



HAL
open science

Inter-reflections in computer vision: importance, modeling & application in spectral estimation

Rada Deeb

► **To cite this version:**

Rada Deeb. Inter-reflections in computer vision: importance, modeling & application in spectral estimation. Other [cond-mat.other]. Université de Lyon, 2018. English. NNT: 2018LYSES035 . tel-02305680

HAL Id: tel-02305680

<https://theses.hal.science/tel-02305680>

Submitted on 4 Oct 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



N°d'ordre NNT : 2018LYSES035

THESE de DOCTORAT DE L'UNIVERSITE DE LYON
opérée au sein de
Université Jean Monnet – Saint-Étienne

Ecole Doctorale N° 488
Sciences, Ingénierie, Santé

Spécialité de doctorat :
Image

Soutenue publiquement le 04/10/2018, par :
Rada Deeb

**Inter-reflections in Computer Vision :
Importance, Modeling & Application in
Spectral Estimation**

Devant le jury composé de :

Van de Weijer, Joost Professeur. Universitat Autònoma de Barcelona. Président

Hardeberg, Jon Professeur. Norwegian University of Science and Technology.
Rapporteur

Hanselaer, Peter Professeur. KU Leuven. Rapporteur

Vanrell Martorell, Maria Professeure associée. Universitat Autònoma de Barcelona.
Examinatrice

Hébert, Mathieu Maître de Conférences Université. Jean Monnet. Examineur

Muselet, Damien Maître de Conférences Université. Jean Monnet. Examineur

Trémeau, Alain Professeur. Université Jean Monnet. Directeur de thèse



University of Jean Monnet
Hubert Curien Laboratory
Image Department

Interreflections in Computer Vision Importance, Modeling & Application in Spectral Estimation

PHD THESIS

RADA DEEB

Supervisors: Alain Tremeau
Professor

Mathieu Hebert
Professor

Damien Muselet
Assistant Professor

Reporters: Jon Yngve Hardeberg
Professor

Peter Hanselaer
Professor

Saint Etienne, 2018

In this thesis, we study an optical phenomenon often ignored in computer vision, the interreflection phenomenon. Interreflections, which can also be found in the literature under the name *mutual illumination* happen whenever a concave surface is illuminated. As the name tells, a light ray coming from the light source and hitting a surface point will reflect toward some other point, then another, and so on, before reaching the camera sensor or the eye. Hence, a ray does inter-reflect between the different points of a concave surface.

Interreflections lead to color variations, or color gradients, all over the concave surface. These variations are more or less pronounced depending on many factors including, but not limited to, the surface reflectance and its geometry. We will show in this manuscript that these color variations hold some important information which is worth to be used in computer vision. They also play an important role in perception leading to a better color constancy in human vision as demonstrated in our experiments.

In order to be able to efficiently use interreflections in some computer vision applications, a spectral infinite-bounce model of interreflections is introduced in the manuscript. This radiometric model allows us to define the relation between the raw RGB values corresponding to the concave surface in the image on one side, and the spectral reflectance and the geometry of this surface, the spectral power distribution of the light and the spectral responses of the camera sensors on the other side. Thanks to this model, we were able to study some applications of interreflections in spectral estimation. We show that a task, such as spectral reflectance estimation from a single RGB image, which is almost impossible without learning even under known illuminant and spectral responses of the camera, is made possible thanks to interreflections. Our results show that, spectral reflectance estimation of a folded surface gives a similar accuracy and sometimes a better one when compared to the state of the art approaches that need three different images of the flat surface taken under three different illuminants. Moreover, interreflections help in proposing a more concrete application of spectral reflectance estimation where a standard light SPD can be used and no pre-calibration for the acquisition settings is needed.

Later, we show that interreflections are useful in some applications which need color charts such as camera characterization. The nature of interreflections leads to special colors resulted from raising the spectral reflectance to multiple powers. Using these colors along with the interreflection model helps in introducing some non linearly-related information and thus in obtaining a better spectral characterization. Hence, using 3D color charts is more beneficial than adding new colors to 2D color charts.

Finally, we train a convolutional neural network on simulated images of interreflections in order to get an estimation of both the spectral reflectance and the SPD of light from a single RGB image. The experimental results show that our approach is able to get both spectra with a very good accuracy compared to other approaches. In addition, this approach performs very well on real images thanks to the added noises in the training process.

Keywords

Interreflection; Spectral Reflectance Estimation; Color Constancy; Deep Learning, Camera Characterization

Dans cette thèse, nous étudions un phénomène optique souvent ignoré en vision par ordinateur : les inter-réflexions. Les inter-réflexions, qui peuvent être trouvées dans l'état de l'art sous le nom « illumination mutuelle », se produisent quand une surface concave est illuminée. Dans ce cas, un rayon lumineux venant de la source de lumière vers un point de la surface, va réfléchir vers d'autres points de la même surface plusieurs fois avant d'arriver à nos yeux, ou aux capteurs de l'appareil photo. Donc, un rayon inter-réfléchi entre les différents points de la surface concave, d'où le nom « inter-réflexions ». Les inter-réflexions conduisent aux variations de couleurs, ou gradients de couleurs, sur la totalité de la surface concave. Ces variations sont plus au moins prononcées selon plusieurs facteurs comme la réflectance de la surface et sa géométrie. Dans ce manuscrit, nous allons montrer que ces variations de couleurs contiennent en elles des informations importantes qui méritent d'être utilisées en vision par ordinateur. Ces mêmes variations jouent un rôle important dans la perception ce qui permet à l'être humain une meilleure constance de couleur, comme montré par nos résultats empiriques. Dans l'objectif d'utiliser efficacement les inter-réflexions pour quelques applications en vision par ordinateur, nous introduisons dans ce manuscrit un modèle spectral d'inter-réflexions prenant en compte une infinité de rebonds. Ce modèle construit sur des bases radiométriques nous permet de définir la relation entre les valeurs RGB brut correspondant à la surface concave dans l'image d'un côté, et la réflectance spectrale et la géométrie de cette même surface, la distribution de puissance spectrale de l'éclairage (SPD), et les courbes des réponses spectrales de l'appareil photo de l'autre côté. Grâce à ce modèle, nous sommes capables d'étudier plusieurs applications d'inter-réflexions en estimation spectrale. Nous montrons que l'estimation de la réflectance spectrale à partir d'une seule image RGB, une tâche qui est quasi-impossible sans apprentissage même sous un éclairage connu, est devenue possible grâce aux inter-réflexions. Nos résultats ont démontré que l'estimation de la réflectance spectrale d'une surface concave donne une précision similaire, et même parfois meilleure, en comparaison avec les approches de l'état de l'art qui ont besoin de trois images de la même surface prises sous trois différents éclairages. De plus, les inter-réflexions nous ont aidés à proposer une application plus concrète de l'estimation de la réflectance spectrale dans laquelle il est possible d'utiliser un spectre d'un éclairage standard sans nécessiter un pré-calibrage pour les paramètres de l'acquisition. Par la suite, nous démontrons que les inter-réflexions sont aussi utiles dans des applications qui utilisent des mires de couleurs, comme par exemple la caractérisation de l'appareil photo. La nature de l'inter-réflexion sur une surface d'une seule couleur conduit aux couleurs spéciales qui sont les résultats des multiplications de la réflectance spectrale avec elle-même de multiples fois. Utiliser ces couleurs avec notre modèle d'inter-réflexion aide à introduire de la non-linéarité sur les mires de couleurs et donc à obtenir une meilleure caractérisation spectrale. Par conséquent, utiliser des mires de couleurs 3D est plus bénéfique qu'ajouter des nouvelles couleurs aux mires 2D. Finalement, nous entraînons un réseau neuronal convolutif sur des images simulés d'inter-réflexions dans le but d'estimer à la fois la réflectance spectrale de la surface et la SPD de l'éclairage d'une seule image RGB. Nos résultats expérimentaux démontrent que notre approche est capable

d'estimer les deux spectres avec une très bonne précision en comparaison avec les autres approches. De plus, cette approche fonctionne très bien sur les images réelles grâce aux niveaux de bruits ajoutés dans le processus d'apprentissage.

Mot Clés

Inter-réflexions; Estimation de Réflectance Spectrale; Constance de Couleur; Apprentissage Profond ; caractérisation de l'appareil photo

To my family and friends.

*To my home country, Syria, and all the parts of the world suffering from wars, poverty
and neo-colonization.*

Acknowledgements

I would like to thank my supervisors for guiding me through this PhD. I appreciated our discussions and learned a lot from each one of them. I would also like to thank Joost Van de Weijer for hosting me in CVC. This experience was very enriching, I enjoyed working there and learned much from him.

A special thanks for my family and friends who accompanied me through these years and were there for me whenever I needed help and support.

Rada Deeb

Abstract	i
Acknowledgements	vii
1 Introduction	1
1.1 Research problem: from colors to spectral information	2
1.2 The motivation: interreflections & their particularity	3
1.3 Structure of the manuscript	6
2 Ground Concepts on Light and Color	13
2.1 Radiometry	13
2.1.1 Geometrical concepts	13
2.1.2 Radiometric definitions	14
2.1.3 Lambert’s law	15
2.2 Light acquisition and color imaging	16
2.3 Color vision and colorimetry	17
2.4 Color constancy, color appearance & color naming	19
3 Interreflections: Mechanism & Roles	21
3.1 The mechanism of interreflections	21
3.2 The effects of interreflections in images	22
3.3 Interreflections’ role in perception	24
3.4 Previous use in computer vision	29
4 Physics-Based Interreflection Model	33
4.1 The Continuous Model	33
4.2 The discrete model	34
4.3 Some existing applications of the interreflection models	34
4.3.1 Interreflections removal based on shape from shading	35
4.3.2 Interreflection removal via light field	35
4.3.3 Interreflections removal without light field	36
4.4 Reviewing the discrete model	38
4.4.1 Relation with Nayar’s model [1]	39
4.4.2 Generalization to spectral model	39
4.4.3 Self-interreflection model	41
4.5 Scene geometry & geometrical kernel calculation	42
4.5.1 The problem	43
4.5.2 Monte Carlo Estimation	43
4.6 Direct irradiance modeling	45
4.6.1 Directional lighting	45
4.6.2 Diffuse lighting	45

4.7	Interreflection model in practice: accuracy & discussions	46
4.7.1	Self-interreflection	47
4.7.2	Interreflection between two surfaces with different colors	50
4.7.3	Interreflection with different geometries and surfaces	50
4.7.4	Two-bounce vs. Infinite-bounce models	52
4.7.5	Sources of error	56
4.8	Conclusion	57
5	The Inverse Approach: Interreflections in Spectral Estimation	59
5.1	Interreflection in direct spectral reflectance estimation	59
5.1.1	State of the art	59
5.1.2	Direct spectral estimation with the infinite-bounce interreflection model	64
5.1.3	Experiments & results	66
5.1.4	Discussion	71
5.2	Interreflections in camera characterization	77
5.2.1	The problem	78
5.2.2	Interreflection-based approach	78
5.2.3	Results	80
5.3	Conclusion	84
6	Interreflections in Learning-Based Spectral Estimation	87
6.1	State of the art	87
6.1.1	Kernel-based regression approaches	87
6.1.2	Hyperspectral recovery approaches	88
6.1.3	Deep learning use in physics-based vision	89
6.2	Interreflection-based approach	90
6.2.1	Datasets	90
6.2.2	Network structure & loss functions	91
6.3	Experimental results & discussions	92
6.3.1	Simulated data	93
6.3.2	Real data	93
6.3.3	Validation of network structure	96
6.3.4	Generalization to other angles	96
6.4	Conclusion	98
7	Conclusion	99
7.1	Future Work	100
7.1.1	Generalization of the interreflection model	100
7.1.2	Toward a less constrained spectral estimation	101
7.1.3	Applications in color constancy	101

Appendices	105
------------	-----

Bibliography	115
--------------	-----

Humans, as well as other beings, use different senses in order to understand their surroundings. Vision is one important sense: with vision one can understand the geometry of the scene, localize himself, and get to know the distances and the movements around. In addition, our vision allows us to identify the objects and to have many clues about their characteristics (form, texture, size, material, etc..). Color is one essential element in human vision. However, describing colors is not as simple as it might seem to be. The concept of color goes beyond the color names we learn at a young age. Human perception of colors starts from a physical stimulus which happens when a light carrying different levels of energy at each wavelength hits a surface. The surface, depending on its physical properties, sends back some energy at some wavelengths. This energy is then received by the eyes' cones and rods which react to it based on different sensibilities to each wavelength. The brain, later, treats the stimuli sent by the eyes, and as a result, we see colors. However, our perception of an object's colors does not depend only on this physical stimulus, but it is also relative to the object's surrounding. Moreover, humans generally are capable of seeing behind colors and understanding some physical characteristics of the scene. For example, one can in most cases tell what is the color of the light in the scene, and what is the original color of the object as if it is seen under a white light. This leads us to see the color of an object relatively constant under varying lighting conditions. This feature of human color perception is called *color constancy*.

With the immense rise of technology, one important area of study focuses on giving computers the abilities human beings have. Thus, many researchers in image processing and computer vision domains are interested in how to get the technologies see and understand scenes the closest possible to ourselves. Colors in an image are the results of camera sensors interacting to the received light signal. Most digital cameras have three types of sensors with different sensibilities to each wavelength, they give as an output the red, green and blue values which are used to obtain the captured color in a color mixing system. Thus, colors in computer vision are generally defined by only three values. There exist several standard color spaces and color attributes used in order to have effective color naming, quantification, and color differences techniques. However, the colors in the image, as opposite to colors seen by humans, are simply the results of how the camera sensors react to the physical stimuli, and treating them as they are does not provide enough information about the scene behind. Therefore, retrieving the physical characteristics behind these values is essential for computer vision to get closer to human vision.

In this thesis, we focus our attention on retrieving some physical characteristics about the scene from the color values in an image using a well known, but often ignored, physical phenomenon: the interreflection. Diffuse Interreflections happen whenever a light hits a concave surface. The concavity will not let the light exists the surface directly but instead it will inter-reflect between the different points of the surface. As a result of this phenomenon, some color variations will be visible on the surface. It is these color variations that we are going to study in this thesis to show their importance in spectral estimation.

1.1 Research problem: from colors to spectral information

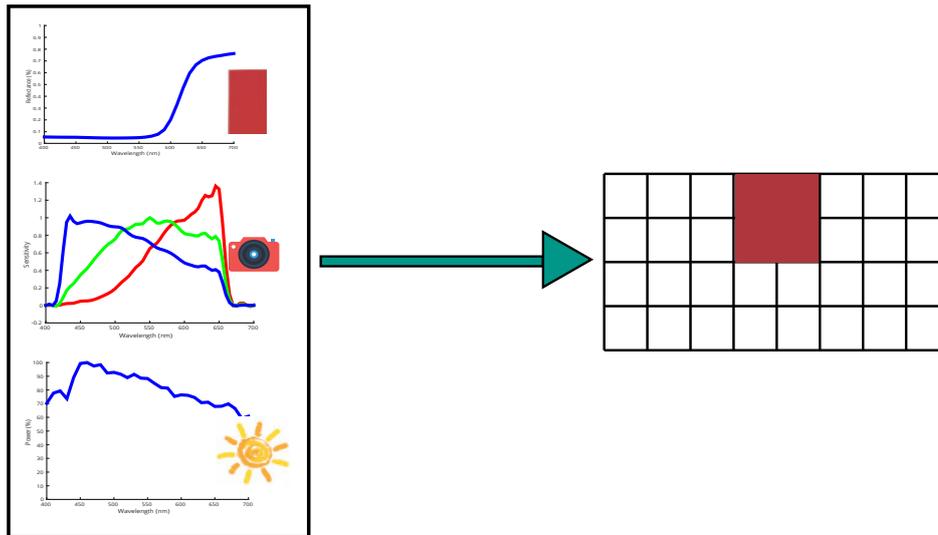
Light is what makes vision possible to human eyes. It is a spectrum carrying different levels of energy at each wavelength. However, colors are not only related to light, but they are the product of several spectral functions: the one of the light often called the *spectral power distribution (SPD)*, the one of the surface or its spectral reflectance for reflective surfaces, and the ones of the sensors, or their spectral responses. When the radiation is received by one sensor, it is going to be associated to a single value based on the spectral responses of this sensor. This leads to the red, green and blue (RGB) values in images taken with the commonly used RGB cameras.

In computer vision, a pixel is generally represented by tristimulus values, whether they are in RGB space or in any other color spaces. However, these values depend on the acquisition device, on the illuminants and on any preprocess involved in the imaging system such as white balancing. A surface of a specific physical characteristics may have different tri-values depending on the used camera and on the light sources in the scene which makes it difficult to retrieve the exact color of the object based on these values. In addition, two surfaces with different physical characteristics may have the exact same tri-values in an image, this is called **Metamerism**. Hence, the tristimulus values in an image do not hold enough information about the color of the objects in the scene, this makes even basic image processing processes such as segmentation prone to errors.

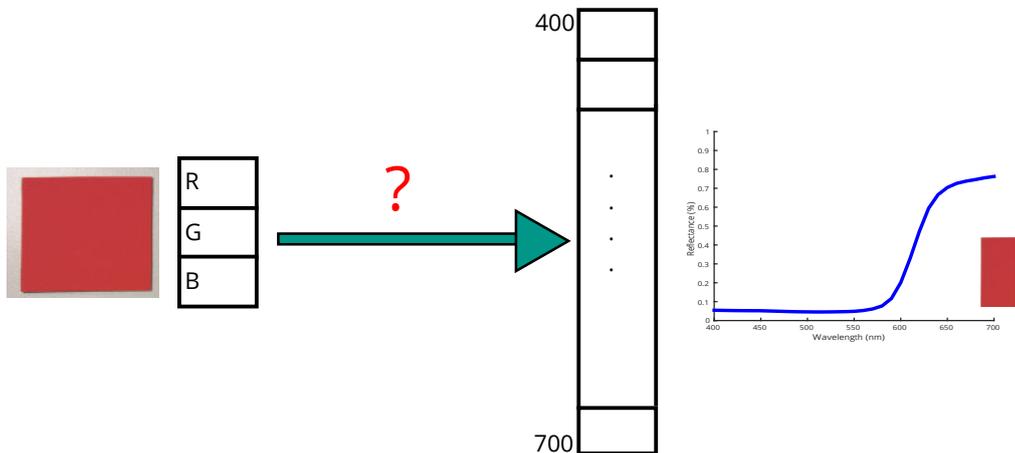
Spectral description of the surface is often needed instead of tristimulus values as tri-values may not be sufficient in applications such as medical imaging, paint selection, cosmetics industry, fruit quality assessment, telemedicine and eHeritage. These applications require spectral acquisition of the light signal reflected by the object in order to have more information about the scene so that precedent cases such as metamerism are detected and avoided. Other spectral information may be needed also depending on the application: recovering the spectral responses of the camera is needed for camera calibration and characterization.

In order to obtain spectral data, it is possible to use spectral cameras. These devices give very accurate results, but they are expensive, and therefore not affordable for the broadest range of customers. As an alternative to these spectral cameras, the use of RGB images was proposed in the literature in order to estimate some spectral functions which lie behind the formation of the RGB colors. However, recovering the spectral information from images taken with conventional RGB cameras is a very difficult and almost impossible task if no other process or knowledge is included. For example, retrieving the spectral reflectance of a surface is impossible even if the illumination spectrum is known. The three measurements provided by the RGB camera are insufficient to recover the spectral data whose dimensionality is high (see Figure (1.1)).

In order to make the spectral recovery possible, many methods tried to reduce the dimensionality of the problem by representing the spectral data with a set of basis functions. It has been shown in [2] that the 1257 reflectance spectra of the Munsell Book of Color's chips can be presented to a good approximation by 8 characteristic spectra. However, eight is still high compared to the three measurements. Thus, the majority of approaches in



(a) A color is the product of the spectral functions of the light, the surface and the camera sensors.



(b) An example of the research problem: how to retrieve the spectral reflectance of a surface from its RGB value in the image.

Figure 1.1: The research problem: from colors to spectral information.

spectral reflectance estimation use several images taken under different acquisition settings, for example different lighting conditions, in order to increase the number of equations needed to accurately estimate the basis coefficients of the spectral reflectance. Other approaches use different surfaces with distinct and known physical properties to estimate the spectral responses curves of the camera or the spectral power distribution of the lighting supposing a single illuminant in the scene. All these approaches will be explained in more details in the following chapters. However, it is remarkable that very few methods in the literature tried to explore the use of some optical phenomena in spectral estimation.

1.2 The motivation: interreflections & their particularity

The interreflection phenomenon, which is going to be described, modeled and used in different contexts during this manuscript, happens whenever a scene containing concave

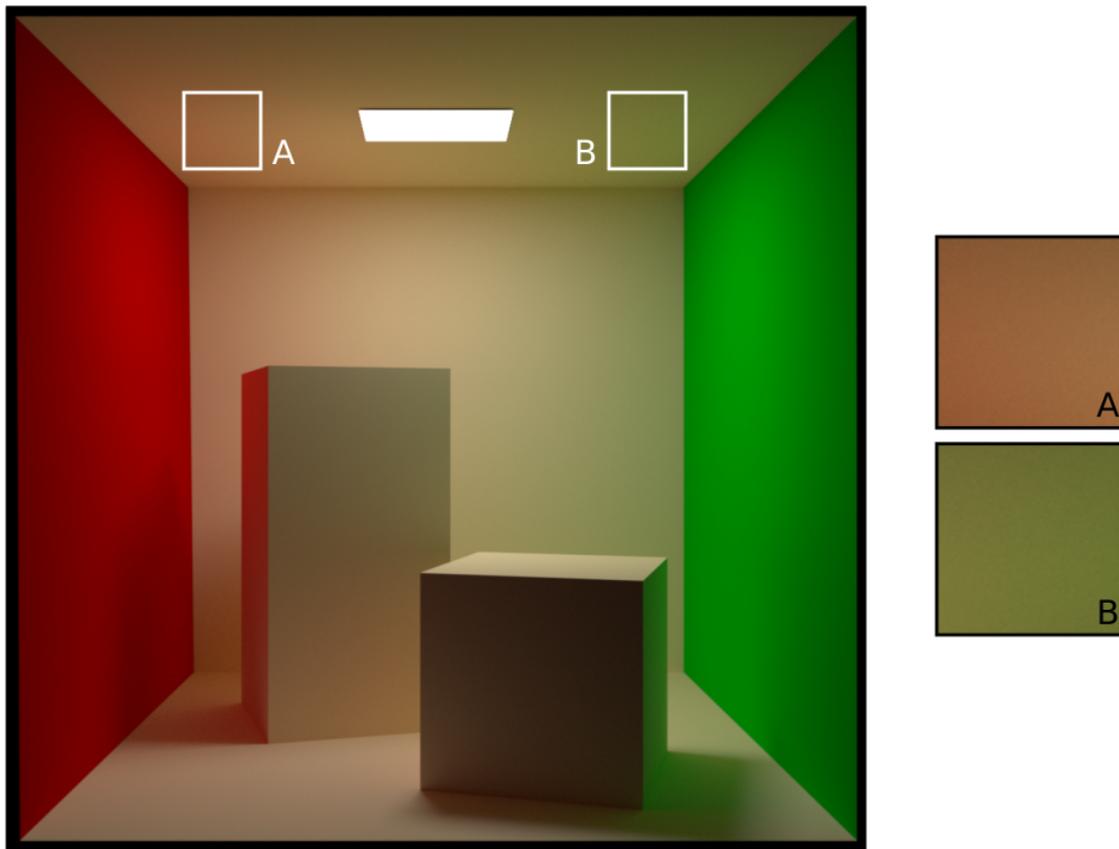


Figure 1.2: An example of a scene with interreflection: a Cornell box simulated with LuxRender. Two walls of the box are colored: one is red and the other is green. The two smaller boxes and the other walls of the main box are all of the same white color.

surfaces is illuminated. Interreflections create particular color variations in the area of the scene corresponding to the concave object. Examples of this phenomenon are multiple. If a room has two adjacent walls, one is white and the other is red, one can see reddish gradients on the white wall close to the red one. The latter would also show variations of red color, getting more saturated the closest to the white wall. These color changes are often called color bleeding in computer graphics community. Figure (1.2) shows a commonly used example of interreflections in the computer graphics community, the Cornell box. One can observe the changes of color on the white roof due to interreflections. Moreover, the boxes in the middle have all their facets of the same white color. The facets facing the colored walls appear green and red only because of interreflections as they are the only sources of light which hits these facets.

Interreflections happen also on uniformly colored concave surfaces. A green sofa for example, would show more saturated green where its sitting and its back join. This case of interreflections is called self-interreflections.

The reason behind these color variations can be understood by exploring how light rays travel in the scene. A light ray coming from the light source and hitting a point on the concave surface will strike successively between the surface points before reaching the eye or the camera sensor. Each time the ray reflects from a surface point is called a

bounce. If a ray hitting a point reflects directly toward the camera, then it encounters only one bounce. If it hits only one other point before reflecting toward the camera then it encounters two bounces, and so on. A ray may bounce few times, many times, or almost infinitely depending on many factors: with each bounce, the energy carried with the ray is reduced, depending on the spectral reflectance of the surface point it hits and on the geometrical relation between the two points the ray is traveling between. The sooner this energy vanishes the less number of bounces the ray does. The energy will reduce slowly on wavelengths where the surface reflectance is high, but it will vanish quickly where the surface reflectance is low. In addition, points which are closer to each other will exchange many rays, whereas farther points will exchange few ones. This is the reason behind the observation that color variations are more visible next to the shared edges in Figure (1.3). Thus, interreflections do not depend only on the color of the surface, but rather on its spectral reflectance.

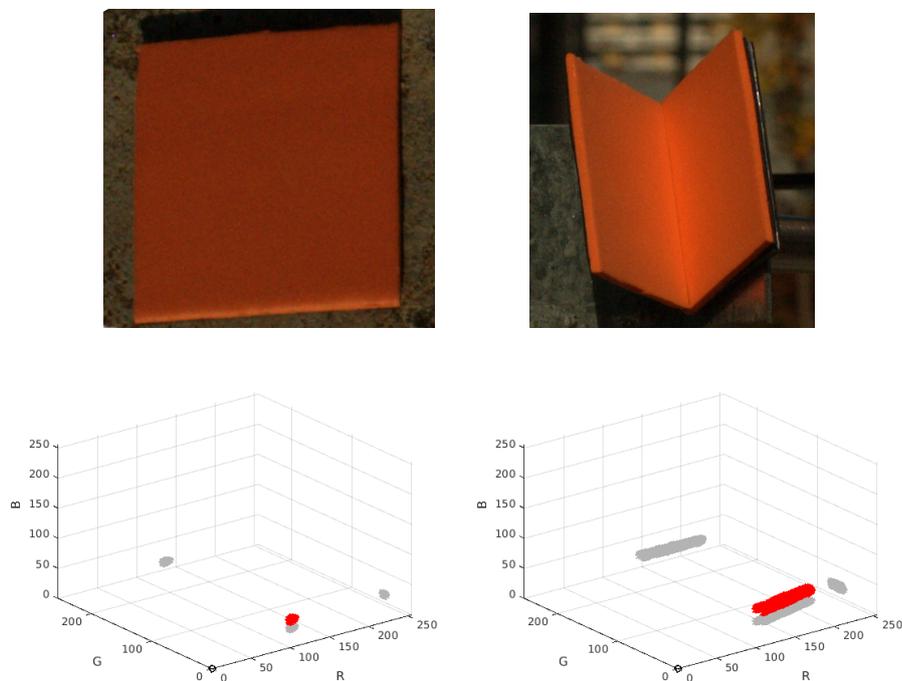


Figure 1.3: Difference of dispersion of the RGB values in an image between a flat and a folded orange textile sample.

Interreflections depend also on the geometry of the illuminant which determines, in the first place, how much energy is received from the light source at each point of the surface. It is this energy then which is going to be reduced with each bounce of inter-reflected light.

The color variations that interreflections create can be an importance source of information about the scene. In Figure 1.3 one can see the dispersion of these color variations in a concave surface when compared to a flat one with the same optical characteristics. These particular color gradients happening in the concavity are rich in information related to both the geometry and the physics of the scene. This information can be useful to retrieve some of the spectral curves lying behind the colors such as the spectral reflectance of the surface

(see Figure 1.4), the lighting conditions in the scene, etc. This is true in computer vision, as it is going to be demonstrated in this thesis. We think that it is also true in human vision even though our visual system is not capable of evaluating the spectral properties of an object (we even get trapped in case of metamerism and see the same colors from two physically different flat objects). However, it is highly probable that one can get a hint about the spectral information by observing the color gradients due to interreflections in the concavities. We show in this thesis that the human color constancy is better when observing concavities with interreflections compared to flat surfaces. Moreover, and as these color gradients are characteristic of the spectral reflectance of the surface, we argue that two surfaces which show metamerism when they are flat, have almost no chance to show the same color gradients when they are folded into a concave shape (see Figure 1.5).

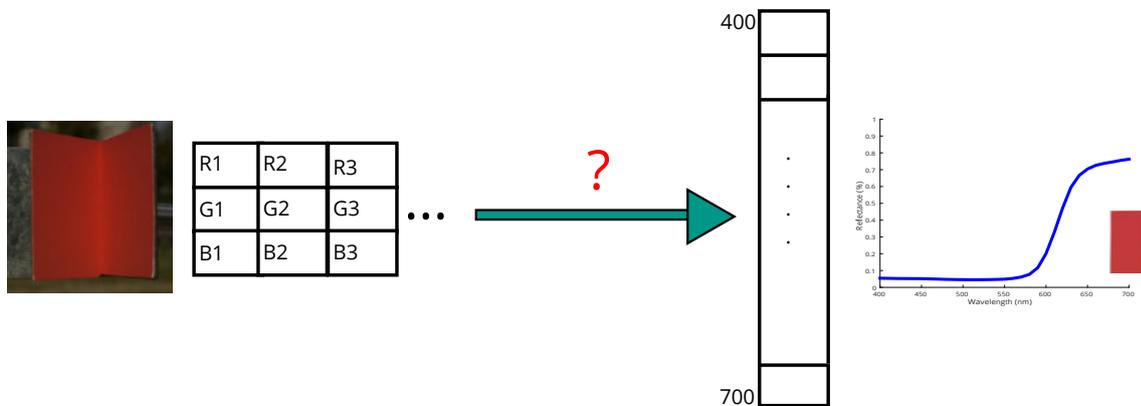


Figure 1.4: *The motivation: color gradients due to interreflections provide more information to retrieve the spectral reflectance of a surface from its RGB value in the image.*

This finding constitutes probably the most important discovery in this thesis and the basis of the spectral estimation techniques we present here in the domain of computer vision. Even though interreflections are rarely visible to the point that they get our attention, they exist almost everywhere and it is then essential to take them into account in image rendering in order to obtain realistic scenes. In the domain of computer vision on the other hand, interreflections were studied and used in some approaches mainly in shape-from-shading and object recoloring domains. However, few approaches in the literature treated interreflections as sources of information about the scene, and even when they were used in spectral estimation, only two values of the color gradients were taken into account.

1.3 Structure of the manuscript

In the following chapters, interreflection phenomenon is studied, modeled, and used in different approaches to retrieve some spectral curves with and without learning.

- In **Chapter 2** a reminder of some ground concepts in radiometry, color vision and colorimetry is given. These concepts are essential for the understanding of the later chapters.

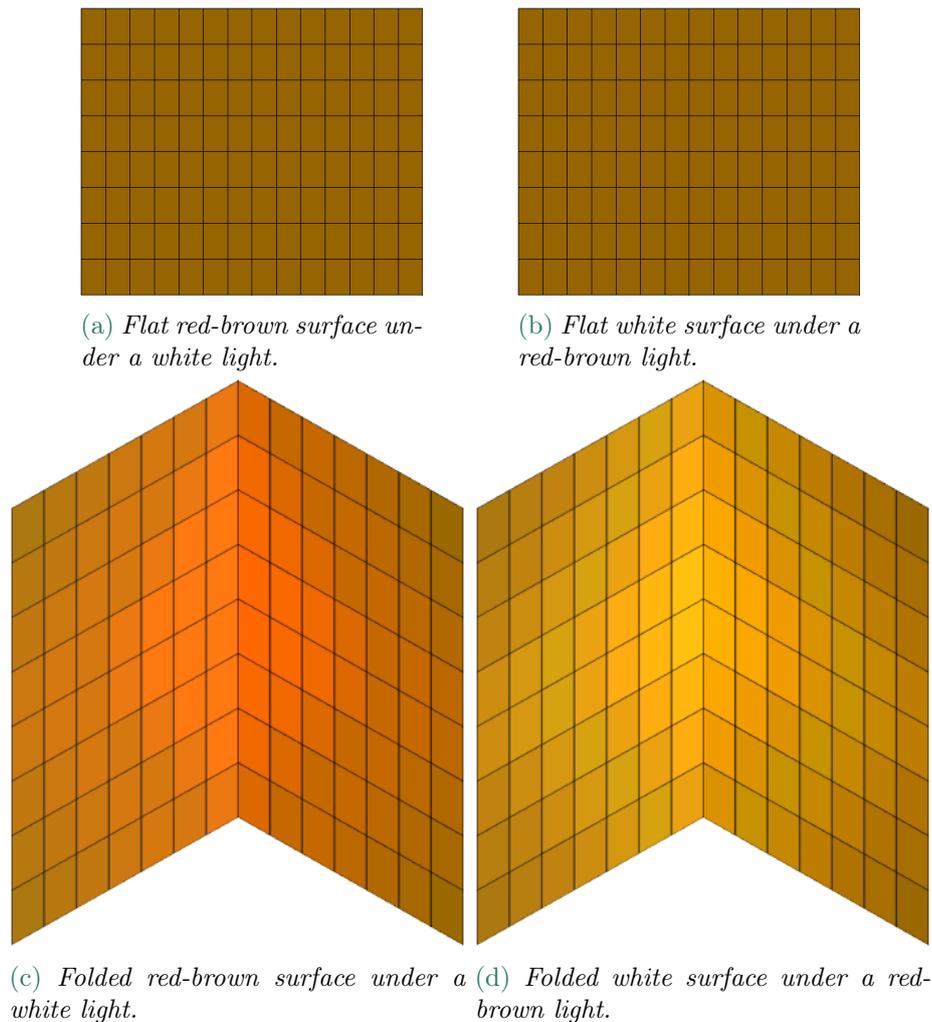


Figure 1.5: *Interreflections and metamerism: two Metameric flat surfaces show no metamerism when folded with an angle of 45° .*

- In **Chapter 3** the mechanism of the interreflection phenomenon is explained, and its role in perception and computer vision in addition to the state of the art of its use in computer vision are detailed. In this chapter, our first contribution regarding the role of interreflections in human color constancy [3] is explained in details.
- In **Chapter 4** the state of the art of interreflection modeling and our spectral infinite bounce model [4] are detailed. Discussions on direct irradiance modeling, interreflection model accuracy with number of bounces and geometrical kernel calculation are also given.
- In **Chapter 5** the proposed interreflection model is used in an inverse approach in spectral estimation. First, the surface spectral reflectance is estimated from a single RGB image of a uniformly colored object under direct sunlight [5]. No learning is used in this approach and camera spectral response curves and light SPD are considered known. Our results are discussed and compared to the state of the art approaches presented in the same chapter. Discussions regarding the nature of the solution and the effects of the different parameters are also given. Second, our model is also used

in retrieving the spectra responses of the sensors in camera characterization. We show that the non-linearity of interreflections's effects help in estimating the camera spectral response curves with a good accuracy [6]. We also demonstrate that using interreflections gives better results than adding new color patches to the color charts commonly used for this sake.

- In **Chapter 6** the proposed interreflection model is used to build datasets of simulated images to be used to train a convolutional neural network (CNN). The network is used to estimate both the spectral reflectance of the surface and the SPD of the light. Our results are discussed and compared to the state of the art approaches presented in the same chapter. Discussions regarding the network structure and the possible generalization of the presented approach are also given. This work is submitted to JOSA A journal.
- In **Chapter 7** we first conclude the work presented in the previous chapters. Afterwards, some of the possible future work perspectives and research directions are discussed.

