2011. 59

[75] R. Moreno, M. Grañ anda, and A. d'Anjou, "An image color gradient preserving color constancy," in *Fuzzy Systems (FUZZ), 2010 IEEE International Conference on*, pp. 1 –5, july 2010. 60, 83

[76] S. Skaff, T. Arbel, and J. J. Clark, "A sequential bayesian approach to color constancy using non-uniform filters," *Computer Vision and Image Understanding*, vol. 113, pp. 993–1004, Sept. 2009. 61

[77] O. Lézoray and C. Charrier, "Color image segmentation using morphological clustering and fusion with automatic scale selection," *Pattern Recognition Letters*, vol. 30, pp. 397–406, Mar. 2009. 61

[78] H. Cheng, X. Jiang, Y. Sun, and J. Wang, "Color image segmentation: advances and prospects," *Pattern Recognition*, vol. 34, pp. 2259–2281, Dec. 2001. 61, 70

[79] V. Osma-Ruiz, J. I. Godino-Llorente, N. Sáenz-Lechón, and P. Gómez-Vilda, "An improved watershed algorithm based on efficient computation of shortest paths," *Pattern Recogn.*, vol. 40, pp. 1078–1090, March 2007. 70, 72

[80] A. Bieniek and A. Moga, "An efficient watershed algorithm based on connected components," *Pattern Recognition*, vol. 33, no. 6, pp. 907 – 916, 2000. 70, 72

[81] F. Hachouf and N. Mezhoud, "A clustering approach for color image segmentation," in *Advanced Concepts for Intelligent Vision Systems* (J. Blanc-Talon, W. Philips, D. Popescu, and P. Scheunders, eds.), vol. 3708 of *Lecture Notes in Computer Science*, pp. 515–522, Springer Berlin / Heidelberg, 2005. 70

[82] G. Li, C. An, J. Pang, M. Tan, and X. Tu, "Color image adaptive clustering segmentation," in *Image and Graphics, 2004. Proceedings. Third International Conference on*, pp. 104–107, 2004. 70

[83] J. Fan, M. Han, and J. Wang, "Single point iterative weighted fuzzy c-means clustering algorithm for remote sensing image segmentation," *Pattern Recognition*, vol. 42, pp. 2527–2540, Nov. 2009. 70, 71

[84] K. Fu and J. Mui, "A survey on image segmentation," *Pattern Recognition*, vol. 13, no. 1, pp. 3–16, 1981. 70

[85] N. R. Pal and S. K. Pal, "A review on image segmentation techniques," *Pattern Recognition*, vol. 26, pp. 1277–1294, Sept. 1993. 70, 74

[86] D. Kim, K. H. Lee, and D. Lee, "A novel initialization scheme for the fuzzy c-means algorithm for color clustering," *Pattern Recognition Letters*, vol. 25, pp. 227–237, 2004. 71

[87] M. Oussalah and S. Nefti, "On the use of divergence distance in fuzzy clustering," *Fuzzy Optimization and Decision Making*, vol. 7, pp. 147–167, June 2008. 71

[88] Y. Xia, D. D. Feng, T. Wang, R. Zhao, and Y. Zhang, "Image segmentation by clustering of spatial patterns," *Pattern Recognition Letters*, vol. 28, pp. 1548–1555, Sept. 2007. 71

[89] M. Grana, R. Moreno, and F. X. Albizuri, "Convex coordinates based on lattice independent sets as pattern features," in *Proc. IEEE Int Fuzzy Systems Conf*, pp. 225–230, 2006. 71

[90] G. X. Ritter, G. Urcid, and M. S. Schmalz, "Autonomous single-pass endmember approximation using lattice auto-associative memories," *Neurocomput.*, vol. 72, pp. 2101–2110, June 2009. 71

[91] H. Struve and R. Struve, "Lattice theory and metric geometry," *Algebra Universalis*, vol. 58, pp. 461–477, 2008. 71

[92] A. H. And, *Mathematical Morphology in the L*a*b* Colour Space*. 71

[93] A. Hanbury and J. Serra, *Mathematical Morphology in the HLS Colour Space*. 2001. 71, 83

[94] A. Hanbury and J. Serra, "Colour image analysis in 3d-polar coordinates," in *In Proceedings of the DAGM'03 conference*, pp. 124–131, Springer-Verlag, 2003. 71

[95] J. Serra and Angulo, "Mathematical morphology in color spaces applied to the analysis of cartographic images," *International Workshop on Semantic Processing of Spatial Data GEOPRO 2003*, vol. (ISBN 970-, pp. 59–66. 71

[96] D. Wang, "A multiscale gradient algorithm for image segmentation using watershelds," *Pattern Recognition*, vol. 30, no. 12, pp. 2043 – 2052, 1997. 72, 75

[97] I. Dagher and K. E. Tom, "Waterballoons: A hybrid watershed balloon snake segmentation," *Image and Vision Computing*, vol. 26, no. 7, pp. 905 – 912, 2008. 72, 86

[98] H. Zhang, J. E. Fritts, and S. A. Goldman, "Image segmentation evaluation: A survey of unsupervised methods," *Computer Vision and Image Understanding*, vol. 110, no. 2, pp. 260 – 280, 2008. 72

[99] D. Martin, C. Fowlkes, D. Tal, and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," in *Proc. 8th Int'l Conf. Computer Vision*, vol. 2, pp. 416–423, July 2001. 79, 88, 90

[100] J. Angulo and J. Serra, "Modelling and segmentation of colour images in polar representations," *Image and Vision Computing*, vol. 25, pp. 475–495, Apr. 2007. 83

[101] Y. Tarabalka, J. Chanussot, and J. A. Benediktsson, "Segmentation and classification of hyperspectral images using watershed transformation," *Pattern Recogn.*, vol. 43, no. 7, pp. 2367–2379, 2010. 83, 86