UNIVERSITÀ DEGLI STUDI DI PADOVA

Scuola di Dottorato in
Ingegneria dell'Informazione
XX Ciclo

Indirizzo Elettronica e Telecomunicazioni

Acquisition and processing of multispectral data for texturing 3D models

Anna Paviotti

Il Direttore	**Il Tutore**
Prof. Matteo Bertocco	Prof. Guido Maria Cortelazzo

AA. AA. 2005–2008

Abstract

This thesis deals with three problems concerning the use of a multispectral imaging spectrograph in applications of cultural heritage. The multispectral camera is part of an instrument developed within the project “Shape&Color” (CARIPARO, 2003-2005), coupling the spectrograph with a 3D laser scanner. Although the issues we have addressed arose from the characteristics of this specific instrument, they can be regarded as general problems concerning multispectral imaging, and are therefore of broader interest.

The first part relates on the characterization of the spectrograph performance in measuring spectral reflectance under different illumination conditions. Four different illumination setups have been used to acquire a set of colored calibrated tiles. The system performance has been evaluated through a metrologically-inspired procedure, using as descriptors the average error (AE) and the average error standard deviation (AESTD), calculated by means of error propagation formula. The best results have been obtained with a metallic iodide lamp and an incandescence lamp used in a sequence, juxtaposing the spectral reflectance measured with the metallic iodide lamp in the 400-600 nm interval and that obtained with the halogen lamp in the 600-900 nm interval.

The second presented issue concerns the problem of separating spectral illumination and spectral reflectance from the acquired color signal (the global radiation signal reflected by a target object). Since the latter can be considered as the product of illumination and spectral reflectance, this is an ill-posed problem. Methods in the literature estimate the two functions apart from a scale factor. The proposed solution attempts at the recovery of this scale factor using a statistical-based approach. The core of the algorithm consists of the estimation of the illumination intensity through a modification of the RANSAC algorithm, using relations derived from the physical constraints of the illumination and the spectral reflectance. The spectral reflectance is subsequently computed from the measured color signal and the estimated illumination function. The algorithm has been tested on four case studies, representing artworks of different pictorial techniques, color characteristics and dimensions. The results are good in terms of mean relative error, while the infinity norm of the relative error sometimes assumes high values.

The last problem we have dealt with is that of using the multispectral images acquired with the Shape&Color scanner to texturize uncalibrated 3D data. What makes the problem worth addressing is that the spectral camera is not pinhole, but can be classified as a cylindrical panoramic camera. In this thesis, the general problem of estimating the extrinsic parameters of the camera from a known set of 3D-2D correspondences has been considered. The chosen approach is the classical reprojection error minimization procedure. As the projection operator is nonlin-

ear, the objective function has a very complicated structure. Due to this and to the high dimensionality of the problem, the minimization results are strongly sensitive to the choice of the initial parameter values. This work proposes a way of finding a reliable initial point for the minimization function, so as to lower the risk of being trapped into local minima.

Sommario

La presente tesi affronta tre problematiche relative all'uso di uno spettrografo per l'acquisizione di immagini multispettrali in applicazioni di beni culturali. La camera multispettrale fa parte di uno strumento sviluppato all'interno del Progetto "Forma&Colore" (CARIPARO, 2003-2005), che, oltre allo spettrografo, comprende un laser scanner a tempo di volo. Sebbene le problematiche affrontate siano sorte dalle specifiche caratteristiche di questo strumento, esse possono tuttavia essere considerate come problemi più generali relativi all'imaging multispettrale, e rivestono pertanto un interesse più ampio.

La prima parte della tesi riguarda la caratterizzazione delle prestazioni dello spettrografo nel misurare riflettanze spettrali al variare delle sorgenti di illuminazione. Un set di tavolette calibrate è stato acquisito usando quattro diversi setup di illuminazione. Le prestazioni della ricostruzione della riflettanza sono state valutate con un procedimento metrologico, usando come descrittori l'errore medio e la sua deviazione standard, calcolata tramite formule di propagazione degli errori. I migliori risultati sono stati ottenuti usando in sequenza una lampada a ioduri metallici e una lampada ad incandescenza, ovvero giustapponendo le riflettanze misurate con la lampada a ioduri metallici da 400 a 600 nm e quelle misurate con la lampada ad incandescenza tra 600 e 900 nm.

Il secondo argomento presentato riguarda il problema di separare i contributi di illuminazione e riflettanza spettrale dal segnale di colore (la radiazione globale riflessa dall'oggetto acquisito). Essendo il segnale di colore dato dal prodotto di illuminazione e riflettanza, quello affrontato è un problema mal posto. I metodi presenti in letteratura separano i due contributi a meno di un fattore di scala. La soluzione proposta in questa tesi stima il fattore di scala tramite un metodo statistico. Il cuore dell'algoritmo consiste nella stima dell'intensità dell'illuminazione tramite una modifica dell'algoritmo RANSAC, applicato a relazioni derivate dai vincoli fisici a cui sono soggette illuminazione e riflettanza. La riflettanza spettrale viene successivamente calcolata come rapporto tra il segnale di colore misurato e l'illuminazione stimata. L'algoritmo è stato testato su quattro casi di studio nell'ambito dei beni culturali, un affresco e tre dipinti. I risultati sono buoni in termini di errore relativo medio, mentre la norma infinito dell'errore relativo assume talvolta valori elevati.

L'ultima tematica affrontata riguarda l'utilizzo delle immagini multispettrali per texturizzare modelli 3D non calibrati. L'interesse del problema risiede nel fatto che la camera multispettrale non risponde al modello pinhole, ma va classificata come una camera panoramica cilindrica. In questa tesi, si è affrontato il problema generale di stimare i parametri estrinseci della camera a partire da corrispondenze 3D-2D. Il problema è stato formulato in maniera classica come quello di minimizzare l'errore di riproiezione tra le corrispondenze. Poichè l'operatore

proiezione è non lineare, la funzione obiettivo risultante ha una struttura particolarmente complessa. L'alta dimensionalità del problema comporta pertanto che il risultato della minimizzazione sia estremamente sensibile alla scelta del valore iniziale dei parametri. Il lavoro proposto consiste nella ricerca di valori iniziali dei parametri che riducano la probabilità di rimanere intrappolati in un minimo locale.

A Nicola e Giacomo

Acknowledgements

First of all, I would like to thank my tutor, Prof. Cortelazzo, for giving me the opportunity of doing my PhD in the LTTM group, and for his decisive support during critical moments of my professional and personal life. I would also like to thank Dr. Luca Poletto, of the LUXOR group, for his material and intellectual support of my research and for his stimulating collaboration during these past years. I am also deeply grateful to all the LUXOR collaborators who have shared with me long and sometimes tiresome hours of multispectral acquisitions: Tommaso Fincato, Matteo Caldon, Inglis Modolin, Filippo Ratti and Carlo Tona. I am also much obliged to Dr. Fabio Remondino for giving us the chance of acquiring parts of the Castello del Buonconsiglio in Trento, and to Elena Rebollo San Miguel, Barbara Dal Bianco and Prof. Renzo Bertoncello for their precious insights into the chemical analysis of works of art. I also wish to thank Pietro and Luca for sharing with me ups and downs during these years, and Valeria, Matteo, Maurizio, Enrico, Gabriele, Nicola and Emanuele for livening up my life in the lab.

I am deeply grateful for the material and spiritual support my parents and sister have given me, and I would especially like to thank my mother, who commuted between Udine and Padova for one year to babysit. Lastly, all my love and thankfulness go to Nicola, without whom I would never have survived this PhD, and to Giacomo, who welcomed his mother home with a smile even when she dedicated most of her time to far less important matters than him.

Contents

List of Figures

List of Tables

List of Algorithms

Introduction

In the last fifteen years, multispectral imaging has gained increasing importance in the Signal Processing community dealing with cultural heritage applications [1]. A first reason is that acquiring the spectral reflectance of an object is the most reliable asset for faithful color reproduction, where faithfulness is measured in terms of independence from illumination and acquisition devices. Indeed, traditional (R,G,B) representation depends both on the acquisition device and the environmental conditions [2]. Independence from the former can be achieved by means of device-independent color spaces but, in any case, the acquisition will be influenced by the environment, mainly by lighting [3]. For this reason, a number of cultural heritage institutions have carried out the construction of "digital museums" [3], i.e. the digitization of their collections, using multispectral sensors. A second advantage of multispectral measurements is that they permit monitoring of the conservation status of paintings. Indeed, multispectral color acquisition is objective, hence repeatable and furthermore, digital images are virtually eternal, since they do not degrade over time. A third important consideration is that the non-visible parts of the spectrum can render valuable service both in the non-invasive diagnosis of the conservation status of a work of art and in the reconstruction of its history. For example, infrared (IR) reflectography has been used for some decades for the detection of underdrawings in paintings, while ultraviolet (UV) fluorescence highlights former restoration interventions. Spectral signatures of color pigments used in the painting process can also be found in the non-visible spectrum. Unobtrusive diagnosis can therefore lead to virtual restoration planning. The VASARI, MARC and CRISATEL projects were pioneering projects funded by the European Commission which made use of multispectral data for the acquisition and monitoring of paintings [4–21].

Despite its many advantages, multispectral imaging poses a number of technical challenges which have been only partially overcome. First of all, how many and which bands one needs to acquire is still the subject of debate [3, 22] and, clearly, depends on the application. IR reflectography calls for different illumination sources and sensors than visible spectroscopy or UV fluorescence. When selecting the appropriate spectral window for a chosen cultural heritage application, non-invasiveness issues must also be kept in mind, as it is notorious that some illumination sources can damage the pictorial layer. Achieving independence from the illumination is another non-trivial task. Multispectral imaging sensors typically measure the *color signal*, that is, the global radiation reflected by the object. This radiation is the product of very complex interactions between the illumination source light and the object's surface reflectance properties. The spectral composition of the illumination, the orientation of the target surface with respect to the illumination source and the observer, the surface material and color and the characteristics of the sensing device are all factors influencing the composition of the color signal. Illumination independence in multispectral imaging is generally achieved by subtracting a *reference white signal* from the color signal. The reference white signal is mostly obtained by acquiring a reference white surface in the same position, and subject to the same illumination, as the target object. It is worth noticing that the resulting spectral reflectance is indeed independent of the spectral content of the illumination, but not from the mutual positions of illumination, surface and observer.

All these technical challenges are amplified when dealing with multispectral imaging for texturing 3D models. Extending multispectral imaging to 3D modeling with multispectral textures seems rather natural. Indeed, all the above mentioned advantages of multispectral acquisition can be imagined to be profitably applied to inherently three-dimensional targets such as frescoed chapels, statues, ornamented vases or ancient situlas. However, this extension to 3D poses more challenging problems. As previously mentioned, multispectral acquisition requires accurate control of illumination and of sensor setup in order to recover the spectral reflectance from the raw data. The dependence of the color signal on the geometry of the object, the illumination position and the point of view is much more relevant when dealing with three-dimensional scenes. When acquiring a painting, the multispectral sensor is usually kept normal to the painting surface and acquisition is performed on small areas under uniform illumination [23]. These favorable conditions cannot always be reproduced when acquiring three-dimensional objects of complex geometry. While controlling acquisition parameters to such a degree is still possible when handling small manufactures, it becomes really difficult when dealing, for example, with medium-sized architectural volumes such as chapels or ancient buildings.

This thesis deals with three problems related to the use of a 3D scanner coupling a time-of-flight laser scanner and a multispectral imaging sensor. The instrument was built within the project “Forma e Colore” (Shape&Color) [24], funded by CARIPARO, in 2003-2005. The ambitious aim of the project was that of designing an instrument capable of acquiring 3D models of small architectural volumes with multispectral textures for monitoring and conservation purposes. The resulting scanner is one of the only two existing instruments (the other is described in [25]) simultaneously acquiring 3D shape and multispectral textures. While using the instrument in real-world cultural heritage applications, several challenges have been revealed. This thesis presents and attempts a solution to three of these problems. Although they are clearly related to the specific features of the Shape&Color scanner, they can also be regarded as general problems concerning multispectral imaging, and are therefore of broader interest.

The first presented problem concerns the choice of the best illumination setup for faithfully acquiring spectral reflectances. As the nominal spectral window of the imaging spectrograph is 400-1000 nm, an ideal illumination source should have a flat spectrum across the whole of this band. Unfortunately, no such illumination source exists. Three different lamps have therefore been used, alone or combined, to find the best reconstruction results. The reconstruction performance has been evaluated through a metrologically-inspired approach. A set of calibrated colored tiles and a reference white tile have been acquired with multiple independent measurements. The error has been defined as the difference between the tile tabulated spectra and the recovered spectral reflectances, considered as a function of wavelength. The average mean error (the mean error over different measurements averaged over the different tiles) has been one of the parameters used to evaluate the reconstruction performance. The other parameter has been chosen as the average error uncertainty (standard deviation), calculated from the statistics of the input measured signals through error propagation formula.

The second presented issue concerns the problem of separating spectral illumination and spectral reflectance directly from the color signal, without acquiring a reference white signal. In a simplified framework, the color signal can be considered as the product between spectral illumination and spectral reflectance. Separating the two components is therefore an ill-posed problem. The illumination-reflectance separation problem has been studied in the literature, often as a means of achieving color constancy. However, the proposed solutions avoid estimating the scale factor of illumination and reflectance, and thus ultimately fail to solve the ill-posed separation problem. In the presented thesis, a solution to estimate this scale factor is proposed. The presented illumination-reflectance separation problem is a simplified version of the most general case, as we can expect the illumination spectral content to be constant over space. Therefore, the illumination estimation coincides with the recovery of the illumination scale factor (intensity). The proposed solution relies on a probabilistic approach, and exploits the long-time known observation that illumination and spectral reflectance have a different behavior over space. Specifically, the illumination is a slowly-varying signal, while reflectance presents abrupt variations at the joints of different color patches. Our algorithm starts by estimating the illumination intensity through a modification of the RANSAC (Random Sample Consensus) algorithm [26], applied to a relation linking the maximum (over wavelength) of the color signal with the product of the illumination coefficient and the maximum (over wavelength) of the spectral reflectance. Spectral reflectance is subsequently recovered from the measured color signal and the estimated illumination function. The estimated illumination and spectral reflectance satisfy their physical constraints, without the need of resorting to constrained estimation techniques. The algorithm has been tested on real-world data of cultural heritage applications.