La vision est un sens très important : avec la vision, on peut identifier les objets, se localiser et comprendre les distances et les mouvements autour. La couleur est un élément essentiel dans la vision humaine. Cependant, le concept de la couleur est plus compliqué que ce qu'il paraît. Notre perception commence par un stimulus physique arrivant quand une lumière portant différents montants d'énergie pour chaque longueur d'onde touche une surface. La surface, selon ces propriétés physiques renvoie de l'énergie vers nos yeux. Mais, notre perception ne s'arrête pas ici, mais elle est dépendante de l'entourage de l'objet. De plus, on est capable de comprendre des caractéristiques physiques de la scène. Par exemple, une personne est capable dans la plupart des cas de comprendre la couleur de l'éclairage dans la scène, et donc les couleurs des objets comme s'ils sont vus sous un éclairage blanc. Cette fonctionnalité est nommée "constance des couleurs". Avec la montée des technologies, un des domaines importants est focalisé sur la vision par ordinateur. L'idée est de donner aux machines des capacités visuelles qui leur permettent à analyser et comprendre les scènes comme nous. Les images traitées dans ce domaine sont très souvent prises par des appareils photo RGB ayant trois capteurs, rouge, vert et bleu, chacun avec des sensibilités différentes selon la longueur d'onde. Les couleurs dans une image sont simplement les résultats de la réaction des capteurs au stimulus physique et donc elles portent très peu d'informations sur la scène à l'opposé de la vision humaine. Donc, il est essentiel pour la vision de l'ordinateur de récupérer des caractéristiques physiques derrière ces couleurs pour qu'elle s'approche de la vision humaine. Dans cette thèse, nous concentrons sur la récupération des caractéristiques physiques de la scène à partir des données couleurs en utilisant un phénomène, bien connu mais souvent ignoré, l'inter-réflexion.

La problématique

La limitations des informations contenues dans les valeurs de couleurs dans un image rendent difficile même des tâches essentielles dans la vision par ordinateur comme la segmentation. Donc, beaucoup d'applications ont besoin d'exploiter des informations de plus haut niveau comme les informations spectrales pour avoir un fonctionnement correcte. Les exemples de ces applications sont multiples dans les domaines des la télémédecine, e-héritage, etc. Cependant, l'accès aux informations spectrales n'est pas toujours simple. Ils existent des caméra spectrale mais très chères. L'autre solution est d'extraire les informations spectrales à partir des couleurs RGB. Ce problème est mal définie comme on a le droit à très peu de valeurs pour extraire des informations de haute dimensions. Il était nécessaire pour les approches de l'état de l'art d'utiliser plusieurs éclairage, des mires de couleurs ou de l'apprentissage pour pouvoir résoudre ce problème. Cela était souvent accompagné d'une réduction de la dimensionnalité du problème en utilisant des fonctions de base.

La motivation

L'inter-réflexion signifie qu'un rayon lumineux rebondit entre les points de la surface concave avant de quitter. Le nombre de rebonds qu'un rayon peut avoir dépend de la réflectance des points qu'il touche et de la relation géométrique entre ces points. Cela laisse apparaître des variations de couleur sur la surface qui sont donc la signature de la réflectance de ses points et de sa géométrie. Ces variations de couleurs sont autour de nous mais souvent elles n'attirent pas notre attention : sur le dossier d'un canapé, sur les murs, dans les plis des vêtements, etc. Nous proposons d'exploiter ces variations de couleurs dans l'estimation spectrale. Nous montrons qu'elles contiennent en soi des informations spectrales, ce qui rend la tâche d'estimation spectrale possible sous des conditions estimées impossibles auparavant.

Le plan de la thèse

- **Chapitre 2** : dans ce chapitre nous introduisons les notions et les notations qui sont essentielles pour la compréhension de la suite du manuscrit. Cela inclut les notions de radiométrie et de couleur.
- **Chapitre 3** : dans ce chapitre nous présentons en détail les mécanismes d'inter-réflexion et ses précédentes utilisations en vision par ordinateur. Ce chapitre contient l'une de nos premières contributions qui porte sur l'importance de l'inter-réflexion dans la constance de couleur chez les humains. Cette contribution a été publiée à AIC en 2017 [3].
- **Chapitre 4** : dans ce chapitre nous traitons le problème de modélisations des inter-réflexions. Nous commençons par l'état de l'art, puis nous présentons notre modèle spectral prenant en compte une infinité de rebonds. Nous montrons par la suite l'importance de ce modèle. Ce travail a été publié dans le journal JMIV [4].
- **Chapitre 5** : dans ce chapitre, nous utilisons notre modèle dans une approche inverse pour estimer la réflectance spectrale d'une surface uniformément colorée où les réponses spectrales de l'appareil photo sont sans apprentissage. Notre modèle d'inter-réflexion nous a permis d'estimer la réflectance spectrale d'une seule image RGB sans pré-calibrage. Ce travail a été publié dans le journal Applied Optics [5]. Notre modèle nous a aussi permis de proposer des mûres couleurs 3D qui augmentent l'efficacité de l'estimation des réponses spectrales de l'appareil photo. Ce dernier travail a été présenté dans la conférence BMVC [6].
- **Chapitre 6** : dans ce chapitre nous introduisons une approche d'apprentissage où un réseau neurone convolutif a été entraîné sur des images simulées avec notre modèle d'inter-réflexion. Cela nous a permis de relâcher la condition d'avoir un éclairage connu et d'être plus robuste aux bruits dans l'image. Ce travail a été soumis au journal JOSA A.

- **Chapitre 7 :** nous donnons dans ce chapitre notre conclusion en ouvrant la porte à quelques travaux futur basés sur notre travail.

Color imaging technologies are based on the acquisition of light signals by the camera sensors. Image processing is then based on interpreting these light signals in order to find some physical attributes of the objects in the scene. Hence, in this chapter we recall some concepts in radiometry, light acquisition and color formation which are important to understand the manuscript. We provide also some ground concepts in colorimetry which are important to understand color vision, color spaces and color differences.

2.1 Radiometry

Light is an electromagnetic radiation (Figure 2.1) which can be classified by its frequency, or equivalently by its wavelength. Only part of the radiation is visible to human eyes. This is called the visible spectrum and ranges between $380nm$ and $760nm$.

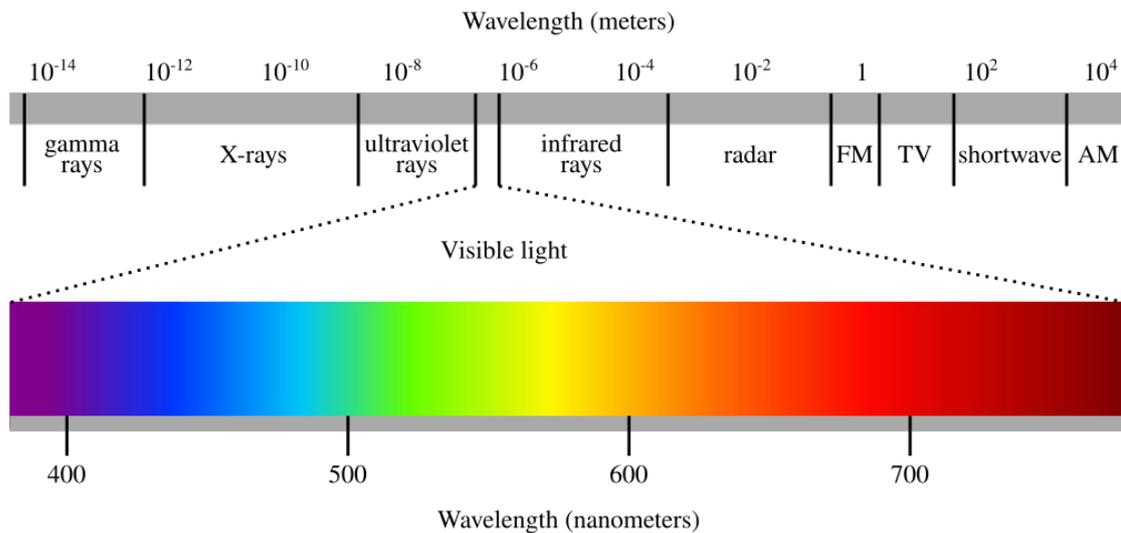


Figure 2.1: *Electromagnetic radiation of light.*

Light can be characterized by either photometric quantities when speaking about human vision, or by radiometric quantities which are the physical ones. In this thesis, our work focuses on computer vision and not on human vision, thus we are particularly interested in radiometric quantities.

2.1.1 Geometrical concepts

Geometry has an important role in describing the transport of light between light sources, objects and sensors in the scene. A light ray is used to represent the path the photons follow between two points. A point with a small area around it is called a **surface element**. It may receives photons from any direction. A subset of these directions is called a **solid angle**. An infinitesimally small solid angle, $d\omega$, defines a single direction which is generally characterized by its polar and azimuth angles (θ, φ) in spherical coordinates, and can be written as:

$$d\omega = \sin\theta d\theta d\varphi. \quad 2.1$$

The **geometrical extent**¹ describes the geometry of light rays propagating between two surfaces. Given two surface elements dP_i and dP_j , around two points, P_i and P_j , distant enough from each other, the elemental geometrical extent can be defined as [7]:

$$d^2G = \frac{dP_i \cos\theta_i dP_j \cos\theta_j}{\Delta^2}, \quad 2.2$$

where Δ is the euclidean distance between the centroids of the two surface elements, θ_i and θ_j are the angles between $P_i P_j$ and the normals of the two surface elements (see Figure 2.2).

The geometrical extent can be also written in terms of solid angle as:

$$d^2G = dP_i \cos\theta_i d\omega_j, \quad 2.3$$

or,

$$d^2G = dP_j \cos\theta_j d\omega_i, \quad 2.4$$

where $d\omega_j$ and $d\omega_i$ are the elemental solid angles corresponding to dP_j and dP_i respectively.

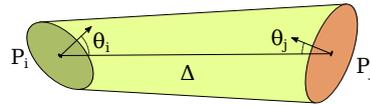


Figure 2.2: Geometrical extent between two surface elements.

2.1.2 Radiometric definitions

The **radiant flux** F , also called the radiant power, is the energy radiated per unit of time. It is independent of the direction and of the position in the space. The **spectral flux**, F_λ is the flux per unit of wavelength, it defines the spectral distribution of the radiation as:

$$F_\lambda = \frac{dF}{d\lambda} \quad 2.5$$

The **radiant intensity** I is a directional quantity which is not related to surfaces. It is obtained by taking the radiant flux per unit of solid angle. However, the word intensity is often found in computer vision under another meaning: it is used to refer to the gray level values in the image.

The radiant flux per unit of geometrical extent, G [8], is called the **radiance**, denoted as L :

$$L = \frac{d^2F}{d^2G}. \quad 2.6$$

¹Also called "étendue"

The light signals received by the eyes or the camera sensors are characterized by the spectral radiance defined as the radiance per unit of wavelength.

The **irradiance**, E , is defined as the radiant flux received by a unit of area, A :

$$E = \frac{dF}{dA}. \quad 2.7$$

Based on these definitions, the irradiance can be written in terms of radiance as follows:

$$dE = \frac{Ld^2G}{dA}. \quad 2.8$$

The irradiance received at a point P_i is denoted as $E(P_i)$, and the radiance from P_i to a second point P_j is denoted as $L(P_i, P_j)$. Then, and based on the definition of geometrical extent (see Equation 2.2), the previous equation can be written as:

$$\begin{aligned} dE(P_i) &= \frac{L(P_j, P_i)dP_i \cos\theta_i dP_j \cos\theta_j}{\Delta^2 dP_i} \\ &= \frac{L(P_j, P_i) \cos\theta_i dP_j \cos\theta_j}{\Delta^2}. \end{aligned} \quad 2.9$$

When a surface receives a light signal, it may reflect some amount of light, absorb some other, and transmit the rest, thus we talk about the reflectance and the transmittance of an object. Here we are interested only in the reflective properties of a surface. The **reflectance**, R of a surface of a specific material characterizes its effectiveness in reflecting radiant energy. When taking the reflectance at all wavelengths, the corresponding curve is called the *spectral reflectance*. However, how a surface reacts to light does not only depend on its reflective (or transitive) properties but also on the direction of incident light. In computer graphics and computer vision, it is common to use the **bidirectional reflectance distribution function**(BRDF) to characterize how light is reflected by an opaque surface. It is a function of the incoming light direction and the outgoing direction. It gives back the ratio of reflected radiance along the outgoing direction to the incident irradiance on the surface received from the incoming direction.

2.1.3 Lambert's law

As seen before, surfaces reflect the received light differently based on the incident direction of the light rays. However, some surfaces have the capacity of diffusing light uniformly over the hemisphere around it while sending the same amount of energy in every direction independently of the incident direction. Such a perfectly diffuse surface is called **Lambertian** as it diffuses light according to Lambert's law. It is in this type of surfaces that we are interested in this thesis.

For Lambertian surfaces, the radiance is independent of the propagation direction, then the radiance of P_i is denoted as $L(P_i)$, and, according to Lambert's law, it can be defined as:

$$L(P_i) = \frac{1}{\pi} E(P_i) R_i, \quad 2.10$$

where R_i is the reflectance of the point P_i .

Along this manuscript, direct irradiance received from the light source, and radiance after the first bounce of direct light are denoted as E_0 and L_1 respectively. The notations \mathbf{L} and \mathbf{E} refer to, respectively, the vectors of radiance and irradiance on various surface elements: $\mathbf{L} = [L(P_1) \ L(P_2) \ \dots]^T$ and $\mathbf{E} = [E(P_1) \ E(P_2) \ \dots]^T$.

Actually, the terms $L(P_i)$, $E(P_i)$ and R_i are all wavelength dependent, when used in this manuscript they are taken for only a single wavelength. However, when talking about spectral functions, the terms $R(\lambda)$, $L(\lambda)$ and $E(\lambda)$ are used for spectral reflectance, spectral radiance, and spectral irradiance respectively.

2.2 Light acquisition and color imaging

In computer vision and color science, light is generally characterized by its *spectral power distribution (SPD)*. It describes the power per unit of area per unit of wavelength of an illuminant. This quantity is equivalent in radiometry to the irradiance $E(\lambda)$ which is received by a unit of area of the surface assuming no energy loss.

What sensors receive as a radiation does not only depend on the properties of the light but also on the optical properties of the surfaces in the scene. As seen previously, a reflective surface will send back light rays according to its spectral reflectance or more accurately to its BRDF. The received light signal by one sensor is going to be associated with a single value based on the responsivity of this sensor. The responsivity is a spectral function and is the output of the photodetector divided by the radiant energy input at each wavelength. This term can be applied to both human eyes and physical detectors. For cameras, a spectral response curve is associated to each category of sensors, which are generally red, green and blue ones.

The color formation equation describing the raw (RGB) outputs of the camera at pixel, i , without taking into consideration the non-linearity of the sensors is generally written as:

$$\rho_i^k = \int_{\Omega} C^k(\lambda) L_i(\lambda) d\lambda, \quad 2.11$$

where Ω is the wavelength range over which the camera is sensitive, ρ_i^k is the k^{th} output of the camera (ρ_i^1 , ρ_i^2 and ρ_i^3 represent R_i , G_i and B_i respectively), and $C^k(\lambda)$ is the spectral response of the k^{th} sensor of the camera, and $L_i(\lambda)$ is the spectral radiance.

The radiance is a function of both the surface and the light spectral properties (see Equation (2.10)). In the general case, it depends also on the direction of incident and output light rays, thus on the BRDF of the surface. However, let us take into consideration a special case of surfaces, the Lambertian ones. Then, the raw (linear) RGB outputs of a camera, observing a Lambertian surface with spectral reflectance $R_i(\lambda)$, under the illuminant SPD $E(\lambda)$ is given by:

$$\rho_i^k = \frac{1}{\pi} \int_{\Omega} C^k(\lambda) E(\lambda) R_i(\lambda) d\lambda, \quad 2.12$$

The process of color formation and all the spectral functions implied in it are illustrated in Figure (2.3).

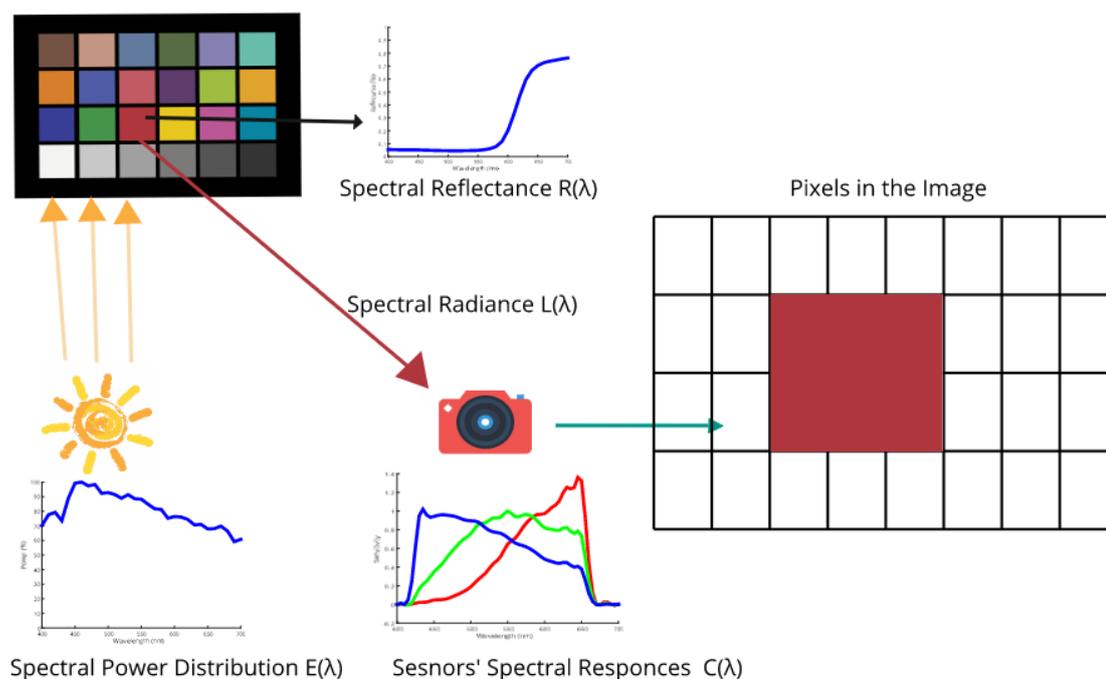


Figure 2.3: *Color Formation Process*

A non-fluorescent surface is not able to send back more light than it receives, thus the spectral reflectance function $R(\lambda)$ has its values in the range $[0, 1]$. The spectral responses functions $C(\lambda)$ of the camera also have their values in the range $[0, 1]$ as they define the extent to which the sensors are sensitive to the received light signal at each wavelength. However, the irradiance of light, $E(\lambda)$, can take whatever positive real value.

2.3 Color vision and colorimetry

Color is a physiological sensation and not a physical quantity, thus it is not possible to describe it in a physically quantifiable way. However, many experiments have been conducted in order to find a way to describe colors. As a result of these experiments, it has been found that three parameters suffice to classify colors, for example, the hue, chroma and saturation. In addition, the color of any light signal can be reproduced by mixing three colored lights, for example red, green and blue ones. Thus, a three dimensional mathematical model is able to describe color.

In order to be able to spectrally characterize human vision in a standard way, CIE defined RGB **color matching functions** for two different observation angles, 2° and 10° . These color matching functions are the amounts of the red, green and blue stimuli needed

to match a monochromatic stimuli for each wavelength. There exist, then, two sets of standard color matching functions: the CIE 2° RGB also called CIE 1931, and the CIE 10° RGB ones also called CIE 1964.

Based on the color matching functions, the seen color can be characterized in RGB color system as:

$$\begin{aligned} R &= \int_{\Omega} \bar{r}(\lambda)L(\lambda)d\lambda \\ G &= \int_{\Omega} \bar{g}(\lambda)L(\lambda)d\lambda \\ B &= \int_{\Omega} \bar{b}(\lambda)L(\lambda)d\lambda, \end{aligned} \tag{2.13}$$

where \bar{r} , \bar{g} and \bar{b} are the RGB color matching functions, and $L(\lambda)$ is the spectral radiance.

However, these color matching functions have some negative values. This happens when the targeted color cannot be matched by mixing the three primary colored lights because of the high saturation of the targeted monochromatic light. Then, in order to lower the saturation, the targeted light is mixed with one of the primary colored lights (which leads to the negative values) and matched with the other two. In order to avoid these negative values, new color matching functions XYZ can be obtained by a simple linear conversion. The seen color can be characterized in XYZ color system as:

$$\begin{aligned} X &= \int_{\Omega} \bar{x}(\lambda)L(\lambda)d\lambda \\ Y &= \int_{\Omega} \bar{y}(\lambda)L(\lambda)d\lambda \\ Z &= \int_{\Omega} \bar{z}(\lambda)L(\lambda)d\lambda, \end{aligned} \tag{2.14}$$

where \bar{x} , \bar{y} and \bar{z} are the XYZ color matching functions.

Equations (2.13) and (2.14) are very similar to Equation (2.11), the only difference is that the spectral response functions of the camera are replaced by the color matching functions. Therefore, the RGB values presented here have nothing to do with the RGB values in the image obtained by the camera sensors.

One problem of RGB and XYZ color spaces is that they are not uniform making it difficult to calculate the differences between colors. Uniform color spaces as CIELUV and CIELAB have been proposed and can be obtained easily from XYZ values. The idea is to be able to calculate the color difference, ΔE as an euclidean distance in these uniform spaces. However, it has been found that these color spaces are still not totally uniform, and further adjustments of color difference calculation depending on the position of the colors in the color space are still needed. There exist many formulas to calculate the color difference, one of the popular ones is called CIEDE2000 or CIEDE00 [9].

2.4 Color constancy, color appearance & color naming

Color constancy is the ability that humans have in perceiving the colors of objects almost unchanged under changing lightings [10, 11]. Thanks to the chromatic adaptation, objects' colors are seen as unchanged even when the spectral power distribution of the light changes, as for example when moving from indoor environment to outdoor one. Computational color constancy, on the other hand, groups a set of approaches with the goal of retrieving illumination independent surface descriptors from an image taken under an unknown lighting [12]. This is usually done by estimating the illumination color first. Based on this, the color of the objects which is invariant to the illuminant color can be retrieved.

CIE system of colorimetry has many advantages with the capacity of accurately predicting color matches for an average observer. However, it also suffer from some limitations as tristimulus values do not contain any information about the appearance of the color [13]. In order to describe color appearance, one needs more sophisticated scales regarding the attributes of visual perception: hue, brightness, lightness, colorfulness and chroma.

In color appearance terminology, a hue describes the similarity between an area and one of the perceived colors: red, green, blue, and yellow, or a combination of two of them. The brightness defines if an area is perceived to emit less or more light, whereas the lightness represents the perceived variation of brightness relatively to a similarly illuminated area. Colorfulness describes to which extent an area appears to be chromatic. Finally, the chroma is the relative colorfulness compared to a similarly illuminated area.

Color naming is a way to describe the color appearance using linguistics terms. For example, the common color names that one uses are hue descriptors. color naming and categorization is an active field of research. Color names can be culturally dependent [14, 15]. Many color naming methods and models exist to help categorize and understand the color perception [16]. The elementary color naming method [17] describes the hue as a mix of one or two hues from the four hues (red, green, blue, yellow). The brightness is described by the percentage of whiteness and blackness, whereas the colorfulness is related to the amount of chromaticness.

In this chapter, we are going to take a closer look at interreflections in order to understand their mechanism. The particularity of the resulted color gradients are going to be demonstrated and the roles they play both in human perception and in images are going to be discussed. It is these gradients which lead to the importance of interreflections in computer vision. At the end, all the previous uses of interreflections in the domain of computer vision are going to be listed.

3.1 The mechanism of interreflections

When a light ray hits a point on a concave surface, it may not exit the surface directly, instead it may bounce almost infinitely between the surface's points, see Figure (3.1). Then, the total irradiance at a given point P_i in the scene is the sum of the direct irradiance, $E_0(P_i)$, which is received from the light source, and indirect irradiances, $E_1(P_i), E_2(P_i), \dots$, after the first bounce of light, the second one, the third one and so on:

$$E(P_i) = E_0(P_i) + E_1(P_i) + E_2(P_i) + \dots$$

3.1

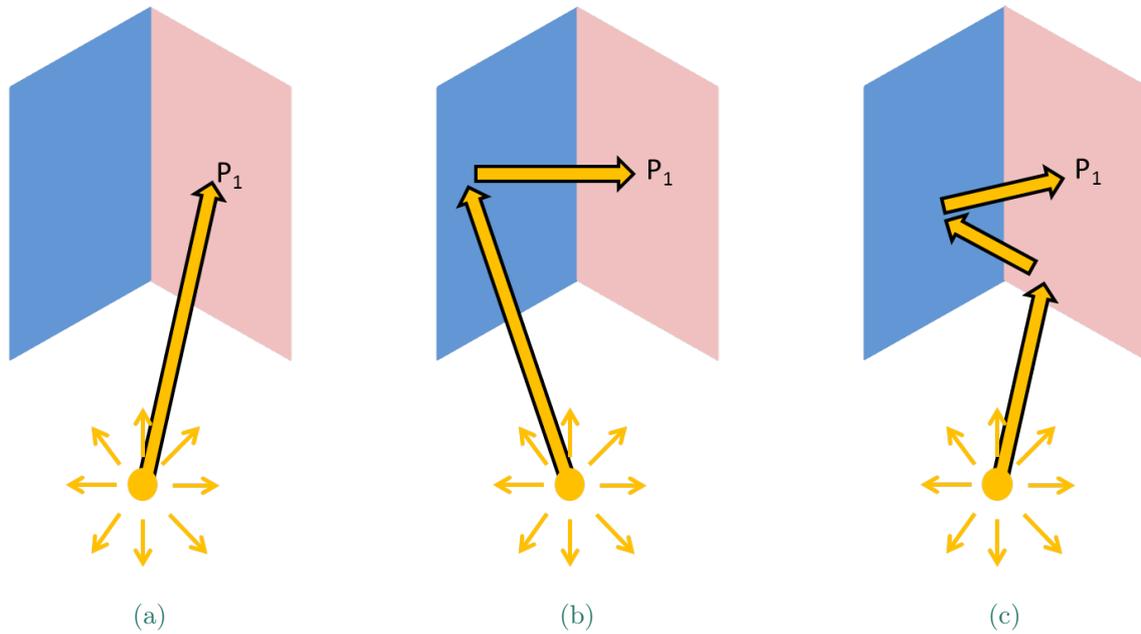


Figure 3.1: Decomposition of the irradiance received by P_1 : (a) Direct light, (b) first bounce of interreflection corresponding to what we call a two-bounce model, (c) second bounce of interreflection. For the sake of clarity, the sums over all the pixels (multiple arrows) are not illustrated here.

For a Lambertian surface, the irradiance in a given position P_i , received after one bounce of light from all other points P_j of the surface, can be written as:

$$E_1(P_i) = \int_{P_j \in \mathcal{S}} R_j \frac{E_0}{\pi} K(P_i, P_j) dP_j,$$

3.2

where $K(P_i, P_j)$ is the geometrical kernel (or interreflection kernel) subtended by any pair of points P_i and P_j in the scene, defined as:

$$K(P_i, P_j) = \frac{(\mathbf{N}_i \cdot \mathbf{P}_i \mathbf{P}_j)(\mathbf{N}_j \cdot \mathbf{P}_j \mathbf{P}_i)V(P_i, P_j)}{\Delta^4}, \quad 3.3$$

where \mathbf{N}_i and \mathbf{N}_j are the normals of surface elements i and j respectively, the dot operator "." is the dot product between two vectors, Δ is the euclidean distance between the two points, and $V(P_i, P_j)$ is a visibility function giving 1 when the two points can see each other and 0 otherwise. This function can be defined as follows:

$$V(P_i, P_j) = \frac{\mathbf{N}_i \cdot \mathbf{P}_i \mathbf{P}_j + |\mathbf{N}_i \cdot \mathbf{P}_i \mathbf{P}_j|}{2|\mathbf{N}_i \cdot \mathbf{P}_i \mathbf{P}_j|} \cdot \frac{\mathbf{N}_j \cdot \mathbf{P}_j \mathbf{P}_i + |\mathbf{N}_j \cdot \mathbf{P}_j \mathbf{P}_i|}{2|\mathbf{N}_j \cdot \mathbf{P}_j \mathbf{P}_i|}. \quad 3.4$$

Equation (3.3) is driven from the definition of the geometrical extent, (see Equations (2.2) and (2.9)). The reason behind that Δ is raised to the power 4 instead of the power 2, is the use of dot products instead of $\cos\theta_i$ and $\cos\theta_j$ in the numerator.

Similarly, taking into consideration the second bounce of light from all points $P_{j'}$, we can write:

$$E_2(P_i) = \int_{P_j \in \mathcal{S}} \int_{P_{j'} \in \mathcal{S}} R_j R_{j'} \frac{E_0}{\pi^2} K(P_i, P_j) K(P_j, P_{j'}) dP_j dP_{j'}, \quad 3.5$$

and for the third bounce:

$$E_3(P_i) = \int_{\mathcal{S}} \int_{\mathcal{S}} \int_{\mathcal{S}} R_j R_{j'} R_{j''} \frac{E_0}{\pi^3} K(P_i, P_j) K(P_j, P_{j'}) K(P_{j'}, P_{j''}) dP_j dP_{j'} dP_{j''}, \quad 3.6$$

and so on.

One can observe from these equations that oppositely to highlights, the spectral radiation of a ray is changed with each bounce depending on the spectral reflectance of the point it hits. Thus, each bounce can be considered as a new light source with different spectral properties. In addition, the number of bounces exchanged between a pair of points, P_i and P_j is related to their relative geometrical relation defined by the term $K(P_i, P_j)$. Close face to face points have big geometrical kernel values, thus they would exchange a high number of rays, while far points will exchange much less having a small value in the corresponding geometrical kernel. However, even for two close points with a high kernel value, the energy carried with a light ray reflecting from P_j toward P_i will be low at a given wavelength if R_j is low at this wavelength.

3.2 The effects of interreflections in images

It is, then, this complex relation that inter-reflected light rays have with the surface reflectance and its geometry which leads to color gradients in the image that are rich in photometric cues. It was observed in the literature that shape from shading methods failed to retrieve the accurate shape because of these color gradients. The retrieved shape when not accounting interreflections was less concave than the real one [1].

In order to better observe and understand the effects of interreflections on images, let us take into consideration the following case study: a V-shaped surface with a uniform spectral reflectance, illuminated by a frontal collimated light source. In Figure (3.2) a real camera output of a folded red paper with an approximative angle of 45° is shown, we can see that the red becomes more saturated when getting closer to the fold edge. Figure (3.3) shows a simulation of interreflections in a green paper folded with an angle of 45° also. Arrows show the direction of augmented effect of interreflections due to the geometrical factors.

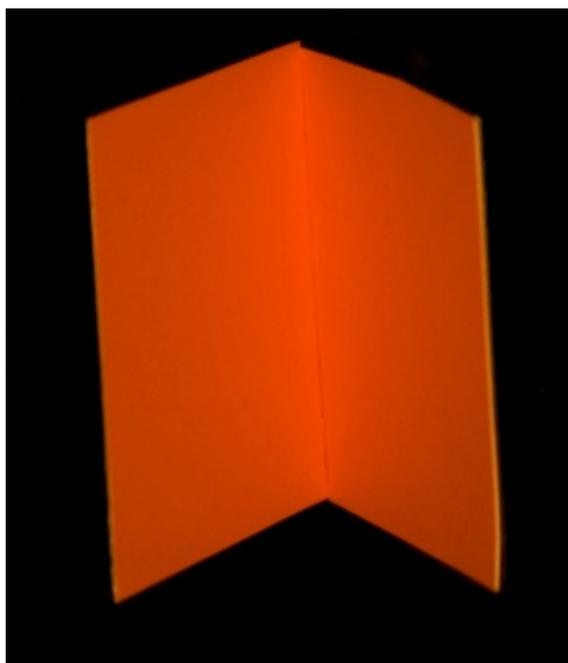


Figure 3.2: Camera output of a red V-shaped surface with an angle of 45°

Figures (3.4, 3.5) show 3D plots of raw RGB values of V-shaped surfaces with different angles when using a Sigma SD-10 Camera. A D65 light source was used where each point is assumed to receive the exact same amount of light. When analyzing these figures, we can see that the effects of interreflections increase when the surface spectral reflectance has high values. Low reflective surfaces show a very small dispersion of RGB values (Figure 3.4 middle)), as opposite to high reflective surfaces (Figure 3.4 top)). In addition, the biggest variance is seen on the axes that correspond to the reflective properties of the surface. A red surface will show the highest variance on the red axis while a green one on the green axis.

In order to have a better idea of color changes due to interreflections, we used CIE 1964 X_{10} , Y_{10} and Z_{10} color matching functions instead of the camera spectral response curves. Then, using the obtained XYZ values, we can get lightness, chroma and hue values in CIELUV color space¹.

Tables 3.1, 3.2 and 3.3 show the maximum lightness, chroma, hue and color differences calculated in terms of euclidean distances between colors due to interreflections of V-shaped

¹It is possible to use CIELAB instead of CIELUV. We chose CIELUV in order to be able to obtain the saturation values.

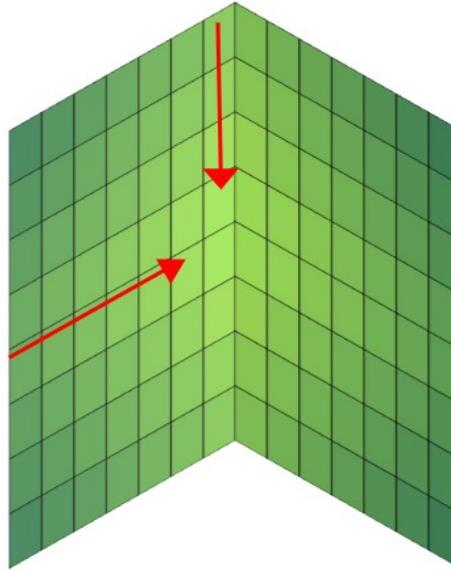


Figure 3.3: Simulated output of a green V-shaped surface with an angle of 45° . Arrows show the direction of increase amounts of interreflections' effects

surfaces of different angles for the same patches of the color checker presented in Figures (3.4) and (3.5). By observing the values in these tables, we can see that achromatic patches (white, gray and black ones) show the lowest change in hue. The yellow patch, on the other hand, shows the highest hue change. This relatively high hue change is due to the fact that the differences in its spectral reflectance between wavebands corresponding to the green color and those corresponding to the red color are accentuated with interreflections. One can notice also that the color differences go up with decreasing angle. They also increase when the spectral reflectance of the considered patch has high values, i.e. when the surface is highly reflective.

Table 3.1: Maximum color distances between the colors created by interreflections with an angle of 90°

Patch	90°			
	ΔL^*	ΔC^*	ΔH^*	ΔE_{uv}^*
White	16.60	2.39	0.12	16.77
Black	0.16	0.010	0.001	0.16
Yellow	9.81	17.96	6.63	21.52
Red	2.78	16.48	0.50	16.72
Green	2.75	5.83	0.41	6.46
Blue	0.91	5.86	0.78	5.98

3.3 Interreflections' role in perception

In this section, the effects of interreflections on human perception are explained based on psycho-visual experiments.

Harding et al. [18] tested shape perception in scenes that contained illumination gradi-

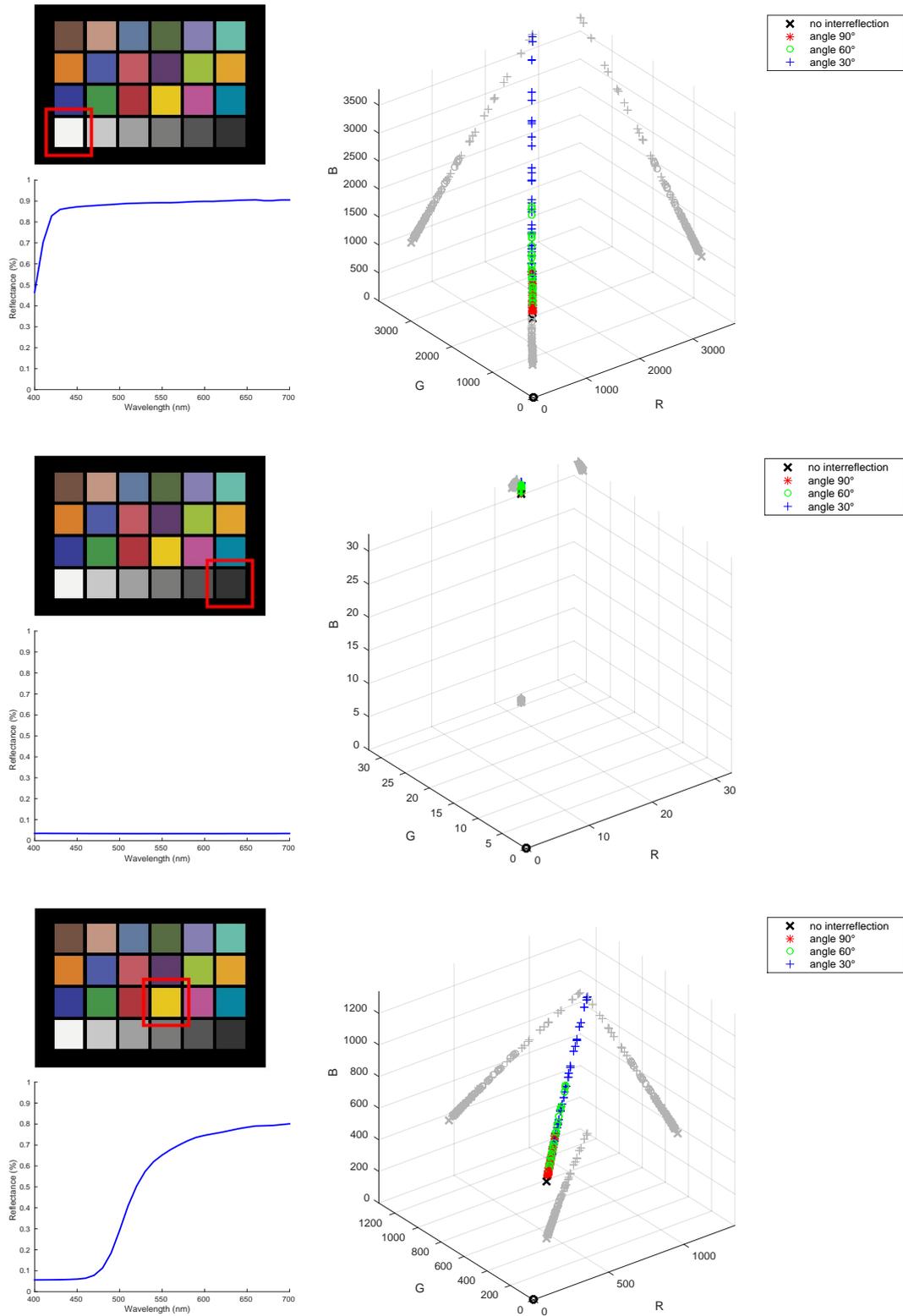


Figure 3.4: Simulated output in raw RGB space of a V-shaped surface of white, black and yellow Color Checker patches with different angles

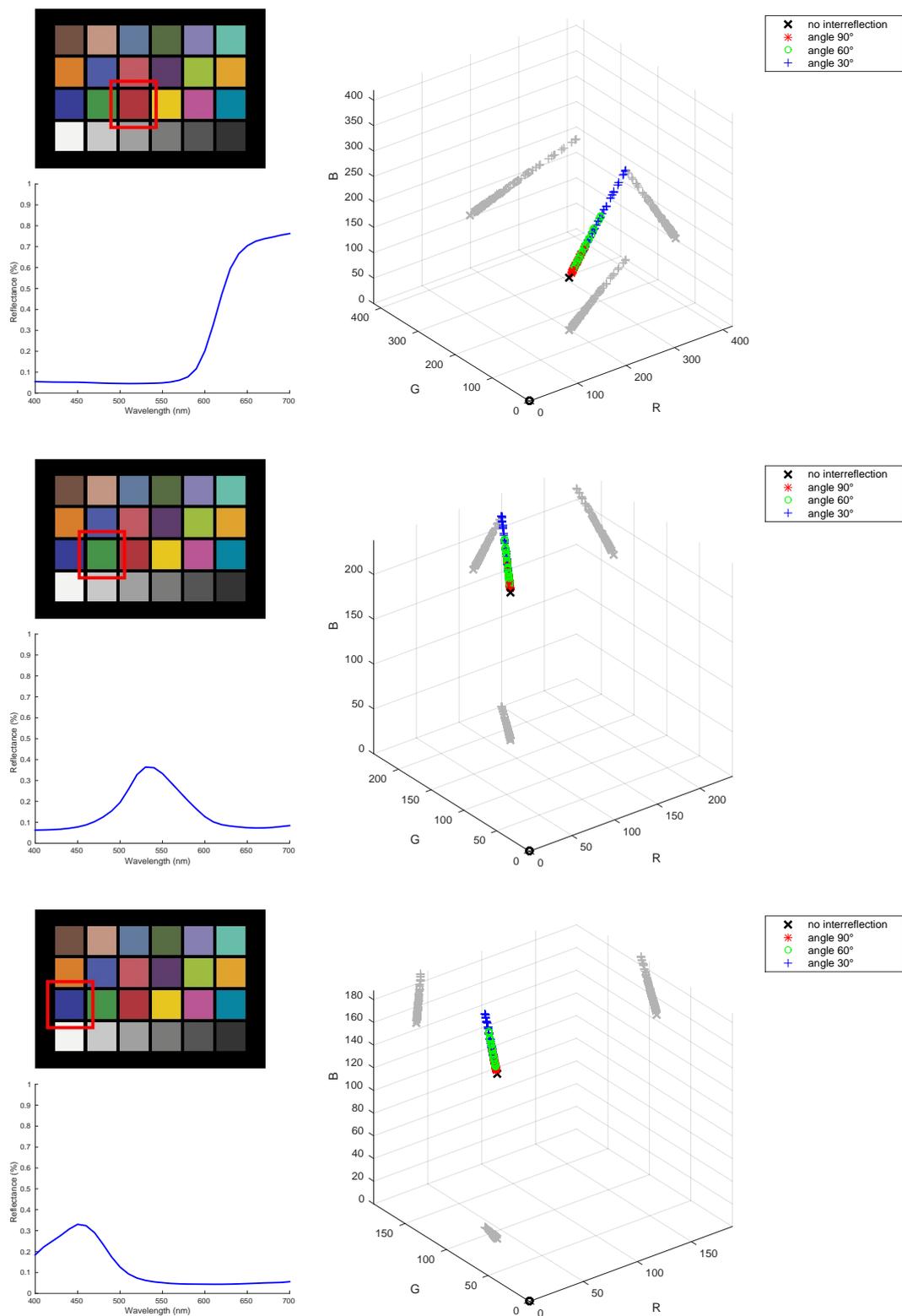


Figure 3.5: Simulated output in raw RGB space of a V-shaped surface of red, green and blue Color Checker patches with different angles

Table 3.2: Maximum color distances between the colors created by interreflections with an angle of 60° .

Patch	60°			
	ΔL^*	ΔC^*	ΔH^*	ΔE_{uv}^*
White	32.88	5.77	0.28	33.39
Black	0.25	0.016	0.001	0.25
Yellow	17.94	34.15	13.20	40.77
Red	4.99	30.10	0.94	30.52
Green	4.47	9.59	0.70	10.60
Blue	1.46	9.51	1.27	9.70

Table 3.3: Maximum color distances between the colors created by interreflections with an angle of 30° .

Patch	30°			
	ΔL^*	ΔC^*	ΔH^*	ΔE_{uv}^*
White	61.41	15.59	0.76	63.36
Black	0.29	0.019	0.001	0.29
Yellow	28.00	57.62	23.62	68.28
Red	7.56	46.41	1.53	47.04
Green	5.73	12.48	0.94	13.77
Blue	1.84	12.16	1.63	12.40

ent such as shading and interreflections. They showed that observers are able to use these gradients to make consistent shape judgments. Likewise, in [19], Bloj et al. demonstrated that the human visual system incorporates knowledge of interreflections at an early stage in color perception. They used a concave folded card with a magenta paper on one side and a white one on the other side. Interreflections lead to significant shift in the perceived color of the white side which goes from a desaturated pink in the corner to a more saturated magenta in the roof.

In this thesis, we propose to test if humans are better in color constancy when looking at flattened surfaces (without interreflections) than when looking at folded surfaces (with interreflections).

Our idea is to ask observers to look at colored papers (through one aperture) under unknown lighting conditions and to name the color of this paper using the elementary color naming method. Two geometric conditions were tested and compared: flattened paper, where no interreflections can be observed, and folded paper, where interreflections appear between the two paper sides (self-interreflections). For the psycho-visual experiment, 16 observers were asked to name the colors of 4 different colored papers: red, green, yellow and blue under 3 different colored lightings: red, green and blue. The observers stated the amount in percentage of blackness, whiteness and chromaticness of the colored paper they perceived. Subsequently, if they perceived chromaticness in it, its hue "red-yellow", "yellow-green", "green-blue" or "blue-red", must be evaluated in percentage. A hue angle was calculated according to estimated percentages of red, green, yellow and

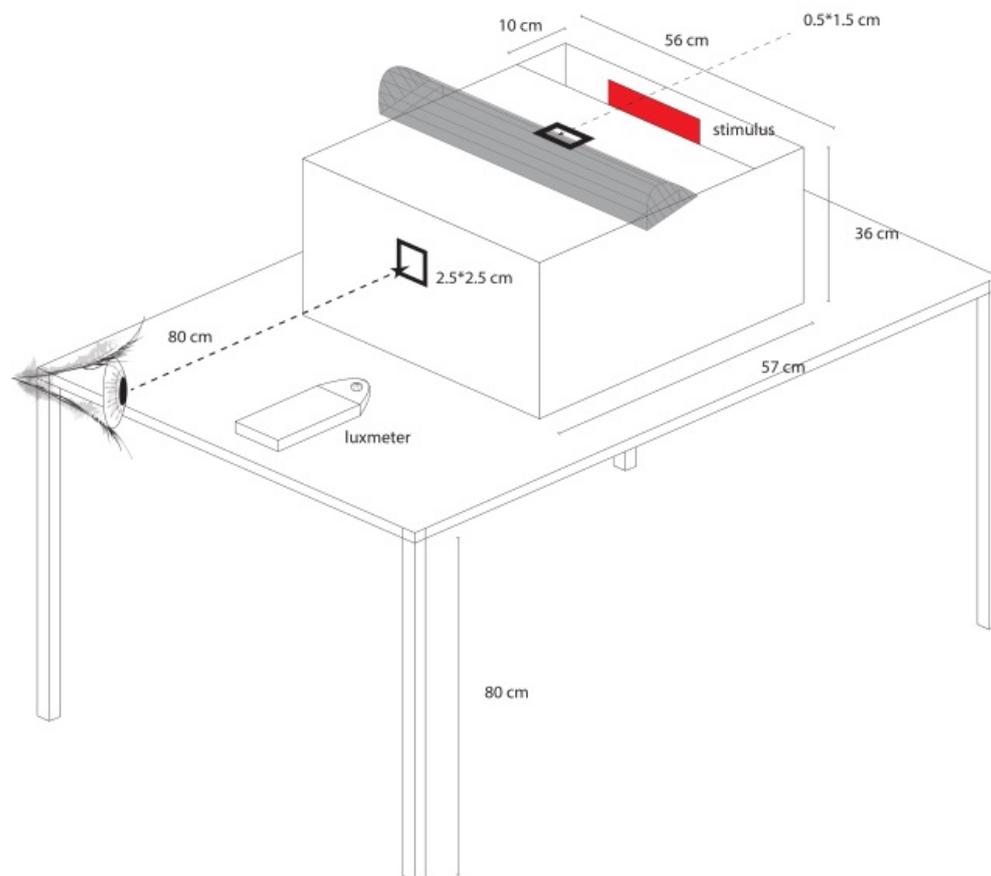


Figure 3.6: *Experimental configuration.*

blue. Subsequently, xy coordinates were computed according to hue angle and percentage of chromaticness.

The physical configuration of the experimental setup is illustrated in Figure 3.6. We have created a box in which we placed some (flattened or folded) colored papers. This box was in a dark room and we put a light source with colored filters just above the box, so that the observer could not see this light source but only the reflection of this source on the paper through an aperture. The observer was 80cm from the aperture and was asking to name the color of the paper he was seeing. If the observing conditions allow color constancy, the observer should name the current paper color without impact from the colored light source.

Figure 3.7 shows the results obtained from our psycho-visual experiment. Since we have used 4 different colored surfaces, we provide 4 different plots. The elementary color naming method provides a chromaticity point for each tested stimulus. Since we have 16 observers, we have 16 points for each tested stimulus. Instead of displaying all these points, we prefer to show ellipses that represent the distribution of these 16 color naming. The dashed ellipses represent the color naming results for the flattened papers (without interreflections) and the solid ellipses represent the results for the folded papers (with interreflections). On

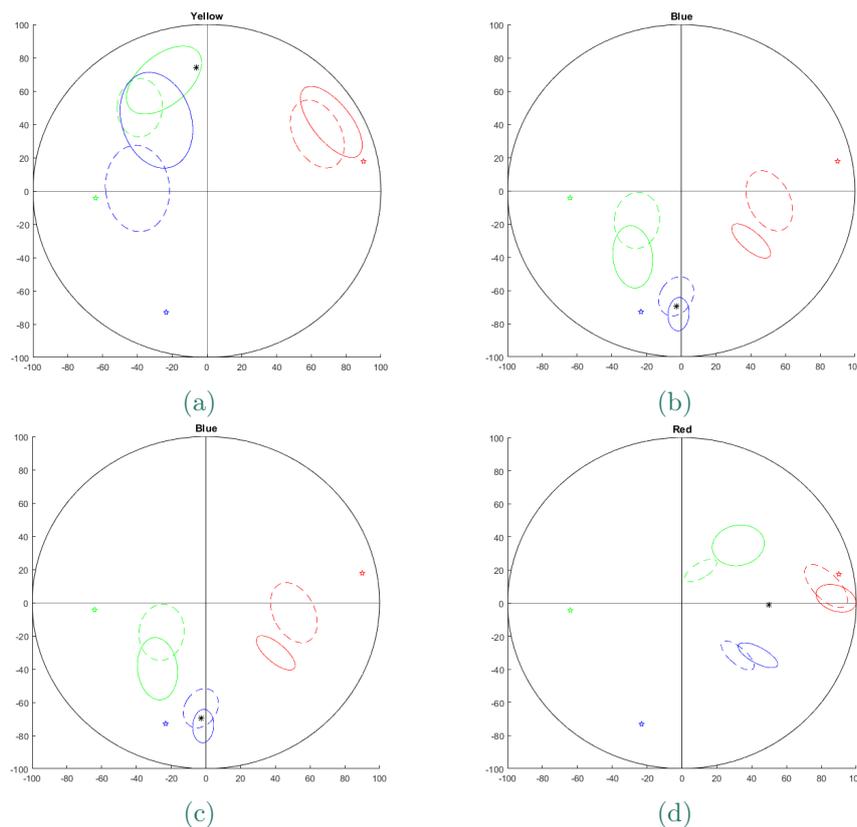


Figure 3.7: Chromatic shift due to interreflections, from dashed to solid ellipses. See text for details.

each plot, the black asterisk is the chromaticity point of the considered paper under white light. This is the point where all the color naming should be in case of perfect color constancy. The green, blue and red stars represent the chromaticity of the light source used to construct the green, blue and red ellipses, respectively.

Our experimental results clearly show that the observers are much better in estimating the color of the papers with interreflections than without. Indeed, most of time, the solid ellipses are closer to the black asterisk than the dashed ellipses. This is an experimental proof that interreflections help for human color constancy.

3.4 Previous use in computer vision

As seen before, interreflections plays an important role both in perception and in images. Interreflections gained a lot of attention in computer graphics. This process has been simulated iteratively by using different approaches in order to calculate the color values of each surface element of the scene to be displayed on a screen [20, 21, 22]. On the other hand, interreflections have been also studied in computer vision but to a lower extent. In this section a brief description of all the approaches that used interreflection in computer vision is given. A detailed description of some of these approaches will be given later in the following chapters.

Handling interreflection is important for many applications in computer vision. Some

of the state of the art approaches focused on removing the effects of interreflections, i.e. separating the direct radiance from indirect one (which is due to interreflections). This is mainly the case for shape-from-Shading approaches in order to be able to retrieve the correct shape, and in 3D-scanning using structured light methods. Other approaches tried to use interreflections as an extra information in domains such as color constancy, spectral imaging and photometric stereo. Meanwhile, in domains such as object recoloring, interreflections' effects need to be detected and re-applied after recoloring in order to give realistic rendering. This process is generally user-aided.

Other than the particularity of the domain of application, state of the art approaches differ in many other points. Some methods propose to model the interreflections, others do not. Some models are built on proper physical basis, others used a simplified modeling prone to errors in some cases. Some approaches take into consideration infinite number of bounces of inter-reflected light, others takes only one ray of inter-reflected light. Moreover, depending on the application and the method, prerequisites and preprocessing differ largely between one approach and the other.

Inspired by radiosity equation in computer graphics [20], Koenderink and Van Doorn [23] introduced an interreflection equation for Lambertian surfaces based on a definition of a geometrical kernel. This provided a rigorous model to handle interreflections in computer vision applications while taking into consideration infinite number of inter-reflected light rays. Different approaches in the literature, especially in shape from shading research field, are based on this equation. An analytical iterative solution to this equation exists in case of uniformly colored surfaces [23]. A more general solution to this equation is proposed by Nayar *et al.* [1]. It consists of sampling the surface into facets with a constant radiance and reflectance over each one of them. This discrete model was then used to iteratively correct, for gray surfaces, a pseudo shape and pre-estimated reflectance values obtained by applying a shape from shading method. Later, this approach was extended to handle colored surfaces [24] by taking into consideration three wavebands corresponding to RGB channels instead of a single wavelength. Seitz *et al.* [25] provided a theoretical proof of the existence of a linear matrix transformation to obtain the first-bounce radiance from the final radiance for any scene. An application is provided in case of Lambertian surfaces based on light field. This approach works in any lighting condition, even challenging ones like flashlight and non uniform light sources. In addition, it is robust in case of occlusions. However, the method needs the construction of Impulse Scatter Functions (ISFs) based on laser beams, which is not always easy or available for everyday applications. In [26], the authors described a method to separate n -bounced light rays in order to remove the effect of interreflections. However, no light field or ISF matrix is needed. The only preconditions are: a single colored surface, and the possibility to illuminate the scene with various light sources of different SPDs. This method is suitable for real-time applications and can be applied for all scenes whatever the geometry is, and under any static light sources. This approach can be extended in order to separate more bounces under the condition of adding a lighting with a different SPD for each additional bounce needed. However, this algorithm is applicable only to surfaces having a uniform reflectance. In addition, two calibration steps are needed using a gray-scale sphere and a non-concave surface of the same material

as the surfaces in the scene. Later, Fu *et al.* [27] extended this approach using fluorescent surfaces. They proposed to remove the effects of interreflections by a bi-spectral observation of fluorescent and reflective components. Their approach is based on a single image, under the assumption of available fluorescent materials.

Other Radiosity-based approaches are also worth to be listed, Shimshoni *et al.* [28] used an interreflection model similar to [24] coupled with a reflectance model for specular objects to recover the shape of polyhedral objects. Stewart and Langer [29] adapted a similar interreflection model to be used for shape recovery under diffuse lighting. It has also been shown by Chandraker *et al.* [30] that interreflections can resolve generalized bas-relief ambiguity in uncalibrated photometric stereo as interreflections are distance-dependent phenomena.

In another work, Funt and Drew [31] focused their attention on analyzing the RGB color space to separate the color with no interreflection from the color after one bounce of interreflection. They consider two homogeneous colored flat surfaces having different spectral reflectance functions. The light is assumed to be constant regarding its spectral power distribution but its irradiance can vary with position. The main idea behind this work is that a plane in RGB space can be formed from two colors: the color after one bounce of light, i.e. without taking into consideration interreflections, called one-bounce color, and the color after two bounces of light, called two-bounce color. Then, the intersection between the two planes corresponding to the two surfaces gives the two-bounce RGB color.

In color constancy and spectral estimation domains, some approaches [32, 33, 34] in the literature focused on the use of interreflections as extra information in order to obtain surface spectral reflectance and illuminant spectral power distribution. Each of these spectral curves is modeled with three basis functions. The used methods were based on adjacent panels having different spectral reflectance functions, and only one bounce of inter-reflected light was considered. The authors assume the existence of a point not affected by interreflections at each side of the surface leading to 9 equations to solve for the 9 unknowns. In addition, due to some used approximations, these approaches cannot handle highly reflective surfaces.

Based on the dichromatic reflectance model proposed by Shafer [35], Bajcsy *et al.* [36, 37] analyzed interreflections by dividing them into four components, interface-interface, body-interface, interface-body and body-body reflections. The authors proposed a method that can handle and separate some interreflections in a metric space of intensity, hue and saturation in order to perform color image segmentation for dielectric materials. In [38], the authors located image areas affected by interreflections in the chromaticity sphere. Given that object colors are already estimated, they found the interreflection color by plane intersection.

In spectral imaging, Tominaga *et al.* [39, 40] proposed to use interreflections between fluorescent objects in order to characterize the spectral image component from the observed spectral radiance. Their method is based on a two-bounce bi-spectral interreflection model.

In the domain of object recoloring, handling interreflections can be necessary in order to give realistic rendering. A common approach for Lambertian scenes is to estimate two intrinsic images: reflectance intrinsic image related to surface reflectance, and illumination

intrinsic image related to light's SPD. Then, in order to take into consideration diffuse interreflections, Carroll *et al.* [41] proposed a user-assisted method to separate the illumination intrinsic image into direct and multiple indirect components. In another work, Beigpour and Van de Weijer [42] proposed to estimate Multi-illuminant Dichromatic Reflection model for a pre-segmented single-colored object in order to deal with multiple light sources, and indirect light sources such as interreflections.

Interreflections, as well as other indirect illumination, have been also studied in the domain of 3D-scanning using structured light. Nayar *et al.* [43] presented a method to separate the radiance into direct and indirect components by using high frequency illumination patterns. Based on this work, many researchers [44, 45, 46] have proposed to perform scanning under indirect illumination using structured light. However, these methods may fail in case of strong interreflections, in addition they suffer from noise and need a high number of images compared to the traditional techniques. In order to better handle interreflections, Couture *et al.* [47] used band-pass unstructured patterns, but a large number of images is still required in this approach. Recently, Gupta *et al.* [48] analyzed the error in structured lights introduced by indirect illumination and proposed to combine different binary structured light patterns, each resilient to an individual indirect illumination in order to obtain a 3D scanning system applicable in the presence of a broader range of indirect illumination while needing fewer images compared to the previous approaches.

In the previous chapter, we have discussed the importance of interreflection phenomenon and its possible applications in computer vision. One can also argue that having an interreflection model which is constructed on the basis of physics is highly important in many applications. Such a model should be able to accurately predict the RGB coordinates of a pixel corresponding to a small area of a surface with interreflections.

Even though some previously presented approaches did not try to provide a model for interreflections, many of the existing approaches are based on a physical definition of interreflections. In the following, these approaches are detailed and discussed. Afterwards, we introduce our model that aims at expressing the RGB values in terms of the physical and optical properties of the surface, the lighting in the scene, and the acquisition device. Finally, the performance and the accuracy of this model are discussed.

4.1 The Continuous Model

Inspired by radiosity equation in computer graphics [20], Koenderink and Van Doorn [23] introduced an interreflection equation for Lambertian surfaces based on a definition of a geometrical kernel. Different approaches in the literature, especially in shape from shading research field, are based on this equation. Based on the works of Koenderink and Van Doorn [23], Forsyth and Zisserman [49, 50, 51] and Nayar *et al.* [1], the interreflection equation of the final radiance at point P_i , $L(P_i)$, can be written in terms of direct radiance, $L_1(P_i)$ as:

$$L(P_i) = L_1(P_i) + \frac{R_i}{\pi} \int_{P_j \in \mathcal{S}} K(P_i, P_j) L(P_j) dP_j, \quad 4.1$$

where \mathcal{S} is the area containing all surface elements of the scene, and $K(P_i, P_j)$ is the geometrical kernel defined in Equation 3.3.

Equation (4.1) does not have a straightforward solution for $L(P_i)$. However, an analytical solution to this equation exists in case of uniformly colored surfaces [23], by modeling infinite bounces of direct radiance, using Neumann series:

$$L(P_i) = L_1(P_i) + \sum_{b=1}^{\infty} R^b \int_{P_j \in \mathcal{S}} K_b(P_i, P_j) L_1(P_j) dP_j, \quad 4.2$$

where b denotes the number of bounces,

$$K_b(P_i, P_j) = \int_{P_k \in \mathcal{S}} \frac{K(P_i, P_k)}{\pi} K_{b-1}(P_k, P_j) dP_k, \quad 4.3$$

and,

$$K_1 = \frac{K}{\pi}, \quad 4.4$$

The solution, valid only in the case of Lambertian surfaces with a uniform reflectance, can be computed iteratively.

4.2 The discrete model

A more general solution to the interreflection equation is proposed by Nayar *et al.* by sampling the surface into m facets [1]. Both radiance and reflectance are considered to be constant over each facet. Considering the facets to be infinitesimally small, a facet i is represented by a point P_i and its area is denoted as ΔP_i . Then, equation (4.1) can be written as:

$$L(P_i) = L_1(P_i) + \frac{R_i}{\pi} \sum_{j \neq i} L_j K_{ij}, \quad 4.5$$

where:

$$K_{ij} = K(P_i, P_j) \Delta P_j. \quad 4.6$$

Then, the authors introduced a matrix form for the geometrical kernel:

$$\mathbf{K} = \frac{1}{\pi} \begin{bmatrix} 0 & K_{12} & \cdot & \cdot & K_{1m} \\ K_{21} & 0 & \cdot & \cdot & K_{2m} \\ \cdot & \cdot & 0 & \cdot & \cdot \\ K_{m1} & \cdot & \cdot & \cdot & 0 \end{bmatrix}. \quad 4.7$$

Note that this matrix is always symmetric, i.e., $K_{ij} = K_{ji}$.

Moreover, as each facet has its own spectral reflectance, those for all the facets can be gathered into a diagonal matrix, denoted as \mathbf{R} :

$$\mathbf{R} = \begin{bmatrix} R_1 & 0 & \dots & 0 \\ 0 & R_2 & \dots & 0 \\ \cdot & \cdot & \cdot & \dots \\ \cdot & \cdot & \cdot & \dots \\ 0 & 0 & \dots & R_m \end{bmatrix}. \quad 4.8$$

Then, equation (4.5) can be written as:

$$\mathbf{L} = \mathbf{L}_1 + \mathbf{RKL}, \quad 4.9$$

or:

$$\mathbf{L} = (\mathbf{I} - \mathbf{RK})^{-1} \mathbf{L}_1, \quad 4.10$$

where \mathbf{L} and \mathbf{L}_1 are vectors containing the values on each facet corresponding to respectively the final radiance as perceived by the observer, and the radiance after the first bounce of direct light.

4.3 Some existing applications of the interreflection models

These interreflections models were generally used to remove the effect of interreflections based on different techniques.

4.3.1 Interreflections removal based on shape from shading

Nayar *et al.* [1] used Equation (4.10) to iteratively correct, for gray surfaces, a pseudo shape and pre-estimated reflectance values obtained by applying a shape from shading method.

The authors' first approach handled only one wavelength. Later, they extended it to three wavebands corresponding to RGB channels [24]. They considered that the image produced by a sensor with a narrow-band filter represents reflections and interreflections of almost a monochromatic light. The shape from shading method and the recovery of actual shape using interreflections are applied separately for each color channel.

In these approaches, both geometry and spectral reflectance are unknown and need to be estimated. Surfaces with different reflectances per position can be handled, but this method requires that the surface has no discontinuity.

4.3.2 Interreflection removal via light field

In the work by Seitz *et al.* [25], a theoretical proof of the existence of a linear matrix transformation to obtain the first-bounce radiance from the final radiance for any scene is presented. An application of this theoretical model is also presented for Lambertian surfaces for any lighting condition and without the use of a shape from shading method.

The existence of interreflection cancellation operators enables the computation of n -bounce images (i.e. images that would be obtained by capturing only light rays being bounced n times), for each n , by matrix multiplication. This stands for any scene geometry and surface BRDF.

The total radiance between two points written as:

$$L(P_i, P_j) = L_1(P_i, P_j) + L_2(P_i, P_j) + L_3(P_i, P_j) + \dots, \quad 4.11$$

can be expressed as a light transport equation:

$$L(P_i, P_j) = L_1(P_i, P_j) + \int_{k \in \mathcal{S}} A_{P_j}(P_k, P_i) L(P_k, P_i). \quad 4.12$$

The term $A_{P_j}(P_k, P_i)$ is a general term: it defines the proportion of irradiance from point P_k to P_i that gets transported as radiance towards P_j . It is a function of the surface BRDF and of the visibility between the two points. Thus, this equation is similar to the interreflection equation (4.1) but it is more general as it can handle non-Lambertian surfaces.

By presenting the previous equation in matrix form and developing it, one obtains:

$$\mathbf{L} = (\mathbf{I} - \mathbf{A})^{-1} \mathbf{L}_1. \quad 4.13$$

Equation (4.10) is a special case of this equation (4.13).

Then, for any scene, one can write:

$$\mathbf{L}_1 = \mathbf{C}_1 \mathbf{L}, \quad 4.14$$

where \mathbf{C}_1 is defined as $(\mathbf{I} - \mathbf{A})$. Therefore, it is possible to remove interreflections by multiplying the radiance vector with the linear matrix \mathbf{C}_1 .

Moreover, the authors proved that any bounce of light can be found iteratively by treating indirect light coming from a point as if it was a direct light.

Later, the authors proposed to relate equation (4.13) to image intensities: for a ray r_i with a unit radiance illuminating a point P_i of the scene, the resulting light field \mathbf{t}_i captured by the sensor represents the full light transport in response to that ray. Vector \mathbf{t}_i is called Impulse Scatter Function (ISF). For a surface represented by m points, and by concentrating the m corresponding rays, we can define a matrix \mathbf{T} as $[\mathbf{t}_1 \ \mathbf{t}_2 \ \dots \ \mathbf{t}_m]$.

The authors considered the light to be linear, then the cancellation operator can be written as:

$$\mathbf{C}_1 = \mathbf{T}_1 \mathbf{T}^{-1}, \quad 4.15$$

where \mathbf{T}_1 is the matrix containing Impulse Scatter Function (ISF) due to 1-bounce reflection (with no interreflections).

Note that, although \mathbf{T} is measurable, \mathbf{T}_1 is generally not, so this is only a theoretical proof. However, for Lambertian scenes and based on images taken at a single camera viewpoint, the cancellation operator is unique, and \mathbf{T}_1 is a diagonal matrix. In addition, we can write:

$$\mathbf{T}_1[i, i] = \frac{1}{\mathbf{T}^{-1}[i, i]}. \quad 4.16$$

In the case of Lambertian surfaces, the incident and outgoing light fields are both two dimensional, ISF matrix can be captured by scanning a narrow beam of light (Laser) over the surface and capturing an image for each position of the beam.

4.3.3 Interreflections removal without light field

Another method to remove interreflections is proposed in [26]. The authors describe a method to separate n -bounced light rays in order to remove the effect of interreflection. However, no light field or ISF matrix is needed. The only preconditions are: a single colored surface, and the possibility to illuminate the scene with various light sources of different SPDs.

The main idea in this approach is that, for Lambertian surfaces with uniform reflectance, whenever a light ray is bounced, its corresponding intensity in the image is attenuated by the factor R corresponding to the albedo of the surface. Thus, the measured intensity value can be written as a polynomial function of R , and the contribution of an n -bounced ray can be expressed as R^n .

Based on a similar definition of the geometrical kernel as introduced before, and considering only Lambertian surfaces with a homogeneous color, the authors rewrote equation (4.2) as:

$$L(P_i) = \frac{E_0(P_i)}{\pi} R + \sum_{n=1}^{\infty} R^n \int_{j \in \mathcal{S}} K_n(P_i, P_j) \frac{R E_0(P_j)}{\pi} dP_j. \quad 4.17$$

This equation can be seen as a polynomial function of R , and can be rewritten as:

$$L(P_i) = \frac{E_0(P_i)}{\pi} R + \sum_{n=2}^{\infty} \left(\int_{j \in \mathcal{S}} K_{n-1}(P_i, P_j) \frac{E_0(P_j)}{\pi} dP_j \right) R^n. \quad 4.18$$

By defining:

$$\begin{aligned} C_1(P_i) &= \frac{E_0(P_i)}{\pi}, \\ C_2(P_i) &= \int_{j \in \mathcal{S}} K_1(P_i, P_j) \frac{E_0(P_j)}{\pi} dP_j, \\ &\dots \\ C_n(P_i) &= \int_{j \in \mathcal{S}} K_{n-1}(P_i, P_j) \frac{E_0(P_j)}{\pi} dP_j, \end{aligned} \quad 4.19$$

one can write:

$$L(P_i) = \sum_{n=1}^{\infty} C_n(P_i) R^n. \quad 4.20$$

By considering only the first two bounces of light, the previous equation becomes:

$$L(P_i) = C_1(P_i) R + C_2(P_i) R^2. \quad 4.21$$

Note that the terms C_n depend upon direct irradiance and scene geometry. When the radiance $L(P_i)$ is measured twice ($L^{(1)}(P_i), L^{(2)}(P_i)$) for the same lighting and scene geometry but for two different known albedos of the surface (R_1, R_2), the following two equations (with two unknowns C_1 and C_2) can be written as:

$$\begin{aligned} L^{(1)}(P_i) &= C_1(P_i) R_1 + C_2(P_i) R_1^2, \\ L^{(2)}(P_i) &= C_1(P_i) R_2 + C_2(P_i) R_2^2. \end{aligned} \quad 4.22$$

There are two major difficulties with these equations. The first one is how to change surface albedo easily. The second is how to determine the surface albedo.

For the first issue, the authors proposed to simulate the change of surface albedo by varying the light color, based on the idea that albedo is light spectrum dependent. However, C_1 and C_2 depend upon direct irradiance, so the previous equation solving holds only when lighting is the same. In order to keep only two unknowns in the equation, the authors calibrated the light source to calculate the ratio between the irradiance received from different light sources. For this purpose, they use a gray convex object having a wavelength-independent albedo for any light spectrum.

Assuming that the irradiance with the second lighting is α times the irradiance with the first one, i.e., $E_0^{(2)}(P_1) = \alpha E_0^{(1)}(P_1)$, and considering that the ratio between the two albedos, β , is known, the equations in (4.22) become:

$$\begin{aligned} L^{(1)}(P_i) &= C_1(P_i)R_1 + C_2(P_i)R_1^2, \\ L^{(2)}(P_i)/\alpha &= \beta C_1(P_i)R_1 + \beta C_2(P_i)R_1^2. \end{aligned} \quad 4.23$$

Knowing α and β , the terms $C_1(P_i)R_1$ and $C_2(P_i)R_1^2$ can be calculated.

Another calibration step is needed to obtain β by using a non-concave object of the same material, thereby having the same reflectance, as the object of the scene. This calibration step is used to calculate the ratio between image intensities corresponding to the used lights.

By observing the previous approaches, one can see that all of them wrote the final radiance in terms of direct radiance, with the objective of canceling the effects of interreflection or of correcting an already estimated shape. None of these approaches proposed a spectral model and some of them used only one bounce of inter-reflected light.

In the following, we propose an original spectral infinite-bounce model which is more general than the ones presented earlier. It relates directly a physical model of interreflection to image values.

4.4 Reviewing the discrete model

To better understand the interreflection model, let us start by expressing each bounce of inter-reflected light. As seen in Chapter 3, the total irradiance at a given point P_i in the scene can be written as the sum of direct irradiance, $E_0(P_i)$, received from the light source, and indirect irradiances, $E_1(P_i), E_2(P_i), \dots$, after the first bounce of light, the second one, the third one, etc.

By sampling the surface into a m infinitesimal facets as proposed by Nayar *et al.* [1], the irradiance after first, second and third bounces can be written as:

$$E_1(P_i) = \sum_{j \in \mathcal{S}} R_j \frac{E_0}{\pi} K(P_i, P_j) dP_j, \quad 4.24$$

$$E_2(P_i) = \sum_{j \in \mathcal{S}} \sum_{j' \in \mathcal{S}} R_j R_{j'} \frac{E_0}{\pi^2} K(P_i, P_j) K(P_j, P_{j'}) dP_j dP_{j'}, \quad 4.25$$

$$\begin{aligned} E_3(P_i) &= \sum_{j \in \mathcal{S}} \sum_{j' \in \mathcal{S}} \sum_{j'' \in \mathcal{S}} R_j R_{j'} R_{j''} \frac{E_0}{\pi^3} K(P_i, P_j) K(P_j, P_{j'}) \\ &\quad K(P_{j'}, P_{j''}) dP_j dP_{j'} dP_{j''}. \end{aligned} \quad 4.26$$

Again, both irradiance and reflectance are considered to be constant over each facet.