The last addressed problem is that of using the multispectral images acquired with the Shape&Color scanner to texturize uncalibrated 3D models. What makes the problem worth addressing is that the spectral camera is not pinhole, but can be classified as a *cylindrical panoramic camera* [27]. This camera model has been extensively considered in the literature, mostly as concerns stereo reconstruction from panorama. An application to texture mapping has also been considered, but the proposed approach is closely dependent on the specific application [28]. In this thesis, the general problem of estimating the extrinsic parameters of the camera from a known set of 3D-2D correspondences has been considered. The chosen approach is the classical reprojection error minimization procedure. As the projection operator is nonlinear, the objective function has a very complicated structure. Due to this and to the high dimensionality of the problem, the minimization results are strongly sensitive to the choice of the initial parameter values. This thesis proposes a way of finding reliable initial parameter values for the minimization function, so as to lower the risk of being trapped into local minima.

The thesis is organized as follows. Chapter 1 describes the Shape&Color scanner, focusing on the strengths and weaknesses which were most relevant to inspire the following work. Chapter 2 relates on the characterization of the performance of the imaging spectrograph under different illumination sources. Chapter 3 describes the proposed illumination-reflectance separation algorithm, the results of which are presented in Chapter 4. Chapter 5 presents the extrinsic parameter calibration problem for the multispectral camera. Lastly, some conclusions are drawn.

Chapter 1

Shape&Color: a 3D scanner acquiring multispectral texture information

This chapter describes the Shape&Color scanner, an instrument for the acquisition of 3D data and multispectral textures of small architectural volumes. The first prototype of the instrument was developed as part of the project "Forma & Colore"(Shape&Color), funded by CARIPARO, in 2003-2005. Some hardware improvements have been recently performed [29]. The work of this PhD thesis originated by some issues arisen while using the scanner in real-world applications. Therefore, we will review the instrument characteristics, focusing on its strengths and weaknesses for its target application.

1.1 State-of-the art instrumentation for the acquisition of cultural heritage artifacts

A major point of interest of the presented system is that at present there are no commercial range finders for the joint acquisition of geometry and multispectral textures for volumes of this size. To the authors' knowledge, the only other existing instrument acquiring 3D shape and multispectral textures is that developed by Mansouri *et al.* [25]. The instrument is a 3D multispectral scanner for the acquisition of small objects such as pottery or small statues. It couples an LCD projector with a camera equipped with a rotating wheel containing seven interference filters and a free hole. The filters span the visible spectrum (380 nm to 775 nm), while the free hole is used to acquire a monochromatic image of the object. The system is spectrally calibrated by recovering the correspondence between the seven-valued vector of the filter outputs and the eighty-valued vector of the current point spectral reflectance. The learning stage consists of the acquisition of a series of known reflectances and the inverse problem of estimating the input spectra from the camera output is solved by means of neural networks. In this case, the independence from illumination and device is achieved only in a pre-defined working volume and, each time the acquisition setup is changed, the calibration procedure has to be repeated. The acquisition volume used to illustrate the performance of the system was about 50 cm × 50 cm × 20 cm. This system was used for visualization purposes in order to acquire faithfully textured 3D models of a jug and of some small statues and to relight them with various illuminants in a virtual museum application.

As for other state-of-the art acquisition systems, there exist commercial range finders which acquire shape and trichromatic color data of small volumes, with side up to a couple of meters. There also exist range finders which operate at very large distances for topographic [30] or photogrammetric applications [31], but these instruments only acquire geometric information. There are no commercial products targeting small architecture volumes, i.e., of side 3-10 meters, although such dimensions are of interest for many architectural environments. For instance, in [32–34] the authors report on a robotized system for the automatic 3D modeling of typical industrial architecture. Commercial devices for 3D modeling small volumes are also available [35–38].

Several laser scanning methods are available for distance and displacement measurements [39]. For medium distance ranges, the most common techniques are geometric triangulation and time-of-flight laser scanning. Geometric triangulation allows for good accuracy with simple and inexpensive implementations. Any geometric triangulation system is based on the conversion of an angular measurement into the desired distance information. Their performance is therefore determined by the angular measurement absolute error, and the distance measurement error increases with the target distance. As a consequence, while the relative error grows linearly, the absolute error follows a square law [40].

Time-of-flight methods measure the time a light pulse emitted by a laser takes to hit the target and return back to the receiver. If the target is at distance L from the instrument and if ΔT denotes the time difference between transmitted and received pulses, we have $L = c \cdot \Delta T/2$, where c denotes the speed of light. Therefore, the absolute error of the distance measurement is constant. This technique allows to achieve good and constant accuracy over a wide range of distances [41]. These considerations concern the basic implementation schemes and neglect the influence of propagation trough the atmosphere.

In recent years, the acquisition of multispectral images of paintings has become quite popular in the imaging community dealing with cultural heritage applications [1]. A first reason is that acquiring the spectral reflectance is the most reliable asset for faithful color reproduction, where faithfulness is measured in terms of independence from illumination and acquisition devices. Indeed, (R,G,B) representation depends both on the acquisition device and the environmental conditions [2]. Independence from the former can be achieved by means of device-independent color spaces but, in any case, the acquisition will be influenced by the environment, mainly by lighting [3]. For this reason, a number of cultural heritage institutions have carried out the construction of “digital museums” [3], i.e. the digitization of their collections, using multispectral sensors. A second advantage of multispectral measurements is that they permit monitoring of the conservation status of paintings. Indeed, multispectral color acquisition is objective, hence repeatable and furthermore, digital images are virtually eternal, since they do not degrade over time. A third important consideration is that the non-visible parts of the spectrum can render valuable service both in the non-invasive diagnosis of the conservation status of a work of art and in the reconstruction of its history. For example, infrared (IR) reflectography has been used for some decades for the detection of underdrawings, while ultraviolet (UV) fluorescence highlights former restoration interventions. Spectral signatures of color pigments used in the painting process can also be found in the non-visible spectrum. Unobtrusive diagnosis can therefore lead to virtual restoration planning. The VASARI, MARC and CRISATEL projects were pioneering projects funded by the European Commission which made use of multispectral data for the acquisition and monitoring of paintings [4–21].

VASARI and MARC considered the multispectral acquisition of paintings in order to derive reliable CIELAB coordinates [6, 9].

A remarkable multispectral imaging project involving the CRISATEL scanner is the "Mona Lisa project". In 2004, an unprecedented number of techniques were applied to the analysis of the Mona Lisa [42]. The methodologies adopted included radiography, X ray fluorescence, Raman spectrometry, digital photography under raking light, infrared reflectography, multispectral imaging and 3D reconstruction. The study was aimed at assessing and recording the conservation status of the painting and at gaining an insight into Leonardo's pictorial technique. Multispectral digitization was performed by the French Center of Museum Research with the camera developed within the CRISATEL project [43]. This multispectral camera comprises 13 filters with a bandwidth of 40 nm spanning the spectrum from 380 nm to 1050 nm. Multispectral measurements were used to identify the various colors of the painting through luminance and color segmentation, to relight the painting with different illuminants and to identify restorations from the most superficial layers (near UV) to the deepest ones (near IR).

Generally speaking, two different non-invasive reflectance spectroscopy techniques are currently employed for cultural heritage applications: Fiber Optics Reflectance Spectroscopy (FORS) and Imaging Spectroscopy (IS) [44]. The former uses fiber optics to convey light (typically exiting a monochromator) to the target object. Reflectance information is recovered in two possible ways: when a dual-beam approach is used, the incident light beam is split before being directed onto the object, and the unreflected beam is used as a reference to calculate the reflection coefficient. In the single-beam approach, the absolute reflected light intensity is measured. In this case, reflectance can be calculated after acquiring a reference white object. Dual-beam fiber optics spectrometers are very accurate but not easily portable. On the contrary, single-beam fiber optics spectrometers are more compact and easily transportable, but can lead to greater measurement errors.

FORS instrumentation operates in a point-wise fashion. On the contrary, IS determines spectral reflectance data for each pixel in a spatial image. Imaging spectrometers can be divided into two classes according to the way multispectral information is recovered. A first approach consists in putting a set of filters in front of the detector [4, 43, 45–47]. The number of filters used in state-of-the-art solutions varies from 7 to 32, and the spectral resolution from 10 nm (narrow-band filters) to 40 nm (wide-band filters) [1]. Another approach consists in using a dispersive element to separate the different light components, which are then detected by a sensor (typically a CCD). In this case, higher spectral resolution (<1 nm) can be achieved, but the received signal intensity is considerably lower than in the filter-based approach.

Imaging spectroscopy allows the recovery of the spectral reflectance of the whole target object (or parts of it). Therefore, not only a faithful color reproduction of the object can be achieved, but also more interesting tasks, such as the comparison of the spectral content of different parts of the object or image segmentation based on spectral features, can be performed. IS methods are generally less accurate than FORS ones, but allow a more exhaustive inspection of the spectral content of the target. It could be claimed that IS methods are sufficiently accurate if they allow the detection of significant spectral features in certain spots of an object, to which more accurate analysis can be restricted. In this logic, IS and FORS can be considered as complementary. The reliability of the spectral content analysis clearly depends on the accuracy of the spectral reflectance reconstruction, which in turn is determined by several factors, such as the instrument's spectral resolution, the measurement noise and the choice of the illumination source. The Shape&Color scanner features an imaging spectrograph by SPECIM, the Imspector V10 (400-1000 nm).

Extending multispectral imaging to 3D modeling with multispectral textures seems rather natural. Indeed, all the above mentioned advantages of multispectral acquisition can be imagined to be profitably applied to inherently three-dimensional targets such as frescoed chapels, statues, ornamented vases or ancient situlas. However, this extension to 3D poses some challenging problems. Multispectral acquisition requires accurate control of illumination and of sensor setup in order to recover the spectral reflectance from the raw data. Indeed, the light reflected by an object depends on the geometry of the object, on the materials it is made of and on the light illuminating it. The situation is further complicated by the fact that the light reflected by the object varies with both the illumination angle and the point of view. When acquiring a painting, the multispectral sensor is usually kept normal to the painting surface and acquisition is performed on small areas under uniform illumination [23]. These favorable conditions cannot always be reproduced when acquiring three-dimensional objects of complex geometry. While controlling acquisition parameters to such a degree is still possible when handling small manufactures, it becomes really difficult when dealing, as in our case, with medium-sized architectural volumes such as chapels or ancient buildings.

The acquisition instrument we present fills a current instrumental gap, since it collects geometry information with millimetric precision in the volume range of 3-10 meter side by a time-of-flight range finder (Section 1.3), and spectral information in the range 400-1000 nanometers by an imaging spectrograph integrated with the range finder. The imaging spectrograph will be presented in Section 1.2, and further discussed in Chapter 2. The complete instrument and the mutual calibration of the two scanners will be described in Section 1.4.

The 3D modeling procedure will be presented in Section 1.5. The ambitious aim of the Shape&Color project was that of providing the instrument with a fully automated 3D modeling procedure, so as to allow repeatable acquisitions (and hence monitoring over time), and ease of use for non-technical personnel.

Some experimental results are presented in Section 1.6, while Section 1.7 discusses some strengths and weaknesses of the instrument.

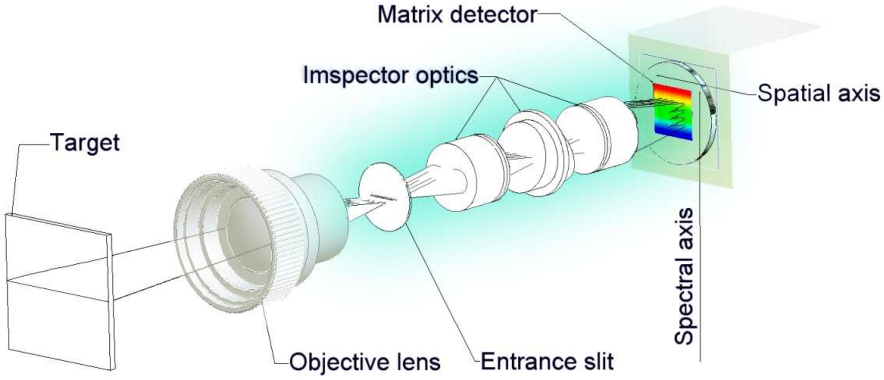


Figure 1.1: The SPECIM Imspector V10 spectrograph.

1.2 The imaging spectrograph

As mentioned in section 1.1, the imaging spectrograph is the SPECIM Imspector V10. ImSpector is a direct sight imaging spectrograph provided with a dispersive element that can be quickly combined with a broad range of industrial and scientific monochrome area cameras to form a spectral camera. Compared to conventional color cameras and other filter-based imaging systems, ImSpector produces full contiguous spectral information with high quality spectral and spatial resolution. It can cover a broad spectral range over which it enables flexible wavelength selections via software.

A schematic drawing of the imaging spectrograph is shown in Fig. 1.1. The objective lens focuses the image of the target to be acquired on the plane of the input slit of the spectrograph. The light coming from a rectangular narrow strip conjugated with the slit enters into the spectrograph, is dispersed by a dispersive element (prism or grating) and focused on the plane of the 2D detector. With reference to Fig. 1.1, the horizontal axis of the camera is the spatial axis, while the vertical axis is the spectral axis. The light coming from the same point on the target but with different wavelengths is focused on the sensor on the same column (i.e. the same spatial position) but in different rows (i.e. different spectral positions). A 2D spectral image can be recovered using a series of monochromatic images of a 2D region on the target obtained by scanning it in the direction perpendicular to the slit [48]. The choice of the objective focal length is imposed by the required field-of-view (FOV) in the direction parallel to the slit. Several objectives are currently available. We normally use a Canon f=16 mm, f/# 1.4, that matches the f/number of the spectrograph, or a Canon f=25 mm, f/# 1.4 objective. The nominal spectral window of the spectrograph is 400 to 1000 nm. The entrance slit size is 9.8 mm × 25 μm. The sensor is the Hamamatsu C8484-05G. It is a progressive scan interline CCD with micro-lenses, 1024 (spectral) × 1344 (spatial) pixel format, 6.45 μm × 6.45 μm pixel size, 6.6 mm × 8.7 mm active area. The CCD is moderately cooled to reduce the thermal noise and increase the sensitivity. The camera has 10 e-rms readout noise, dynamic range of 1800:1 and quantum efficiency in the 0.35-0.70 range in the 400-750 nm spectral interval. Its performances are definitely higher than a conventional TV camera, as that used

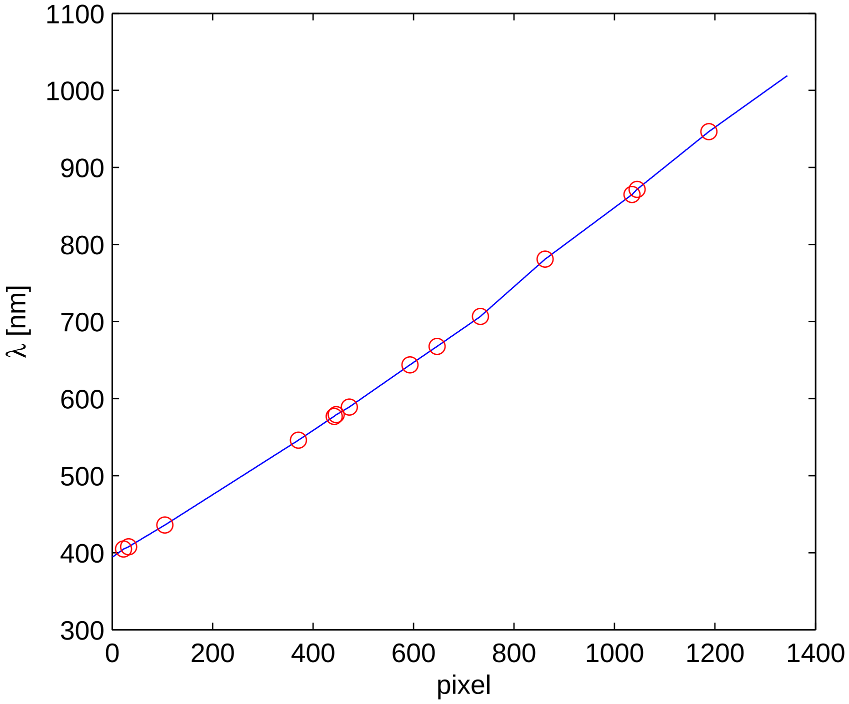


Figure 1.2: Spectral calibration curve.

in the first prototype [24]. The FOV in the direction parallel to the slit is limited by the number of pixels in the spatial direction. A spatial sampling of 2 mm gives a FOV of 2.7 m at a distance of about 5 m with the f=16 mm objective. The scanning in the direction perpendicular to the slit is performed by a rotation of the spectrograph by means of a rotating stage (Physik Instrumente M-062). To illuminate the target, several lamps can be chosen. This aspect will be further discussed in Chapter 2. The spectral reflectance of the target for any pixel $R_{m,n}$ is calculated by

$$R_{m,n} = \frac{S_{m,n} - D_{m,n}}{W_{m,n} - D_{m,n}} \tag{1.1}$$

where m, n are the coordinates of the pixel within the sensor, S_{mn} is the actual signal, D_{mn} is the dark signal of the camera and W_{mn} is the standard white signal.

1.2.1 Spectral calibration

The spectral calibration of an imaging spectrograph consists in determining the correspondence between pixel indexes and wavelengths. The calibration was performed by measuring the spectra of low pressure gas lamps which emit very narrow spectral lines characteristic of the gas in the lamp. The correspondence between pixels and wavelengths has been found to be nearly linear (see Fig. 1.2). The spectral dispersion is 72.1 nm/mm and the spectral resolving element is 0.465 nm/pixel.

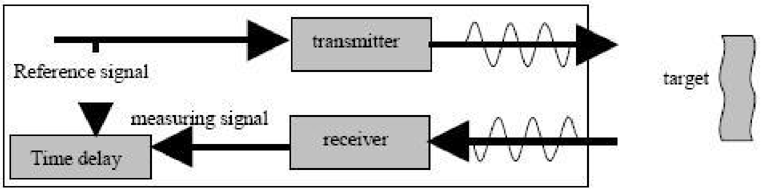


Figure 1.3: The time-of-flight method.

1.3 The Laser scanner

As mentioned in Section 1.1, the most used techniques in the field of medium range laser scanning are geometric triangulation and time-of-flight scanning. Geometric triangulation allows high accuracy for small distances, but the measurement error increases with distance. On the contrary, time-of-flight laser scanning exhibit a measurement error independent of the target distance. For these reasons, triangulation-based scanners are generally used to scan relatively small objects, such as handmade objects, ceramics and statues, with high precision. On the contrary, time-of-flight techniques are widely used for large volumes, such as monuments, buildings and walls.

Given the large range of operation of the instrument (2-20 m), the large area to be scanned (up to $10\,\mathrm{m}^2$ in a single scan) and the constant accuracy over the whole volume (~ 1 mm independently of target distance), we chose to use a time-of-flight system. We recall that time-of-flight distance measurement is based on the measure of the time that a light pulse emitted by a laser takes to hit the target and return back to the receiver (see Fig. 1.3). Namely, if the target is at distance L from the instrument and T denotes the time difference between transmitted and received pulse, it is $L = c \cdot T/2$, where c denotes the speed of light. In this way, the absolute error of the distance measurement is constant. The laser scanner perform a distance calculation on a single point, so the laser beam has to be deflected by a suitable system to scan the whole area to be acquired. This is usually performed by rotating mirrors. In particular, our scanner has to acquire shape data on the same area that is simultaneously acquired by the spectrograph, i.e. a narrow strip parallel to the entrance slit. A single rotating mirror performs such operation.

The core of the scanner is the Noptel CM2-30 rangefinder. The pulsing rate can be up to 6 kHz depending on the device version, baud rate and processor capacity. A 4 KHz pulsing rate has been commonly employed in on-field applications. The single-shot precision of the range-finder is 520 mm rms, which is too bad for our purposes. Precision has therefore been improved by averaging several measurements. The plot of precision vs. number of averaged measurements is shown in Fig. 1.4 for a white target placed at a distance of 3.5 m. These measurements have been repeated for different distances and we observed that the average over 10-20 points gives a precision better than 2 mm rms.

Furthermore, a set of measurements has been carried out in order to quantify the resolution of the range-finder, i.e. the minimum displacement that is detectable. A white target has been placed on a motorized translation stage and moved at known distances from the range-finder. For each distance, the target

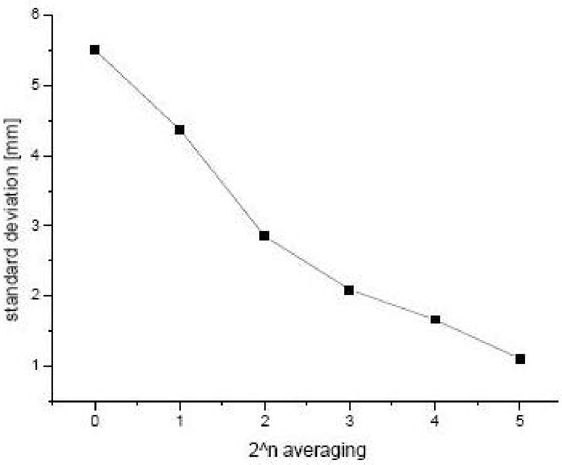


Figure 1.4: Precision vs. number of averaged measurements. The target is a white tile placed at a distance of 3.5 m.

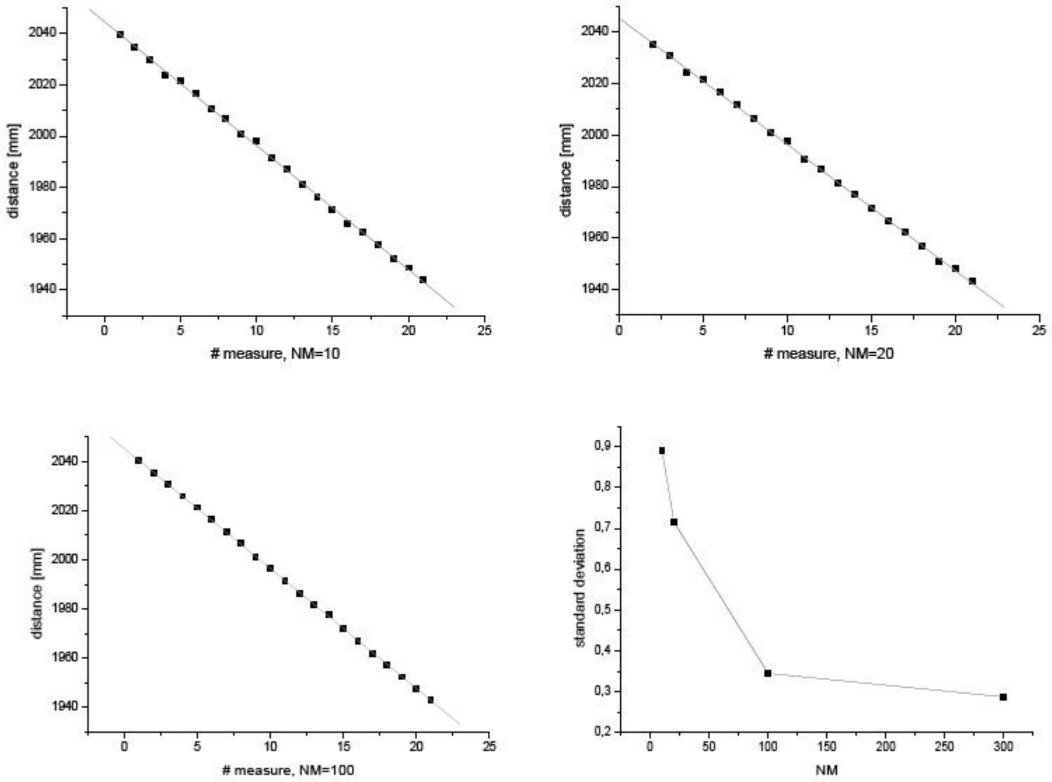


Figure 1.5: Resolution vs. number of averaged samples with the target placed at a distance of 2 m: a) average over $N = 10$; b) average over $N = 20$; c) average over $N = 100$; d) rms error (mm) vs. N.

has been moved in 21 positions with a constant step of 500 μm, resulting in a total translation of 10 cm around the actual distance. We show in Fig. 1.5 some plots obtained with the samples placed at a distance of 2 m. As expected, the higher the number of averaged samples, the higher the resolution.

1.3.1 Laser calibration procedure

A calibration procedure has been carried out to take into account the variability of target surfaces. In fact, one of the main problems of a time-of-flight rangefinder is that the amplitude of the received signal greatly varies depending on the measurement distance and the reflectivity and orientation of the target, with a dynamic range that can be up to 1:1000 or even more. If the amplitude of the received pulse changes, a walk error Δt is generated due to the change in the timing of the stop signal. In order to reduce this error, sophisticated techniques such as automatic gain control on a received pulse or constant fraction discriminator have been developed [49]. In these methods, the timing point is generated by comparing an attenuated pulse and a delayed pulse, so that their crossing point defines a constant fraction of the pulse. The timing point (i.e. the crossing point) is then insensitive to variations in the amplitude of the signal.

The accuracy of the range-finder, i.e. the absolute difference between the measured and real distance, is very good for highly reflective targets, but gets worse when the reflected signal intensity decreases. The measurements have thus to be corrected on the basis of the target's reflectivity and distance. The Noptel rangefinder includes an amplitude pulse proportional output that can be used to improve measurement accuracy. In the following, we will briefly describe the procedure scheme that has been included in the acquisition software. The parameters of the correction have been calculated using six reference samples with different reflectivity in the 0.03-0.90 range (i.e. six colored tiles from black into white through a grey scale) placed at several distances ranging up to 15 m. At each sampled distance, the signal from the six tiles has been acquired. Since the detected intensity changes with the target reflectivity, the acquired distance measurements were not constant. As a general rule of thumb, it can be assessed that the higher the signal, the more precise the distance measurement. Then, the "true" distance can be identified as that measured from the white tile, whose reflectance is close to unity. The other measurements can be thus corrected accordingly.

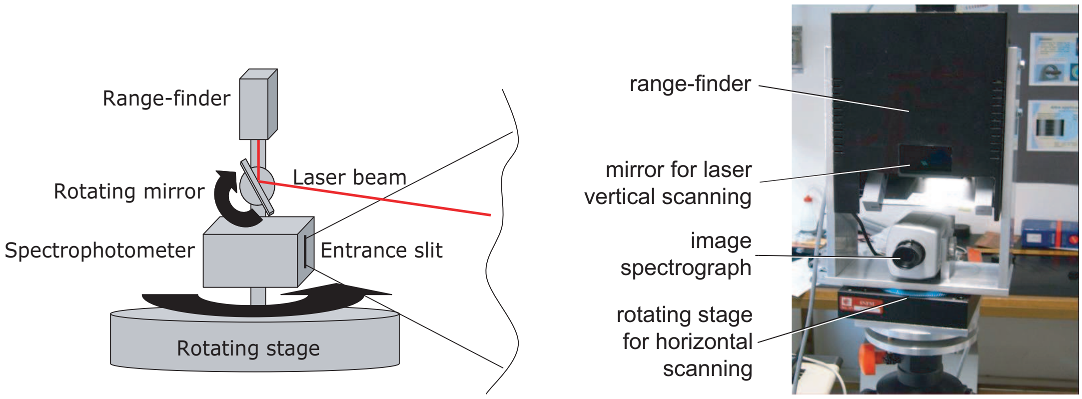
By repeating this procedure for several distances, a look-up table has been obtained to correct the actual measurements of the range-finder. The input values for this table are the distance and intensity measured by the range-finder, the output is the correction of the distance. The outcome of the calibration procedure reveals that:

1. measurements are affected by a systematic error for targets with different reflectivity placed at distance lower than 10 m; the error for a 3% reflecting target can be as high as 10 cm;
2. for target distances higher than 10 m, no correction is needed.

It can be concluded that the presented range-finder has an accuracy better than 1 mm rms over the whole 1-20 m interval.

1.4 System architecture

The acquisition instrument integrates the 2D imaging spectrograph of Section 1.2 and the laser range-camera of Section 1.3 on the same platform, as shown by the



(a) (b)

Figure 1.6: Schematic (a) and picture (b) of the instrument.

scheme of Fig. 1.6(a), so that they can simultaneously measure the multispectral color and geometry information of a scene.

As shown in Fig. 1.6(a), the imaging spectrograph is mounted on a horizontal rotating stage with the entrance slit oriented along the vertical direction. The range-finder is positioned on top of the spectrograph so that the vertical sweep of the range-finder coincides with a vertical line of the spectrograph . Scanning along the horizontal direction is performed by jointly rotating both instruments around the vertical axis using the rotating stage. Fig. 1.6(b) shows a picture of the instrument. The system is mounted on a support that can be manually rotated around a horizontal axis to point the optical axis of the instrument (that is the axis of the spectrograph) at any inclination between 0 and 90 degrees with respect to the horizontal plane. In such a way, it is easy to acquire images both from walls and from ceilings or vaults.

As previously said, the acquisition instrument simultaneously produces a cloud of N points $\mathbf{P} = (X, Y, Z)^T$ and a multispectral image $S(x, y, \lambda)$, where each $\mathbf{P}$ position in space corresponds to a $\mathbf{Q} = (x, y)^T$ position in the multispectral image coordinate system. Wavelength λ is generally sampled using 112 spectral values, corresponding to a spectral resolution of 5 nm/pixel.