Note that $K(P_i, P_j)$ is the element (i, j) of matrix \mathbf{K} presented earlier in equation (4.7). Note also that R_i is the element (i, i) of the matrix \mathbf{R} presented previously in equation

(4.8). Then, the irradiance vector, corresponding to the surface irradiance in its m facets, after n bounces of light can be written as:

$$\mathbf{E} = \sum_{b=0}^{b=n} (\mathbf{KR})^b \mathbf{E}_0. \quad 4.27$$

Note that the division on π has been included inside \mathbf{K} . This sum is a geometric series, which, when n goes to infinity, converges to:

$$\mathbf{E} = (\mathbf{I} - \mathbf{KR})^{-1} \mathbf{E}_0. \quad 4.28$$

The convergence is guaranteed for any non-fluorescent surface having no zeros in its spectral reflectance (see Appendix A for proof).

This is a general equation of irradiance after infinite bounces of light for a Lambertian scene. This equation is a function of wavelength.

The total irradiance \mathbf{E} is unknown and cannot be measured easily, in opposition to the total radiance. Hence, equation (4.28) can be written in terms of total radiance as:

$$\begin{aligned} \mathbf{L} &= \frac{1}{\pi} \mathbf{R}(\mathbf{I} - \mathbf{KR})^{-1} \mathbf{E}_0 \\ &= \frac{1}{\pi} (\mathbf{R}^{-1} - \mathbf{K})^{-1} \mathbf{E}_0. \end{aligned} \quad 4.29$$

4.4.1 Relation with Nayar's model [1]

At first sight, equation (4.29) seems different from (4.10) proposed by Nayar *et al.* [1] (cf. section 4.2). However, by developing equation (4.10) in order to write it in term of irradiance, we can show that they are equivalent:

$$\frac{1}{\pi} \mathbf{R} \mathbf{E} = (\mathbf{I} - \mathbf{RK})^{-1} \frac{1}{\pi} \mathbf{R} \mathbf{E}_0. \quad 4.30$$

Then,

$$\begin{aligned} \mathbf{E} &= \mathbf{R}^{-1} (\mathbf{I} - \mathbf{RK})^{-1} \mathbf{R} \mathbf{E}_0 \\ &= ((\mathbf{I} - \mathbf{RK}) \mathbf{R})^{-1} \mathbf{R} \mathbf{E}_0 \\ &= ((\mathbf{R}^{-1} (\mathbf{I} - \mathbf{RK}) \mathbf{R}))^{-1} \mathbf{E}_0 \\ &= (\mathbf{I} - \mathbf{KR})^{-1} \mathbf{E}_0 \end{aligned} \quad 4.31$$

4.4.2 Generalization to spectral model

As seen before, in a vision system, the radiance is captured by camera sensors, and is converted to image intensity values based on sensors' response functions as follows:

$$\rho_{\lambda}^j = \int_{\lambda_1}^{\lambda_2} C^j(\lambda) L(\lambda) d\lambda. \quad 4.32$$

By sampling the spectrum, this equation can be written in a matrix form:

$$\boldsymbol{\rho} = \mathbf{C}\mathbf{L}_\lambda, \quad 4.33$$

where $\boldsymbol{\rho}$ is a column vector containing the sensor values of a pixel, \mathbf{C} is the camera response matrix of size $s \times l$ where s is the number of channels and l is the number of wavelengths, and \mathbf{L}_λ is a column vector containing the radiance values at all different wavelengths.

Let us note that the previous equation is defined for one pixel and all wavelengths. However, the equation of radiance with interreflection introduced earlier is defined for one wavelength and for all pixels in order to take into consideration the geometrical kernel. Thus, in order to be able to combine the two equations, we need to extend the matrices to take into consideration both all pixels and all wavelengths.

We propose to extend the vector $\boldsymbol{\rho}$ to the vector $\boldsymbol{\rho}_{ext}$ whose size is $ms \times 1$, where m is the number of pixels in the image corresponding to the concave surface, and whose form is:

$$\boldsymbol{\rho}_{ext} = \left[\rho_1^1 \quad \dots \quad \rho_m^1 \quad \rho_1^2 \quad \dots \quad \rho_m^2 \quad \dots \quad \rho_1^b \quad \dots \quad \rho_m^b \right]^T. \quad 4.34$$

The vector of irradiance \mathbf{E}_0 is extended in a similar way. The extended form, $\mathbf{E}_{0,ext}$, has the size of $ml \times 1$ by concatenating the \mathbf{E}_0 vector for each wavelength one after the other.

The reflectance matrix \mathbf{R} is extended to the square diagonal matrix \mathbf{R}_{ext} whose size is $ml \times ml$. As \mathbf{R} is defined for one wavelength, we can name it \mathbf{R}_λ , then the extended form is a concatenation of all diagonal matrices, \mathbf{R}_λ , on the diagonal of the new matrix:

$$\mathbf{R}_{ext} = \begin{bmatrix} \mathbf{R}_1 & 0 & \dots & 0 \\ \cdot & \mathbf{R}_2 & \dots & \cdot \\ \cdot & \dots & \cdot & 0 \\ 0 & \cdot & \dots & \mathbf{R}_l \end{bmatrix}. \quad 4.35$$

The geometrical kernel matrix \mathbf{K} is extended to a $ml \times ml$ block matrix whose blocks on the diagonal are \mathbf{K} and all others are zero block matrices:

$$\mathbf{K}_{ext} = \begin{bmatrix} \mathbf{K} & 0 & \dots & \cdot & \cdot & \cdot & \cdot & 0 \\ \cdot & \mathbf{K} & \cdot & \dots & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & \cdot & \cdot & \cdot & \dots & \mathbf{K} \end{bmatrix}. \quad 4.36$$

Finally, the camera response matrix is extended to the size $ms \times ml$ to take into consideration all the pixels. In order to explain the form of matrix \mathbf{C}_{ext} , let us introduce the matrix \mathbf{C}_λ^i , for a channel, i , of size $m \times m$:

$$\mathbf{C}_\lambda^i = \begin{bmatrix} c_{\lambda_1}^i & 0 & \dots & \cdot & \cdot & \cdot & \cdot & 0 \\ 0 & c_{\lambda_1}^i & \cdot & \dots & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & \cdot & \cdot & \cdot & \dots & c_{\lambda_1}^i \end{bmatrix}. \quad 4.37$$

Then, the extended matrix form of \mathbf{C} can be written as:

$$\mathbf{C}_{ext} = \begin{bmatrix} \mathbf{C}_{\lambda_1}^1 & \mathbf{C}_{\lambda_2}^1 & \cdots & \cdot & \cdot & \cdot & \cdot & \mathbf{C}_{\lambda_l}^1 \\ \mathbf{C}_{\lambda_1}^2 & \mathbf{C}_{\lambda_2}^2 & \cdots & \cdot & \cdot & \cdot & \cdot & \mathbf{C}_{\lambda_l}^2 \\ \cdot & \cdot \\ \mathbf{C}_{\lambda_1}^s & \mathbf{C}_{\lambda_2}^s & \cdots & \cdot & \cdot & \cdot & \cdot & \mathbf{C}_{\lambda_l}^s \end{bmatrix}. \quad 4.38$$

After the introduction of these new matrices, a generalized spectral form of the equation 4.29 can be written as:

$$\boldsymbol{\rho}_{ext} = \frac{1}{\pi} \mathbf{C}_{ext} (\mathbf{R}_{ext}^{-1} - \mathbf{K}_{ext})^{-1} \mathbf{E}_{0_{ext}}. \quad 4.39$$

4.4.3 Self-interreflection model

By focusing on the special case of a uniformly colored concave surface, a simplified model of interreflections can be obtained.

First, equation (4.29) can be written as:

$$\mathbf{L} = \frac{1}{\pi} \left(\frac{1}{r} \mathbf{I} - \mathbf{K} \right)^{-1} \mathbf{E}_0, \quad 4.40$$

where \mathbf{I} is the identity matrix of size $m \times m$ and r the reflectance of the surface at the considered wavelength, assumed to be non-zero.

Assuming a discretization into equal size facets, \mathbf{K} is then symmetric and the formal expression can be rewritten based on *Eigendecomposition* as:

$$\mathbf{L} = \frac{1}{\pi} \left(\frac{1}{r} \mathbf{I} - \mathbf{Q} \mathbf{G} \mathbf{Q}^{-1} \right)^{-1} \mathbf{E}_0, \quad 4.41$$

where \mathbf{Q} contains the eigenvectors of \mathbf{K} , and \mathbf{G} is a diagonal matrix containing the eigenvalues of \mathbf{K} .

By developing this equation, we can write:

$$\mathbf{L} = \frac{1}{\pi} \mathbf{Q} \left(\frac{1}{r} \mathbf{I} - \mathbf{G} \right)^{-1} \mathbf{Q}^{-1} \mathbf{E}_0 \quad 4.42$$

As explained in the previous section, the extended version of matrix \mathbf{K} , given by Equation (4.4.2), is organized into blocks on the diagonal. Then, the inverse of matrix $(\mathbf{R}_{ext}^{-1} - \mathbf{K}_{ext})$, denoted as \mathbf{M} , can be written as:

$$\mathbf{M} = \begin{bmatrix} \mathbf{Q} \left(\frac{1}{r_1} \mathbf{I} - \mathbf{G} \right)^{-1} \mathbf{Q}^{-1} & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \mathbf{Q} \left(\frac{1}{r_m} \mathbf{I} - \mathbf{G} \right)^{-1} \mathbf{Q}^{-1} & \cdot \end{bmatrix}, \quad 4.43$$

which can be further simplified as:

$$\mathbf{M} = \mathbf{Q}_{ext} \begin{bmatrix} (\frac{1}{r_1} \mathbf{I} - \mathbf{G})^{-1} & \dots & 0 \\ \cdot & \dots & \cdot \\ 0 & \dots & (\frac{1}{r_m} \mathbf{I} - \mathbf{G})^{-1} \end{bmatrix} \mathbf{Q}_{ext}^{-1}, \quad 4.44$$

where

$$\mathbf{Q}_{ext} = \begin{bmatrix} \mathbf{Q} & 0 & \dots & 0 \\ \cdot & \mathbf{Q} & \dots & \cdot \\ \cdot & \dots & \cdot & 0 \\ 0 & \cdot & \dots & \mathbf{Q} \end{bmatrix}, \quad 4.45$$

and,

$$\mathbf{Q}_{ext}^{-1} = \begin{bmatrix} \mathbf{Q}^{-1} & 0 & \dots & 0 \\ \cdot & \mathbf{Q}^{-1} & \dots & \cdot \\ \cdot & \dots & \cdot & 0 \\ 0 & \cdot & \dots & \mathbf{Q}^{-1} \end{bmatrix}. \quad 4.46$$

Equation (4.39) can be rewritten as:

$$\rho_{ext} = \frac{1}{\pi} \mathbf{A} \mathbf{B}, \quad 4.47$$

where

$$\mathbf{A} = \mathbf{C}_{ext} \mathbf{Q}_{ext} \text{diag}(\mathbf{Q}_{ext}^{-1} \mathbf{E}_0), \quad 4.48$$

and

$$\mathbf{B} = \begin{bmatrix} \text{vec}(\frac{1}{r_1} \mathbf{I} - \mathbf{G})^{-1} \\ \cdot \\ \cdot \\ \text{vec}(\frac{1}{r_l} \mathbf{I} - \mathbf{G})^{-1} \end{bmatrix}, \quad 4.49$$

where vec is the vectorization operator.

Note here that the calculation of the values in vector \mathbf{B} is straightforward: as both \mathbf{G} and \mathbf{I} are diagonal matrices, the inverse of the matrix $(\frac{1}{r_l} \mathbf{I} - \mathbf{G})$ can be obtained by a simple division.

Now that we have defined the spectral infinite bounce model of interreflections in its general and self-interreflection case, let us now discuss the scene geometry and the calculation of the geometrical kernel.

4.5 Scene geometry & geometrical kernel calculation

In this thesis, we generally consider V-shaped surfaces in our study of interreflections. A V-shaped surface, illustrated in Figure 4.1, consists of two planar Lambertian surfaces \mathcal{S}_1 and \mathcal{S}_2 having an edge contact with a known angle θ between them. Each surface has

its own 2D coordinates system: X_1Y_1 and X_2Y_2 for \mathcal{S}_1 and \mathcal{S}_2 respectively. By defining a global coordinates system XYZ as shown in Figure 4.1, and by performing the coordinates system changes, the geometrical kernel defined in Equation 3.3 can be easily obtained after discretizing each surface into a finite set of facets. However, there is an issue in the calculation of the geometrical kernel which is going to be explained and solved in the following.

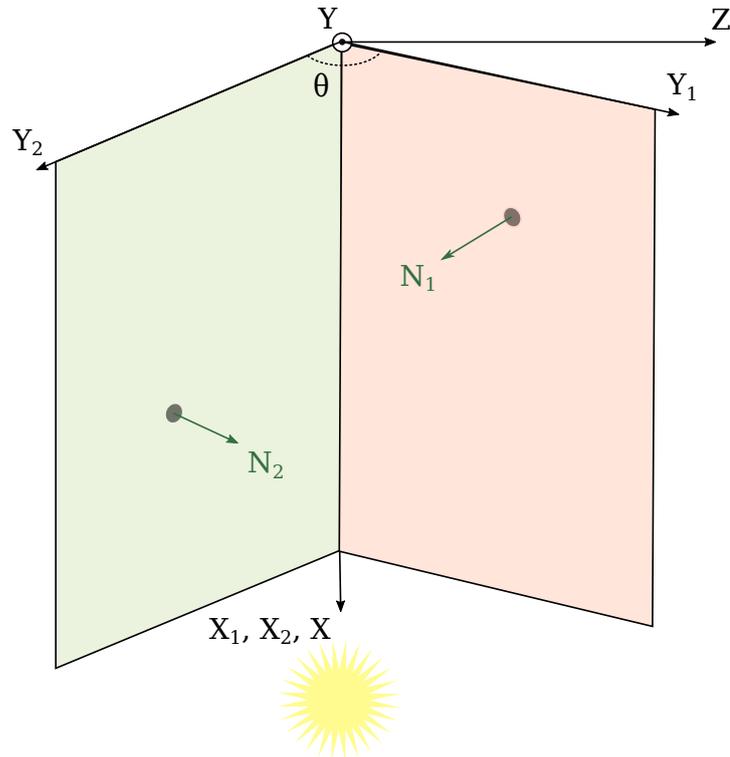


Figure 4.1: Case study: a V-shaped surface consisting of two adjacent planar surfaces with a known angle θ between them

4.5.1 The problem

The calculation of the discrete geometrical kernel presented previously (see Equation (3.3)) gives good approximation only in case that the facets i and j are infinitesimal and distant from each other. However, in vision systems, we are limited to the resolution of the sensors when choosing the size of facets.

As the distance between the facets becomes smaller, especially when the facets are adjacent (see Figure 4.2), the error of calculation becomes important. Thus, calculation based on this discrete geometrical kernel model becomes false. This problem was identified by Nayar *et al.* [1], but no solution has been proposed.

4.5.2 Monte Carlo Estimation

As introduced in Chapter 3, the definition of geometrical kernel, K , is driven from the definition of geometrical extent, G , which is defined between two surfaces \mathcal{S}_1 and \mathcal{S}_2 and

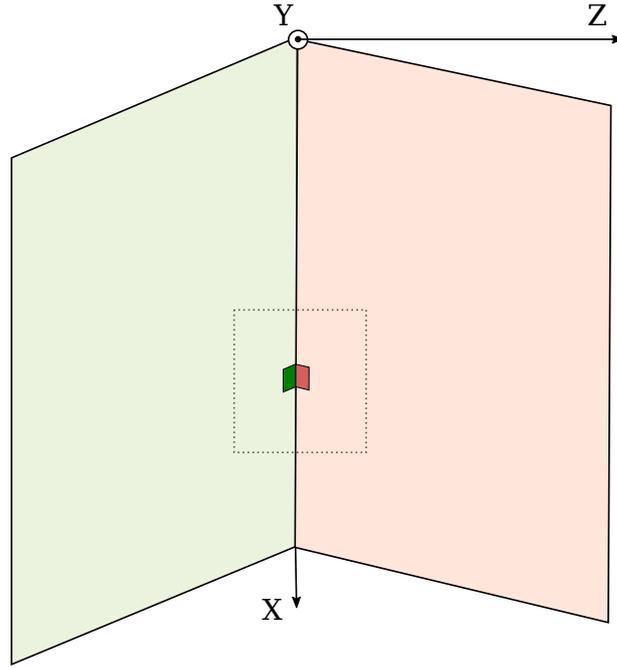


Figure 4.2: Geometrical Kernel Problem: in the case of these two facets (in dark green and dark red), the error of calculation is important.

can be written in terms of solid angle as:

$$d^2G = dS_1 \cos\theta d\omega, \quad 4.50$$

where $dS_1 \cos\theta$ is the apparent area of the infinitesimal surface dS_1 , and $d\omega$ is an infinitesimal solid angle.

The function we need to estimate is then:

$$G = \int_{P_1 \in \mathcal{S}_1} \int_{\Omega} \cos\theta d\omega dP_1, \quad 4.51$$

where Ω is the set of direction where \mathcal{S}_2 is visible to \mathcal{S}_1 , this can be found by projecting \mathcal{S}_2 on the unit hemisphere centered on P_1 .

In order to resolve this problem we propose to estimate this integration with Monte Carlo estimator. This is done by sampling the surface of each facet and the hemisphere around the sample point: one million samples were chosen randomly over the facet, and from each sample point a ray with a random direction inside the hemisphere was traced to see if it hits the other facet. Finally, the geometrical kernel was obtained by dividing the geometrical extent on the area of the facet.

Note that, the previous definition of geometrical extent based on the solid angle is chosen for our estimator instead of the definition based on the distance between the two facets. This choice is motivated by the fact that using the distance for adjacent facets is prone to error.

4.6 Direct irradiance modeling

Beside the geometry of the scene, interreflection phenomenon is strongly tied with the geometry of lightings in the scene. The amount of light firstly received by a surface point from direct light determines the energy carried with each following bounce of inter-reflected light. For this reason, knowing the amount of light received at each point of the concave surface is mandatory to model interreflections. These amounts are the irradiance values at each surface point and they constitute the vector $\mathbf{E}_{0\text{ext}}$ in Equation (4.39). Supposing a uniform light that covers the whole concave surface, let us introduce briefly some models of the lighting geometry. For more details concerning direct irradiance modeling and how it affects interreflections, we invite the readers to check the work described in [52].

4.6.1 Directional lighting

One of the simplest models of illuminant is the frontal collimated lighting, parallel to the Y axis (see Figure 4.1) covering uniformly all the area of the concave surface. Each ray emitted from the source is parallel to the direction vector, $d = [0 \ 1 \ 0]^T$. This means that each surface point receives the same exact amount of light independently to its location on the surface.

However, if this illuminant is oblique, parallel to another direction vector d_2 , each point will receive an amount of light depending on its surface normal and the direction of the light. Moreover, it is possible in this case that a part of a panel is not illuminated because of shadowing by the other panel. Shadows can be handled by setting to 0 the entries of the irradiance vector $\mathbf{E}_{0\text{ext}}$ in Equation (4.39) corresponding to the non-illuminated facets. This can be done for example by multiplying this vector with a diagonal matrix containing zeros for non-illuminated facets and ones for the other facets.

4.6.2 Diffuse lighting

Another type of lighting geometry is the perfectly diffuse light, or Lambertian illumination. Every point of a horizontal surface receives light from every direction in the hemisphere, leading to the same amount of irradiance at each point under this lighting. However, in the case of V-surfaces, the panels are not homogeneously illuminated because of shadowing. The points near the edges are more illuminated than the ones near the fold.

In this thesis, we focus mainly on the frontal collimated lighting. This is motivated by two reasons. First, interreflections' effects are more visible under this light geometry as shadows do not happen. This allows us to better analyze and understand this phenomenon in spectral estimation. Second, this scenario can be tested simply by putting the V-shaped surfaces under direct sunlight. Even though ambient light does exist in this case, but we assume that it can be neglected without an important loss of accuracy.

4.7 Interreflection model in practice: accuracy & discussions

We performed our experiments with two objectives in mind. The first, is to study the accuracy of our model in interreflection simulation in comparison to real camera outputs. The second objective is to study the effect of using infinite bounces in comparison to using only two bounces. As seen in the state of the art, using only two bounces is very common to simplify the interreflection model [32, 31, 26].

We divide the experiments in several sets to cover different cases of interreflections. The extensive study is performed for cases of self-interreflections and interreflections between two different surfaces of V-shaped configurations. In order to be able to perform these experiments, we assume that the geometrical kernel, the direct irradiance, the surface spectral reflectance, and the camera response functions are all known. Then, the model can be tested for its ability to accurately simulate the camera sensor values in case of interreflections. A Canon EOS 1000D camera was used to capture images, in RAW format, of the used patches under the same configuration as in simulation in order to compare simulated images to real ones. The spectral response curves of this camera were obtained using an integrating sphere with a monochromator in a dark room. For every wavelength between 400 nm and 700 with a step of 5 nm, a photo was taken and the radiance was measured with a Konica Minolta CS – 1000 Spectroradiometer. The spectral reflectances of the used samples were measured using the Minolta X-Rite Color i7 with a 10 nm step. Therefore, all spectral functions in these experiments are taken between 400 nm and 700 nm with a discretization step of 10 nm. The user chose the area corresponding to a surface in the image, then this area is discretized in the same way as in the simulation. The mean value over each discrete facet is calculated and compared with the corresponding simulated value. An example of the camera output we uses in our experiments is shown in Figure (4.3).

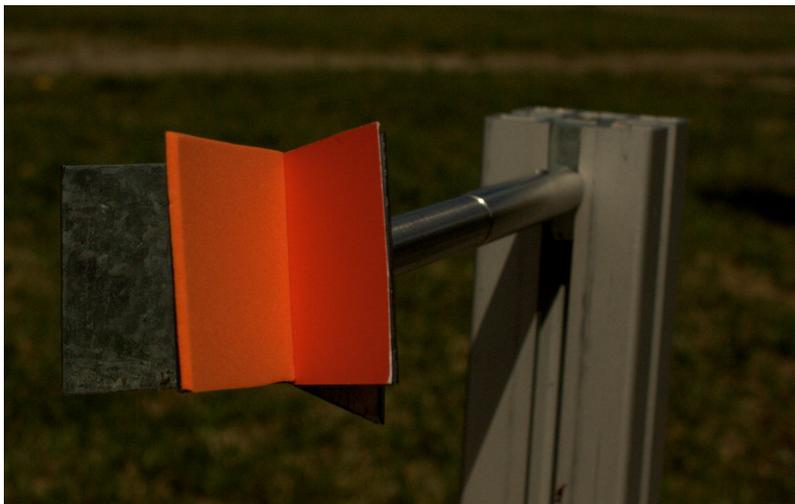


Figure 4.3: An example of a camera output we used in our experiments: an orange and a Munsell red patches with an angle of 45°

4.7.1 Self-interreflection

We chose to start with the case of self-interreflection, where interreflection happens between surfaces having the same spectral reflectance. In this case, the effect of interreflection can be easily observed and compared.

In this experience, we used six uniform colored papers, one of them was a Red Munsell paper (5R 6/12), and the other five were textile samples (see Figure (4.4)). For each sample, we have a set of two patches of size 4×4 cm. The spectral reflectances of these samples were measured using the Minolta X-Rite Color i7. Using a platform to fix two papers on planar surfaces and to set various angles between them, we were able to capture images of interreflections between two surfaces with angles of 45° and of 60° between them.



Figure 4.4: *The set of used samples: the first sample is a Munsell paper (5R 6/12 Mat). The other samples are textile ones.*

Two different configurations were used to perform the experiments. The first is in a dark room where we used a light source to illuminate directly the platform of interreflection. The spectral power distribution of this light was measured using the Minolta cs-1000 spectroradiometer (see Figure 4.5). The irradiance received from direct light at each point is needed to be known. For this purpose, we assume having a homogeneous light source assuring that every point receives approximately the same amount of light. The second configuration is under direct sunlight in the early afternoon. We did not use a measured SPD for this light but instead we used CIE Standard Illuminant D50 to simulate it. The main motivation behind this choice was to study our model under uncontrolled settings where the exact SPD of light is not known. A more detailed discussion about these two configurations and their effects on the results will be provided later.

Even though we did not control all the settings in our acquisition, we were able to evaluate the ability of our model to simulate interreflections by comparing it to RGB values from real camera output.

We used our model to simulate infinite-bounces images for two patches of the same color with a known angle between them (60° and 45°). In addition, to be able to see the effect of using an infinite number of bounces instead of only two, we used our model to simulate images with only two bounces of interreflections for the two patches under the

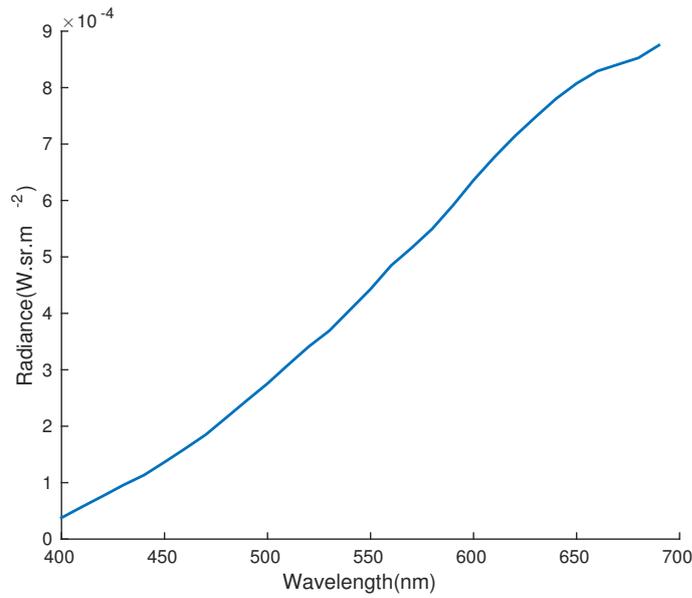


Figure 4.5: *Spectral Power Distribution of the lighting used in the indoor experiments.*

same configuration. This process was then repeated for all the patches in our sample set. The number of facets used in simulation is 8×8 , where each facet is of size 5×5 mm. The number of pixels corresponding to one facet varies between one photo and another as we did not use fixed acquisition settings. In some images a facet corresponded to approximately 2000 pixels. This discretization size was sufficient to demonstrate the performance of our model. However, a more detailed discussion about the facet size and its effects on the results will be provided later.

Results. Root mean square error (RMSE) values between the vector of all RGB values of camera outputs and that of the simulated image using our model with two bounces and infinite number of bounces are given in Tables 4.1 and 4.2 for angles of 45° and 60° respectively for the case of indoor direct light, and in Tables 4.3 and 4.4 for angles of 45° and 60° respectively for the case of direct sunlight.

In these tables, significant differences in RMSE values, which are around 2% or more, are highlighted with green for the smallest error and with red for the biggest one. Values of the error are generally bigger for the channels corresponding to the color of the patch. We can notice that all the significant enhancements of error values happen when an infinite number of bounces is taken into consideration. In addition, these enhancements would generally occur in the channels that correspond to the color of the patch, for example a red patch would show enhancements in RMSE values in the red channel. One can notice also that a higher number of significant error amelioration happens when the angle is smaller (see Tables 4.1 and 4.3 compared to Tables 4.2 and 4.4). This is due to the fact that a smaller angle leads to more inter-reflected light and thus to a less accurate two-bounce model.

In Figure 4.6, we show an example of obtained RGB values in the case of two pieces of Munsell sheets, of the same color, with an angle of 45° between them. The values obtained

Table 4.1: *RMSE values between the simulated image values using 2 and ∞ bounces and the camera output, for self-interreflection under **indoor** configuration with an angle of 45° . Significant differences (± 1.9) are highlighted in green when the error is reduced and in red otherwise*

RMSE (%)	2 bounces			∞ bounces		
	R	G	B	R	G	B
Red Munsell	8.6	3.4	0.3	5.7	2.7	0.4
Red	7.9	4.1	0.3	8.7	3.1	0.3
Orange	9.6	5.9	0.2	7.2	4.0	0.3
Blue	2.5	3.6	2.4	2.4	3.5	2.5
Cyan	1.2	7.2	3.7	1.5	4.0	2.1
Gray	4.2	7.3	1.7	4.2	5.6	1.5

Table 4.2: *RMSE values between the simulated image values using 2 and ∞ bounces and the camera output, for self-interreflection under **indoor** configuration with an angle of 60° . Significant differences (± 1.9) are highlighted in green when the error is reduced and in red otherwise.*

RMSE (%)	2 bounces			∞ bounces		
	R	G	B	R	G	B
Red Munsell	7.3	3.5	0.6	4.6	2.8	0.6
Red	6.7	3.5	0.6	5.4	3.0	0.7
Orange	7.2	5.3	0.3	8.0	4.3	0.3
Blue	3.2	4.0	3.2	3.1	3.9	3.4
Cyan	1.7	6.2	6.2	1.6	5.4	2.1
Gray	3.9	6.4	1.5	3.7	5.3	1.4

Table 4.3: *RMSE values between the simulated image values using 2 and ∞ bounces and the camera output, for self-interreflection under **outdoor** configuration with an angle of 45° . Significant differences (± 1.9) are highlighted in green when the error is reduced and in red otherwise*

RMSE (%)	2 bounces			∞ bounces		
	R	G	B	R	G	B
Red Munsell	10.6	6.5	1.3	4.7	4.5	0.9
Red	6.6	1.2	2.3	3.1	1.8	1.4
Orange	6.1	2.3	1.6	5.3	3.6	0.6
Blue	10.0	2.9	9.9	9.9	3.1	9.7
Cyan	3.9	5.4	5.1	3.1	1.9	4.4
Gray	3.6	2.6	3.0	3.8	3.2	3.5

from the camera output, and from simulation with infinite bounces and with two bounces are shown. It is observable that the simulated image with infinite number of bounces is more faithful to the real image, especially in the part closer to the joint edge.

Table 4.4: *RMSE values between the simulated image values using 2 and ∞ bounces and the camera output, for self-interreflection under **outdoor** configuration with an angle of 60° . Significant differences (± 1.9) are highlighted in green when the error is reduced and in red otherwise*

RMSE (%)	2 bounces			∞ bounces		
	R	G	B	R	G	B
Red Munsell	11.2	6.7	3.2	7.7	5.4	2.1
Red	8.2	5.9	1.6	5.4	4.5	1.0
Orange	7.4	4.1	1.8	3.4	2.3	0.8
Blue	27.4	7.0	34.1	27.4	6.9	33.9
Cyan	2.4	2.5	3.2	2.0	3.5	3.3
Gray	6.1	4.4	7.3	6.1	3.9	7.4

4.7.2 Interreflection between two surfaces with different colors

Now, and after studying the case of self-interreflections with different angles and using different illuminants, we study the case of interreflection between two surfaces with different spectral reflectance functions in order to show how well our model generalizes to a more complex scenario. The acquisitions were done under direct sunlight only. The SPD of this light is considered that of Standard Illuminant D50 in the simulation. The angle between the two surfaces is set to 45° only. We did two different experiments, the first was for interreflection between two colored surfaces. For this case, the patches, the sizes, and all the other settings were exactly the same as explained in the previous section. The second was for interreflection between one colored surface and another white one. Both surfaces were of the size of 9×6 cm. For simplicity's sake, each surface is discretized into 100 facets (instead of 64) of the size of 9×6 mm. For the colored surfaces we used the same set of patches as before. The white patch is the White Balance Chart of the ColorChecker Passport.

Results. Results in terms of RMSE values between camera output and simulation are shown in Table 4.5 for interreflection between two-colored surfaces and between one white and one colored surfaces.

As before, significant differences in RMSE values are highlighted in green when there is an amelioration and in red when there is a degradation. One can notice that most of the positive significant differences happen when an infinite number of bounces is taken into consideration instead of two bounces. However, some exceptions can be seen sometimes and specifically in the red channel.

4.7.3 Interreflection with different geometries and surfaces

In this section, we present some examples of interreflection simulations including scenes with different geometries and with more than two surfaces with different spectral reflectances. The first example consists of two planar surfaces having each two colors with an angle of 45° between them (see Figure (4.7)). The acquisition was done under direct

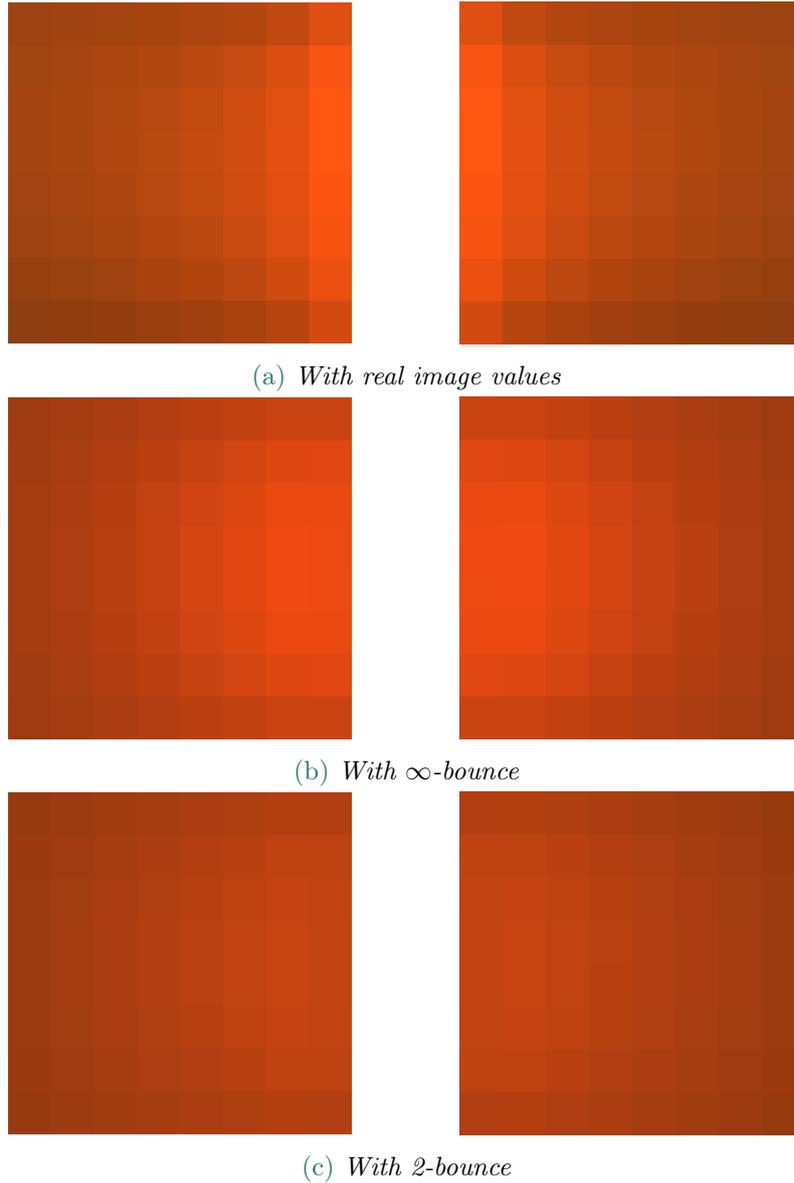


Figure 4.6: Real and constructed image values of two Munsell papers with an angle of 45° in indoor environment under direct light.

sunlight, and the SPD of the illuminant was considered the one of CIE D50 standard illuminant. Each surface was of the size of 8×4 cm and was discretized into 100 facets of size 8×4 mm. The shared edge between the surfaces was along the longest dimension.

The second example was about a geometry of the shape W where the central surfaces are shallower than the exterior ones (see Figure (4.8)). The exterior surfaces were both of a size of 10×8 cm and are discretized each to 200 facets of the size of 10×4 mm. The interior surfaces were of a size of 10×4 cm and were discretized each to 100 facets of the size of 10×4 mm. The shared edge between the surfaces was along the longest dimension. The acquisitions here also were done under direct sunlight simulated as CIE D50.