The registration between texture data and geometry data is based on the fact that each point $\mathbf{P} = (X, Y, Z)^T$ of a given vertical line X imaged by the range finder corresponds to the point (x, y) of the imaging spectrograph according to the following relationship:

$$\begin{bmatrix} y' \\ w \end{bmatrix} = \begin{bmatrix} \frac{f}{l_x} & c_x \\ 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & t_1 \\ r_{21} & r_{22} & t_2 \end{bmatrix} \begin{bmatrix} Y \\ Z \\ 1 \end{bmatrix}, \tag{1.2}$$

where the y coordinate can be found as $y = y'/w$, since eq. (1.2) is in homogeneous coordinates. Parameters f, l_x and c_x are the focal length of the camera, the x size of the CCD pixel in mm and the horizontal coordinate of the central point, respectively. r_{ij} and t_i (with $j = \{0, 1\}$, $i = \{0, 1\}$) are the elements of the two-dimensional roto-translation between the reference systems of the laser scanner and that of the spectrograph. Eq. (1.2) accounts for the fact that the relationship between coordinates x and X is trivial because of the vertical alignment of the two instruments; w is a scale factor. The calibration parameters

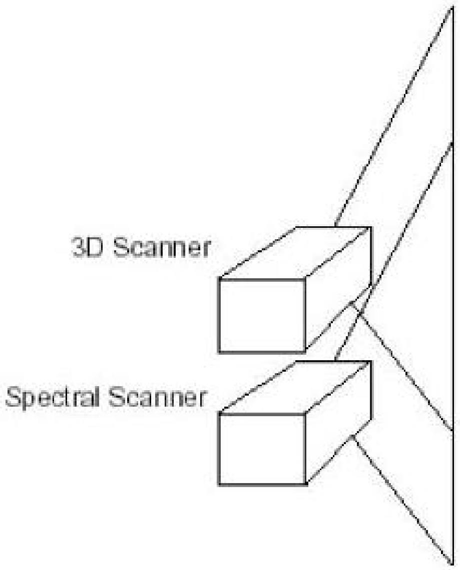


Figure 1.7: Spectral scanner and 3D scanner.

$(a_x = f/l_x, c_x, \mathbf{r}_1, \mathbf{r}_2, \mathbf{t})$ can be computed from the correspondences obtained by acquiring a number of spots projected by the laser scanner on a given surface with the spectrograph. Since there are five degrees of freedom, at least 5 correspondences are needed by the DLT algorithm [50]. Clearly, the more correspondences are given, the better the calibration parameters are estimated. In this case, the calibration has been generally obtained by projecting 15 range-finder spots over the same surface in different positions.

The spectrograph and the range finder are positioned at a distance of 99 mm, therefore in principle some regions can be seen from the spectrograph but not from the range-finder, and viceversa (see Fig. 1.7). However, dealing with wall surfaces, the potential parallax problem does not create any trouble, since the two scanners are positioned very close (only few centimeters apart) when compared to the size of the imaged surfaces (several meters). In any case this problem, if present, can be solved by performing several acquisitions in different positions.

The regions acquired by the two sensors depend on the angle spanned by the range-finder deflector and on the viewing angle of the spectrograph. The two areas cannot be exactly the same, since their relative position changes according to the distance from the target. The viewing angle of the spectrograph cannot be modified, once the objective has been chosen (16 mm or 25 mm in our case). Therefore, the measurement angle of the range-finder has been usually set to cover an area slightly wider than that seen by the spectrograph at a representative measurement distance. In this way, some portions of the acquired 3D views may not have a multispectral texture counterpart. However, new acquisitions above or below such scans can cover all the region of interest.

1.5 3D modeling procedure

In this Section we will summarize the procedure to build a textured 3D model from the single textured scans. The author gave no original contribution to the modeling procedure, so it will be only briefly outlined. For further details, see [24].

The algorithm to integrate multiple textured scans into a single 3D model is described in Algorithm 1.

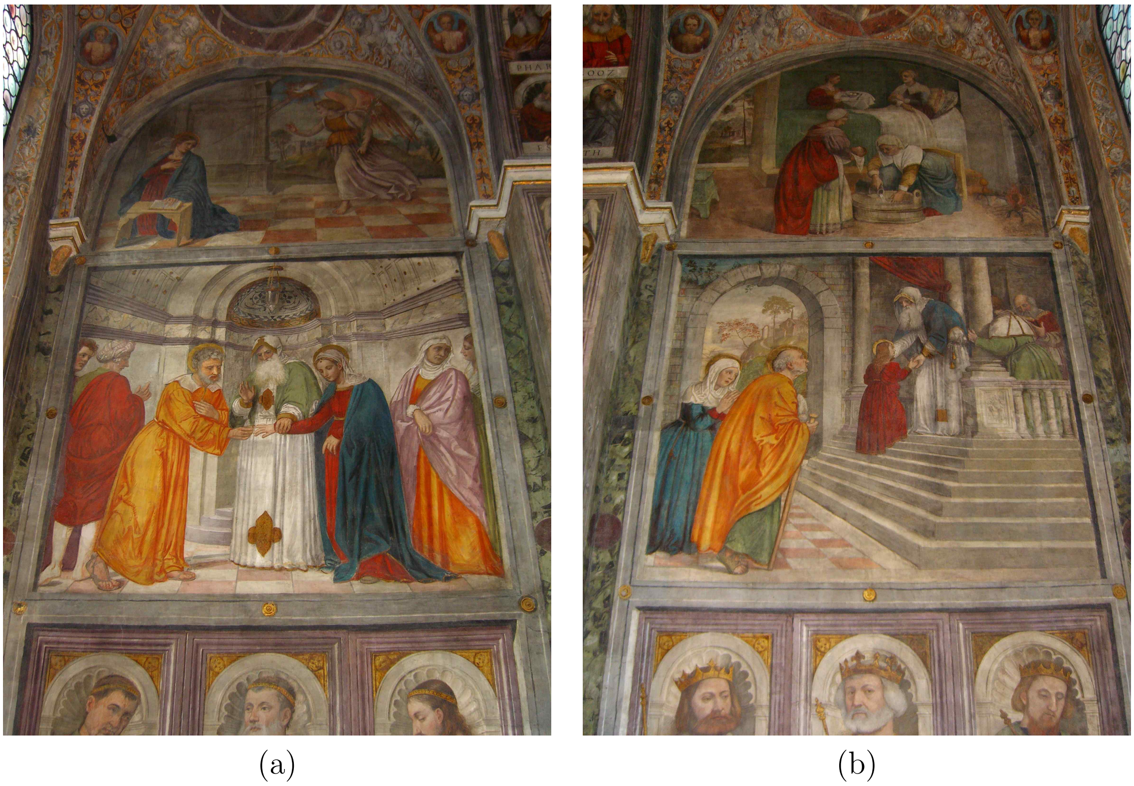
```
Input:
a set of textured 3D scans;
Output:
an integrated textured 3D model;
begin
    automatic detection of a common region between overlapping 3D views
    by a variation of the method of textured spin-images [51];
    pairwise registration by the method of [52];
    global registration by the methods of [53];
    multispectral texture construction by an extension of the method
    of [24, 54];
    surface fusion by the method of [55];
end
```

Algorithm 1: The 3D modeling pipeline

1.6 Acquisition results: an example

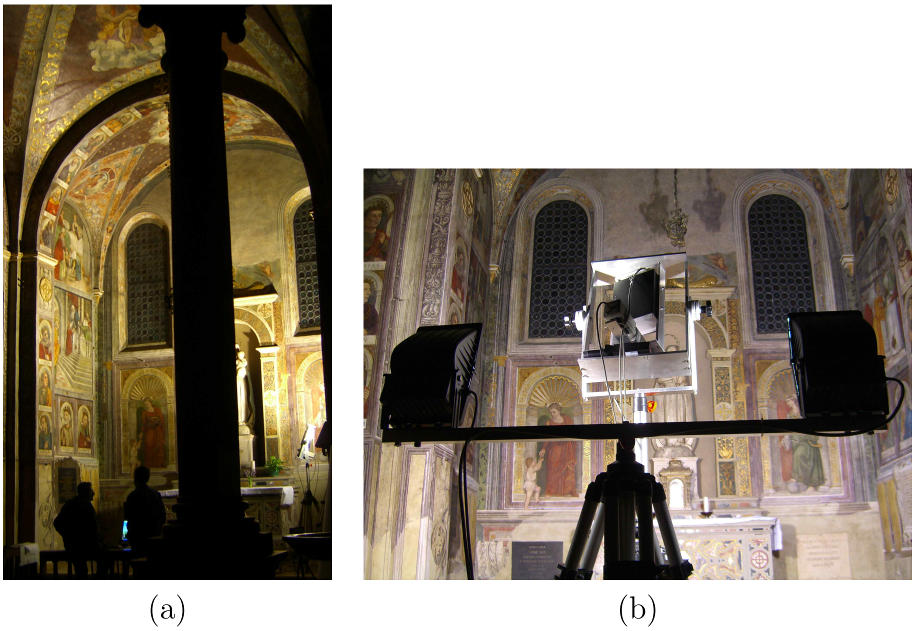
We present here some experimental results relative to the deployment of the multispectral range camera in the Chapel of Charity, a frescoed lateral chapel in the Church of S. Francesco in Padua (Italy). The walls and the ceiling were frescoed by Girolamo Tessari in 1523-1526 with stories of the life of Virgin Mary. The chapel is approximately 5 m (width) × 3 m (depth) × 8 m (height). Two images of the frescoes are shown in Fig. 1.8.

Two pictures of the acquisition set-up are shown in Fig. 1.9. Two metallic iodide lamps have been used to illuminate the chapel. The 400-800 nm visible range has been acquired in 45 spectral bands with 0.1 s exposure time. Simultaneously, the 3D shape data have been acquired. The reference spectrum to be used in eq.(1.1) for the normalization has been acquired for each scan by placing a white screen in front of the frescoed walls. The acquisition of the whole vault required 25 scans of area (1.5-2.5 m) × (3-6 m) and took approximately 15 hours. Figs. 1.10 and 1.11 show an example of the output of the spectrograph and of the range-finder while acquiring the chapel vault. Fig. 1.12 shows the overall output of the multi-spectral range camera, i.e. the intensity and the (R,G,B) renditions of the associated multispectral image registered together in a multispectral 3D view. Other examples are shown in Fig. 1.13.



(a) (b)

Figure 1.8: Images of two episodes from the cycle of Stories of Mary on the lateral walls of the Chapel of Charity (Church of S. Francesco, Padua): (a) presentation to the Temple; (b) marriage of Mary.



(a) (b)

Figure 1.9: Set-up for the acquisition of the images in the Chapel of Charity: (a) view of the chapel; (b) picture of the instrument (center) and the two lamps (side) during a scan of the central wall.

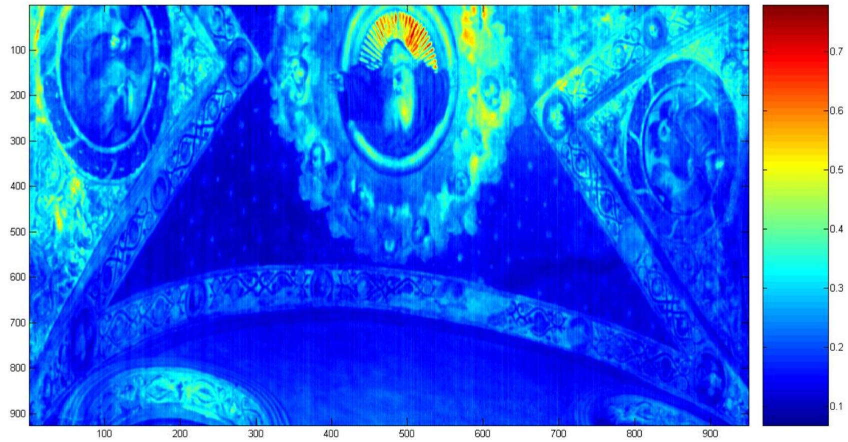


Figure 1.10: Multi-spectral image of the vault (mean spectral reflectance).

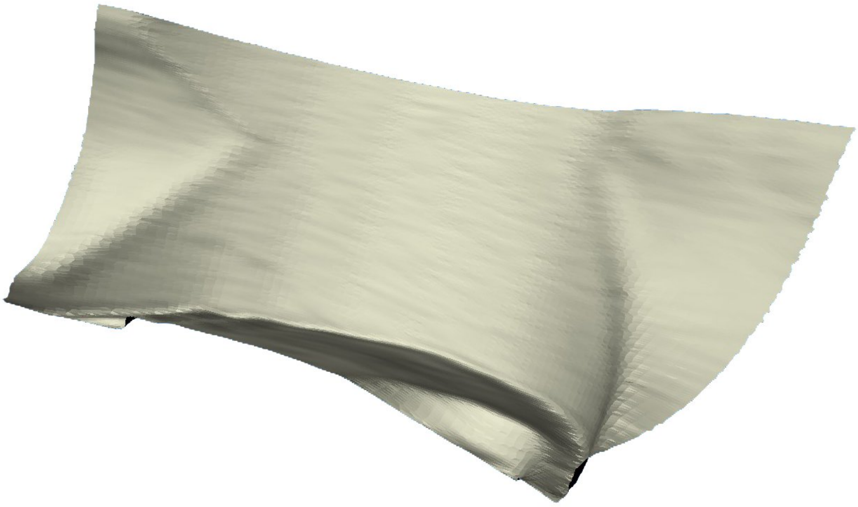


Figure 1.11: 3D scan of the vault.

Figure 1.12: Multispectral image of Fig. 1.10 registered on the 3D scan of Fig. 1.11.

1.7 Discussion

To the author's knowledge, no commercial instruments currently exist for the joint acquisition of geometry and multispectral texture information for small ar-

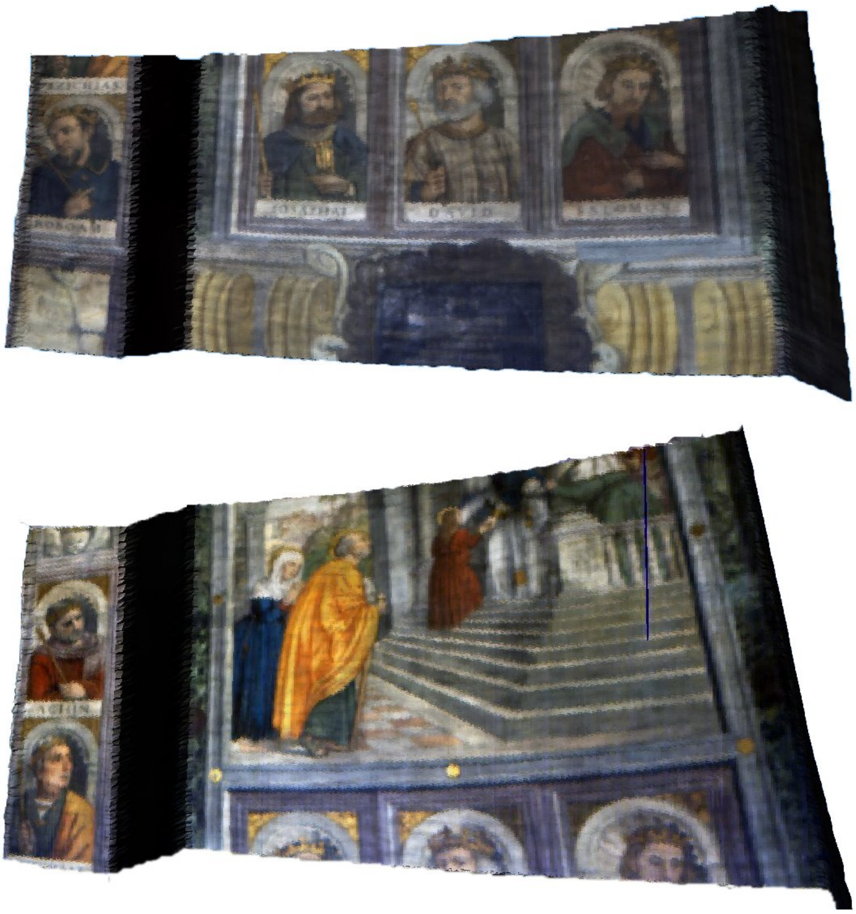


Figure 1.13: Multispectral 3D images of the walls.

chitectural volumes.

The automation of the 3D modeling procedure guarantees procedural repeatability, and is therefore particularly suitable for monitoring purposes. Moreover, it makes the modeling procedure accessible to non technical personnel. However, as observed in [56], the 3D modeling of complex environments often forces a choice between automation and reconstruction precision. In all on-field applications, manual adjustments were always necessary to obtain the final textured 3D model.

As for practical issues, the acquisition procedure is not very fast as compared with the performance of state-of-the-art commercial 3D scanners (30 kHz). However, the acquisition of multispectral data partially compensates for this disadvantage.

The joint acquisition of geometry and spectral reflectance offers an improvement in the analysis of frescoed surfaces, highlighting the interrelation between wall and painting modifications over time. Another strong point of the proposed methodology is that it is completely non-invasive. Multispectral textures not only permit to recover a faithful reproduction of a fresco's color, but also potentially allow to monitor its conservation status by analyzing the infrared content of its spectral reflectance. Unfortunately, chemical elements are usually identified by analyzing their spectral signatures in the near or mid infrared spectrum [44], that is at longer wavelengths that those detected by the Specim

Imspector V10($\lambda > 1000\,\mathrm{nm}$). We have recently undergone a collaboration with the Chemistry Department of the University of Padova to study what kind of information on the composition and conservation status of the pictorial layer can be extracted from our current spectral window.

An instrument such as the Shape&Color scanner can really fulfil its purpose only if metrological guarantees are given on the acquired geometry and reflectance data. As discussed in Chapter 2, the illumination setup used in the presented on-field validations cannot guarantee sufficient precision across the 400-1000 nm interval. Other illumination setups have thus been investigated.

Another critical point is that the acquisition of the white signal gives unsatisfactory results for illumination subtraction both in uneven areas (where it is impossible to reproduce subtle changes in illumination and shadows) and in topmost areas (where the white signal cannot be acquired, and the database white signals poorly reproduce illumination conditions in such complex environments). The method described in Chapter 3 was conceived to try to overcome these problems, separating reflectance and illumination contributions directly from the color signal.

Another weak point of the proposed acquisition method is that it does not account for reflectance variations with the viewing and illumination angles (see Section 3.1). This may not be a problem when acquiring diffuse surfaces that approximately exhibit a Lambertian behavior, but certainly becomes so when dealing with specular patches such as the golden parts of a fresco. In this case, significant comparisons between subsequent acquisitions can be obtained only by placing the instrument and the illumination sources in matching positions.

Chapter 2

Performance characterization of the imaging spectrograph

During the first acquisitions with the Shape&Color scanner (see Chapter 1), it was soon realized that the (R,G,B) colors computed from the measured spectral reflectances looked darker and slightly distorted with respect to the originals. A spectral validation of the instrument performance had been performed (see [24]) by acquiring a set of calibrated tiles of known reflectance. However, quantitative results had been given only for the (L^*,a^*,b^*)coordinates [57] of the tabulated and measured reflectance values. We believed that this performance characterization procedure, though commonly accepted in the color science community, was not fit for an instrument whose aim was to faithfully reproduce spectral reflectance data. Therefore, we have re-acquired the calibrated tiles and evaluated the system performance under different illumination sources, so as to select the best illumination setup for our acquisitions. We have defined the measurement error as a function of wavelength, and calculated the error uncertainty with a metrology-inspired procedure.

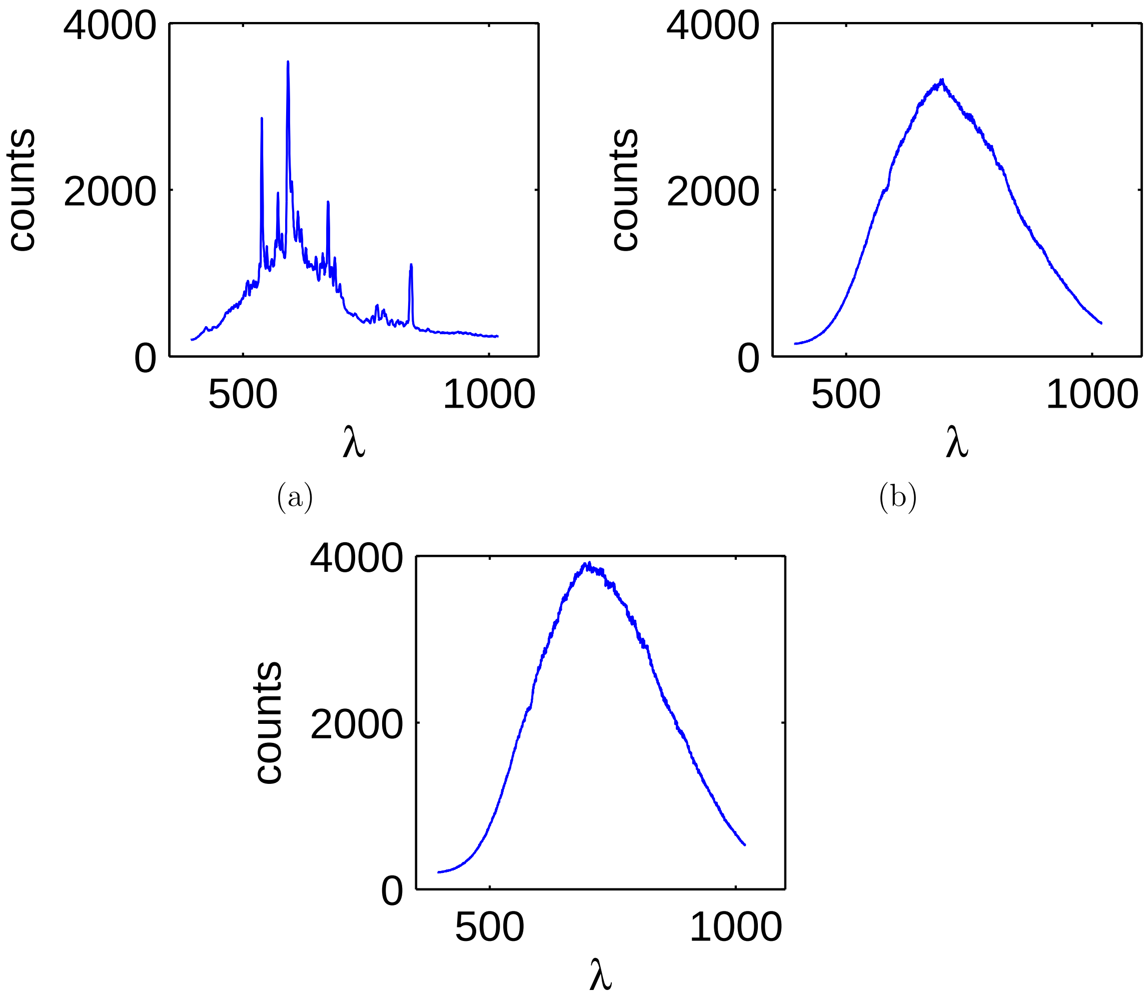
2.1 Reflectance measurement

We recall that our imaging spectrograph measures absolute light intensities, therefore it is necessary to acquire a reference white signal under the same illumination conditions as the target object in order to calculate its spectral reflectance. In our experimental trial, the reference white signal has been obtained by acquiring a Labsphere white tile with a flat 80% spectral reflectance over the interval 250-2500 nm [58].

The reflectance of the target for any pixel $R_{m,n}$ has been then calculated as:

$$R_{m,n} = 0.8 * \frac{S_{m,n} - D_{m,n}}{W_{m,n} - D_{m,n}}, \tag{2.1}$$

where $m = 1, \ldots, 1024$ and $n = 1, \ldots, 1344$ are the coordinates of the pixel within the sensor, $S_{m,n}$ is the acquired signal, $D_{m,n}$ is the dark signal of the camera and $W_{m,n}$ is the white signal. The subtraction of the dark signal accounts for the presence of the dark current signal in the CCD. The dark current signal is constant (apart from a random noise term) for a given temperature and exposure



(a) (b)

(c)

Figure 2.1: Spectrum of the lamps used in our study: (a) metallic iodide lamp (b) halogen lamp (c) incandescence lamp.

time. Therefore, it can be measured by acquiring a dark frame at the same exposure time and in the same temperature conditions as the actual frame [59].

Looking at eq. (2.1), it is clear that the spectral reflectance at a spatial point m on the target can be obtained as:

$$R_m = [R_{m,n}]_{n=1,\dots,1344}, \tag{2.2}$$

with the corresponding wavelength array $\lambda = [\lambda_n]_{n=1,\dots,1344}$ given by the spectral calibration procedure.

2.2 The experimental trial

The light reflected by an object depends both on the object's surface reflectance characteristics and on the composition of the light illuminating the object. An ideal illumination source for the measurement of an object's spectral reflectance should thus exhibit a uniform spectrum over the whole desired wavelength range. Unfortunately, no existing illumination apparatus satisfies this condition over the visible and near-infrared spectrum, so that sub-optimal solutions must be adopted.

Moreover, other considerations concerning the spectral emission of a lamp should be taken into account. For example, when acquiring a painting or fresco, it is often unadvisable to use lamps with a strong emission in the infra-red spectrum as their heating may cause damages to the pictorial layer. As our spectral camera was primarily conceived for the acquisition of frescoed environments (see Chapter 1), we usually utilized metallic iodide lamps due to their low emission both in the NIR and UV spectra. However, as the instrument can be employed in situations where such precautions may not be necessary, in the presented study we have looked for other kinds of lamps that would provide a sufficiently strong emission in the long-wavelength visible and NIR spectrum. Three different lamps have been used in our experiments:

1. a Disano 250 W metallic iodide lamp (with the spectrum shown in Fig. 2.1(a));
2. a Cixi Zhongfa Lamps 500 W halogen lamp (Fig. 2.1(b));
3. an Osram 60 W incandescence lamp (Fig. 2.1(c)).

It can be observed that the metallic iodide lamp has strong emission in the blue-green region, while the halogen and incandescence lamps mainly emit in the red-NIR spectrum. However, it can easily be seen that the emission of all the lamps definitively drops around 850-900 nm. Therefore, all reflectances have been calculated in the 400-900 nm interval.

Our experimental trial has consisted in the acquisition of known reflectance objects under different illumination conditions. Four lamp configurations have been used:

1. Metallic iodide lamp alone. We have chosen this setup as it is the only one which may be extended to the acquisition of painted surfaces; as the metallic iodide lamp has low emission in the red-NIR region, a correction function has been derived for the measured spectral reflectances.
2. Metallic iodide lamp *together with* halogen lamp. This setup has the advantage of combining the emissions of the two lamps without doubling the acquisition time.
3. Metallic iodide lamp and halogen lamp in a sequence. The reflectance measurement has been obtained by juxtaposing the spectral reflectance measured with the metallic iodide lamp in the 400-600 nm interval and that obtained with the halogen lamp in the 600-900 nm interval.
4. Metallic iodide lamp and incandescence lamp in a sequence, obtaining the spectral reflectance as in the previous case. The incandescence lamp has been used instead of the halogen lamp because it has a slightly stronger emission in the NIR spectrum and because its spectrum is smoother.

The acquired target has been the Labsphere Pastel Color Standard set [58]. This set is made of eight highly diffusive tiles with colors spanning the visible spectrum. Their spectral reflectance is tabulated between 380 nm and 830 nm. No quantitative error evaluation has therefore been possible in the 830-900 nm interval.

For each target acquisition, five measurements have been done to average down measurement noise. The camera exposure time has been set so as to maximize the white signal dynamic without saturating the camera (saturation value is 4000 counts).

2.3 Performance evaluation

A way of characterizing the accuracy of reflectance measurements could be that of calculating the CIELAB distances [57] between tabulated and measured reflectances [23]. However useful this approach may be for colorimetry applications, it does not account for accuracy variations with wavelength. Therefore, we have chosen to characterize our system's performance by the error and the error standard deviation calculated as functions of wavelength and averaged over the tile set.

Remembering eqs.(2.1) and (2.2), we have that the mean reflectance for each tile can be calculated as:

$$\begin{aligned} R_m^i(\lambda) &= \frac{\frac{1}{5}\sum_{j=1}^5 S_m^{i,j} - \frac{1}{5}\sum_{j=1}^5 D_m^{i,j}}{\frac{1}{5}\sum_{j=1}^5 W_m^{i,j} - \frac{1}{5}\sum_{j=1}^5 D_m^{i,j}} = \\ &= \frac{\bar{S}_m^i - \bar{D}_m^i}{\bar{W}_m^i - \bar{D}_m^i} = \\ &= f(\bar{S}_m^i, \bar{W}_m^i, \bar{D}_m^i), \end{aligned} \tag{2.3}$$

where $i = 1, \ldots, 8$ is the tile index and $j = 1, \ldots, 5$ the acquisition index. From now on, we shall drop the pixel index m, with the understanding that we will consider the central pixel of each tile.

The reflectance error for each of the eight colors is therefore:

$$e^i(\lambda) = R^i(\lambda) - R_{tab}^i(\lambda), \tag{2.4}$$

where $R_{tab}^i(\lambda)$ is the tabulated spectrum of the i-th tile.