Results. RMSE values are shown in Tables 4.6 and 4.7 for simulation with infinite bounces and with two bounces for each uniformly colored surfaces of the first and second

Table 4.5: *RMSE values between the simulated image values using 2 and ∞ bounces and the camera output, for interreflection between two different surfaces under **outdoor** configuration with an angle of 45° . Significant differences (± 1.9) are highlighted in green when the error is reduced and in red otherwise*

RMSE (%)	2 bounces			∞ bounces		
	R	G	B	R	G	B
Red Munsell	15.9	2.2	1.7	4.1	1.2	1.9
Orange	5.7	7.6	3.1	9.5	5.6	2.8
Red Munsell	8.2	3.0	4.3	5.8	2.5	5.0
Cyan	1.1	2.5	3.6	1.1	3.7	3.1
Red Munsell	4.7	3.1	2.3	5.9	3.1	2.8
Blue	5.3	5.4	8.9	5.9	5.7	8.6
White	0.8	7.3	7.9	2.6	5.5	4.4
Cyan	2.2	10.1	10.4	2.3	5.9	6.0
White	1.7	4.8	4.4	5.8	3.6	3.5
Red	18.8	4.8	4.6	10.5	5.4	4.9

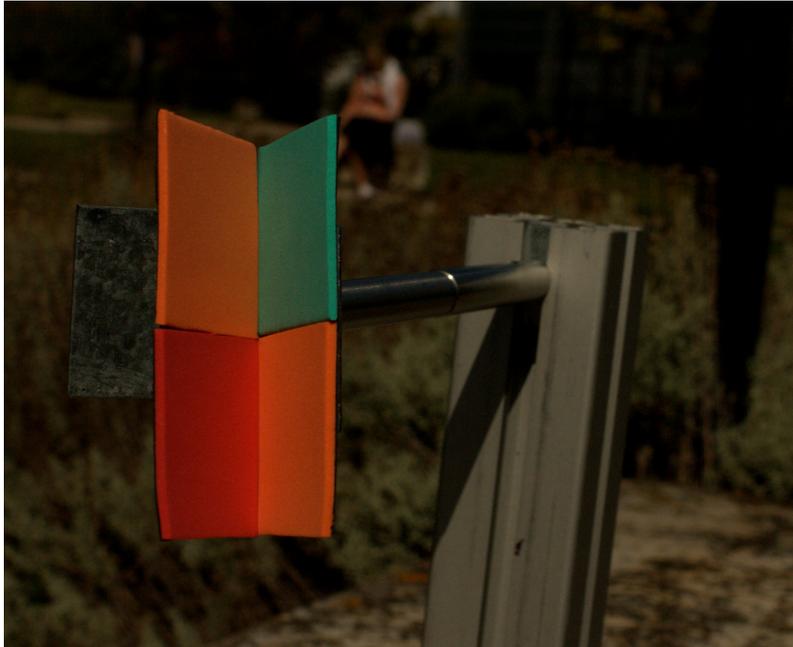


Figure 4.7: *The camera output we used in our experiments for surfaces with multiple spectral reflectances: orange-red and cyan-orange surfaces with an angle of 45°*

examples respectively. Similarly as before, significant differences in RMSE values between the infinite-bounce and two-bounce models are highlighted. It is noticeable here again the majority of significant enhancements happen when using the infinite bounce model.

4.7.4 Two-bounce vs. Infinite-bounce models

From the previous tables, we can observe that there is sometimes no significant enhancement when using an infinite-bounce model in comparison to a two-bounce model. However, the smaller the angle the stronger the enhancement. This is especially true in



Figure 4.8: The camera output we used in our experiments for surfaces with geometry showing occlusions: W shape with red, orange, cyan and white surfaces. The angles are approximately 45° .

Table 4.6: RMSE values between the simulated image values using 2 and ∞ bounces and the camera output, for interreflection between two different surfaces each of two colors under **outdoor** configuration with an angle of 45° . Significant differences (± 1.9) are highlighted in green when the error is reduced and in red otherwise

RMSE (%)	2 bounces			∞ bounces		
	R	G	B	R	G	B
Orange 1	4.4	12.0	1.1	8.7	9.7	1.4
Red	19.9	1.8	4.7	8.0	2.4	4.9
Cyan	2.3	4.6	2.5	1.5	3.4	1.9
Orange 2	8.3	8.9	0.6	7.4	6.7	0.6

case of outdoor configuration where the homogeneity of illuminant can be "guaranteed". In the case of Lambertian surfaces, which is almost the case for the Munsell sheets, RMSE values in the red channel are always significantly smaller when using our infinite-bounce model. The only exception to this observation is when interreflection happens between this Munsell patch and the blue one. In this case, energy carried by the inter-reflected rays is very low, because of the nature of the spectral reflectances of both patches (see Figure (4.9)). Even though, using infinite bounces model shows to be better than using only two

Table 4.7: RMSE values between the simulated image values using 2 and ∞ bounces and the camera output, for interreflection between surfaces of a **W** shape under **outdoor** configuration. Significant differences (± 1.9) are highlighted in green when the error is reduced and in red otherwise

RMSE (%)	2 bounces			∞ bounces		
	R	G	B	R	G	B
Red	18.5	7.1	5.2	10.7	5.1	5.3
Orange	11.7	2.9	5.3	6.0	4.4	2.7
Cyan	1.8	9.5	5.6	1.8	7.1	4.4
White	3.9	14.2	8.5	2.4	8.6	5.8

bounces in most of the cases, the enhancement can be still seen as non-significant in some cases. Thus, at first sight, the assumption that it is enough to take into consideration only two bounces of light, thus only one bounce of inter-reflected light, could be valid. However, if we look at the RMSE values taken for the closet facets to the joint edge between the patches, the difference between RMSE values when simulating with infinite bounces and with two bounces becomes important. In Table 4.8, the evolution of RMSE values with number of bounces in indoor configuration is shown. In this table, RMSE values are calculated using only the first column of facets; it is the closest to the shared edge, thus the most subject to interreflection. The effect of interreflections is very important in this area and decreases when getting farther from the joint edge. The RMSE values are shown for the channel where the error is the highest. For Munsell sheets, the error is tripled when using only two bounces. The error decreases when adding extra bounces.

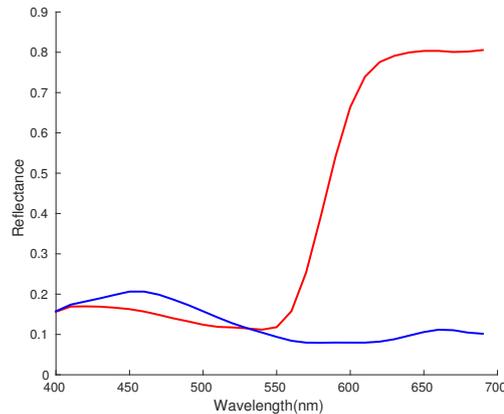


Figure 4.9: Spectral reflectance functions: Munsell patch in red, blue patch in blue.

Notice that this observation is true only in case of a surface with a high reflectance (the surface spectral reflectance function shows high values). For surfaces with low reflectance, the error difference is negligible. For example, for the blue patch, the error in all cases is small (see Table 4.8). The blue patch has a much lower reflectance than the Munsell patch, as shown in Figure 4.9.

Figures (4.10, 4.11) show the evolution of RMSE values with column number starting from the shared edge between two Munsell sheets having an angle of 45° and 60° ,

Table 4.8: *RMSE between the simulated image values of the column closest to the joint edge and those of the camera output with an angle of 45° for indoor configuration*

RMSE (%)	Color Channel	2 Bounces	3 Bounces	4 Bounces	∞ Bounces
Munsell	R	20.4	12.0	8.8	6.7
Red	R	22.2	14.9	12.1	10.3
Orange	R	21.4	11.7	7.6	4.5
Blue	B	2.0	1.7	1.7	1.7

respectively, between them.

We also compared the radiances after two bounces and infinity of bounces. Figure (4.12a) shows that radiance after infinite bounces of interreflection carries more energy compared to that after one bounce of interreflection for a facet close to the shared edge. When a facet which is far from this edge, the difference between the two radiances becomes small, (4.12b).

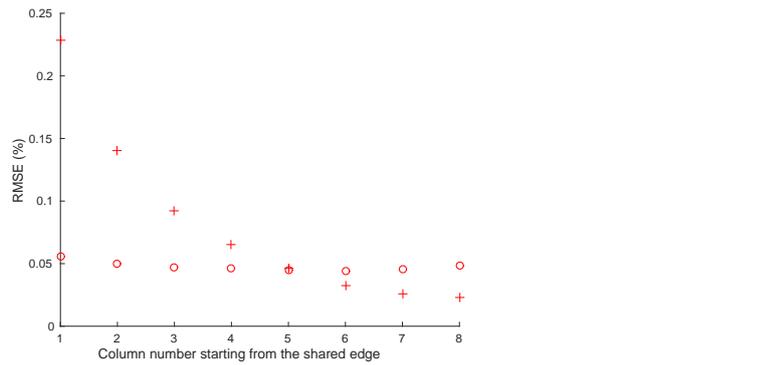


Figure 4.10: *RMSE values with column number starting from the shared edge between two Munsell red patches with an angle of 45° : circle for infinite bounce, cross for 2 bounces.*

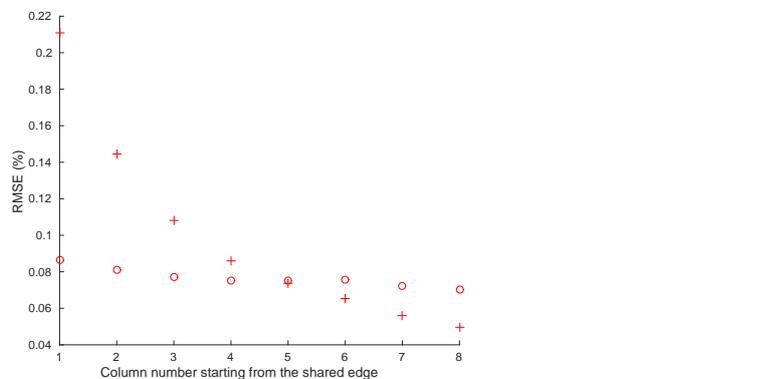
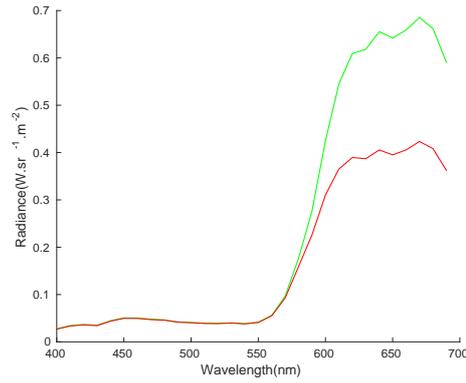
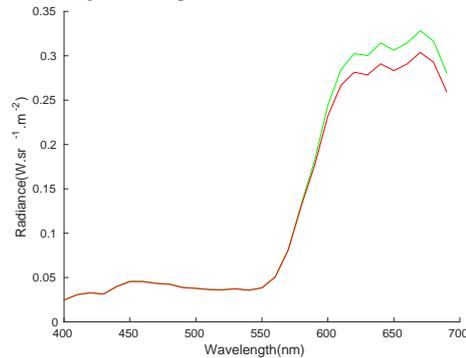


Figure 4.11: *RMSE values with column number starting from the shared edge between two Munsell red patches with an angle of 60° : circle for infinite bounce, cross for 2 bounces.*



(a) Center facet in the column closest to the joint edge



(b) Center facet in the column farthest to the joint edge

Figure 4.12: Radiance after two bounces, in red, and infinite number of bounces, in green of interreflection for Munsell red patch with an angle of 45° .

4.7.5 Sources of error

The model we proposed is driven from a physical background which guarantees its accuracy in the continuous domain. We have seen in Tables 4.1 to 4.5 that the RMSE values in simulation with infinite number of bounces compared to the real image is less than 6% in most of the cases. Although this error value could be considered as negligible in most of the computer vision applications, it is important to have an idea about the reasons behind this error. In the following, a discussion about some possible sources of this error is presented.

Non homogeneity of light. The illuminant used in our indoor experiments is an optical fiber behind a diffusing disk, which illuminates directly the platform. Even though we have considered the light to be perfectly diffuse and homogeneous, in reality it is not the case. Therefore, using a flat white surface, we performed a study on the homogeneity of this illuminant.

We measured the spectral power distribution of light on different places on a region of the white surface corresponding to the size of our samples. The coefficient of variation of the RMSE between different measurements is 1.4%. However, as the direct irradiance received on a surface element depends on its distance from the light source, variations in

irradiance per surface element are important especially in our experiments where surfaces are close to the light source. We calculated the coefficient of variation of RMSE values for two different distances between the white surface and the light source. The value we obtained is 14%. The two distances correspond to the closest and farthest edges of the used interreflection platform. In addition, we calculated the RMSE values between image intensities taken for the white surface and its mean value. The RMSE values are 0.025, 0.031 and 0.013 for R, G and B channels respectively. However, these values do not correspond solely to the non-homogeneity of lighting, they are related also to image noise introduced by the camera sensors. Thus, these values may vary if there is a change in camera settings.

The non-homogeneity of the indoor illuminant explain that the results in outdoor configuration are sometimes better. However, ambient light is not taken into consideration in outdoor simulation, which plays a role in increasing the error in this case. In addition, the used light SPD in outdoor configuration is a standard one and not a measured one, which might also contribute to the obtained error. However, we did measure the SPD of this light for one of our acquisition sessions but the use of this measured SPD did not provide a significant enhancement on the results.

Surface defects. The textile samples we have used are not fully Lambertian neither totally homogeneous. They show textures over their surfaces which introduces an error in the measurements. They also show some specularly at grazing incidence of light. Moreover, neither the size nor the angle are totally controlled, which contributes to the error we obtained.

Facets size. The physical model is accurate in continuous space. When passing to discrete space, the error is increased with the size of the facets. However, in our case, doubling the number of facets reduced the error of simulation by only 0.2%. This probably due to the noisy nature of the used camera outputs. The choice of the number of facets is an important trade-off between the processing time, the accuracy, and the sensors noise. For most vision application, reasonably big facets can be the best choice. Regarding the computation time, on a computer with Intel(R) Core(TM) i7-6700HQ CPU @ 2.60GHz processor, and with a brute-force implementation using **Matlab(R)** software by Mathworks, the model takes 3 seconds to be computed having two surfaces with 64 facets each, 11 seconds with 100 facets each, and 132 seconds with 256 facets each.

4.8 Conclusion

In this chapter, we have presented the earlier attempts to model interreflections in the domain of computer vision. These earlier works were destined to shape form shading applications where the biggest concern was to reverse the effects of interreflections in order to retrieve the accurate shape of the object. Later, we proposed a spectral infinite-bounce model of interreflections to be used in this thesis in spectral estimation. The problem of geometrical kernel calculation was explained and a solution to this problem was proposed.

The accuracy of this model is then studied and the importance of a model that takes into account an infinite number of bounces is demonstrated. However, although the accuracy is not perfect due to the possible sources of errors explained earlier, we will show in the following chapters that the use of this model helps in outperforming the state of the art approaches in spectral estimation.

This work has been published in the *Journal of Mathematical Imaging and Vision* [4]. It is the first survey that has been published in the domain regarding interreflection modeling. Our proposed model shares its basis with some other physics-based models of interreflections. However, it is the first model that relates the raw RGB values in an image with an infinite number of bounces of interreflections. The experiments presented in this chapter were chosen as they are the most pertinent in demonstrating the performance of our model and the differences between a two-bounce model and an infinite-bounce one.

One of the possible limitations of the proposed model is the rapid augmentation of the number of facets needed to simulate interreflections with a good accuracy especially when treating surfaces with multiple spectral reflectances and more complex geometries. This leads to extremely big matrices, and thus to a high memory usage and processing time. One possible solution to this problem can be to use sparse matrices. This solution is possible in our case due to the fact that the used matrices in the model contain a large number of zeros.

Our infinite-bounce interreflection model explained in the previous chapter gives the RGB values in the image based on the SPD of the light, the geometry and the spectral reflectance of the surface, and the spectral responses of the camera. In this chapter, we study the use of this model in an inverse approach. The model can be used in order to estimate either the surface spectral reflectance, the surface geometry, the light SPD, the camera spectral responses, or a any combination of those from the RGB values. However, we focus in this chapter on two applications: spectral reflectance estimation from a single RGB image, and camera characterization based on 3D color charts.

5.1 Interreflection in direct spectral reflectance estimation

The state of the art in spectral reflectance estimation of surfaces from RGB images can be divided into two categories: *direct methods* as called in [53, 54], also named *observational-model based methods* in [55], and *indirect methods* as called in [53, 54], also named *learning-based methods* in [55]. Direct methods group the set of approaches where no learning is needed but knowing the spectral power distribution of light and the sensor sensitivity curves is often necessary. On the other hand indirect approaches do use a learning step in order to be able to get the spectral reflectance with less known parameters. In this chapter, we draw our attention to direct methods. We first provide the state of the art concerning these approaches. We, then, show how interreflections give a boost to these approaches leading them to more concrete application.

5.1.1 State of the art

One of the earliest approaches in spectral reflectance recovery were based on using colored filters in front of camera lens [56, 57, 58, 53]. Later, more affordable and easier to use solutions were proposed. A common approach of direct methods is to combine trichromatic imaging with multiple light sources which helps in obtaining sufficient information in order to accurately estimate the spectral reflectance [59, 60, 61, 62, 63, 64, 65].

Park *et al.* [59] obtained spectral information by using a RGB camera and a cluster of light sources with different spectral power distributions. Their approach is based on the assumption that most real-world spectral reflectance functions can be well approximated by a small set of basis functions in a low-parameter linear model. These basis functions are obtained by eigendecomposition of a matrix containing 1257 Munsell chips' spectral reflectances as proposed in [2]. The spectral reflectance, $R(\lambda)$, can then be written as:

$$R(\lambda) = \sum_{j=1}^v \sigma_j R_j(\lambda), \quad 5.1$$

where σ_j are scalar coefficients which are to be estimated, $R_j(\lambda)$ are the known basis functions, and v is the number of parameters needed in the model.

Then, the raw linear camera response of the channel k at a pixel can be written as:

$$\rho^k = \frac{1}{\pi} \sum_{j=1}^v \sigma_j \int_{\Omega} C^k(\lambda) E(\lambda) R_j(\lambda) d\lambda, \quad 5.2$$

where $E(\lambda)$ is the irradiance of light and $C^k(\lambda)$ is the camera spectral response of the k -th sensor.

Suppose that n different illuminants are used in the acquisitions, then the set of knowns in the right side of this equation can be grouped in a matrix \mathbf{F} of size $pn \times v$ where p is the number of channels ($P = 3$ for color cameras). Then, the previous equation can be rewritten in a matrix form as:

$$\boldsymbol{\rho} = \mathbf{F}\boldsymbol{\sigma}. \quad 5.3$$

Knowing that all the spectral reflectances do not have negative values and that natural reflectances tend generally to be smooth, the problem of spectral estimation is then formulated as a constrained minimization:

$$\hat{\boldsymbol{\sigma}} = \underset{\boldsymbol{\sigma}}{\operatorname{argmin}} \left(|\mathbf{F}\boldsymbol{\sigma} - \boldsymbol{\rho}|^2 + \alpha \left| \frac{\delta^2 R(\lambda)}{\delta \lambda^2} \right|^2 \right), \quad 5.4$$

subject to $\mathbf{A}\boldsymbol{\sigma} \geq 0$,

where the matrix A groups the basis functions $R_j(\lambda)$, α and the term $\left| \frac{\delta^2 R(\lambda)}{\delta \lambda^2} \right|^2$ are respectively the smoothness parameter and the smoothness constraint.

The authors used 8 basis functions, leading to 8 coefficients to be estimated, and three multiplexed illuminants of a compound light source composed of LEDS of white, red, green, blue and amber colors.

This approach received little attention mainly because of the need of several illuminants, which is incompatible with common situations where the surface is illuminated by ambient light.

Later, Jiang et al. [64] proposed a more convenient solution, still based on various illuminants. They recover the spectral reflectance by estimating 6 coefficients from two images taken under two different but commonly available lighting conditions. These lighting conditions can be for example daylight at different times of a day, fluorescent or tungsten lights. In order to relax the constraint that the SPD of light and the spectral response curves of the camera should be known, the authors proposed to estimate the matrix \mathbf{F} in Equation 5.3 by a calibration step using six samples with known reflectances. An image of these samples is captured under each lighting condition. Then, the \mathbf{F}^{-1} can be estimated by minimizing either the reflectance difference or the color difference. This calibration step is also important in order to handle variations in color temperature and spectral properties of the used illuminants.

More recently, Khan et al. [65] proposed the use of a portable screen-based lighting setup in order to estimate the spectral reflectance of the considered surface. The portable screen was used to give three lightings with red, green and blue colors. The capturing device

is an RGB camera; therefore, based on the three different spectral responses, the spectral reflectance of the surface is expressed by nine coordinates in a basis of nine spectra. Instead of using the same basis functions obtained by eigendecomposition of spectral reflectances of Munsell color chips as in earlier approaches, the authors propose to discard the part of the spectra that corresponds to the sensor-illuminant null-space before performing the eigendecomposition. This leads to more suitable basis functions for the used configuration.

It can be observed from these approaches that the use of several light sources is mandatory to increase the number of equations needed to get the spectral values; together with the use of basis functions in order to decrease the dimensionality of the problem. The number of the used lights is correlated to the number of used parameters to model the spectral reflectance. The higher the number is, the better the accuracy of the estimated spectral function.

Instead of using multiple lightings, there exists very few approaches that tried to use some properties of interreflections in order to find the spectral reflectance of the surface while at the same time estimating the spectral power distribution of an ambient light.

Funt, Drew and Ho [32, 33, 34] proposed to use interreflections as additional information to find illumination spectral power distribution and surface spectral reflectance from RGB data. Considering that some parts of the surface are not affected by interreflections, they combine equations with and without interreflections to solve them for the unknown parameters. Their work is based on modeling the spectral reflectance and the spectral power distribution of lights by finite dimensional linear models using a set of basis functions. Then, they solve a set of equations with equal number of unknown parameters. They demonstrate that it is possible to use three basis functions for surfaces and illumination with no need for more than three sensor classes.

Considering diffuse irradiance $E(\lambda)$, and two lambertian surfaces whose spectral reflectances are $R^{(1)}(\lambda)$ and $R^{(2)}(\lambda)$ respectively, radiances, $L^{(1)}(\lambda)$ and $L^{(2)}(\lambda)$, reflected from surface S_1 and S_2 , without taking into consideration interreflections, can be written as:

$$\begin{aligned} L^{(1)}(\lambda) &= \frac{1}{\pi} E(\lambda) R^{(1)}(\lambda), \\ L^{(2)}(\lambda) &= \frac{1}{\pi} E(\lambda) R^{(2)}(\lambda). \end{aligned} \tag{5.5}$$

By modeling the illumination using u basis functions denoted as E_i , the spectral power distribution can be written as:

$$E(\lambda) = \sum_{i=1}^u \epsilon_i E_i(\lambda). \tag{5.6}$$

Similarly, the two surfaces' spectral reflectances are modeled with v basis functions

denoted as R_j :

$$\begin{aligned} R^{(1)}(\lambda) &= \sum_{j=1}^v \sigma_j^{(1)} R_j(\lambda), \\ R^{(2)}(\lambda) &= \sum_{j=1}^v \sigma_j^{(2)} R_j(\lambda). \end{aligned} \quad \text{5.7}$$

Using these equations, we can rewrite (5.5) as:

$$\begin{aligned} L^{(1)}(\lambda) &= \frac{1}{\pi} \sum_{i=1}^u \sum_{j=1}^v \epsilon_i \sigma_j^{(1)} E_i(\lambda) R_j(\lambda), \\ L^{(2)}(\lambda) &= \frac{1}{\pi} \sum_{i=1}^u \sum_{j=1}^v \epsilon_i \sigma_j^{(2)} E_i(\lambda) R_j(\lambda). \end{aligned} \quad \text{5.8}$$

Notice that, in these equations, the unknown parameters are ϵ_i , $\sigma_j^{(1)}$ and $\sigma_j^{(2)}$, so we have $(2v + u)$ unknown parameters.

In order to have more equations, the authors propose to take into consideration regions affected by interreflections. In computer graphics, diffuse interreflections are modeled using configuration factors which give the fraction of light from a surface that reaches the other surfaces. Let us consider α_{12} the fraction of light signal reflected from a surface S_1 to a surface S_2 , and α_{21} the fraction of light signal reflected from surface S_2 to surface S_1 . Let L_m^j be radiance reflected in every direction by one of the two surfaces ($j = 1$ or 2) in the interreflection area, thus it can be written as:

$$\begin{aligned} L_m^{(1)}(\lambda) &= L^{(1)}(\lambda) + \alpha_{21} L_m^{(2)}(\lambda) R^{(1)}(\lambda), \\ L_m^{(2)}(\lambda) &= L^{(2)}(\lambda) + \alpha_{12} L_m^{(1)}(\lambda) R^{(2)}(\lambda). \end{aligned} \quad \text{5.9}$$

By developing these equations, we obtain:

$$\begin{aligned} L_m^{(1)}(\lambda) &= \frac{L^{(1)}(\lambda) + \alpha_{21} L^{(2)}(\lambda) R^{(1)}(\lambda)}{1 - \alpha_{12} \alpha_{21} R^{(1)}(\lambda) R^{(2)}(\lambda)}, \\ L_m^{(2)}(\lambda) &= \frac{L^{(2)}(\lambda) + \alpha_{12} L^{(1)}(\lambda) R^{(2)}(\lambda)}{1 - \alpha_{12} \alpha_{21} R^{(1)}(\lambda) R^{(2)}(\lambda)}. \end{aligned} \quad \text{5.10}$$

In order to make these equations easier to solve numerically, the authors assume that the dominator is equal to one. They suggest that this can be done without significant loss in accuracy in most cases. That's because both α and $R(\lambda)$ are inferior to one. Thus, the equations can be rewritten as:

$$\begin{aligned} L_m^{(1)}(\lambda) &= L^{(1)}(\lambda) + \alpha_{21}L^{(2)}(\lambda)R^{(1)}(\lambda), \\ L_m^{(2)}(\lambda) &= L^{(2)}(\lambda) + \alpha_{12}L^{(1)}(\lambda)R^{(2)}(\lambda). \end{aligned} \quad 5.11$$

However, for highly reflective materials (high values of $R(\lambda)$), and for geometric configurations with high values of α , this approximative model is far from being accurate.

Moreover, with this latter assumption, only one bounce of interreflection is considered, further bounces being ignored; therefore, the previous equations represent a two-bounce model.

From the previous equations, knowing that:

$$L^{(1)}(\lambda)R^{(2)}(\lambda) = L^{(2)}(\lambda)R^{(1)}(\lambda), \quad 5.12$$

the ratio between α_{12} and α_{21} can be written as:

$$\frac{\alpha_{21}}{\alpha_{12}} = \frac{L_m^{(1)}(\lambda) - L^{(1)}(\lambda)}{L_m^{(2)}(\lambda) - L^{(2)}(\lambda)}. \quad 5.13$$

This ratio can be known as all the radiances are considered to be measurable, thus, only one unknown parameter is added with two additional equations (5.11).

In order to solve the set of equations (5.8) and (5.11), let us first consider the sensor sensitivity function $C^k(\lambda)$ for $k = 1 \dots p$ where p is the number of sensors. Then, by integrating according to the wavelength over the visible spectrum of light, we can define the following pre-calculated tensor:

$$g_{ijk} = \frac{1}{\pi} \int_{\Omega} E_i(\lambda)R_j(\lambda)C^k(\lambda)d\lambda. \quad 5.14$$

By rewriting equations (5.8) and (5.11) in terms of camera response, we can write the k -th sensor response, ρ_k , as follows:

$$\rho^{k(1)} = \sum_{i=1}^u \sum_{j=1}^v \epsilon_i \sigma_j^{(1)} g_{ijk}, \quad 5.15$$

$$\rho^{k(2)} = \sum_{i=1}^u \sum_{j=1}^v \epsilon_i \sigma_j^{(2)} g_{ijk}, \quad 5.16$$

for the response of the camera to the radiance reflected after one light bounce on surfaces S_1 and S_2 , and:

$$\rho_m^{k(2)} = \rho^{k(2)} + \alpha_{12} \sum_{i=1}^u \sum_{j=1}^v \sum_{j'=1}^v \epsilon_i \sigma_j^{(1)} \sigma_{j'}^{(2)} h_{ijj'k}, \quad 5.17$$

for the camera response in the mutual reflection area taking into consideration only one

bounce of interreflection, where:

$$h_{ijj'k} = \frac{1}{\pi} \int_{\Omega} E_i(\lambda) R_j(\lambda) R_{j'}(\lambda) C^k(\lambda) d\lambda. \quad 5.18$$

Now, the set of equations (5.15), (5.16) and (5.17) for the unknown parameters: ϵ_i , $\sigma_j^{(1)}$, $\sigma_j^{(2)}$ and α needs to be solved. Representing light spectral power distribution and spectral reflectances using three basis functions for each, will lead to 10 unknowns. However, having three sensors, and three equations, we have only 9 equations. Then, in order to be able to solve this problem and use only three sensors, the authors impose that $\epsilon_1 = 1$, thus neglecting the brightness of the surfaces.

This approach gives an approximation of the spectral reflectance of two surfaces and the spectral power distribution of light based only on one RGB image to an approximation of 3 basis functions for each spectrum. Only ambient light is considered (isotropic diffuse illumination). This assumption is made in order to consider the irradiance at each point of both surfaces to be the same. The method considers two convex Lambertian surfaces with uniform spectral reflectances each. The existence of at least one point on each surface with no interreflection is needed, which can be in some cases a strong assumption. The authors suppose that taking into consideration only two light bounces is sufficient, which is not the case for all scenes. In addition, modeling each spectrum with only three basis functions is not enough to obtain accurate estimations. The approach also cannot handle highly reflective surfaces due to some simplifications applied on the mathematical model.

Later, this work has been extended to a variational approach in order to relax some constraints, especially those related to diffuse illuminant and to the number of surfaces. However, the complexity of the algorithm has considerably increased and the model still handles only one bounce of interreflection [66].

In the following section, we present our interreflection-based spectral reflectance estimation. Infinite number of bounces is taking into consideration, and no preconditions on the surface to have a point without interreflections or a low reflection is considered. In addition, we show that taking into account infinite number of bounces helps in avoiding the use of basis function to model the spectra.

5.1.2 Direct spectral estimation with the infinite-bounce interreflection model

Now, we present our approach for direct spectral reflectance estimation based on the infinite-bounce interreflection model explained in Chapter 4.

In Equation (4.39) we have related RGB values of pixels in an image of a concave surface (vector $\boldsymbol{\rho}_{ext}$), to the spectral reflectance (matrix \mathbf{R}_{ext}) and the geometry (matrix \mathbf{K}_{ext}) of the surface, the illuminant SPD (vector \mathbf{E}_{0ext}), and the sensor response functions (matrix \mathbf{C}_{ext}). Here, we propose to use this interreflection model in an inverse approach in order to estimate the spectral reflectance of a surface, i.e. matrix \mathbf{R}_{ext} , assuming that all other matrices and vectors in Equation (4.39) are known.

The problem can be formulated as a constrained minimization one, and the objective

function can be written as:

$$\widehat{\mathbf{R}}_{ext} = \underset{\mathbf{R}_{ext}}{\operatorname{argmin}} (|\boldsymbol{\rho}_{img} - (\frac{1}{\pi} \mathbf{C}_{ext} (\mathbf{R}_{ext}^{-1} - \mathbf{K}_{ext})^{-1} \mathbf{E}_{0ext})|^2 + \alpha |\frac{\delta^2 r(\lambda)}{\delta \lambda^2}|^2), \text{ subject to } \forall r, 0 < r \leq 1, \quad 5.19$$

where $\delta^2 r(\lambda)/\delta \lambda^2$ is a smoothness constraint controlled by the smoothness parameter α . This constraint helps to remove numerical noise and to obtain a spectrum that has a smoother shape. This choice is widely adopted in optimization, and is motivated by the observation that natural reflectances tend to be smooth [59].

As seen before, existing approaches in the state of the art need to use multiple light sources, and to reduce the dimensionality of the problem by using basic functions in order to estimate spectral reflectance from RGB values. In our case, the use of a single image of the surface under one illuminant is made possible thanks to interreflections. The effects of these interreflections are more or less pronounced in the different areas of the surface. Thus, the camera output of a single folded surface contains many different colors which provides a number of equations sufficient to solve for the unknowns even without the use of basis functions.

Thus, one of the main differences between our approach and the state of the art ones is that interreflections create a collection of different color values that are related in a non linear way to the spectral reflectance of the surface and can be used to solve an equation of a high dimension. Our model needs to analyze the whole area of the surface, whereas in approaches relying on flat surfaces, analyzing the color of one pixel (or a mean value taken over a small area) suffices.

Since this problem is constrained and non-linear, it cannot be defined mathematically as an inverse problem [67]. The number of unknowns to estimate is related to the number of different spectral reflectances contained in the surface, if the latter is not uniform, and to the number of wavebands considered in the visible spectrum. Solving this problem in its general form is not an easy task for multiple reasons. Firstly, estimating the number of unknowns related to this equation is not straightforward given the non-linear nature of the model. Secondly, if different facets in the surface have different spectral reflectances, then there is no unique solution for Equation 5.19. For these reasons, we propose to focus on a special case where the surface has a homogeneous spectral reflectance over its whole area.

In this case, since the surface is uniform, its spectral reflectance can be fully characterized by a single vector \mathbf{R}_λ of size l , where l is the number of wavelengths. Then, based on the self-interreflection model presented in Equation (4.47), the minimization problem formulated in Equation (5.19) becomes:

$$\hat{\mathbf{R}}_\lambda = \underset{\mathbf{R}_\lambda}{\operatorname{argmin}} \left(|\boldsymbol{\rho}_{img} - \frac{1}{\pi} \mathbf{A} \mathbf{B}|^2 + \alpha \left| \frac{\delta^2 r(\lambda)}{\delta \lambda^2} \right| \right),$$

subject to $\forall r, 0 < r \leq 1$.

5.20

We remind the reader that the matrix \mathbf{A} is defined in Equation (4.48) and is related to the eigenvectors of the matrix \mathbf{K} as well as the spectral responses of sensors and the SPD of light. The vector \mathbf{B} on the other hand is defined in Equation (4.49) and is related to the eigenvalues of the matrix \mathbf{K} and to the spectral reflectance.

5.1.3 Experiments & results

In order to verify the validity of the proposed model and its capacity to provide accurate spectral reflectances of Lambertian surfaces, we carried out various tests, first based on simulated images in order to estimate the optimal performance of the method, then based on real images in order to see the influence of image noise and other geometrical imperfections on the accuracy of the estimated reflectance.

Case study

For our tests on both synthetic and real camera data, we considered the following case study: a scene consisting of a Lambertian surface of uniform spectral reflectance, folded into a V-shaped surface of two similar planar square panels \mathcal{S}_1 and \mathcal{S}_2 with an angle of 45° between them. Each panel is discretized into 100 equally-sized facets. The geometrical kernel is then calculated using Monte Carlo estimation. Although, the results presented here are for a specific configuration, the effect of the angle and the discretization size will be studied later in Section 5.1.4. The V-cavities are illuminated frontally by collimated lighting parallel to the bisecting plane of the two panels, thus every facet is assumed to receive the same amount of lighting.

Implementation

The form of the minimization problem described by Equation (5.20) is very handy: only one matrix inverse (\mathbf{Q}^{-1}) is needed and can be calculated before starting the minimization process. Our problem is still constrained and non-linear, but the number of variables to be estimated depends only on the number of wavebands considered. We can assume that this problem is convex, and that a local minimum optimization algorithm is able to give a good solution. A more detailed study of the nature of this problem will be explained in Section 5.1.4.

The minimization method we chose is the interior point algorithm [68, 69]. This method is implemented under the function *fmincon* in **Matlab(R)** software by Mathworks. The minimization is done iteratively.

Evaluation

The performance of the spectral reflectance estimation is evaluated using the root mean square error (RMSE) calculated between the estimated spectral reflectance $\hat{\mathbf{R}}$ and the ground truth one \mathbf{R} ¹:

$$RMSE(\mathbf{R}, \hat{\mathbf{R}}) = \sqrt{|\mathbf{R} - \hat{\mathbf{R}}|^2/q}, \quad 5.21$$

where q is the number of wavebands.