We now want to calculate the variance of the error from the variances of the experimentally observed variables $\{X^{i,j}\}_{j=1,\ldots,5}$, $X \in \{S, W, D\}$. The latter can be estimated as [60]:

$$\sigma_{X^i}^2(\lambda) = \frac{1}{4}\sum_{j=1}^{5}(X^{i,j}(\lambda) - \bar{X}^i(\lambda))^2, \tag{2.5}$$

with the understanding that each observation is given by $X^{i,j}(\lambda) = \tilde{X}_i(\lambda) + w_{i,j}(\lambda)$, where $\tilde{X}_i(\lambda)$ is the "true" spectrum and $w_{i,j}(\lambda)$ the realization of a random noise.

The variance of the derived variable $e^i(\lambda)$ can be calculated as (see [60]):

$$\sigma_{e^i}^2 = \sum_{k=1}^{3}\left(\frac{\partial f}{\partial \bar{X}_k^i}\right)^2 \sigma_{X_k^i}^2 + 2\sum_{k=1}^{3}\sum_{l=k+1}^{3}\frac{\partial f}{\partial \bar{X}_k^i}\frac{\partial f}{\partial \bar{X}_l^i}\sigma_{X_k^i, X_l^i} + \sigma_{R_{tab}}^2, \tag{2.6}$$

where $X_1 = S,\ X_2 = W,\ X_3 = D$, $\sigma_{X_k^i, X_l^i}$ is the covariance between the variables X_k^i and X_l^i, $\sigma^2_{R_{tab}}$ the variance of the tabulated reflectance, and the dependence by λ has been omitted for the sake of clarity. As the variance of the tabulated reflectance is less than $2.5 \cdot 10^{-5}$ [58], it will be neglected in the following.

Looking at eq. (2.3), the partial derivatives can be easily calculated as:

$$\begin{aligned}
\frac{\partial f}{\partial \bar{S}^i}(\lambda) &= \frac{1}{\bar{W}^i(\lambda) - \bar{D}^i(\lambda)}, \\
\frac{\partial f}{\partial \bar{W}^i}(\lambda) &= -\frac{\bar{S}^i(\lambda) - \bar{D}^i(\lambda)}{(\bar{W}^i(\lambda) - \bar{D}^i(\lambda))^2}, \\
\frac{\partial f}{\partial \bar{S}^i}(\lambda) &= \frac{\bar{S}^i(\lambda) - \bar{W}^i(\lambda)}{(\bar{W}^i(\lambda) - \bar{D}^i(\lambda))^2}.
\end{aligned} \tag{2.7}$$

Lastly, the covariances $\sigma_{X_k^i, X_l^i}$ are given by:

$$\sigma_{X_k^i, X_l^i}(\lambda) = \frac{1}{4}\sum_{j=1}^{5}(X_k^{i,j}(\lambda) - \bar{X}_k^i)(X_k^{i,j}(\lambda) - \bar{X}_l^i). \tag{2.8}$$

Once we have obtained the error and error variance for each tile, we can average them over the eight tiles to obtain the average error (AE) and the average error variance as:

$$e(\lambda) = \frac{1}{8}\sum_{i=1}^{8} e^i(\lambda), \quad \sigma_e^2(\lambda) = \frac{1}{8}\sum_{i=1}^{8}\sigma_{e^i}^2(\lambda). \tag{2.9}$$

However, as the standard unit of uncertainty is standard deviation [60], we have eventually used the average error standard deviation (AESTD) instead of the variance. The AESTD is given by:

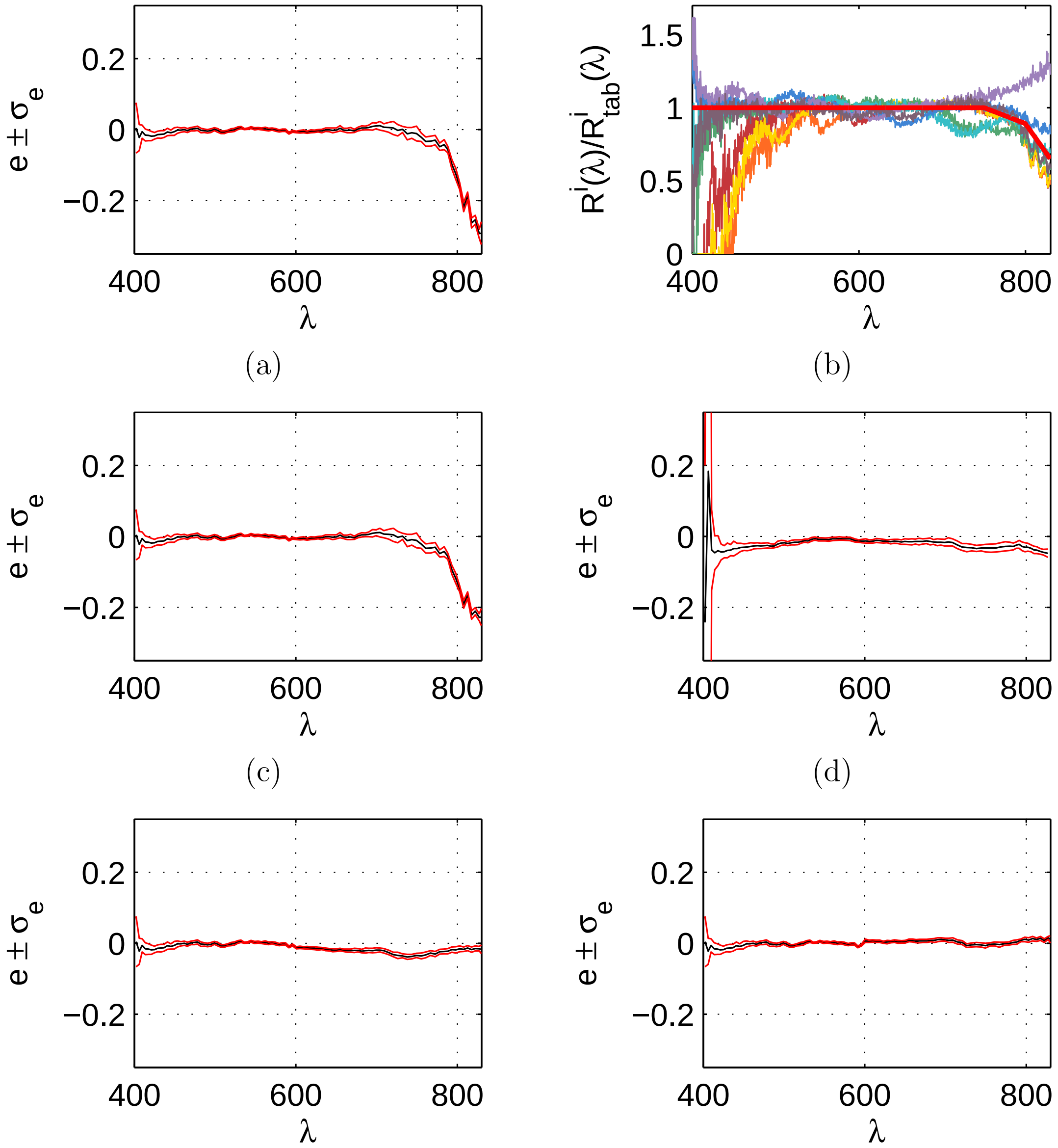
$$\sigma_e(\lambda) = \frac{1}{8}\sum_{i=1}^{8}\sigma_{e^i}(\lambda), \tag{2.10}$$

which is close to, but does not coincide with the square root of the variance expressed in eq.(2.9).

2.4 Results

2.4.1 Metallic iodide lamp

The AE and AESTD of the spectral reflectances acquired using the metallic iodide lamp can be seen in Fig. 2.2(a). It can be noted that the AE greatly increases (in absolute value) after 750 nm, in correspondence with the drop in the lamp emission. On the contrary, the AESTD increases (up to 0.08 between 400 and 420 nm. The reason is that there are great differences in the reconstructed reflectances in the blue region. Namely, the measured reflectances for the red, yellow and orange tiles are zero over most of this interval. The reason is that in these cases three different phenomena combine: a drop in the lamp emission, a



(a) (b) (c) (d) (e) (f)

Figure 2.2: (a) average error (AE) and average error standard deviation (AESTD) using the metallic iodide lamp (b) correction function for the spectra (c) AE and AESTD after the correction (d) AE and AESTD with the metallic iodide lamp and the halogen lamp used together (e) AE and AESTD with the metallic iodide lamp and the halogen lamp used in a sequence (f) AE and AESTD with the metallic iodide lamp and the incandescence lamp used in a sequence.

drop in the CCD sensitivity and a drop in the material reflectance. As a result, the measured reflected signal falls below the dark signal.

For the reasons stated in section 2.2, we have tried to correct the measured spectra defining a mean correction function. To do so, we have considered the ratio between measured and tabulated spectra, which should be one over the entire spectrum for perfectly reconstructed reflectances. We have then defined a continuous piecewise linear function approximating the global reflectance behavior in the least-square sense, and have used it to correct each reflectance. As can be seen in Fig. 2.2(b), there is indeed a global trend in the ratios, with the notable exception of violet in the red-NIR region. The AE and AESTD of the corrected reflectances can be seen in Fig. 2.2(c). It can be observed that, although slightly improved, the error still reaches -0.2 in the NIR region.

2.4.2 Metallic iodide lamp in combination with halogen lamp

As can be seen in Fig. 2.2(d), the combined use of the metallic iodide and halogen lamps has brought an improvement in the red-NIR region, but has also caused a decrease in the reconstruction performance in the blue region. The reason is that in order to use the two lamps simultaneously we have had to halve the camera exposure time with respect to the previous case, so as to avoid camera saturation. This has meant that in the spectral regions where just one of the two lamps has strong emission (the blue and red-NIR regions), we have obtained an “average” performance between the two lamps, rather than the best of the two. This has suggested using the two lamps in a sequence, instead of illuminating the target with both at the same time.

2.4.3 Metallic iodide lamp and halogen lamp in a sequence

We recall that in this case the resulting spectral reflectance has been obtained by juxtaposing the spectral reflectance measured with the metallic iodide lamp between 400 and 600 nm, and that acquired with the halogen lamp between 600 and 830 nm. Looking at Fig. 2.2(e), one can see that the AE is indeed reduced to less than 0.05 in absolute value. However, confronting Figs. 2.2(a) and 2.2(e), it can be observed that the metallic iodide lamp outperforms the halogen lamp between 600 and 750 nm. Rather than changing the juxtaposition threshold to a less intuitive value than the center of the considered spectrum, we have decided to try another light source emitting in the red-NIR region.

2.4.4 Metallic iodide lamp and incandescence lamp in a sequence

Using the incandescence lamp instead of the halogen lamp (Fig. 2.2(f)), the AE is reduced to less than 0.02 (in absolute value), with an AESTD of about 0.01 over the whole 420-830 nm spectrum (for the 400-420 nm interval the considerations done in subsection 2.4.1 still hold). Moreover, the AE is more uniform than in the previous case.

2.5 Conclusions

We have assessed the performance of our spectral camera in the measurement of spectral reflectance in the 400-830 nm interval by acquiring a set of colored calibrated tiles under different illumination conditions. We have used a metallic iodide lamp, the metallic iodide lamp together with a halogen lamp, the metallic iodide and halogen lamps one after the other, and the metallic iodide and incandescence lamp in a sequence. In the last two cases, we have obtained spectral reflectances as a juxtaposition of the reflectance measured with the metallic iodide lamp from 400 to 600 nm and that acquired with the other lamp from 600 to 830 nm. To evaluate the system performance, we have defined the error as a function of wavelength and used a metrological procedure to infer the uncertainty of the computed error from the statistics of the measured variables. To describe the reflectance measurement performance, we have used the average error (AE) and the average error standard deviation (AESTD), calculated for each illumination setup and averaged over the eight-tile set. The best results have been obtained with the metallic iodide and the incandescence lamps used in a sequence. In this case, the absolute AE is less than 0.02 over the whole spectrum, and the AESTD less than 0.01 between 420 and 830 nm and less than 0.08 between 400 and 420 nm. The proposed methodology can be generalized to quantify the effects of other influence quantities (e.g. the surface material) onto the accuracy of reflectance measurements, and can also be extended to other types of multispectral sensors.

Chapter 3

A statistical approach to the illumination-reflectance separation problem

3.1 Problem statement

The aim of this study is that of assessing a method for separating a color signal into illumination and surface reflectance components. The light reflected by a surface depends on many factors:

1. surface material (matte, glossy, translucent, transparent);
2. surface spectral reflectance (color);
3. surface geometry;
4. spectral distribution of the illumination;
5. position of the illumination with respect to the target surface;
6. interreflections by the environment;
7. geometry of the environment (shadows).

The problem of illumination-reflectance separation can be formulated as follows. Given the observation of the light reflected by an object (color signal), sampled wrt wavelength in a certain number of values, we want to estimate the surface spectral reflectance and the illumination that give rise to that spectrum.

The dependence of the color signal upon surface reflectance and illumination is very complex. Even considering non-transparent, non-translucent materials, the way an object reflects light at a certain point is a function of wavelength, angular position of the observer and angular position of the light source. The function expressing the dependence of an object reflectance upon these parameters is called *Bidirectional Reflectance Distribution Function (BRDF)* [61] and will be indicated as $\rho(\mathbf{p}, \boldsymbol{\omega}_i, \boldsymbol{\omega}_o, \lambda)$, where λ is wavelength and $\boldsymbol{\omega}_i$ and $\boldsymbol{\omega}_o$ are two-valued vectors containing, respectively, the light and observer direction wrt the surface normal at point $\mathbf{p}$. Ramamoorthi and Hanrahan [62] have observed that, assuming no shadowing and interreflections occur, and considering an homogeneous, isotropic material (with no texture), the light reflected by an object at point $\mathbf{p}$ can be expressed as:

$$S(\mathbf{p}, \boldsymbol{\omega}_o, \lambda) = \int_\Omega I'(\mathbf{p}, \boldsymbol{\omega}_i, \lambda)\rho(\mathbf{p}, \boldsymbol{\omega}_i, \boldsymbol{\omega}_o, \lambda)(\boldsymbol{\omega}_i \cdot \mathbf{n})\, d\boldsymbol{\omega}_i, \tag{3.1}$$

where S is the reflected light, Ω is the unit hemisphere, I' the illumination and $(\boldsymbol{\omega}_i \cdot \mathbf{n})$ stands for the inner product between the surface normal at $\mathbf{p}$ and the direction $\boldsymbol{\omega}_i$. In other words, the reflected light is the result of a convolution between illumination and BRDF in the angular domain.

Measuring the BRDF of a chosen material is a laborious and time-consuming task. It requires to acquire spectral reflectance values (generally at least three samples) for sampled ingoing and outgoing directions, each spanning the whole hemisphere. BRDF measurement is mostly performed for computer graphics applications to increase photorealism [63,64]. In this case, the BRDF of a number of significant materials (sand, stone, brick, grass, fabric) is measured using trichromatic sensors, and then used to render 3D scenes in a realistic manner. Another important field of application is remote sensing, where BRDF is used to identify surface constituents independently from the viewing and illuminating angles [65]. In both cases, a somewhat qualitative knowledge of BRDF is sufficient, whether because only "perceptually right" results are pursued, or because few parameters (e.g. peaks or minima) must be extracted from spectral reflectances.

The presented study concerns the measurements of spectral reflectance of painted surfaces, such as paintings or frescoed walls. Measuring the BRDFs of all the materials constituting these surfaces, in all possible conditions of temperature, humidity and aging, is virtually impossible. Therefore, a number of simplifications have been introduced. Let us assume all surfaces to be Lambertian, i.e. having constant BRDF wrt the viewing angle. In this case, the light reflected by the object surface only depends on the cosine of the angle between illuminating direction and surface normal [61]. Eq.(3.1) changes as:

$$S(\mathbf{p}, \lambda) = R(\mathbf{p}, \lambda) \int_\Omega I'(\mathbf{p}, \boldsymbol{\omega}_i, \lambda) \cos(\theta_i) d\boldsymbol{\omega}_i, \tag{3.2}$$

where $R(\mathbf{p}, \lambda)$ is the surface (spectral) albedo at point $\mathbf{p}$ and $\theta_i = \arccos(\boldsymbol{\omega}_i \cdot \mathbf{n})$. Let us now pose

$$I(\mathbf{p}, \lambda) \triangleq \int_\Omega I'(\mathbf{p}, \boldsymbol{\omega}_i, \lambda) \cos(\theta_i) d\boldsymbol{\omega}_i. \tag{3.3}$$

$I(\mathbf{p}, \lambda)$ can be interpreted as the cumulative illumination resulting from the presence of an extended light source, weighted according to direction. In the unrealistic assumption of dealing with a point source, and maintaining the assumption of no interreflections, we would have $I(\mathbf{p}, \lambda) = I'(\mathbf{p}, \boldsymbol{\omega}_i^0, \lambda) \cos(\theta_i^0)$, with $\boldsymbol{\omega}_i^0$ direction of the source. From now on, we will refer to I simply as "illumination". With the position of eq.(3.3), eq.(3.2) finally becomes

$$S(\mathbf{p}, \lambda) = I(\mathbf{p}, \lambda) R(\mathbf{p}, \lambda), \;\; \mathbf{p} \in \mathbb{P}, \;\; \lambda \in \Lambda, \tag{3.4}$$

where $\mathbb{P}$ is a two-dimensional lattice and Λ is a discrete set of sampled wavelengths. We will call eq.(3.4) the *illumination-reflectance separation equation.*

3.1.1 Ilumination-reflectance separation and color constancy

The task of separating illumination and reflectance from a color signal is sometimes referred to as *color constancy* in the literature [66, 67]. This expression refers to a property of the human eye, which exhibits the capability of recognizing the color of objects independently of the illuminating light [68]. However, we adopt Forsyth's formulation [69], according to which color constancy involves integrating the color signal multiplied by the sensitivity functions of a certain number of receptors. The *color constancy equation* can therefore be expressed as:

$$\begin{aligned} E_k(\mathbf{p}) &= \int \varrho_k(\lambda) I(\lambda, \mathbf{p}) R(\lambda, \mathbf{p}) d\lambda = \\ &= \int \varrho_k(\lambda) S(\lambda, \mathbf{p}) d\lambda, \quad k = 1, \ldots, K, \end{aligned} \tag{3.5}$$

where E_k is the response of the k-th receptor at point $\mathbf{p}$ and $\varrho_k(\lambda)$ is the sensitivity function of the k-th receptor. Generally, one has three receptors corresponding to the red, green and blue regions of the spectrum.

In its most general formulation, the aim of a color constancy algorithm is that of recovering surface properties which are independent of the illumination [69]. Illumination-reflectance separation can be a specific solution to this problem, but is certainly not the only one [69–72]. However, as illumination-reflectance separation and color constancy are closely related problems, and as the former has often been addressed as part of color constancy algorithms, we will review some works concerning color constancy.

3.2 Related work

We will begin by reviewing the literature on spectral illumination-reflectance separation algorithms, which most closely relate to our work.

Ho *et al.* [67] first addressed the problem of spectral illumination-reflectance separation. The authors assumed reflectance and illumination to be a linear combination of a finite number of basis functions, and formalized the estimation problem as the minimization in the least-square sense of the difference between the color signal and the product of illumination and reflectance. They proved that if the functions obtained by multiplying in all possible combinations the reflectance and illumination basis functions form an orthogonal set, it is possible to recover the illumination and reflectance coefficients, under the assumption that the first illumination coefficient is normalized to 1.

Chang and Hsieh [73] observed that Ho's method has the drawback of producing unfeasible solutions, such as negative reflectance values. They therefore added some physical constraints to the minimization problem, forcing reflectance functions to be within the $[0, 1]$ interval and finding a similar constraint for the illumination. They performed the minimization using an efficient implementation of the simulated annealing algorithm [74] based on variable separation.

Ikari *et al.* [75] pointed out that Ho's and Chang's minimization methods are affected by the problem of local minima. They attributed this to the fact that both methods rely on single points to separate illumination and reflectance components. To overcome this problem, they proposed to use multiple points of equal color illuminated by different illumination sources, or points of different colors illuminated by the same light, an approach that can also be found in the color constancy literature [76]. They also proposed to separate the two components using the dichromatic reflection model [**?**], an approach which can also be found in physics-based color constancy algorithms [77].

Among color constancy algorithms, we would like to mention the early work by Horn [78], where the author separates illumination and surface reflectance contributions exploiting the observation that in real-world images, illumination varies far more slowly than reflectance. Horn considers the logarithm of the color signal, which can be expressed as the sum of the logarithms of illumination and surface reflectance. He then proposes to take the Laplacian of the logarithm image, and eliminate all edges whose intensity falls below a certain threshold. In this way, he claims that the illumination contribution, which gives rise to small intensity changes, can be eliminated. Horn's assumption of slowly-varying illumination will be used in our algorithm.

Another interesting approach found in the color constancy literature is that of exploiting the correlation between output color signal and input illumination to estimate the illumination. For example, Forsyth [69] observes that only a subset of all possible colors can be observed under a reference light, and defines color constancy as the problem of finding the (linear, diagonal) transformation mapping the observed image gamut to the canonical one. His method, called "Gamut mapping", has been recently extended in [79], where a useful review of similar illumination estimation methods can be found. All of them are founded upon the assumption that, given an image, some illumination spectra can be assumed to have produced the observed color gamut with greater probability than others. Dealing with spectral functions, we have obviously not considered color gamuts. However, we will infer assumptions on the spectral reflectance dynamic of a pixel from the color signal in a somewhat similar approach.

3.2.1 Illumination-reflectance separation: the curse of the scale factor

All the above mentioned illumination-reflectance separation algorithms have in common that illumination and reflectance are estimated only up to a scale factor. It is easy to see from the illumination-reflectance separation equation that if we multiply reflectance and illumination by reciprocal constants, the resulting color signal does not change. In other words, the illumination-reflectance separation is an ill-posed problem, for which a unique solution cannot, in general, be found.

In [67, 73, 75], fixing the first illumination coefficient to one is equivalent to defining a pixel-dependent scale factor equal to the first illumination coefficient. The approach of dividing the illumination-reflectance separation equation by the first illumination coefficient is also followed in the dichromatic-based color constancy algorithms such as [?, 80], with the additional assumption that the first illumination basis function is a constant multiple of the mean illumination.

In Horn's algorithm, the scale factor has the form of an unknown additive constant to the logarithm of the illumination and spectral reflectance. The author proposes to solve this problem by assuming that the highest color signal ("lightness" in the paper) value corresponds to 100% white. The same assumption is also made in the original paper by Land on Retinex color theory [68, 81]. A generalization of this approach is that of supposing that a single reflectance value in the image is known. In this case, it is sufficient to estimate the ratio between all other pixels and the one of known color to recover the scale factors. Another approach consists in assuming some statistic of the image reflectances is known. For example, the "gray world hypothesis" [82] consists in assuming the mean reflectance to be neutral. It is worth pointing out that assuming that two pixels have the same color (or are illuminated by the same light), as in [75], is not sufficient to recover the scale factor, as only the ratio between the pixel illuminants (colors) can be recovered. This is also why the scale factor must be estimated under the hypothesis of uniform illumination [67, 78, 82].

The main difference between our algorithm and the other illumination-reflectance separation methods present in the literature is that we try to recover the illumination intensity without requiring any knowledge about the reflectance of the image pixels, except that there are multiple reflectance materials in the image. As we will see in Section 3.5, the illumination spectral content will be assumed to be constant over the image. Similarly, we will not assume the illumination to be uniform, nor that some pixels have equal reflectance coefficients (this hypothesis would rest upon color observation, where we want to measure spectral reflectances). The only assumptions we will make are that the illumination varies slowly with respect to reflectance, and that its intensity can be approximated with a regular function of spatial coordinates.

3.3 Preliminary assumptions

Estimating illumination and spectral reflectance from a color signal is a very complex problem, for which no present algorithm has a general solution. A number of preliminary assumptions will therefore be introduced. Some of them are necessary to simplify the illumination-reflectance separation equation, and are those introduced in Section 3.1. They can be summarized as follows:

1. non-translucent, non-transparent materials (BRDF existence [61]);
2. no shadows;
3. no interreflections;
4. Lambertian materials.

Another set of assumptions is related to the physical nature of the problem:

5. spectral reflectance is never greater than one, nor less than zero:

$$0 \leq R(\mathbf{p}, \lambda) \leq 1, \ \forall \mathbf{p} \in \mathbb{P}. \tag{3.6}$$

6. Illumination is never less than the color signal, nor greater than a threshold dependent on the sensor:

$$S(\mathbf{p}, \lambda) \leq I(\mathbf{p}, \lambda) \leq \xi, \ \forall \mathbf{p} \in \mathbb{P}. \tag{3.7}$$

These are very weak assumptions, commonly found in the literature [73].

The last set of assumptions is then related to our specific application:

7. the spectral content of the illumination is constant over space up to a spatially changing scale factor. In other words,

$$I(\mathbf{p},\lambda) = k(\mathbf{p})I_0(\lambda),\ k(\mathbf{p}) \in \mathbb{R},\ \forall \mathbf{p} \in \mathbb{R}. \tag{3.8}$$

8. All surfaces are flat.

9. There are multiple colors in each image.

Assumption 7 cannot generally be applied to the real world, but is satisfied to a good extent in our case. This is due to the fact that in the presented applications, multispectral acquisitions have been performed under controlled illumination. The spectral content of lamp emissions strongly depends on the reflector structure [83]. Roughly speaking, if the reflector is "good" the spectral composition of the light is not modified along the lamp emission solid. Eq.(3.8) has been experimentally tested, and will be further discussed in Section 3.7.

Assumptions 8 and 9 are also related to the presented acquisition setup and specific application.

3.4 Linear model approximation

The problem of illumination-reflectance separation is generally addressed by assuming that reflectance and illumination belong to a finite linear space. Linear bases for both reflectance and illumination of real-world scenarios have been investigated. In our case, only reflectance linear spaces will be considered, as the illumination is assumed to be multiple of a single basis function (see assumption 7 in Section 3.3). For illumination finite linear models, see for example [84].

Let us indicate with $\mathcal{R}$ the set of the actual measured reflectances. Assuming that they belong to a finite-dimensional linear space means that is possible to write

$$R(\mathbf{p},\lambda) = \sum_{i=1}^{M} c_i(\mathbf{p})r_i(\lambda),\ c_i(\mathbf{p})\ \in\ \mathbb{R},\ \forall\ R\ \in\ \mathcal{R}, \tag{3.9}$$

where $\{r_i\}_{i=1,\ldots,N}$ are some known basis functions. Estimating the reflectance is thus equivalent to estimating the coefficients c_i.

Several works have attempted to define the correct dimension and basis for these spaces [85–89]. However, it is most likely that a unique answer does not exist, and that the choice must be done according to the chosen application. Among the factors influencing this choice, the considered spectral band, the target material and the desired accuracy surely play a dominant role.

For our problem we have chosen to use Parkkinen's eight basis functions [88]. They were obtained by performing the Karhunen-Loeve transform [90] on a total of 1257 reflectance spectra from the Munsell color chips [91]. The authors found that 98.4% of the spectra could be approximated using the first eight eigenvectors of the covariance matrix with an error less than 0.002[1], and that the overall mean error was 0.008. The eight basis vectors are plotted in Fig. 3.1.

From eqs.(3.4),(3.8) and (3.9), we have that the measured color signal at each point $\mathbf{p}$ and wavelength λ can be expressed as:

$$\begin{aligned} S(\mathbf{p},\lambda) &= k(\mathbf{p})I_0(\lambda)\sum_{i=1}^{N} c_i(\mathbf{p})r_i(\lambda) = \\ &= \sum_{i=1}^{N} \eta_i(\mathbf{p})I_0(\lambda)r_i(\lambda), \end{aligned} \tag{3.10}$$

where $\eta_i(\mathbf{p}) = k(\mathbf{p})c_i(\mathbf{p})$ and $N = 8$.

We can observe that eq.(3.10) is linear in the parameters η_i. Therefore, they can be estimated as the solution of a linear least-square problem of the form:

$$\hat{\boldsymbol{\eta}}(\mathbf{p}) = \arg\min_{\boldsymbol{\eta}} \left\| S(\mathbf{p},\lambda) - \sum_{i=1}^{N} \eta_i(\mathbf{p})I_0(\lambda)r_i(\lambda) \right\|^2 , \ \forall\ \mathbf{p} \in \mathbb{P} \tag{3.11}$$

where $\boldsymbol{\eta} = (\eta_i)_{i=1,\dots,N}$

One could wonder if there might be multiple solutions to this problem. Ho *et al.* [67] observed that the solution is unique if the functions $t_i(\lambda) = I_0(\lambda)r_i(\lambda)$ are linearly independent. This is always the case in our application.