The performance is also evaluated in terms of color distance by using CIEDE00 distance [9], computed between the ground truth spectral reflectance, \mathbf{R} , and the estimated one, $\hat{\mathbf{R}}$, for a 10° standard observer and a CIE D65 illuminant.

We also added the Pearson distance (PD) to the used metrics:

$$PD(\mathbf{R}, \hat{\mathbf{R}}) = 1 - \frac{\mathbf{R}^T \hat{\mathbf{R}}}{|\mathbf{R}| |\hat{\mathbf{R}}|}. \quad 5.22$$

This distance can also be found in the literature under the name Complementary Goodness of Fit Coefficient (CGFC) [70, 71, 72], and is independent of the magnitude, thus it helps in giving better evaluation based on the shape of the estimated spectral reflectance especially in case of non-calibrated settings.

Synthetic data

We started by testing our method on synthetic data. For this sake, scenes with interreflections were simulated supposing a collimated light source with SPD corresponding to CIE D65 illuminant, a patch with a spectral reflectance of one of the Macth ColorChecker patches, and a given geometry of V-shape. A SD-10 camera is considered whose response curves are similar to the ones used in [65]. Under these configurations, the RGB values for each of the patches of Macbeth ColorCheckers were obtained using our model [see Equation (4.47)]. Afterwards, these RGB values are used to estimate back the surface spectral reflectance.

Our results are compared with the results of two other comparable approaches [59] and [65] relying on three different illuminants. We implemented these approaches and used the first 8 components of the Parkkinen basis as suggested in [59]. For our approach, we argue that there is no more need for reducing the dimensionality of the problem, thus we estimate the spectral reflectance from 400 nm to 700 nm in steps of 5 nm, giving a total of 61 values to be estimated. The size of the planar surfaces is set to 2×2 cm and the smoothness parameter, α , is chosen to be 2.5 which guarantees a good balance between smoothness and accuracy.

The resulted RMSE, CIEDE00 and PD values shown in Table 5.1 indicate that spectral reflectance is as accurately predicted by using the interreflection approach with one light as by using approaches based on several lightings, and it can be even more accurate. It is

¹The ground truth spectral reflectance is the one used in the simulation in case of experiments on synthetic data. For real data, ground truth spectral reflectances are measured using a spectrometer.

worth mentioning that using three different lights means performing the acquisition three times and using three lamps.

Table 5.1: *Reflectance estimation accuracy for synthetic data*

Method	No. Illuminants	ColorChecker24			ColorChecker240		
		RMSE	CIEDE00	PD	RMSE	CIEDE00	PD
Park [59]	3	0.024	0.31	0.006	0.018	0.37	0.004
Khan [65]	3	0.018	0.31	0.003	0.011	0.30	0.002
Our Method	1	0.017	0.29	0.003	0.017	0.26	0.003

Real data

To perform the experiments on real data, we used a platform enabling us to fix two pieces of the same sample in a planar configuration and to set approximately an angle of 45° between them, as explained earlier. We used the six uniformly colored samples presented earlier in Figure 4.4: one of them is a Red Munsell sample (5R 6/12), and the other five are textile samples. For each sample, we have a set of two patches of size 4×4 cm. The ground truth spectral reflectances of these samples were measured using the Minolta X-Rite Color i7. For the acquisition, we used a Canon EOS 1000D camera with known spectral response functions to capture images in RAW format. The photos were taken under a sunny daylight in the early afternoon. Having in mind to propose a practical setup, we did not measure the spectral power distribution of this illuminant, instead we used the spectral power distribution of CIE D50 standard illuminant in the equations to represent the direct sunlight [73]. The intensity of the illuminant was not measured as well and the camera is used with various settings and focal lengths without any extra calibration process.

Under the same settings, we also captured an image of all the available samples but this time without interreflection (flat samples).

Before acquisition, a X-Rite Color Checker was added to each scene. It is going to be used in some experiments to pre-calibrate the images. The set of experiments in which we used the Color Checker are named "with pre-calibration". An example of a camera output used in the experiments is shown in Figure (5.1).

RGB values of the image are extracted from an area selected by the user and corresponding to a whole area of one panel. A discretization size of 4×4 mm is used to divide the selected area into 100 facets, each one attached to an RGB value corresponding to the mean RGB value over its area. Based on this discretization size, the geometrical kernel can be built using Monte Carlo estimation method according to the panels size and the angle between them.

We also tested the methods described in [59] and [65] on images of similar flat samples taken under a single illuminant, CIE D50. As a single light source is used, the spectral reflectance is modeled with 3 basis functions. The RGB value given to these approaches is the mean value over an area selected by the user and corresponding to the surface of the patch.



Figure 5.1: *Example of acquisition: The Munsell V-cavity on the top and the X-Rite Color Checker on the bottom under direct sun light.*

The used camera is not calibrated to the specific acquisition settings, and the intensity of light source is not measured. For these reasons, we separate the experiments into two categories:

- With pre-calibration: we used the 24 patches of the Color Checker chart to learn second degree polynomial curves relating the expected red, green and blue values with the simulated red, green and blue ones respectively. Then, the learned transformation is applied on the image values used in the estimation.
- Without pre-calibration: in this case, no pre-calibration is used to estimate the quantity of light entering the camera lens. Instead, we propose to normalize image values and simulated values on their respective sums in order to assume that they have similar amounts of energy. In order to apply these normalization steps, we reimplemented the approaches [59] and [65] using the same optimization algorithm as ours.

Tables 5.2 and 5.3 show the results of our approach compared to [59] and [65]. Ta-

ble 5.2 shows the mean values over all the patches of RMSE, DE00 and PD in case of pre-calibration. Table 5.3 shows the mean values over all the patches of RMSE, DE00 and PD when no pre-calibration is used. The results show that our approach performs better than the approaches [59] and [65] in both non-calibrated and pre-calibrated cases. Moreover, our results without pre-calibration seems comparable to the results of the two other approaches after pre-calibration. This shows that one can take a camera whose spectral response functions were measured previously in a lab, and use it in a complete non-calibrated settings, and use a standard SPD of light instead of measuring it and still get good spectral estimations with our model.

Table 5.2: *Reflectance estimation accuracy on real images taken under daylight **after pre-calibration***

Method	RMSE	CIEDE00	PD
Park [59]	0.060	3.76	0.009
Khan [65]	0.059	3.87	0.009
Our Method	0.046	3.82	0.008

Table 5.3: *Reflectance estimation accuracy on real images taken under daylight **without pre-calibration***

Method	RMSE	CIEDE00	PD
Park [59]	0.143	8.02	0.012
Khan [65]	0.138	7.99	0.012
Our Method	0.061	5.62	0.012

Table 5.4 shows the results of our approach compared to [59] and [65] for each of the used samples. On the one hand, the results of our approach are obtained without any pre-calibration. On the other hand, when applying the approaches [59] and [65], the pre-calibration step using the Color Checker chart was applied. By observing Table 5.4, one can see that in most cases our approach performed better even though no pre-calibration is performed in our case. An exception can be seen for two particular patches, the gray and the blue ones, where the surface is weakly reflective. Actually, interreflections help in spectral reflectance estimation only in case of surfaces with high reflectance in some wavebands. This is due to the fact that interreflections are very weak when the surface reflectance is low as explained in Chapter 4.

The estimated spectral reflectances compared to the ground truth ones are shown in Figure (5.2). In (5.2a), the graph shows that our approach without pre-calibration outperformed the approaches [59] and [65] after pre-calibration for the case of Munsell patch. However, in the case of low reflectance surfaces, for example the blue one, our approach failed to accurately estimate its spectral reflectance, see Figure (5.2b).

But why interreflections help in case of non-calibrated settings? To answer this question, let us consider two different cases. In the first case, consider that the power of light used in the acquisition is multiplied by a factor f_1 , then by a factor f_2 . After normaliza-

Table 5.4: Details of reflectance estimation accuracy on real images taken under daylight

Patch	Metric	Park [59]	Khan [65]	Our Method
		w/ pre-calibration	w/ Pre-calibration	w/o pre-calibration
Munsell Red	RMSE	0.078	0.078	0.055
	CIEDE00	5.03	5.38	4.74
	PD	0.006	0.007	0.006
Orange	RMSE	0.076	0.069	0.061
	CIEDE00	3.53	3.71	3.71
	PD	0.004	0.004	0.006
Cyan	RMSE	0.088	0.076	0.065
	CIEDE00	3.44	3.40	5.89
	PD	0.022	0.016	0.006
Red	RMSE	0.088	0.10	0.069
	CIEDE00	6.31	6.78	5.80
	PD	0.019	0.026	0.012
Blue	RMSE	0.009	0.008	0.084
	CIEDE00	2.88	2.77	8.74
	PD	0.002	0.002	0.039
Gray	RMSE	0.022	0.021	0.035
	CIEDE00	1.35	1.18	4.82
	PD	0.001	0.001	0.004

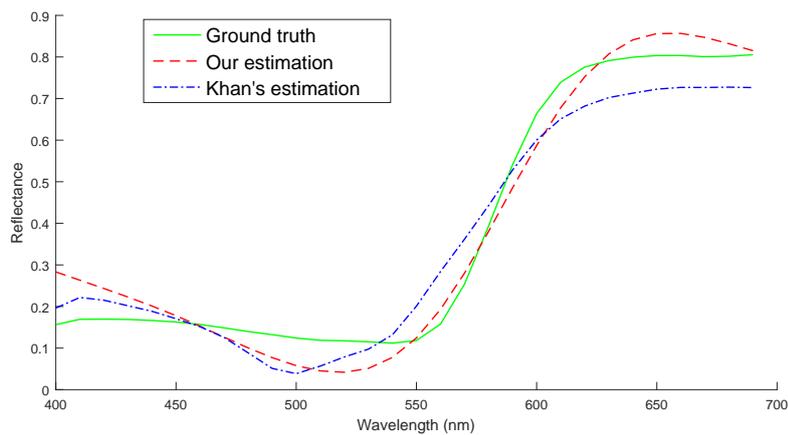
tion, RGB values with and without interreflections will be the same in both cases as the factors f_1 and f_2 are canceled by the normalization step. In the second case, consider instead two surfaces where the spectral reflectance of the first surface is f times the spectral reflectance of the other one. Here, in case of flat surfaces, RGB values for both surfaces will be the same as the factor f is canceled by the normalization step. However, in case of interreflections, and due to the fact that reflectance is raised to different powers with each bounce, the factor f cannot be canceled by the normalization step and RGB values will differ between one surface and the other. Thus, interreflections help in distinguishing changes in spectral reflectance from changes in light power, and the variations of RGB values over the surface help in retrieving the correct spectral reflectance.

5.1.4 Discussion

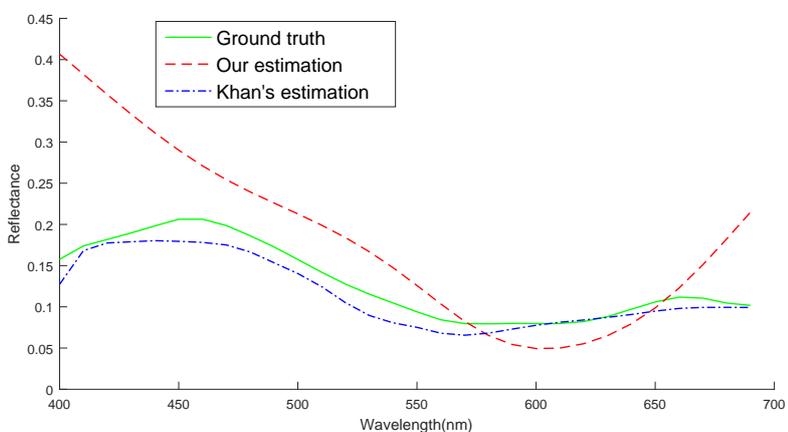
In this section, we propose a detailed study of the influence on the estimation accuracy of the angle between the two panels and the chosen facet size. In addition, an analysis of the mathematical nature of the problem that we solve and the uniqueness of the solution is given.

Study of the effect of angle and facet size

We performed an experimental study on the effect of the angle between the two panels, on the accuracy of the spectral reflectance estimation using our method. Angles of 90° , 60° , 45° and 30° are used. Another parameter that may have an important effect on the results is the spatial sampling of the surface (size and number of facets). Discretization



(a) A comparison of estimation for Munsell red patch



(b) A comparison of estimation for blue patch

Figure 5.2: Comparison between the estimated spectral reflectances and the ones measured with a spectrometer of some surfaces: in green the ground truth, in red the estimation based on our method without pre-calibration, in blue the estimation based on Khan et al. approach [65] after pre-calibration.

into 64, 100 and 256 facets on each planar surface is used. Results in term of RMSE and CIEDE00 are shown in Table 5.5.

Table 5.5: Comparison of reflectance estimation accuracy for our method when using different angles and facet sizes

Angle	30°			45°			60°			90°		
No. facets	64	100	256	64	100	256	64	100	256	64	100	256
RMSE	0.016	0.0158	0.0160	0.0189	0.0172	0.0170	0.0189	0.0181	0.0182	0.0250	0.0246	0.0244
CIEDE00	0.281	0.287	0.265	0.295	0.292	0.305	0.321	0.308	0.316	0.423	0.391	0.387

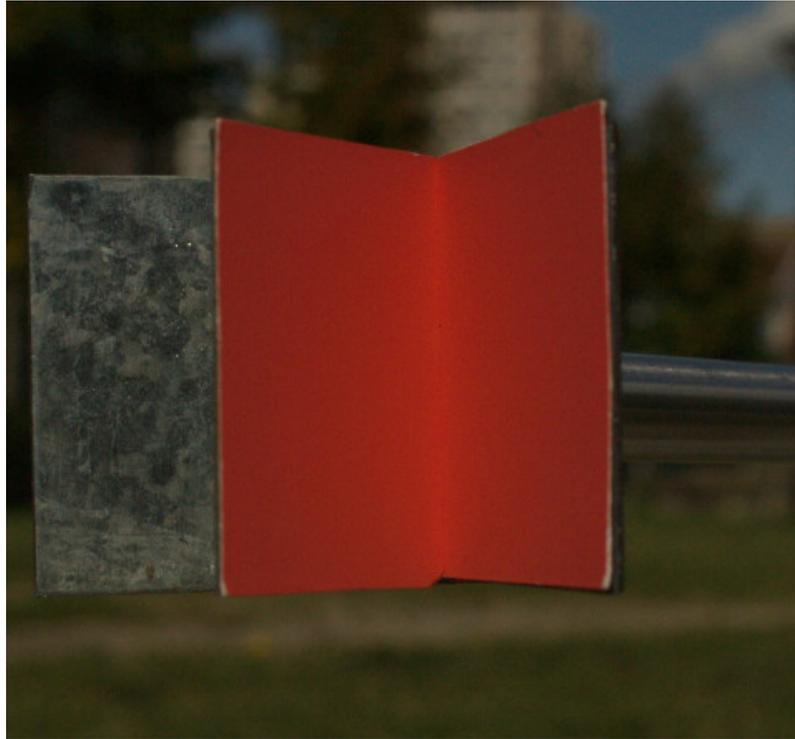
By observing changes of error values with angle in Table 5.5, it is clear that the error gets lower as the angle decreases. A smaller angle emphasizes the interreflection effects, especially near the joint border of the panels where light has more chance to bounce multiple times (see Figure 5.3). In order to take advantage of this, a small discretization step is preferable to better observe the changes in RGB values per facet.

Nonetheless, when observing the changes of errors with facet sizes in Table 5.5, it is noticeable that using different facet sizes did not significantly affect the errors in general. An interpretation of this observation may be related to the dimensionality of the problem: enough information can be obtained when the two surfaces get closer to each other without the need of a very small facet size. In other words, using more facets may not necessarily help in adding significantly different RGB values to help the optimization process. This depends on the choice of the angle and on the number of unknowns to be estimated. This question becomes more relevant in real situations, as a coarse discretization may be needed to avoid noise in the image and to reduce the computation time. If more accuracy is needed, an adaptive facet size can be adopted. However, the matrix \mathbf{K} would be no more symmetric and the form proposed for the self-interreflection case in Equation 4.47 would not be valid anymore.

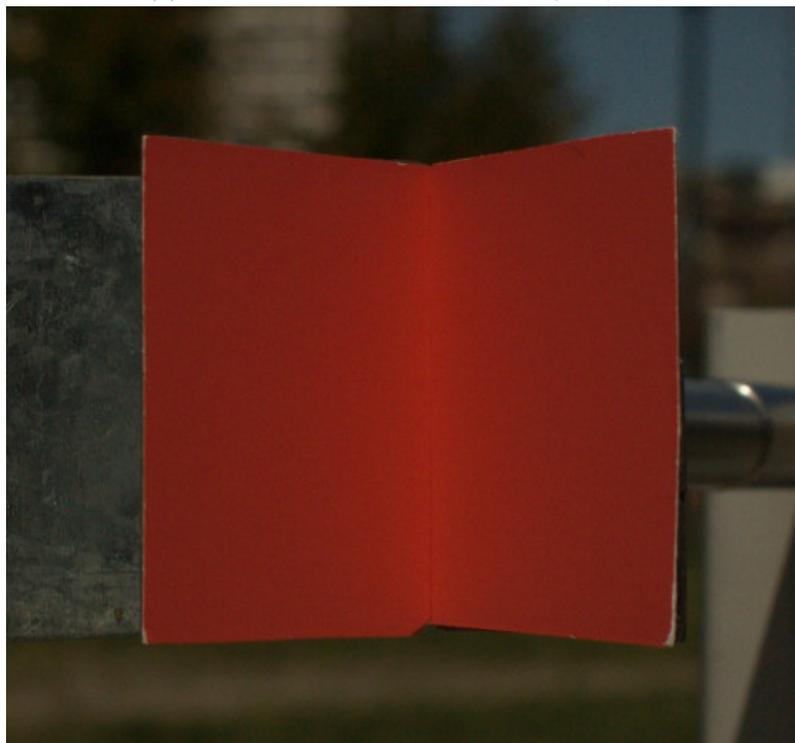
Practically speaking, angles between 45° and 60° are good choices. They are small enough for significant interreflections to happen, and big enough to let the surface receive sufficient amount of light. The number of pixels corresponding to the panel should be at least the number of the used facets, but it is recommended to have more in order to reduce the sensibility to image noise. The size of panels is optional: interreflections depend only on the relative size of the panels.

Study of the uniqueness of solution

The experiments reported in this study showed, empirically, the capacity of our inter-reflection based method to provide rather good estimate of the spectral reflectance from one RGB image. However, since this method is neither a direct spectral reflectance measurement nor a direct combination of spectral measurements yielding a spectral reflectance, but the result of a mathematical optimization from RGB values, we must question the capacity of this operation to provide the expected spectral information, regardless of the experimental uncertainties. The question is not on the objective function that is optimized [see Equation 5.20], which relies on a well-established physical ground through the inter-reflection equation. The question is rather on the convergence of the optimization process, even though we could verify in all our tests that the iterative optimization algorithm is not sensible to the starting values. In absence of mathematical proof of convergence, which looks to be a hard mathematical problem, we cannot exclude the possibility that the optimization provides a spectrum totally different from the expected spectral reflectance. This is why we prefer the term "estimation" rather than "prediction" or "measurement". However, physical arguments help us to claim that the results of the optimization process is most often strongly correlated with the surface's spectral reflectance, which means that the range of spectral reflectances that the method would systematically fail to retrieve with



(a) *Two Munsell sheets with an angle of 45°*



(b) *Two Munsell sheets with an angle of 60°*

Figure 5.3: *Effect of angle change on image RGB values*

reasonable tolerance is rather limited, and excludes most real materials.

In order to understand why the method has good chance to converge towards the expected result, let us remind how interreflections intrinsically bring spectral information

into the color image, especially in the case of V-shaped surfaces where color gradients are well visible (actually, under directional illumination, color gradients are always displayed except in the very special case of integrating spheres where the geometrical extent between every point and every other point is constant, which results in a homogeneous color over the whole sphere and therefore an absence of color gradient). The color gradient comes from the fact that the radiances perceived from the different points of the surface contain different populations of photons having bounced different numbers of times in average in the concavity before reaching the sensor: near the external edges of the panels, photons have more chance to escape the concavity after one bounce (these photons form a flux proportional to the incident spectral irradiance and the reflectance of the surface), whereas near the joint-edge, they have more chance to strike the surface again after each bounce, therefore to bounce many times (the corresponding flux is proportional to the incident spectral irradiance and a combination of the reflectance, its square, its cube, and so on. . .). Hence, the radiance $L(\lambda)$ viewed from a given point P can be approximated as:

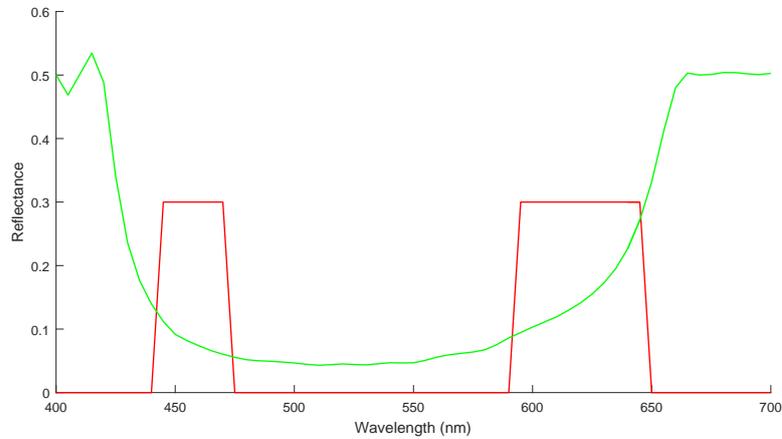
$$L_P(\lambda) = \sum_{j=1}^n q_j r^j(\lambda) \quad \text{5.23}$$

where $r(\lambda)$ denotes the surface reflectance, q_j are factors related to the surface geometry, and n is the highest significant number of bounces undergone by the photons which transit from point P to the sensor. The number n increases from the external edges where it is typically around 1 – 2, to the joint edge between the panels where it can reach extremely high values. Since the spectral radiance is a non-linear function of n , the collection of spectral radiances viewed over the surface form a spectral space with appreciable dimensionality (which explains why there is no need with this method to use various illuminants or decrease the dimensionality of the spectral space of the considered samples).

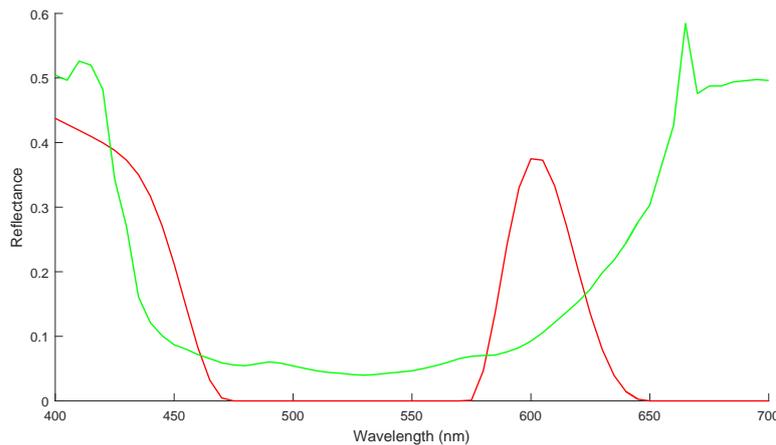
Once captured by the sensors, the collection of spectral radiances is converted into a collection of RGB values, all correlated with the spectral radiance of the surface. If we modify the spectral reflectance of the surface, it is possible that the RGB values given by the sensors in one point are not modified (case of metamerism), but it is almost impossible that the R, G, and B values remain unmodified for all the points, even in very special cases of spectral reflectances as shown in Figure (5.4a) through a multi-rectangular function. The two experimental spectral reflectance functions in Figure (5.4a) give exactly the same camera response with the SD-10 camera under CIE D65 illuminant. Using our model, we obtained the RGB values corresponding to each of these reflectances for a surface bent with an angle of 60° . Then, we used our objective function to find back the spectral reflectances. This is done without using the smoothness term in order to not affect the solution. Figure (5.4b) shows the estimated spectral reflectance for both these spectral reflectances.

Therefore, the collection of RGB values displayed by an image of surface with known geometry, illuminant and sensor, seems to be a signature of the spectral reflectance of the surface: it is almost impossible to find a surface with different spectral reflectance that would display similar set of RGB colors in the same experimental configuration. This means that two metameric flat surfaces cannot be metameric anymore once bent in concave

shape. To the best of our knowledge, this hypothesis was never investigated or stated until now. It might open doors for new research areas in metamerism field.



(a) *Metamerism case: experimental spectral reflectances giving the same RGB values when the surfaces are flat.*



(b) *Interreflections-based spectral reflectance estimation for both spectra: in Red the estimation for the step-shaped spectrum, in green the estimation for the smooth one α*

Figure 5.4: *Experimental study on the uniqueness of the solution.*

5.2 Interreflections in camera characterization

As a second application for our interreflection model in spectral estimation, we propose to study the use of interreflections in camera characterization by finding the spectral response curve of each of its sensors. We show that using 3D patches (see Figure (5.5)) instead of 2D patches such as the color checker helps in enhancing the quality of the estimated spectra thanks to interreflections.

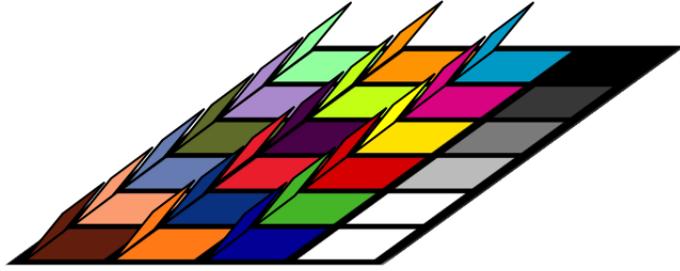


Figure 5.5: A 3D color chart with interreflections.

5.2.1 The problem

In order to estimate the matrix \mathbf{C} from RGB values, more than one pixel with different RGB values are needed. All the n_P pixel color coordinates can be stored in a $3 \times n_P$ matrix $\boldsymbol{\rho}$ expressed as follow:

$$\boldsymbol{\rho} = \frac{1}{\pi} \mathbf{CER}, \quad 5.24$$

where \mathbf{R} is a $n_\lambda \times n_P$ matrix containing on each column i the spectral reflectance of the observed surface S_i .

Similarly to spectral reflectance estimation, starting from this equation and considering that we know $\boldsymbol{\rho}$, \mathbf{E} and \mathbf{R} , we can get an estimate $\hat{\mathbf{C}}$ of the sensor sensitivities by using the Moore-Penrose pseudo-inverse:

$$\hat{\mathbf{C}} = \pi \boldsymbol{\rho} (\mathbf{ER})^T ((\mathbf{ER})(\mathbf{ER})^T)^{-1}. \quad 5.25$$

This least-square solution is the one that minimizes $\|\boldsymbol{\rho} - \frac{1}{\pi} \mathbf{CER}\|^2$. Regularization terms can be added to this equation in order to prevent high norm sensor sensitivities or to get smooth spectral curves [74].

In this case, the unknown spectra can be modeled as a weighted sum of basis functions in order to reduce the number of unknowns [75].

The number of color patches is crucial to get accurate results because the inputs are 3-D RGB and the outputs are N-D spectral curves. Nevertheless, as mentioned earlier, the natural reflectances lie in a low dimensional linear space (8-D), thus adding new patches introduces no or very little improvement on the reconstruction quality [75]. Our idea rather consists in folding the available natural patches in order to create interreflections to add some different inputs.

5.2.2 Interreflection-based approach

From the interreflection model explained earlier, we propose to introduce the matrix $\mathbf{R}_\lambda^{\text{eq}} = (\mathbf{R}_\lambda^{-1} - \mathbf{K})^{-1}$, called hereafter the equivalent reflectance matrix. Thus, Equation (4.29) becomes:

$$\mathbf{L}_\lambda = \frac{1}{\pi} \mathbf{R}_\lambda^{\text{eq}} \mathbf{E}_\lambda. \quad 5.26$$

This equation looks similar to the radiance equation for Lambertian surfaces when no interreflections happen:

$$\mathbf{L}_\lambda = \frac{1}{\pi} \mathbf{R}_\lambda \mathbf{E}_\lambda. \quad 5.27$$

If the kernel matrix is a null matrix $\mathbf{0}$, i.e. no interreflection might happen then $\mathbf{R}_\lambda^{\text{eq}} = (\mathbf{R}_\lambda^{-1} - \mathbf{0})^{-1} = \mathbf{R}_\lambda$. Hence, interreflections change the reflectance matrix \mathbf{R}_λ into \mathbf{R}^{eq} which changes accordingly the output radiance. This means that, if we could have created a flat surface (without interreflections) constituted by facets characterized by the equivalent spectral reflectance provided by the matrix $\mathbf{R}_\lambda^{\text{eq}}$, we would have obtained the same output radiance vector as if we had observed the considered homogeneous concave surface with interreflections. This is why we propose to call this matrix, $\mathbf{R}_\lambda^{\text{eq}}$, the *equivalent spectral reflectance* matrix.

Let us now study the properties of such equivalent reflectances in order to measure their impact on the considered application, i.e. sensor sensitivity or even light SPD estimation.

Non-linearity of interreflections

It is known that the natural reflectances lie in a low dimension linear space [75]. This is a problem when we want to construct a color chart for camera calibration and spectral reconstruction. It has been demonstrated in [75] that adding color patches to the 24 patches of the Macbeth color checker does not provide a significant improvement since these 24 patches already cover a large part of this low-dimensional spectral space. However, in case of interreflection, the equivalent spectral reflectances are the results of natural reflectance multiplications, and they are not completely linearly related to the natural reflectances. In the following experiment, this observation is studied in practice.

We propose to extract n , $n = 1, \dots, 9$, linear basis functions of the 24-patches Macbeth color checker, and to check how well the natural reflectances can be reconstructed from these basis functions, i.e. we project the natural reflectances on these basis and project them back to the reflectance space. The obtained reflectances are compared with the original reflectances thank to the spectral recovery error as recommended by [75]. Figure 5.6 displays the evolution of these errors according to the dimension of the linear space n , for 3 different reflectance sets. The first set is constituted by the 24 reflectances of the Macbeth color checker, the ones used for the basis extraction. The second set is constituted by the 1269 Munsell reflectances, and the last one by the 24 equivalent reflectances provided by the interreflections of the Macbeth color patches. For the plot, we have measured the average reconstruction error over each set. We can see that the reconstruction error of the equivalent reflectances is always higher than the 2 other errors, whatever the dimension used. This means that the equivalent spectra can not be reconstructed easily from the linear basis of the Macbeth color chart, so they should bring complementary information in the spectral reconstruction process.

The idea is illustrated in figure 5.7, where we have displayed one homogeneous color patch discretized with 8×14 facets. When the patch is flat and the illumination is diffuse and far enough from the patch, all the facets have the same color. If the same patch is folded

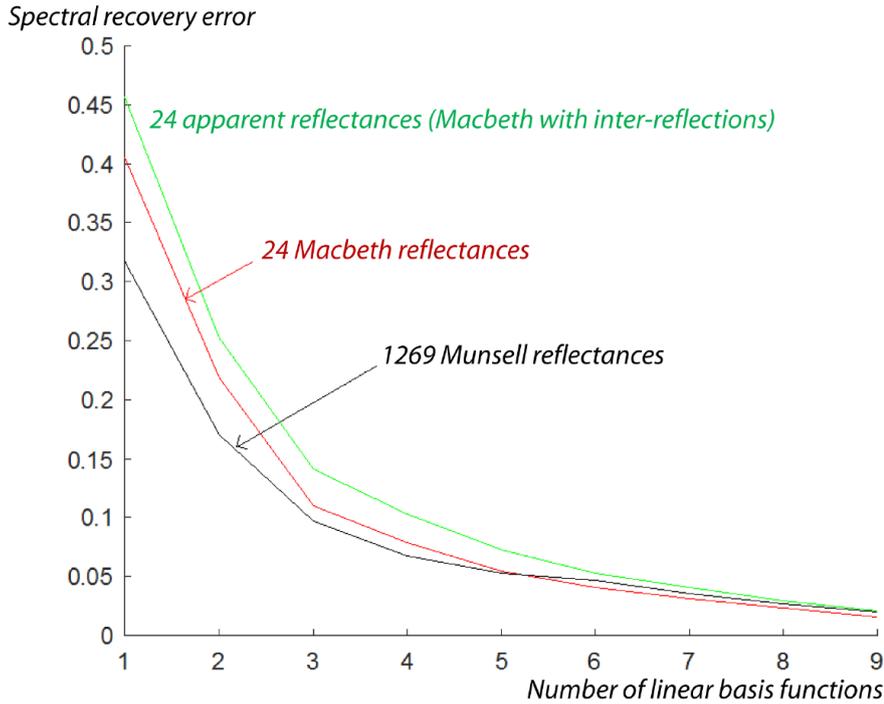


Figure 5.6: *Spectral reflectance reconstruction errors versus dimension.*

as shown in the top-center illustration of this figure, the colors of each facet are changing due to interreflections as explained earlier. The equivalent spectral reflectances of these facets so that they are characterized by the same color but with the patch being not folded, as shown in the top-right illustration of this figure. The equivalent reflectances of these "virtual" facets are displayed on the bottom-right plot and can be compared with their real reflectance on the bottom-left plot. Because they are obtained by multiplication of the same initial reflectances, they are much more narrow-band than the natural reflectances and are able to extract complementary information from the one extracted with the natural reflectances.

5.2.3 Results

In this section, we propose to assess the quality of the spectral sensitivity reconstruction from raw RGB data, with and without accounting interreflections. We first run experiments on synthetic data, as usually done in the literature [75], then we reconstruct spectral curves from real color images.

Synthetic data

In this section, we propose to use a physical model in order to construct a synthetic image and then to evaluate spectral sensitivities from these RGB data. In this aim, we consider one known SPD illuminant \mathbf{E} and one known spectral reflectance \mathbf{R} from which we can simulate a spectral radiance vector $\mathbf{L}^{\text{no-inter}}$ obtained for a flat surface without interreflections thanks to Equation 5.27. This radiance hits the 3 sensors $\mathbf{C}^{\mathbf{R}}$, $\mathbf{C}^{\mathbf{G}}$ and $\mathbf{C}^{\mathbf{B}}$ of the camera that allows us to get 3 color components $R = \sum_{\lambda} \mathbf{C}_{\lambda}^{\mathbf{R}} \mathbf{L}_{\lambda}^{\text{no-inter}}$, $G =$

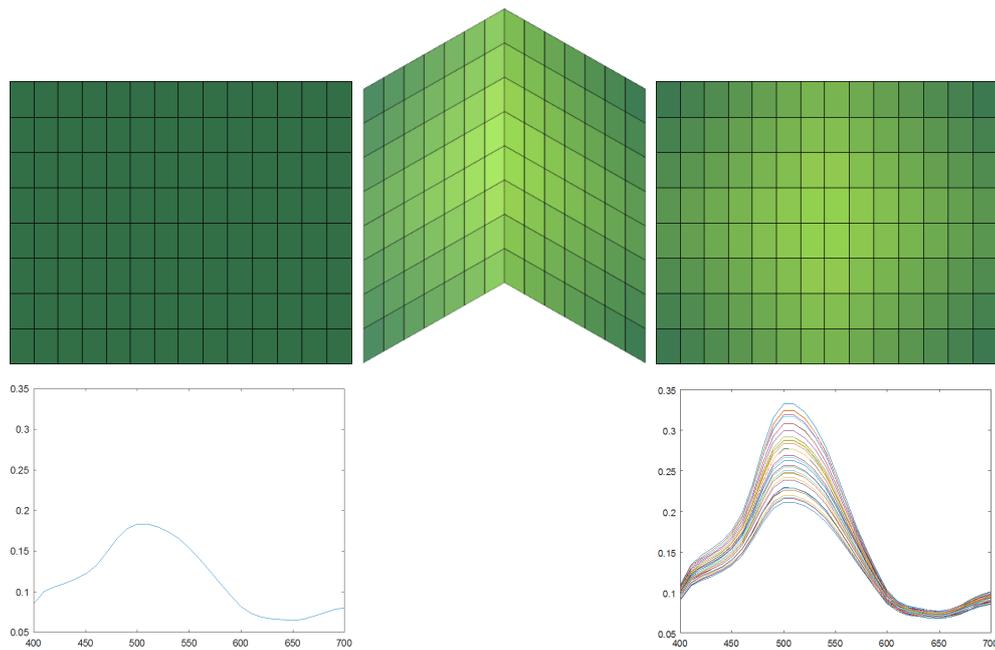


Figure 5.7: Equivalent spectra. Top-left: one green patch extracted from the Macbeth color checker and discretized into 8×14 facets. Bottom-left: the spectral reflectance of this patch. Top-center: The appearance of this patch when folded. We can see how colors are changing for some facets. They appear brighter and more saturated. These colors have been evaluated from the interreflection model of eq.(4.29). Top-right: One 'virtual' flat paper whose facets have the same appearance colors as the ones in the folded green paper. The reflectances displayed on bottom-right are the equivalent reflectances of these virtual facets. We can note that they are more narrow-band than the natural green reflectance because they are the results of reflectance multiplications.

$\sum_{\lambda} \mathbf{C}_{\lambda}^{\mathbf{G}} \mathbf{L}_{\lambda}^{\text{no-inter}}$ and $B = \sum_{\lambda} \mathbf{C}_{\lambda}^{\mathbf{B}} \mathbf{L}_{\lambda}^{\text{no-inter}}$. When creating the synthetic data, we used sensor sensitivities $\mathbf{C}^{\mathbf{R}}$, $\mathbf{C}^{\mathbf{G}}$ and $\mathbf{C}^{\mathbf{B}}$ that are considered unknown at test time. The optimization problem consists in finding the spectral curves $\hat{\mathbf{C}}^{\mathbf{R}}$, $\hat{\mathbf{C}}^{\mathbf{G}}$ and $\hat{\mathbf{C}}^{\mathbf{B}}$ that minimize the following objective function:

$$\sum_i \left\| \sum_{k \in \{R,G,B\}} k_i - \sum_{\lambda} \mathbf{C}_{\lambda}^{\mathbf{k}} \mathbf{L}_{i_{\lambda}}^{\text{no-inter}} \right\|_2 + \alpha \left\| \frac{\sum_{k \in \{R,G,B\}} d^2(C^k)}{d\lambda^2} \right\|_2. \quad 5.28$$

The first term is the euclidean distance between the *RGB* components obtained from the model and the available data. The second term is a regularization one that helps to get smooth sensor sensitivity curves based on the norm of the their second derivative. α is controlling the balance between the two parts of the equation. $\mathbf{L}_i^{\text{no-inter}}$ is the radiance of the i^{th} patch in the image.

As a second test, we consider folded patches and use Equation ((4.29)) in order to obtain the radiance with interreflections. Then we minimize a similar objective function:

$$\sum_i \left\| \sum_{k \in \{R,G,B\}} k_i - \sum_{\lambda} \mathbf{C}_{\lambda}^{\mathbf{k}} \mathbf{L}_{i_{\lambda}} \right\|_2 + \alpha \left\| \frac{\sum_{k \in \{R,G,B\}} d^2(C^k)}{d\lambda^2} \right\|_2. \quad 5.29$$

In this case, we have much more input data compared to the previous one. If we consider the 24 patches of the Macbeth color checker, i will go from 1 to 24 in the equation (5.28) while it will go from 1 to $24 \times n_f$ in equation (5.29), where n_f is the number of facets used to discretize each patch. Thus, for each color patch, we consider n_f different *RGB* triplets in the case of inter-reflections (folded patches) but only 1 triplet per patch without inter-reflection (flat patches).

Since we are working with synthetic data in this section, we propose to add different levels of noise to these data before running the optimization process.

Figure 5.8 illustrates the evolution of the spectral recovery error between real sensor sensitivities and estimated ones with respect to the noise introduced in the *RGB* data. We have added a Gaussian noise whose standard-deviation is increasing along the horizontal axis. By comparing the two cases with and without interreflections, we can see that whatever the noise level is, taking into account interreflections with 3D color charts always improved the accuracy of the results. Indeed, the spectral recovery error is almost twice higher when considering flat classical color chart than when using folded patches. For information, we have added a third curve on this plot, which corresponds to the spectral recovery error that would have been obtained if it was possible to create a very big charts with 1269 color patches that correspond to the 1269 Munsell spectral reflectances. No color chart exists with such number of patches and furthermore we can see that when the noise level is increasing ($> 10^{-7}$), the quality of the reconstruction with these 1269 patches is not better than the one obtained with the 24 folded Macbeth patches. For the sake of clarity, we have displayed the results corresponding to the noise range $[0, 10^{-5}]$, but it is interesting to know that with greater noise such as 10^{-2} , our method get an error of 81.02 while the method without interreflections get 90.67. This confirms the out-performance of our approach for greater noise levels.

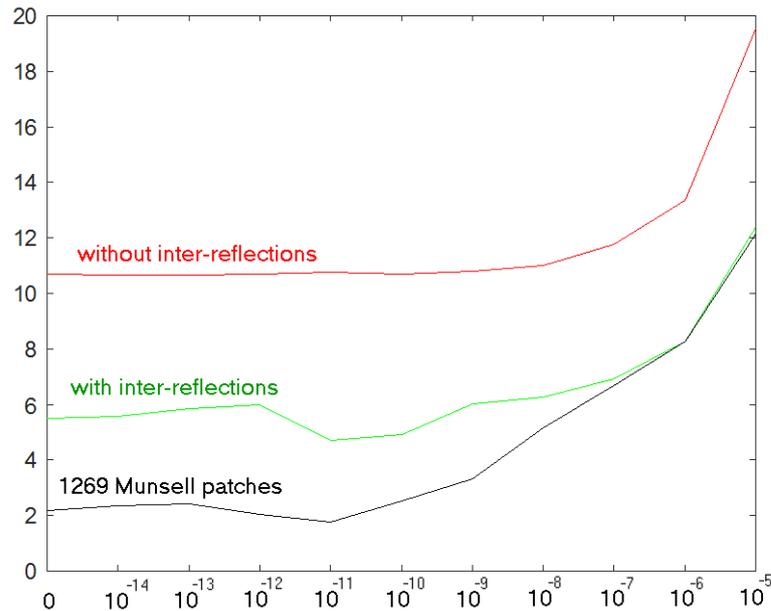


Figure 5.8: *Spectral recovery error versus noise. The horizontal axis represents the evolution of the standard deviation of the added Gaussian noise. Red and green curves are related to the Macbeth color checker*

Furthermore, we have also tested the sensitivity of the results according to the folded angle. For example, with an error of 10° in the angle (test with 35° while using the model for 45°), our results still outperform by more than 16% the ones obtained without interreflections.

Real data

In this section, we have acquired real images with the color camera Canon EOS 1000D and reconstruct the spectral sensitivities of its sensors from the obtained raw RGB values. The same six patches presented in earlier chapters are used for these experiments. They are either acquired flat or folded depending on the experiments we are conducted. The spectral reflectances of these patches have been measured in our lab as well as the SPD of the light that was illuminating the scenes during the acquisitions.

We used the same objective functions as the ones presented in the previous section. We reconstructed the sensor sensitivities either from one color image representing the flat six patches used in earlier chapters, or from six color images representing the same patches but folded. The curves obtained with these two experiments are plotted on figure (5.9). We can see that the spectral reconstructions obtained by accounting interreflections (green curve) are much better than the ones obtained without these interreflections. The associated spectral recovery errors are 13.10% when interreflections are accounted and 23.50% when they are not.

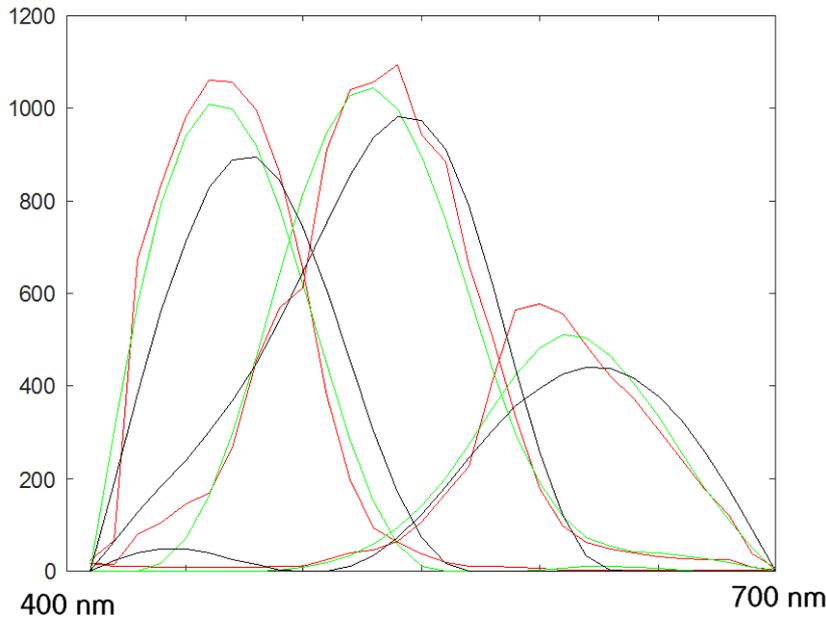


Figure 5.9: *Sensor sensitivities reconstructions with and without interreflections. Red curve: ground-truth, green curve: reconstruction with interreflections, black curve: reconstruction without interreflections.*

5.3 Conclusion

In this chapter, we have presented the state of the art of direct spectral reflectance estimation approaches along with a study of the use of self-interreflections in this domain. Our experiments on synthetic data showed that using interreflections in one RGB image taken under white light gives as accurate spectral estimation as the state of the art approaches which use three RGB images taken under three different lights. The results of our experiments on real data showed that when using a uniformly colored surface, the whole spectral reflectance can be found with a good precision from a single RGB image of a concave surface taken under direct daylight. This performance is achieved without the need of any pre-calibration for the specific acquisition settings. In addition, we presented a detailed study of the effect of the angle and the facet size on the quality of spectral reflectance estimation. Furthermore, the uniqueness of the solution has been analyzed. This work has been accepted in Applied Optics Journal OSA [5].

The proposed solution is valid in case of collimated homogeneous light without taking into consideration the ambient light. This along with image noise explain the degradation of the results obtained on real data when compared to those obtained on synthetic ones. The model can be further enhanced by treating ambient light as well as direct light. Moreover, the surfaces used in the experiments were not fully Lambertian, they show some specularity at grazing incidence of light. Thus, providing a more general form of the interreflections model could be appreciable in order to enhance the accuracy of the method in real scenarios.

Later, we have studied the use of 3D color charts instead of the commonly used 2D ones in camera characterization. We showed that using 3D color charts leads to more accurate

estimation of the spectral responses of the sensors when compared to adding more colors to the 2D color charts. This is due to the fact that while natural reflectances lie in a low dimensional space, interreflections help in adding some non-linearly related information which enhance the spectral estimation. The non-linearity is due to the equivalent spectral reflectances which are the products of raising the natural spectral reflectances to different powers. This work has been published and presented as a poster in BMVC 2017 conference [6]. As far as we know, no other work in the literature proposed to use such 3D color charts for camera characterization. One important question to be investigated is how many 3D patches needed to obtain a spectral accuracy which is comparable to using the 1269 Munsell patches.

Our preliminary experiments demonstrated the potential of our interreflection model in two research areas. Nonetheless, this work, we believe, may open doors for new investigations in other research areas as well.

Chapter 6: Interreflections in Learning-Based Spectral Estimation

In this chapter, we draw our attention to learning-based methods in spectral reflectance estimation. A state of the art concerning these approaches is first given. Afterwards, we investigate the application of a deep network trained on simulated data built using our interreflection model to find the inverse function for interreflections. Deep learning networks are already used in physics-based vision. They have shown an impressive ability to learn non-linear functions for problems where abundant data is available. The advantage of using deep networks for deriving an inverse function is that we only require the image formation equations which are typically simpler than their inverse. In our proposed deep learning approach we loosen the assumption of a known illuminant and leave the task to estimate the SPD of the light also to the network.

6.1 State of the art

As opposite to direct approaches for spectral reflectance estimation, learning-based approaches aim to overcome the necessity of having prior knowledge of the spectral responses of the used camera or the SPD of the lighting system in the scene.

6.1.1 Kernel-based regression approaches

Supposing the color formation equation presented earlier (Equation 2.12) can be written in a matrix form as:

$$\boldsymbol{\rho} = \frac{1}{\pi} \mathbf{C} \mathbf{E} \mathbf{R}, \quad 6.1$$

where $\boldsymbol{\rho}$ is a column vector containing the RGB values, \mathbf{C} is a $3 \times v$ matrix for the v considered wavebands, \mathbf{E} is a diagonal square matrix containing the irradiance at each waveband, and \mathbf{R} is a column vector corresponding to the spectral reflectance.

Let us name the resulting matrix of multiplying \mathbf{C} and \mathbf{E} as \mathbf{W} . In the observational approaches presented in Chapter (5), this matrix is known and used to reconstruct the spectral reflectance from several RGB values. However, in indirect approaches, matrix \mathbf{W} is not known and need to be learned from several examples presented in the training set. Each example contains a spectral reflectance and the corresponding RGB values. In addition, all the examples correspond to the same acquisition conditions, i.e. the same camera and lighting system. Thus, matrix \mathbf{W} can be learned from the set of equations of the form (6.1) whose number is equivalent to the number of examples in the training set. However, as the sensor responses are in a three dimensional space, the reflectance obtained after learning the matrix \mathbf{W} is also in a three dimensional space which leads to inaccurate spectral estimation. In order to solve this issue, Heikkinen *et al.* [67] proposed to add non-linearity to the system by defining the problem as a regression one in reproducing kernel Hilbert space (RKHS) [76, 77, 78]. They proposed to map the RGB values into a high-dimensional feature space and use different kernel functions to represent the inner product in that space. In addition, they used a regularization factor, similar to the one introduced

in Chapter 5, to better cope with measurement noise and numerical instabilities, and to avoid any over-fitting to the data in the training set.

Different subsets of the 1269 Munsell patches were used in training. Kernel regression shows a better results with large subsets. However, the approach still cannot handle the case of metamerism. In this case, the estimated spectral reflectance will depend on the training set.

Later, Heikkinen *et al.* [60] studied the use of two illuminants in kernel regression with Gaussian kernel. They also evaluated the use of a physically motivated $[0, 1]$ constraint on the spectral reflectance which improved their results. They used different standard color sets for their experiments in order to study the effectiveness of the choice of model parameters. The use of multiple lightings helped in achieving improvements on the results in simulated data case. However, for real conditions, that drives the approach to be more sensible to noise, and considering the properties of the used light becomes necessary to cope with this issue.

In [79], the authors focused their attention on testing Matèrn kernel in spectral reflectance estimation. In addition, they tested the use of several link functions in order to obtain physically feasible spectral reflectance values. These link functions helped in improving the accuracy of the estimated spectra, especially when it comes to the spectral shape evaluated by the Pearson Distance (PD). The experiments on both simulated and noisy real data showed that the combination of the Matèrn kernel and the logit link function decreased many spectral errors in comparison to Gaussian kernel regression. More performance evaluations of RKHS regression and Matèrn kernel can be found in [55]. The approach was tested using screen-based adjustable lighting and different datasets classed into weak, moderate and high-quality categories. The use of multiple illuminants helped in improving the accuracy of the estimated spectral reflectance even when a weak training set is used.

6.1.2 Hyperspectral recovery approaches

Recently, some works aimed at recovering the spectral curves of all the pixels of a single RGB image [80, 81, 82, 83]. The idea of [81] is to model the mapping between camera RGB values and scene reflectance spectra with a radial basic function network. Arad *et al.* [80] exploit the prior on the distribution of hyperspectral signatures in natural images in order to construct a sparse hyperspectral dictionary. The projection of this dictionary into RGB provides a mapping between RGB atoms and hyperspectral atoms. Thus, given RGB values that can be decomposed into RGB atoms, their spectral reflectance is obtained by using the same combination of the corresponding hyperspectral atoms. The accuracy and efficiency of this approach was improved by Aeschbacher *et al.* [82] who proposed to learn the dictionary projection between the training RGB and hyperspectral data and to extract anchor points from this projection. At test time, a simple nearest anchor search is run for each RGB triplet in order to reconstruct its spectral curve. Finally, Alvarez *et al.* proposed a convolutional neural network architecture that learns an end-to-end mapping between pairs of input RGB images and their hyperspectral counterparts [83]. They adopt

an adversarial framework-based generative model that takes into account the spatial contextual information present in RGB images for the spectral reconstruction process. For these last works, even if the aim is also to reconstruct spectral functions from RGB data, the objective appears to be different from the task addressed in this paper. Recovering a full resolution hyperspectral image is much more challenging and can not provide results as accurate as the ones provided by approaches that concentrate on a single surface observed under calibrated conditions. For example, the best results in terms of root mean square error (RMSE) reported in [82] are more than 1.00 which is very high in terms of accurate spectral estimation.

6.1.3 Deep learning use in physics-based vision

Deep learning approaches have been applied to some physics based vision problems. In most of the cases, these approaches were used to solve the color constancy problem. For example, since the choice of the pooling between local and global estimates is not an easy task, Fourure et al. [84] proposed a deep network that can choose between different poolings. The output of this network was a global light color estimate. Shi et al. [85] proposed an approach to locally estimate the light color by using two interacting sub-networks. The idea is to use a first sub-network to create multiple hypotheses for each patch and then the best ones of these hypotheses are voted by a second sub-network. Recently, Hu et al. [86] proposed a fully convolutional network to create a confidence map which provides weights for the patches in the image. Patches with the highest weights are considered those which can give the best estimates of the light color.

Deep networks were also applied to intrinsic images decomposition. Narihira et al. [87] proposed to use a convolutional neural network to predict the differences of the perceived surface reflectance between pixels by learning from human judgment on real images. Shi et al. [88] extended this approach to non-Lambertian objects and proposed a CNN that is able to recover diffuse albedo, shading, and specular highlights from a single image of an object. Recently, Janner et al. [89] proposed the Rendered Intrinsic Network that contains two convolutional encoder-decoders: the decoder to decompose the input image into reflectance, shape and lighting and the encoder to reconstruct the input for these resulted images. This network can learn from unlabeled data as it uses an unsupervised reconstruction error as a loss function.

In reflectance estimation, Rematas et al. [90] proposed a convolutional neural network to estimate reflectance maps of specular complex-shaped materials in natural lighting conditions. This is done under the assumptions of a constant material with no shadows nor interreflections, and distant light source and viewer. The reflectance map contains the orientation-dependent RGB radiance values of the fixed material under the specific illumination. Under the same assumptions, Georgoulis et al. [91] proposed a deep convolutional approach to estimate a detailed appearance of the environment from precomputed reflectance maps of multiple unknown materials.

6.2 Interreflection-based approach

In this section, we propose an interreflection-based approach to estimate both the spectral reflectance of the surface and the SPD of the illuminant from a single RGB image. To this end, we propose to use a convolutional neural network trained on simulated data only.

6.2.1 Datasets

In our approach, we use only simulated images in the training sets. These images are built using our infinite-bounce interreflection model for Lambertian surfaces. It is possible to build datasets including any concave shape and using multiple illuminants. However, we consider that each dataset is constructed while taking into consideration only one camera with known spectral responses functions.

In order to evaluate our approach, we created datasets of simulated images of size 10×10 RGB values corresponding to the facets of one side of a V-shaped surface. Each image corresponds to one side of a V-shaped configuration of a surface with a homogeneous spectral reflectance of the 1269 Munsell patches. Given the discretization size 10×10 , and after choosing the angle and the sizes of the planar surfaces, the geometrical kernel values can be obtained using Monte Carlo integration as explained earlier. Then, for the chosen camera with known spectral responses, RGB values can be obtained using Equation (4.47) for each spectral reflectance of Munsell patches, and for every considered illuminants. Only collimated light sources are considered, each parallel to the bisecting plane of the two panels, illuminating the V-cavities frontally.

After obtaining the RGB values using our model, every image is normalized on the sum of its RGB values. This is done in order to avoid the need to calibration to the specific acquisition settings. The images are then pre-processed by applying mean subtraction and standard deviation normalization.

To better cope with image noise, data augmentation is performed by adding different levels of noise to each image. We adopted two types of noise to be added to images:

- Poisson noise is added to each image in order to cope with shot noise which are more dominant in low RGB values.
- Gaussian noise is also added to the images. The standard deviation of the Gaussian function is selected based on a random decision from a set of six variances: $[0, 10^{-6}, 10^{-8}, 10^{-10}, 10^{-12}, 10^{-14}]$. Some images will not have an added Gaussian noise, this is when the selected standard deviation is 0.

However, these noises are added to images at batch level inside the training process. This helps in avoiding the augmentation of the dataset size. In addition, the same image would have different added noises throughout the training. This choice might have a slightly negative effect on the quality of the spectral estimation compared to the ideal case of adding all the possible noises to each image in the dataset before training but it provides a significant gain in training speed.

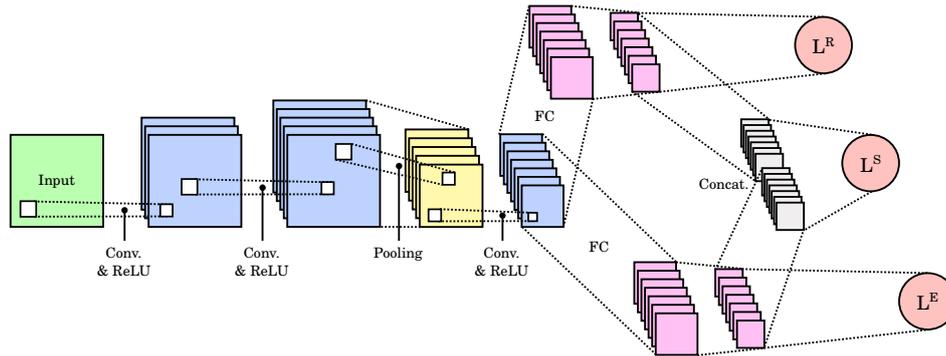


Figure 6.1: Diagram of the proposed network for surface reflectance and illuminant estimation from interreflections.

6.2.2 Network structure & loss functions

For our regression problem, we use the convolutional neural network structure presented in Figure (6.1), in order to obtain both the spectral reflectance of the surface and the spectral power distribution of the lighting from raw image RGB values. The network architecture starts with a shared part which consists of three convolutional blocks. Subsequently, the network is split into two branches; one for the estimation of the surface reflectance and one for the spectral power distribution. Let us take a closer look to the network:

- The first convolutional block is constituted of two convolutional layers of size 3×3 with padding of 1, followed by a ReLU activation layer.
- The second convolutional block is also constituted of two convolutional layers of size 3×3 with padding of 1, followed by a ReLU activation layer.
- A pooling layer of size 2×2 then follows, reducing the input size to 5×5 .
- The third convolutional block is also constituted of two convolutional layers of sizes 1×1 and 3×3 with padding of 0 each, followed by a ReLU activation layer. The second convolutional layer has a stride of size 3×3 .
- Two identical branches of two fully connected layers come after, giving the surface spectral reflectance and the spectral power distribution of light restrictively. Each branch is followed by the corresponding loss function.
- Finally, the resulted spectral values are concatenated, and then passed together to a third loss function.

Let us denote by \mathbf{R}_λ , $\widehat{\mathbf{R}}_\lambda$, the ground truth spectral reflectance vector and the estimated one, respectively, and by $\mathbf{E}_{0\lambda}$, $\widehat{\mathbf{E}}_{0\lambda}$, the ground truth SPD and the estimated one respectively. Each of these vectors contains the values of all the considered wavelengths. For this reason we put the subscript λ to distinguish these vectors from the ones introduced previously.

Three loss functions are used to obtain accurate spectral reflectance and SPD. The first loss function, $L^R(\mathbf{R}_\lambda)$ is regarding the error on the spectral reflectance defined as:

$$L^R(\mathbf{R}_\lambda) = \frac{1}{2}(\mathbf{R}_\lambda - \widehat{\mathbf{R}}_\lambda)^2. \quad 6.2$$

The second loss function is related to the error in the SPD of illuminant and is defined as:

$$L^E(\mathbf{E}_{0\lambda}) = \frac{1}{2}(\mathbf{E}_{0\lambda} - \widehat{\mathbf{E}}_{0\lambda})^2. \quad 6.3$$

In principle the network can be trained with only the L^R and L^E losses. However, we introduce in addition a consistency loss. We would like to have the multiplication of the estimated reflectance and illuminant the closest possible to the real color signal $\mathbf{S}_\lambda = \mathbf{R}_\lambda \odot \mathbf{E}_{0\lambda}$, where \odot is the component-size multiplication operator so that \mathbf{S}_λ is also a spectral signal. This loss enforces consistency between the separately estimated surface reflectance and light SPD. We use the CIE 1931 XYZ color matching functions to obtain an estimation which minimizes a perceptually relevant error because it gives more importance to wavelengths with high visual response [79]. This consistency loss is defined as:

$$\begin{aligned} L^S(\mathbf{S}_\lambda) = & \frac{1}{2}(\bar{\mathbf{X}} \odot \mathbf{R}_\lambda \odot \mathbf{E}_{0\lambda} - \bar{\mathbf{X}} \odot \widehat{\mathbf{R}}_\lambda \odot \widehat{\mathbf{E}}_{0\lambda})^2 \\ & + \frac{1}{2}(\bar{\mathbf{Y}} \odot \mathbf{R}_\lambda \odot \mathbf{E}_{0\lambda} - \bar{\mathbf{Y}} \odot \widehat{\mathbf{R}}_\lambda \odot \widehat{\mathbf{E}}_{0\lambda})^2 \\ & + \frac{1}{2}(\bar{\mathbf{Z}} \odot \mathbf{R}_\lambda \odot \mathbf{E}_{0\lambda} - \bar{\mathbf{Z}} \odot \widehat{\mathbf{R}}_\lambda \odot \widehat{\mathbf{E}}_{0\lambda})^2, \end{aligned} \quad 6.4$$

where $\bar{\mathbf{X}}$, $\bar{\mathbf{Y}}$ and $\bar{\mathbf{Z}}$ are the CIE 1931 XYZ color matching functions.

The network is initialized using Xavier [92]. The learning rate is set to 10^{-4} and is reduced every 20 epochs by a factor 10. The momentum is set to 0.9 and the batch size to 50.

6.3 Experimental results & discussions

In this section, our results on both simulated images and real camera outputs are presented. In our tests, we considered a V-shaped configuration consisting of two planar square surfaces with an angle of 45° between them (see Figure (6.2)). The dataset and the test images contain only the area of one side of the planar surface discretized into 10×10 facets.

Different metrics are used in order to evaluate our approach and to compare our results with those obtained using other spectral reflectance estimation approaches. The root mean square error (RMSE), and the Pearson distance (PD) are used to show the accuracy of spectral estimation. The performance is also evaluated in terms of color distance. We use CIEDE00 distance [73] with CIE D65 lighting and a 2° viewing angle. This distance is computed between the color signal obtained with the ground truth spectral reflectance and the one obtained with the estimated spectral reflectance.

6.3.1 Simulated data

For simulated data tests, the datasets are built using the CIE 1931 XYZ color matching functions as sensor spectral responses. This is done in order to compare our results to the ones in [79] under the exact same configurations. Wavelengths are taken between $400nm$ a $700nm$ with a $5nm$ step, giving a total of 61 wavelengths. Two datasets are considered for this case, the first is built using only the CIE D65 SPD. The network structure in this case is adapted so that the layers corresponding to the SPD of light are deactivated leading to only two loss functions L^R and L^S . In addition, both \mathbf{E}_0 and $\widehat{\mathbf{E}}_0$ are replaced by the SPD of CIE D65 in the calculation of L^S . Following the same process as in [79], the network is trained using 90% of Munsell patches and tested on the remaining 10%. The second dataset is built using 23 different SPDs corresponding to illuminants on the Planckian locus with color temperatures ranging from $4000K$ to $15000K$ with steps of $500K$ between them. In this case, the network whose structure is shown in Figure (6.1) is also trained using 90% of Munsell patches under various illuminants and tested on the remaining 10%.

Our results on both datasets are shown in Table 6.1. They are compared to the best results regarding the average error in [79] obtained using a logit link function with different kernel models. The results are shown in terms of average, max and 95th percentile RMSE and PD values. It can be observed from the results that when training the network with the dataset built using a single illuminant, which represents the setting of reflectance estimation with known illuminant, our approach outperforms the state of the art approaches. In addition, when training using the dataset built with different lightings, which is equal to the setting of unknown scene illuminant, our results are still better than the state of the art ones except for the maximum error values. However, in this case, the SPD of light is learned also with an average RMSE of 0.0154. The average DE00 values are 0.6379 and 0.7279 for the single illuminant dataset and the multiple illuminant dataset respectively. It is worth mentioning that, in this case, our algorithm is able to reconstruct the unknown light SPD while for the other tested methods in Table 6.1, this SPD is a known input. However, one can notice that the RMSE value of the light SPD is bigger than the one calculated for the spectral reflectance: 0.0154 compared to 0.0088. This is due to the fact that the spectra of our 23 lights are different in scales from those of the Munsell patches. In order to better compare the performance of our spectral estimation of light and surface reflectance, Normalized Root Mean Square Errors (NRMSE) can be used where each RMSE is normalized on the mean of the corresponding dataset. NRMSE value calculated for the light SPD is 0.0221 whereas when it is calculated for the spectral reflectance of the surfaces this error goes to 0.0319.

6.3.2 Real data

For real data tests, a new simulated dataset is built using the 1269 Munsell spectral reflectances, the Canon EOS 1000D spectral response functions and the 23 planckian illuminants. The network is trained with this dataset except the green and red patches in order to make sure that no over-fitting is happening. The same camera is then used to take photos of V-shaped surfaces with an approximative angle of 45° of six colored sur-

Table 6.1: Spectral error values for simulated data

Method	RMSE Avg.	RMSE Max.	RMSE 95th	PD Avg.	PD max.	PD 95th
Gaussian kernel	0.0103	0.0508	0.0333	0.00104	0.0154	0.00462
Matérn kernel	0.0092	0.0558	0.0326	0.00088	0.0083	0.00420
TPS kernel	0.0092	0.0556	0.0332	0.00088	0.0083	0.00411
Ours (1 light)	0.0087	0.0460	0.0191	0.00081	0.0068	0.0041
Ours (23 lights)	0.0088	0.0771	0.0213	0.0012	0.0639	0.0041

Table 6.2: Results on real data

Method	Pre-calibration	Illuminant	RMSE	PD	DE00
Park et al. [59]	Yes	Known	0.060	0.009	3.76
Khan et al. [65]	Yes	Known	0.059	0.009	3.87
Interreflection-based (Chapter 5)	Yes	Known	0.046	0.008	3.82
Interreflection-based (Chapter 5)	No	Known	0.061	0.012	5.62
Interreflection with learning	No	Unknown	0.045	0.004	4.12

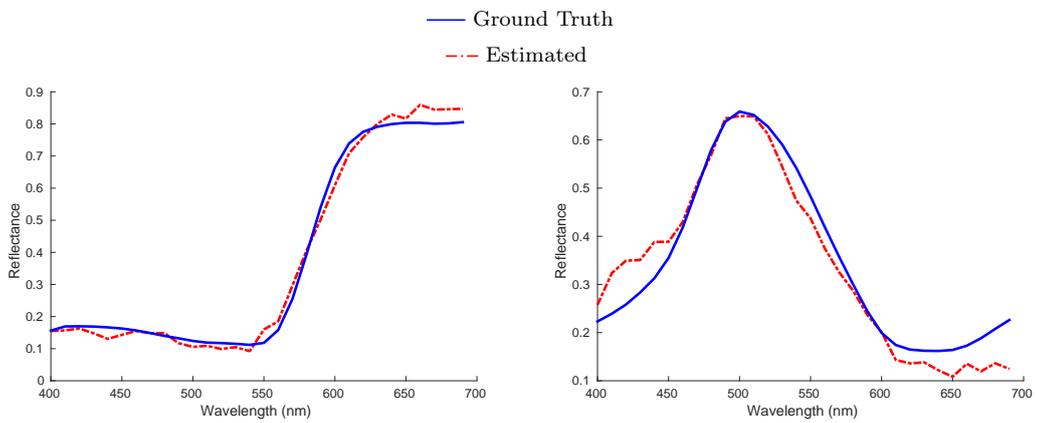
faces: the red Munsell paper, and the five other textile pieces of different colors presented earlier. The photos are taken under direct sunlight in the early afternoon. The area of the photo corresponding to one side of the V-shaped surface is selected manually to be then automatically discretized into 10×10 facets each represented by the mean RGB values over its area. In Table 6.2 our results are compared to the state of the art ones on real images under a known illuminant [59, 65], including our approach presented in Chapter 5. The results show that our learning approach outperforms the other ones in terms of spectral error even when a pre-calibration step using the XRite ColorChecker is performed to help these approaches. This pre-calibration, as detailed in Chapter 5, is mandatory for the classical approaches [59, 65] in order to get accurate reconstructions when used in non-calibrated configuration. It requires to add a color checker in the image, which is a strong constraint. Moreover, when compared to our physics-based interreflection approach, a significant enhancement in the spectral estimation is obtained when no pre-calibration step is used. The SPD of light is not known in our learning approach and is guessed with an average RMSE of 0.1 in comparison to CIE D50 spectral power distribution. Please note that this RMSE value is obtained while comparing the estimated SPD with a default light SPD and not with a measured one.

Figure (6.2) shows the estimated surface spectral reflectance compared to the ground truth one using a single photo taken under daylight for Munsell red paper, and cyan textile piece. The estimated illuminant SPD is also shown next to the one of CIE D50 which was chosen as a representative of direct sunlight. Figure (6.3) shows the estimated spectral reflectance of the Munsell red paper when using deep learning compared with the one obtained with our approach without learning presented in Chapter 5. One can notice that when using deep learning the estimated spectrum is closer to the ground truth one.

Real Images



Surface Spectral Reflectance



Light SPD

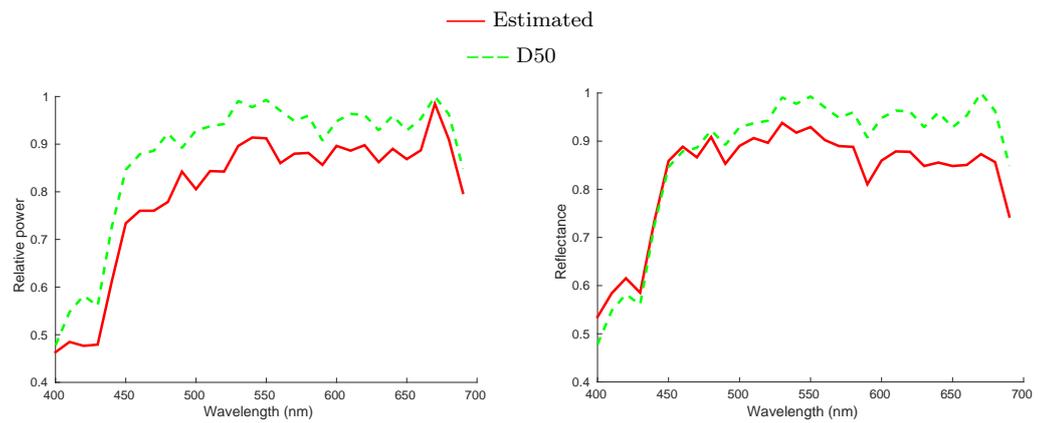


Figure 6.2: Examples of estimated surface spectral reflectance and light SPD using our approach.

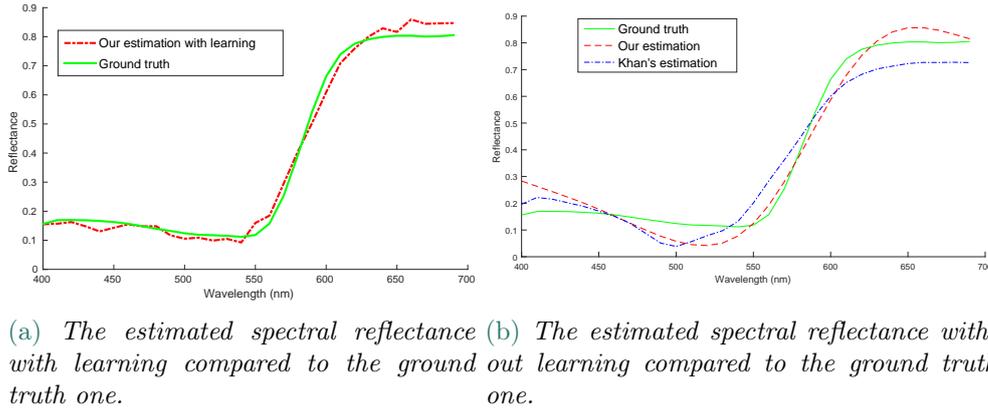


Figure 6.3: Enhancement of spectral reflectance estimation for the Munsell red paper when using deep learning compared to non-learning based approaches.

6.3.3 Validation of network structure

The choice of network structure was done experimentally. The idea was to use a simple network that works well for our regression problem. We found that batch normalization cannot be used in these sort of regression problems where the exact intensity of the output signal is crucial for the quality of the estimation. Moreover, as the output spectral values for both surface and light are in the range $[0, 1]$, one may think to add a sigmoid layer to accelerate the convergence. However, putting a sigmoid layer played the opposite role in our case leading to a non convergence of the network. This is probably due to the fact that our network is relatively shallow, so a sigmoid is being an obstacle in the learning process.

In order to verify the importance of the third loss function, L^S , we performed an ablation study. We trained the network twice under the same settings and on the exact same dataset, the only difference is that we deactivated the consistency loss L^S for one training and activated it for the other. Table 6.3 shows the percentage of enhancement in RMSE, PD and DE00 errors when the consistency loss is activated compared to when it is deactivated.

Table 6.3: Error enhancement when using the consistency loss

	RMSE	PD	DE00
Error enhancement (%)	7.7	11.6	14,8

6.3.4 Generalization to other angles

One important thing to verify is how well our approach generalizes to other geometries. One way to check this can be done by training the network on other V-shaped surfaces with different angles between the two planar surfaces. Thus, we trained the network on new datasets built using V-shaped surfaces of all Munsell patches with angles of 30° , 60° , 90° , 120° and 150° . All the other settings are exactly the same as those used in our tests on simulated data. Table 6.4 shows the RMSE, PD, DE00 errors regarding the spectral

reflectance, as well as the RMSE error regarding the light SPD. One can see that the approach generalizes well whenever the angle is small enough to give significant amount of interreflections. Our results show that both spectra are estimated with a very good accuracy for angles reaching 120° . For much bigger angles, 150° for example, interreflections happen less leading to less important color gradients which are not enough for the network to learn the inverse solution.

However, one possible limitation is that till now we propose to train the network on a specific geometry. Thus, and in order to provide a more concrete application of our approach, we tried to train the network on three angles together, 30° , 60° and 90° . Table 6.5 shows the results in terms of RMSE, PD, and DE00 for spectral reflectance and RMSE of light SPD when the network is tested on one of the angles included in the training set, and also on a new angle which the network did not encounter before. As it can be observed, the approach generalizes very well to training on different geometries even when it is tested on a new geometry. Error values obtained when the network is tested on one of the angles used for training are very close to those obtained when the network is trained only on that specific angle (see Tables 6.4 and 6.5). On the other hand, when the network is tested on a new angle, most of the error values increase significantly but RMSE is still better than the values obtained when using observational approaches (see Table 5.1 in Chapter 5). Moreover, it is noticeable that while all the other errors increase, it is not the case for the PD one. This observation leads us to think that the network is falling in some metamerism traps for some of the patches. This is probably happening when the network is failing to recognize if the color gradients are due to high reflectance with big angle or to lower reflectance with smaller angle. Thus, the network in this case is getting the shape of the spectrum correctly but is not able to accurately identify the exact spectral reflectance. However, training on more angles with smaller steps between them would be a solution for this problem. In this case, more detailed association between the color gradients and the angles would be learned reducing as a consequence the previous case of error.

Table 6.4: Results of training with different angles

Trained angle	Surface reflectance			Light SPD
	RMSE	PD	DE00	RMSE
30°	0.0090	0.0014	0.6976	0.0173
60°	0.0090	0.0014	0.6976	0.0173
90°	0.0091	0.0014	0.6704	0.0160
120°	0.0093	0.0011	0.6775	0.0146
150°	0.0260	0.0073	4.1055	0.1050

Table 6.5: Results of training with different angles

Trained angle	Tested angle	Surface reflectance			Light SPD
		RMSE	PD	DE00	RMSE
30° & 60° & 90°	60°	0.0099	0.0013	0.8382	0.0178
30° & 60° & 90°	45°	0.0134	0.0014	1.1498	0.0203

6.4 Conclusion

In this chapter, a convolutional neural network was trained in order to solve the inverse problem of surface spectral reflectance estimation from a single RGB image of interreflections under unknown lighting. Datasets were built from synthetic images simulated using our infinite-bounce interreflection model. Different noise types and levels were added to the images during the training in order to better cope with noises in camera outputs. Our experiments on simulated data showed that our approach even under an unknown illuminant outperforms the state of the art learning-based spectral reflectance estimation approaches trained for a specific known lighting. In addition, real data results showed that our method gets a better accuracy of spectral reflectance estimation under unknown lighting than physics-based approaches under a light with a known SPD. The improvement of our method over physics-based methods could be explained by the fact that we incorporate realistic noise models in the dataset creation, whereas it is very difficult to propagate the impact of noise through physics-based methods. This work has been submitted to JOSA A journal. In a future work, the approach can be extended to take into consideration interreflection between surfaces with different spectral reflectances and under different geometrical configurations. Also, the dataset can include new models of lighting other than the collimated one used in our case. This can introduce further enhancements on the quality of the estimated spectra from real images as ambient light does exist in the scene but it was ignored in our tests.

We focused in this thesis on studying interreflections and their possible applications in spectral estimation in the domain of computer vision. Interreflections happen in any illuminated concave surface leading to particular color gradients over it. We showed in this manuscript that these color gradients can play an important role in spectral estimation as they are signatures of the spectral reflectance of the surface and of its geometry.

In Chapter 3 we demonstrated that these color gradients play also an important role in perception. Based on our visual experiments, we found that people get more clues about the color of the object and the color of the light when observing concave surfaces with interreflections compared to flat ones. Thus, color constancy is enhanced with interreflections. In the same chapter, we studied the color gradients in images in order to better understand their mechanism and their particularity. We also provided a state of the art for all the precedent uses of interreflections in computer vision. However, most of the previous approaches do not take advantage of the color gradients due to interreflections in images.

After reviewing the previous attempts in interreflection modeling based on a physical background, we introduced our spectral infinite-bounce model in Chapter 4. This model is the first to relate RGB values in the image to the geometry and the spectral reflectance of the surface, the SPD of the lighting and the spectral responses of the camera sensors while taking into account an infinite number of bounces. Later, the accuracy of interreflections simulation using this model compared to real images is discussed, and the importance of taking into consideration an infinite number of bounces instead of only two as common in the state of the art is demonstrated.

In chapter 5, we proposed to use our interreflection model in an inverse approach in two applications. First, we used our model in spectral reflectance estimation from a single RGB image, supposing that both the SPD of light and the spectral responses of the camera are known. Even under these assumptions, it was not possible for the state of the art approaches to accurately retrieve the spectral reflectance from only three values in the RGB space. However, this is made possible with interreflections: the color gradients on a surface with a uniform spectral reflectance over it made it possible to retrieve this spectral reflectance without even the need to use a set of basis functions in order to reduce the dimensionality of the problem. Moreover, having those color gradients which are strongly tied with the spectral reflectance, we showed that there is no more need for a any pre-calibration step for the specific acquisition settings or any measurement of the intensity of light. Later, we introduced a new application of the interreflection model in camera characterization. 2D color charts including patches with known spectral reflectances are generally used in the literature to get the spectral responses of camera sensors, but the accuracy of the estimated spectra is limited. This is due to the fact that colored patches lie in a low dimensional space, thus providing linearly related information which does not help in enhancing the accuracy of the estimated spectra. For the same reason, it has been demonstrated that even adding new color patches would not increase the accuracy of the estimated spectra. However, we show that the nature of interreflections provides in itself

a non-linearly related information that are very helpful in such applications. Therefore, 3D color charts is an alternative to 2D ones, which lead to an important enhancement in spectral estimation.

In order to better cope with image noise and to relax one constraint of our previous spectral reflectance estimation approach which is the need to know the SPD of light, we proposed in Chapter 6 to train a convolutional neural network on data simulated with our interreflection model. Different levels of noises were added to these simulated images including Poisson and Gaussian noises. Our results showed that our approach outperforms the state of the art learning-based spectral estimation ones on simulated data while at same time estimating the SPD of the used light with a good accuracy. In addition, we obtained better results on real data when compared to our previous approach, mainly because image noise is better handled.

7.1 Future Work

The interreflections model and some of its possible uses in spectral estimation for computer vision applications were presented and discussed in this thesis. However, we think that this work merits more future study in order to better generalize it, overcome its limitations, or use it under new applications. In the following, some possible future works are discussed.

7.1.1 Generalization of the interreflection model

As seen in Chapter 4, our spectral infinite-bounce model is limited to Lambertian surfaces. It has been demonstrated by Seitz et al. [25] that, for any type of surfaces, there exists a linear matrix transformation that permits to obtain the first-bounce radiance from the final radiance. However, this matrix is very complicated in case of non-Lambertian surfaces as it does not only depend on the reflectance and the geometry of the surface, but also on the incident and output light directions. One alternative possible solution to generalize the model in a simpler way can be by adding a specular term, independent of interreflections, to the model as it is generally used in dichromatic reflectance models. Specularity is a light-dependent phenomenon and it is not related to the spectral reflectance of the surface. Nonetheless, depending on the geometry of both the light and the surface, some of the specularity can act like a new light source lighting some points of the concave surface, and this should be taken into consideration in the interreflection model.

Moreover, the model can be also generalized to different light geometry, including diffuse light sources and a mix of direct and diffuse light sources. We do also think that interreflections can help in detecting the spectra of both diffuse and direct light sources, as they compete with shadows in this case.

Furthermore, the model can be also enhanced by accounting to non-linearities in camera responses.

7.1.2 Toward a less constrained spectral estimation

It has been shown in Chapter 5 that using interreflections was very useful in direct spectral reflectance estimation. However, one strong assumption that we have and might be an obstacle in front of some applications is the known geometry. Thus, one important future work that has already started during this thesis is about studying the patterns of color gradients due to interreflections, next to a more detailed study of metamerism with interreflections. This can be very useful in order to provide an approach that is able to estimate the spectral reflectance under unknown geometry. Moreover, and in order to be able to estimate the spectral reflectance of a concave surface, one possible future direction is to study the use of mirrors to create interreflections instead of folding the surface.

On the other hand, the learning-based approach presented in Chapter 6 can be farther generalized by training the network on images simulated with different geometries other than the V-shaped ones and with surfaces of multiple colors. In addition, we think that it is important to study the possible generalization of the model to whatever used camera. This can be done by creating datasets built using different spectral responses curves corresponding to several cameras. However, the number of used cameras can be essential for the approach to perform well. Studying the effect of this generalization on the quality of the spectral estimation is another important point in the future work.

7.1.3 Applications in color constancy

Interreflections can be also used in color constancy in order to get the color of the illuminant from a a single image. Generally, in color constancy applications, one uses any available image which is not necessary a raw one. Thus, our interreflection model cannot be used directly in this domain as its outputs are raw RGB values leading to a strict application in the domain of color constancy. However, we have seen that interreflections are related in different ways to the spectral reflectance of the surface and to the SPD of the light: they have non-linear relation with the reflectance and a linear one with the irradiance. We have also seen that interreflections enhanced the color constancy in human vision. These two observations drive us to think that interreflections may have an important role in computer-based color constancy. The patterns of color gradients that interreflections create, and which can be better studied and understood thanks to the interreflection model, can be very useful in providing an interreflection-based application for color constancy.

Nous nous sommes focalisé dans cette thèse à étudier les inter-réflexions et leur possible utilisation dans l'estimation spectrale. Les inter-réflexions arrivent quand une surface concave est éclairée, donnant des variations des couleurs typiques à ce phénomène. Nous avons montré dans ce manuscrit que ces variations de couleurs sont très utiles dans l'estimation spectrale car elles sont des signatures de la réflectance spectrale de la surface et de sa géométrie. Dans le chapitre 3, nous avons démontré que les inter-réflexions jouent un rôle important dans la perception. Nous avons trouvé, en faisant des expériences visuelles, que les gens ont plus d'indices sur la couleur de la surface et la couleur de la lumière en regardant une surface concave qu'en regardant une surface convexe. Cela veut dire que la constance des couleurs chez les humains est améliorée avec les inter-réflexions. Dans ce même chapitre, nous avons étudié plus en détails les variations des couleurs causées par les inter-réflexions et l'ancienne exploitation de ce phénomène dans l'état de l'art. Après avoir expliqué les différents essais dans l'état de l'art pour modéliser les inter-réflexions, nous avons présenté dans le Chapitre 4 notre modèle spectrale en prenant en compte une infinité de rebonds d'inter-réflexions. Ce modèle est le premier à relier les valeurs RGB dans l'image aux différentes fonctions spectrales en jeu et à la géométrie de la surface. Nous avons profité de ce modèle pour montrer l'importance de prendre en compte une infinité de rebonds au lieu de seulement deux, comme il est souvent proposé dans la littérature. En chapitre 5, nous avons proposé d'utiliser notre modèle dans une approche inverse dans deux applications. Premièrement, nous avons proposé une méthode d'estimation de la réflectance spectrale d'une seule image RGB, en considérant connues les fonctions spectrales de la lumière et de l'appareil photo. Il n'était pas possible pour les approches précédents d'avoir une bonne estimation spectrale sous les mêmes conditions en utilisant seulement les trois valeurs RGB. Mais cela a été possible grâce aux inter-réflexions. Les variations couleurs donnent suffisamment d'informations pour une très bonne estimation spectrale sans même avoir besoin d'utiliser les fonctions de bases et sans nécessité d'un pré-calibrage pour les configurations de l'acquisition. Deuxièmement, nous avons démontré que la nature de l'inter-réflexion aide à ajouter de la non-linéarité aux espaces couleurs utilisées pour la caractérisation de l'appareil photo. Ce qui n'est pas possible en ajoutant des nouvelles couleurs aux mires 2D. Donc, utiliser des mires 3D est très bénéfique pour améliorer l'estimation des réponses spectrales de l'appareil grâce aux inter-réflexions. Dans l'objectif de mieux résister aux bruits d'image et de relaxer la contrainte d'éclairage connu, nous avons proposé dans le Chapitre 6 d'entraîner un réseau neurones convolutifs avec des images d'inter-réflexions simulées en utilisant notre modèle. Différents niveaux de bruits ont été ajoutés à ces images. Cette approche a démontré qu'il est possible d'estimer la réflectance spectrale de la surface et la puissance spectrale de la lumière avec une précision qui dépassent les approches de l'état de l'art.

Appendices

As seen earlier in the manuscript, the irradiance vector after n bounces of light can be written as a sum of the direct irradiance vector, and the irradiance vectors after one bounce, two bounces, three bounces,..., and n bounces:

$$\mathbf{E} = \sum_{b=0}^n (\mathbf{KR})^b \mathbf{E}_0 = \sum_{b=0}^n \mathbf{E}_b. \quad 1$$

This sum is a geometric series of the form $a + ar + ar^2 + \dots + ar^n$. Thus, Equation (1) can be written as:

$$\mathbf{E} = (\mathbf{I} - (\mathbf{KR})^n)(\mathbf{I} - \mathbf{KR})^{-1} \mathbf{E}_0. \quad 2$$

Then, when n tends to infinity, the irradiance vector is written as expressed in Equation (4.28):

$$\mathbf{E} = (\mathbf{I} - \mathbf{KR})^{-1} \mathbf{E}_0, \quad 3$$

which can be shown by proving that $\mathbf{I} - (\mathbf{KR})^\infty$ tends to the identity matrix, or that $(\mathbf{KR})^\infty$ tends to the zero matrix.

A nonnegative matrix, \mathbf{A} , is said to be *substochastic* if it satisfies the two following conditions:

- $\sum_j a_{ij} \leq 1$
- There is at least one element in each column that is strictly superior to zero.

It has been proven in [93] that the spectral radius of *substochastic* matrices is strictly less than one and that for a *substochastic* matrix \mathbf{A} :

$$\lim_{n \rightarrow \infty} \mathbf{A}^n = 0 \quad 4$$

In our case, \mathbf{R} is a square diagonal matrix containing the spectral reflectances of the different facets of the surface at a single wavelength. A surface reflectance is necessarily superior or equal to 0 and inferior or equal to 1 in case of non-florescent materials. Thus, the sum of each of its column is inferior or equal to one. However, in order to satisfy the second condition, an assumption that the surface reflectance is strictly superior to zero should be made. This assumption is met in the spectral reflectance of real surfaces. \mathbf{K} is also a square matrix, each of its columns represents the geometrical relation between a facet and all the other facets in the surface. In radiometric terms, this geometrical relation is related to the geometrical extent¹, and when integrated over the hemisphere around an infinitesimal point, this term is equal to π . Each column of \mathbf{K} represents an integration over a part of the hemisphere, and every term in this matrix is divided by π , then the sum of each column of this matrix is inferior or equal to one.

¹The geometrical extent can be also found under the name étendue

Then, both \mathbf{R} and \mathbf{K} are *substochastic* matrices, and:

$$\lim_{n \rightarrow \infty} (\mathbf{KR})^n = 0.$$

5

- [1] S. K. Nayar, K. Ikeuchi, and T. Kanade, “Shape from interreflections,” *International Journal of Computer Vision*, vol. 6, no. 3, pp. 173–195, 1991.
- [2] J. P. Parkkinen, J. Hallikainen, and T. Jaaskelainen, “Characteristic spectra of munsell colors,” *JOSA A*, vol. 6, no. 2, pp. 318–322, 1989.
- [3] P. Katemake, R. Deeb, D. Muselet, M. Hebert, and A. Tremeau, “Mutual illumination and color constancy,” in *13th AIC (International Color Association) International Congress*, 2017.
- [4] R. Deeb, D. Muselet, M. Hebert, and A. Tremeau, “Interreflections in computer vision: A survey and an introduction to spectral infinite-bounce model,” *Journal of Mathematical Imaging and Vision*, pp. 1–20, 2017.
- [5] R. Deeb, D. Muselet, M. Hebert, and A. Trémeau, “Spectral reflectance estimation from one rgb image using self-interreflections in a concave object,” *Applied Optics*, vol. 57, no. 17, pp. 4918–4929, 2018.
- [6] R. Deeb, D. Muselet, M. Hebert, A. Tremeau, J. Van de Weijer, and F. ETIENNE, “3d color charts for camera spectral sensitivity estimation,” in *British machine and vision conference (BMVC)*, 2017.
- [7] M. Hébert, R. D. Hersch, and P. Emmel, “Fundamentals of optics and radiometry for color reproduction,” *Handbook of Digital Imaging*, 2015.
- [8] F. E. Nicodemus, “Radiance,” *American Journal of Physics*, vol. 31, no. 5, pp. 368–377, 1963.
- [9] M. R. Luo, G. Cui, and B. Rigg, “The development of the cie 2000 colour-difference formula: Ciede2000,” *Color Research & Application*, vol. 26, no. 5, pp. 340–350, 2001.
- [10] S. Zeki, *A vision of the brain*. Blackwell Scientific Publ., 1993.
- [11] M. Ebner, *Color constancy*. John Wiley & Sons, 2007, vol. 6.
- [12] K. Barnard, V. Cardei, and B. Funt, “A comparison of computational color constancy algorithms. i: Methodology and experiments with synthesized data,” *IEEE transactions on Image Processing*, vol. 11, no. 9, pp. 972–984, 2002.
- [13] M. D. Fairchild, *Color appearance models*. John Wiley & Sons, 2013.
- [14] K. A. Jameson and N. Alvarado, “Differences in color naming and color salience in vietnamese and english,” *Color Research & Application: Endorsed by Inter-Society Color Council, The Colour Group (Great Britain), Canadian Society for Color, Color Science Association of Japan, Dutch Society for the Study of Color, The Swedish Colour Centre Foundation, Colour Society of Australia, Centre Français de la Couleur*, vol. 28, no. 2, pp. 113–138, 2003.

- [15] K. A. Jameson, "Culture and cognition: What is universal about the representation of color experience?" *Journal of Cognition and Culture*, vol. 5, no. 3, pp. 293–348, 2005.
- [16] A. Mojsilovic, "A computational model for color naming and describing color composition of images," *IEEE Transactions on Image processing*, vol. 14, no. 5, pp. 690–699, 2005.
- [17] M. Ikeda, Y. Mizokami, S. Nakane, and H. Shinoda, "Color appearance of a patch explained by rvsr for the conditions of various colors of room illumination and of various luminance levels of the patch," *Optical review*, vol. 9, no. 3, pp. 132–139, 2002.
- [18] G. Harding, J. M. Harris, and M. Bloj, "Learning to use illumination gradients as an unambiguous cue to three dimensional shape," *PLoS one*, vol. 7, no. 4, p. e35950, 2012.
- [19] M. G. Bloj, D. Kersten, and A. C. Hurlbert, "Perception of three-dimensional shape influences colour perception through mutual illumination," *Nature*, vol. 402, no. 6764, p. 877, 1999.
- [20] M. F. Cohen and D. P. Greenberg, "The hemi-cube: A radiosity solution for complex environments," in *ACM SIGGRAPH Computer Graphics*, vol. 19, no. 3. ACM, 1985, pp. 31–40.
- [21] H. W. Jensen, *Realistic image synthesis using photon mapping*. Ak Peters Natick, 2001, vol. 364.
- [22] M. Pharr, W. Jakob, and G. Humphreys, *Physically based rendering: From theory to implementation*. Morgan Kaufmann, 2016.
- [23] J. J. Koenderink and A. Van Doorn, "Geometrical modes as a general method to treat diffuse interreflections in radiometry," *JOSA*, vol. 73, no. 6, pp. 843–850, 1983.
- [24] S. K. Nayar and Y. Gong, "Colored interreflections and shape recovery," in *Image Understanding Workshop*, 1992, pp. 333–343.
- [25] S. M. Seitz, Y. Matsushita, and K. N. Kutulakos, "A theory of inverse light transport," in *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on*, vol. 2. IEEE, 2005, pp. 1440–1447.
- [26] M. Liao, X. Huang, and R. Yang, "Interreflection removal for photometric stereo by using spectrum-dependent albedo," in *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*. IEEE, 2011, pp. 689–696.
- [27] Y. Fu, A. Lam, Y. Matsushita, I. Sato, and Y. Sato, "Interreflection removal using fluorescence," in *European Conference on Computer Vision*. Springer, 2014, pp. 203–217.
- [28] I. Shimshoni and J. Ponce, "Recovering the shape of polyhedra using line-drawing analysis and complex reflectance models," *Computer Vision and Image Understanding*, vol. 65, no. 2, pp. 296–310, 1997.

- [29] A. J. Stewart and M. S. Langer, "Toward accurate recovery of shape from shading under diffuse lighting," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 9, pp. 1020–1025, 1997.
- [30] M. K. Chandraker, F. Kahl, and D. J. Kriegman, "Reflections on the generalized bas-relief ambiguity," in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, vol. 1. IEEE, 2005, pp. 788–795.
- [31] B. V. Funt and M. S. Drew, "Color space analysis of mutual illumination," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 15, no. 12, pp. 1319–1326, 1993.
- [32] B. V. Funt, M. S. Drew, and J. Ho, "Color constancy from mutual reflection," *International Journal of Computer Vision*, vol. 6, no. 1, pp. 5–24, 1991.
- [33] M. S. Drew and B. V. Funt, "Calculating surface reflectance using a single-bounce model of mutual reflection," in *Computer Vision, 1990. Proceedings, Third International Conference on*. IEEE, 1990, pp. 394–399.
- [34] J. Ho, B. V. Funt, and M. S. Drew, "Separating a color signal into illumination and surface reflectance components: Theory and applications," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 10, pp. 966–977, 1990.
- [35] S. A. Shafer, "Using color to separate reflection components," *Color Research & Application*, vol. 10, no. 4, pp. 210–218, 1985.
- [36] R. Bajcsy, S. Lee, and A. Leonardis, "Color image segmentation with detection of highlights and inter-reflections," in *SPIE/SPSS Symposium, Santa Clara*, 1990.
- [37] R. Bajcsy, S. W. Lee, and A. Leonardis, "Detection of diffuse and specular interface reflections and inter-reflections by color image segmentation," *International Journal of Computer Vision*, vol. 17, no. 3, pp. 241–272, 1996.
- [38] S. Tominaga and T. Okayama, "A method for analyzing color images containing highlight and interreflection effects," *Systems and computers in Japan*, vol. 28, no. 13, pp. 8–18, 1997.
- [39] S. Tominaga, K. Kato, K. Hirai, and T. Horiuchi, "Bispectral interreflection estimation of fluorescent objects," in *Color and Imaging Conference*, vol. 2015, no. 1. Society for Imaging Science and Technology, 2015, pp. 111–115.
- [40] S. Tominaga, K. Kato, K. Hirai, and T. Horiuchi, "Spectral image analysis of mutual illumination between fluorescent objects," *JOSA A*, vol. 33, no. 8, pp. 1476–1487, 2016.
- [41] R. Carroll, R. Ramamoorthi, and M. Agrawala, "Illumination decomposition for material recoloring with consistent interreflections," in *ACM Transactions on Graphics (TOG)*, vol. 30, no. 4. ACM, 2011, p. 43.

- [42] S. Beigpour and J. Van de Weijer, "Object recoloring based on intrinsic image estimation," in *2011 International Conference on Computer Vision*. IEEE, 2011, pp. 327–334.
- [43] S. K. Nayar, G. Krishnan, M. D. Grossberg, and R. Raskar, "Fast separation of direct and global components of a scene using high frequency illumination," in *ACM Transactions on Graphics (TOG)*, vol. 25, no. 3. ACM, 2006, pp. 935–944.
- [44] T. Chen, H. P. Lensch, C. Fuchs, and H.-P. Seidel, "Polarization and phase-shifting for 3d scanning of translucent objects," in *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*. IEEE, 2007, pp. 1–8.
- [45] T. Chen, H.-P. Seidel, and H. P. Lensch, "Modulated phase-shifting for 3d scanning," in *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*. IEEE, 2008, pp. 1–8.
- [46] J. Gu, T. Kobayashi, M. Gupta, and S. K. Nayar, "Multiplexed illumination for scene recovery in the presence of global illumination," in *Computer Vision (ICCV), 2011 IEEE International Conference on*. IEEE, 2011, pp. 691–698.
- [47] V. Couture, N. Martin, and S. Roy, "Unstructured light scanning to overcome inter-reflections," in *Computer Vision (ICCV), 2011 IEEE International Conference on*. IEEE, 2011, pp. 1895–1902.
- [48] M. Gupta, A. Agrawal, A. Veeraraghavan, and S. G. Narasimhan, "A practical approach to 3d scanning in the presence of interreflections, subsurface scattering and defocus," *International journal of computer vision*, vol. 102, no. 1-3, pp. 33–55, 2013.
- [49] D. Forsyth and A. Zisserman, "Mutual illumination," in *Computer Vision and Pattern Recognition, 1989. Proceedings CVPR'89., IEEE Computer Society Conference on*. IEEE, 1989, pp. 466–473.
- [50] D. A. Forsyth and A. Zisserman, "Shape from shading in the light of mutual illumination." *Image and vision computing*, vol. 8, no. 1, pp. 42–49, 1990.
- [51] D. Forsyth and A. Zisserman, "Reflections on shading," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, no. 7, pp. 671–679, 1991.
- [52] D. Saint-Pierre, R. Deeb, D. Muselet, L. Simonot, and M. Hébert, "Light interreflections and shadowing effects in a lambertian v-cavity under diffuse illumination," *Electronic Imaging*, vol. 2018, no. 8, pp. 1–10, 2018.
- [53] Y. Zhao and R. S. Berns, "Image-based spectral reflectance reconstruction using the matrix r method," *Color Research & Application*, vol. 32, no. 5, pp. 343–351, 2007.
- [54] A. Ribes and F. Schmitt, "Linear inverse problems in imaging," *IEEE Signal Processing Magazine*, vol. 25, no. 4, 2008.

- [55] V. Heikkinen, C. Cámara, T. Hirvonen, and N. Penttinen, “Spectral imaging using consumer-level devices and kernel-based regression,” *JOSA A*, vol. 33, no. 6, pp. 1095–1110, 2016.
- [56] J. Y. Hardeberg, F. J. Schmitt, and H. Brettel, “Multispectral color image capture using a liquid crystal tunable filter,” *Optical engineering*, vol. 41, no. 10, pp. 2532–2549, 2002.
- [57] G. D. Finlayson, S. D. Hordley, and P. Morovic, “Colour constancy using the chromagenic constraint,” in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, vol. 1. IEEE, 2005, pp. 1079–1086.
- [58] R. Berns, L. Taplin, and M. Nezamabadi, “Spectral imaging using a commercial colour-filter array digital camera,” *The Conservation Committee of the International Council of Museums (ICOM-CC)*, 2005.
- [59] J.-I. Park, M.-H. Lee, M. D. Grossberg, and S. K. Nayar, “Multispectral imaging using multiplexed illumination,” in *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*. IEEE, 2007, pp. 1–8.
- [60] V. Heikkinen, R. Lenz, T. Jetsu, J. Parkkinen, M. Hauta-Kasari, and T. Jääskeläinen, “Evaluation and unification of some methods for estimating reflectance spectra from rgb images,” *JOSA A*, vol. 25, no. 10, pp. 2444–2458, 2008.
- [61] C. Chi, H. Yoo, and M. Ben-Ezra, “Multi-spectral imaging by optimized wide band illumination,” *International Journal of Computer Vision*, vol. 86, no. 2-3, p. 140, 2010.
- [62] Y. Fu, A. Lam, I. Sato, T. Okabe, and Y. Sato, “Reflectance and fluorescence spectral recovery via actively lit rgb images,” *IEEE transactions on pattern analysis and machine intelligence*, vol. 38, no. 7, pp. 1313–1326, 2016.
- [63] S. Han, I. Sato, T. Okabe, and Y. Sato, “Fast spectral reflectance recovery using dlp projector,” *International journal of computer vision*, vol. 110, no. 2, pp. 172–184, 2014.
- [64] J. Jiang and J. Gu, “Recovering spectral reflectance under commonly available lighting conditions,” in *Computer Vision and Pattern Recognition Workshops (CVPRW), 2012 IEEE Computer Society Conference on*. IEEE, 2012, pp. 1–8.
- [65] R. Khan, J. Van de Weijer, D. Karatzas, and D. Muselet, “Towards multispectral data acquisition with hand-held devices,” in *Image Processing (ICIP), 2013 20th IEEE International Conference on*. IEEE, 2013, pp. 2053–2057.
- [66] M. S. Drew and B. V. Funt, “Variational approach to interreflection in color images,” *JOSA A*, vol. 9, no. 8, pp. 1255–1265, 1992.
- [67] V. Heikkinen, T. Jetsu, J. Parkkinen, M. Hauta-Kasari, T. Jaaskelainen, and S. D. Lee, “Regularized learning framework in the estimation of reflectance spectra from camera responses,” *JOSA A*, vol. 24, no. 9, pp. 2673–2683, 2007.

- [68] M. Kojima, S. Mizuno, and A. Yoshise, “A primal-dual interior point algorithm for linear programming,” in *Progress in mathematical programming*. Springer, 1989, pp. 29–47.
- [69] R. H. Byrd, M. E. Hribar, and J. Nocedal, “An interior point algorithm for large-scale nonlinear programming,” *SIAM Journal on Optimization*, vol. 9, no. 4, pp. 877–900, 1999.
- [70] J. Hernández-Andrés, J. Romero, J. L. Nieves, and R. L. Lee, “Color and spectral analysis of daylight in southern europe,” *JOSA A*, vol. 18, no. 6, pp. 1325–1335, 2001.
- [71] F. H. Imai, M. R. Rosen, and R. S. Berns, “Comparative study of metrics for spectral match quality,” in *Conference on Colour in Graphics, Imaging, and Vision*, vol. 2002, no. 1. Society for Imaging Science and Technology, 2002, pp. 492–496.
- [72] J. Viggiano, “Metrics for evaluating spectral matches: a quantitative comparison,” in *Conference on Colour in Graphics, Imaging, and Vision*, vol. 2004, no. 1. Society for Imaging Science and Technology, 2004, pp. 286–291.
- [73] N. Ohta and A. Robertson, *Colorimetry: fundamentals and applications*. John Wiley & Sons, 2006.
- [74] R. A. Willoughby, “Solutions of ill-posed problems (a. n. tikhonov and v. y. arsenin),” *SIAM Review*, vol. 21, no. 2, pp. 266–267, 1979.
- [75] M. M. Darrodi, G. Finlayson, T. Goodman, and M. Mackiewicz, “Reference data set for camera spectral sensitivity estimation,” *J. Opt. Soc. Am. A*, vol. 32, no. 3, pp. 381–391, Mar 2015.
- [76] G. Wahba, *Spline models for observational data*. Siam, 1990, vol. 59.
- [77] T. Poggio and F. Girosi, “Networks for approximation and learning,” *Proceedings of the IEEE*, vol. 78, no. 9, pp. 1481–1497, 1990.
- [78] T. Evgeniou, M. Pontil, and T. Poggio, “Regularization networks and support vector machines,” *Advances in computational mathematics*, vol. 13, no. 1, p. 1, 2000.
- [79] V. Heikkinen, A. Mirhashemi, and J. Alho, “Link functions and matérn kernel in the estimation of reflectance spectra from rgb responses,” *JOSA A*, vol. 30, no. 11, pp. 2444–2454, 2013.
- [80] B. Arad and O. Ben-Shahar, “Sparse recovery of hyperspectral signal from natural rgb images,” in *European Conference on Computer Vision (ECCV)*, 2016.
- [81] R. M. H. Nguyen, D. K. Prasad, and M. S. Brown, “Training-based spectral reconstruction from a single rgb image,” in *European Conference on Computer Vision (ECCV)*. Springer International Publishing, 2014, pp. 186–201.

- [82] J. Aeschbacher, J. Wu, and R. Timofte, “In defense of shallow learned spectral reconstruction from rgb images,” in *The IEEE International Conference on Computer Vision (ICCV) Workshops*, Oct 2017.
- [83] A. Alvarez-Gila, J. van de Weijer, and E. Garrote, “Adversarial networks for spatial context-aware spectral image reconstruction from RGB,” *CoRR*, vol. abs/1709.00265, 2017. [Online]. Available: <http://arxiv.org/abs/1709.00265>
- [84] D. Fourure, R. Emonet, E. Fromont, D. Muselet, A. Trémeau, and C. Wolf, “Mixed pooling neural networks for color constancy,” in *IEEE International Conference on Image Processing (ICIP)*, 2016.
- [85] W. Shi, C. C. Loy, and X. Tang, “Deep specialized network for illuminant estimation,” in *European Conference on Computer Vision (ECCV)*, 2016, pp. 371–387.
- [86] Y. Hu, B. Wang, and S. Lin, “Fc4: Fully convolutional color constancy with confidence-weighted pooling,” in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 330–339.
- [87] T. Narihira, M. Maire, and S. X. Yu, “Learning lightness from human judgment on relative reflectance,” in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 2965–2973.
- [88] J. Shi, Y. Dong, H. Su, and S. X. Yu, “Learning non-lambertian object intrinsics across shapenet categories,” in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 5844–5853.
- [89] M. Janner, J. Wu, T. Kulkarni, I. Yildirim, and J. B. Tenenbaum, “Self-Supervised Intrinsic Image Decomposition,” in *Advances In Neural Information Processing Systems (NIPS)*, 2017.
- [90] K. Rematas, T. Ritschel, M. Fritz, E. Gavves, and T. Tuytelaars, “Deep reflectance maps,” in *Computer Vision and Pattern Recognition (CVPR), 2016 IEEE Conference on*. IEEE, 2016, pp. 4508–4516.
- [91] S. Georgoulis, K. Rematas, T. Ritschel, M. Fritz, T. Tuytelaars, and L. Van Gool, “What is around the camera?” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 5170–5178.
- [92] X. Glorot and Y. Bengio, “Understanding the difficulty of training deep feedforward neural networks,” in *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*, 2010, pp. 249–256.
- [93] M. Hébert, “Compositional model for predicting multilayer reflectances and transmittances in color reproduction,” Ph.D. dissertation, Ecole Polytechnique Fédérale de Lausanne, 2006.