Summarizing, our problem can be formulated as the estimation of a spatially varying coefficient $k(\mathbf{p})$ for the illumination, and of a set of coefficient $\{c_i(\mathbf{p})\}_{i=1,\dots,N}$ for reflectance, which satisfy:

$$S(\mathbf{p},\lambda) = k(\mathbf{p})I_0(\lambda)\sum_{i=1}^{N} c_i(\mathbf{p})r_i(\lambda) \tag{3.12}$$

s.t.

$$0 \le \sum_{i=1}^{N} c_i(\mathbf{p})r_i(\lambda) \le 1 \tag{3.13}$$

$$S(\mathbf{p},\lambda) \le k(\mathbf{p})I_0(\lambda) \le \xi. \tag{3.14}$$

3.5 Illuminant estimation

As we have seen in Section 3.2, it has often been observed that illumination and reflectance are characterized by different behaviors with respect to spatial frequency. Specifically, illumination has a low-pass behavior, i.e. is slowly-varying in space, while reflectance has a marked high-pass content due to sudden changes between adjacent colors. It can be noted that the assumption of slowly-varying illumination is particularly fit to our application, where exclusively controlled illumination is used, surfaces are flat and there are no shadows. Therefore, the presented algorithm will exploit this different spatial frequency behavior to try to separate the two components.

---

[1] The error was defined as the absolute difference between real and approximated spectra.

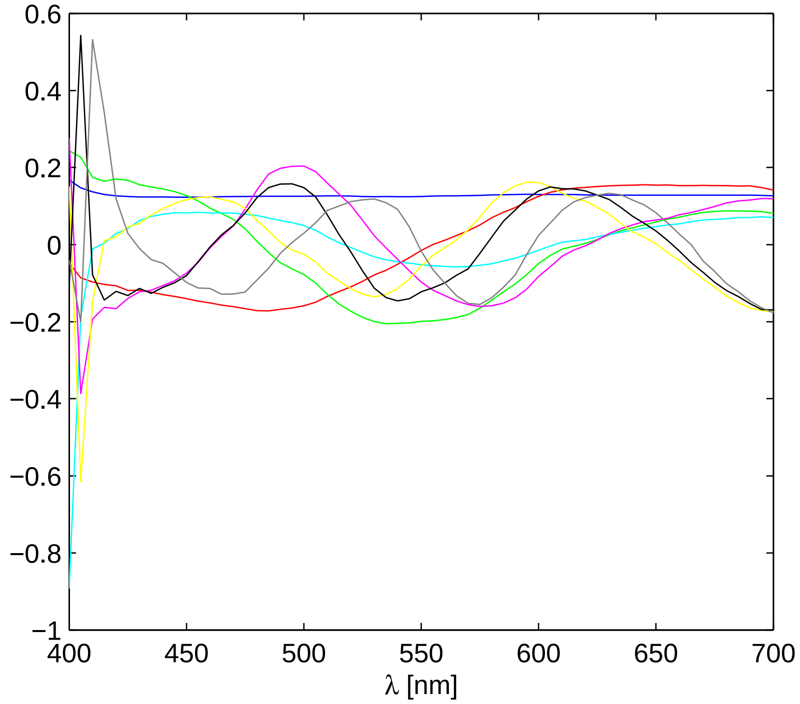


Figure 3.1: Parkkinen's basis vectors for reflectance.

Another useful observation is that the low-pass signal described by $k(\mathbf{p})$ can be approximated as a smooth surface. This surface is the projection of the emission solid of the lamp onto the target surface, i.e. a plane. The emission solid of the lamp is unknown. However, we can form an idea about its shape by observing the reflector. For example, if we think the emission solid can be approximated as a paraboloid, the intensity of the light irradiating a flat surface will be a paraboloid too. In our case, due to the simple surface shape, we will assume the illumination intensity to lie on a plane and we will try to fit a plane to an appropriate function of the observed data.

3.5.1 The lower bound equation

Eq.(3.12) expresses a relation between functions of both spatial position and wavelength. To separate the coefficients of reflectance and illumination, it would be simpler to find a relation between functions of space only. To do so, we reformulate condition (3.14) as:

$$\frac{S(\mathbf{p},\lambda)}{I_0(\lambda)} \leq k(\mathbf{p}) \leq \frac{\xi}{I_0(\lambda)}, \ \forall\lambda \ \in \Lambda, \tag{3.15}$$

where we have assumed $I_0(\lambda) > 0 \ \forall\lambda$.

As eq.(3.15) must hold for every wavelength, we can further rewrite it as

$$\max_{\lambda} \frac{S(\mathbf{p},\lambda)}{I_0(\lambda)} \leq k(\mathbf{p}) \leq \min_{\lambda} \frac{\xi}{I_0(\lambda)} = \frac{\xi}{\max_{\lambda} I_0(\lambda)}. \tag{3.16}$$

The lower bound for $k(\mathbf{p})$ is therefore

$$\max_{\lambda} \frac{S(\mathbf{p}, \lambda)}{I_0(\lambda)}. \tag{3.17}$$

Using again eqs.(3.4),(3.8) we get:

$$\begin{aligned} \mathcal{L}(\mathbf{p}) &\triangleq \max_{\lambda} \frac{S(\mathbf{p}, \lambda)}{I_0(\lambda)} = \max_{\lambda} \frac{k(\mathbf{p}) I_0(\lambda) R(\mathbf{p}, \lambda)}{I_0(\lambda)} = \\ &= k(\mathbf{p}) \max_{\lambda} R(\mathbf{p}, \lambda) = k(\mathbf{p}) \chi(\mathbf{p}), \end{aligned} \tag{3.18}$$

where

$$\chi(\mathbf{p}) \triangleq \max_{\lambda} R(\mathbf{p}, \lambda). \tag{3.19}$$

The equation we will use to estimate the illuminant will therefore be

$$\mathcal{L}(\mathbf{p}) = k(\mathbf{p}) \chi(\mathbf{p}), \tag{3.20}$$

and we will call it the *lower bound equation.* It can be noted that $\mathcal{L}(\mathbf{p})$ can be easily computed from the input color signal and the known illuminant basis vector.

3.5.2 Plane fitting

We would like to estimate the illumination intensity $k(\mathbf{p})$ as the best plane fitting our data in the least-square sense. The problem is that the illumination values are available only multiplied by the variable $\chi(\mathbf{p})$, which satisfies the inequalities

$$0 \leq \chi(\mathbf{p}) \leq 1, \ \forall \mathbf{p} \ \in \ \mathbb{P}. \tag{3.21}$$

Forgetting eq.(3.19), we model $\chi(\mathbf{p})$ as a random variable of some known distribution. For example, in the simplest case, we could assume it to be a uniformly distributed r.v., i.e.

$$\chi(\mathbf{p}) \sim \mathcal{U}(0, 1) \ \forall \mathbf{p} \ \in \ \mathbb{P}.$$

RANSAC [26] is a robust iterative algorithm to estimate the parameters of a model from a set of data with many outliers. The procedure is summarized in Algorithm 2.

The maximum number N of iterations is usually selected as that which guarantees that at least one set of n inliers has been selected with probability p. If we denote by w the probability that any selected data point is within the error tolerance of the model, one can prove that:

$$N = \frac{\log(1 - p)}{\log(1 - w^n)}. \tag{3.22}$$

Similarly, the threshold for the cardinality of the consensus set is chosen according to the supposed percentage of inliers as:

$$t_C = N_D w, \tag{3.23}$$

```
Input:
data - a set of observations;
$\theta$ - a model that can be fitted to data;
$n$ - the minimum number of data required to fit the model;
$N$ - the maximum number of iterations allowed in the algorithm;
$t_\theta$ - threshold to decide if a datum is an inlier;
$t_C$ - threshold to decide if there is a sufficient number of inliers for a model;
Output:
$\hat{\theta}$ - best model fitted to data;
$\hat{\mathcal{C}}$ - best consensus set;
begin
    iter = 0;
    repeat
        choose $n$ random points from the data set;
        compute the model parameters $\theta$;
        foreach point not part of the randomly chosen set do
            if the distance of point from the fitted model is less than $t_\theta$ then
                add point to the consensus set $\mathcal{C}$;
        if $|\mathcal{C}| > |\hat{\mathcal{C}}|$ then
            $\hat{\theta} = \theta$, $\hat{\mathcal{C}} = \mathcal{C}$;
        if $|\hat{\mathcal{C}}| > t_C$ then
            break;
    until iter $\geq N$ ;
end
iter = iter+1;
```

Algorithm 2: The original RANSAC algorithm

with N_D cardinality of the dataset.

The model underlying the RANSAC algorithm is that of a signal corrupted by an additive random noise, and therefore the algorithm cannot be used in its original form in this application. We have then defined a variation of this method, which is described in Algorithm 3, and discussed in the following.

The key difference between the classical RANSAC formulation and our problem is that in the latter case the model cannot be directly fitted to data. In fact, the data that should be fitted are the values of $k(\mathbf{p})$, which are not available. Knowing the true values of $\chi(\mathbf{p})$, and therefore of $k(\mathbf{p})$, would be equivalent to having the solution of our problem. Therefore, the values for $k(\mathbf{p})$ must be derived from eq.(3.20) in some other way. For the moment, we give up the estimation of reflectance and regard $\chi(\mathbf{p})$ as a multiplicative random noise corrupting the signal $k(\mathbf{p})$. We then produce a realization of this noise at each point $\mathbf{p}$ according to its alleged distribution, and derive the supposed value for $k(\mathbf{p})$ as:

$$\hat{k}(\mathbf{p}) = \frac{\mathcal{L}(\mathbf{p})}{\hat{\chi}(\mathbf{p})},$$

where $\hat{\chi}(\mathbf{p})$ is a realization of the random variable. Looking at Algorithm 3, it is possible to see that the problem of zero values for $\hat{\chi}(\mathbf{p})$ is avoided by restricting the interval to $[\epsilon, 1-\epsilon]$, $\epsilon > 0$. This assumption is quite weak, as it is rather safe

```
Input:
data - a set of observations;
$\theta$ - a model that can be fitted to data;
$n$ - a number greater than or equal to the minimum number of data
required to fit the model;
$N$ - the maximum number of iterations allowed in the algorithm;
$t_\theta$ - threshold to decide if a datum is an inlier;
$t_C$ - threshold to decide if there is a sufficient number of inliers for a model;
$(r_k)_{k=1}^{K}$ - $K$ uniform samples of the $[\epsilon, 1-\epsilon]$ interval, $\epsilon > 0$;
Output:
$\hat{\theta}$ - best model fitted to data;
$\hat{\mathcal{C}}$ - best consensus set;
begin
    subsample $\mathcal{L}(\mathbf{p})$;
    iter = 0;
    repeat
        choose $n$ random points $(\mathcal{L}_l)_{l=1}^{n}$ from the subsampled data set;
        choose $n$ random values $(\chi_l)_{l=1}^{n}$ between $\epsilon$ and $(1-\epsilon)$ according to a
        known probability distribution;
        compute $k_l \triangleq \mathcal{L}_l/\chi_l$, $l = 1, \ldots, n$;
        compute the model parameters $\theta$ from $(k_l)_{l=1}^{n}$;
        define $\hat{k}(\mathbf{p}) = \theta^T \mathbf{p}$, $\mathbf{p} \in \mathbb{P}$;
        compute $\hat{\chi}(\mathbf{p}) = \mathcal{L}(\mathbf{p})/\hat{k}(\mathbf{p})$;
        if $\hat{k}(\mathbf{p})$ satisfies (3.14) $\forall \mathbf{p} \in \mathbb{P}$ and $\hat{\chi}(\mathbf{p})$ satisfies (3.13) $\forall \mathbf{p} \in \mathbb{P}$
        then
            foreach point of $\mathbb{P}$ not part of the randomly chosen set do
                foreach $r_k$ do
                    if the distance of point from the fitted model multiplied
                    by $r_k$ is less than $t_\theta$ then
                        add point to the consensus set $\mathcal{C}$;
            if $|\hat{\mathcal{C}}| > t_C$ then
                break;
    until iter $\geq N$ ;
end
iter = iter+1;
```

Algorithm 3: A variation of the RANSAC algorithm

to assert that no natural material has spectral reflectance which is everywhere 0, or somewhere exactly 1.

A rather strong assumption is that the $\chi(\mathbf{p})$s can be treated as independent random variables. This is certainly not the case, as in real images reflectance functions are correlated in space. In order to get closer to the fulfilment of this assumption we therefore subsample the image $\mathcal{L}(\mathbf{p})$. As $k(\mathbf{p})$ has a low-pass behavior, this will not compromise its reconstruction. On the other hand, we will decrease the correlation between the $\chi(\mathbf{p})$ values. We have usually subsampled $\mathcal{L}(\mathbf{p})$ by a factor of 5 along both dimensions.

At this point, we can fit our model to the supposed inliers. Note that in this case a better definition of inlier would be that of a point $\hat{k}(\mathbf{p})$ which is close to the real value $k(\mathbf{p})$. In other words, it is the algorithm itself which creates the outliers by guessing wrongly the maximum reflectance value at a certain point.

Once we have computed a candidate model, the first thing we need to do is check whether it represents a valid solution to our problem. In other words, we have to verify if it satisfies condition (3.14). Moreover, once we have a candidate model we can also compute the candidate values for $\chi(\mathbf{p})$. Therefore, we also need to check whether condition (3.13) holds.

We are now left with the problem of rating a fitted model. The problem is always the same: we only have the illumination intensity data multiplied by the corresponding maximum reflectance value. Intuitively, we could say that a good model can be identified by considering the "cumulative" consensus set of all the points that are close to the model when multiplied by some value between 0 and 1. This is also described in Algorithm 3. To find the cumulative consensus set we sample the feasible interval for $\chi(\mathbf{p})$ with K values. We then multiply the fitted model for each value, and define the k-th consensus set as formed by all the points which differ from the scaled plane less than a given threshold. The final consensus set will be given by the union of all the K partial consensus sets.

Unfortunately, it is difficult to define a suitable threshold, as the threshold is part of the problem itself. This is easily seen by considering that if we knew $k(\mathbf{p})$ and $\chi(\mathbf{p})$, taking K uniformly spaced values of reflectance between ϵ and $(1-\epsilon)$ would be equivalent to quantizing $\chi(\mathbf{p})$ with a certain step Δ. As all the error is granular (see [92]), we would have that

$$|Q[\chi(\mathbf{p})] - \chi(\mathbf{p})| \leq \frac{\Delta}{2},$$

where $Q[\cdot]$ indicates the quantization operator. Therefore,

$$|k(\mathbf{p})Q[\chi(\mathbf{p})] - \mathcal{L}(\mathbf{p})| \leq k(\mathbf{p})\frac{\Delta}{2},$$

i.e. the true threshold is a function of the unknown values $k(\mathbf{p})$. However, using the physically-based inequality (3.14), we can say

$$|k(\mathbf{p})Q[\chi(\mathbf{p})] - \mathcal{L}(\mathbf{p})| \leq k(\mathbf{p})\frac{\Delta}{2} \leq \frac{\xi}{\max_\lambda I_0(\lambda)}\frac{\Delta}{2},$$

or in other words

$$t_\theta \triangleq \frac{\xi}{\max_\lambda I_0(\lambda)}\frac{\Delta}{2}. \tag{3.24}$$

Another critical choice concerns the cardinality of the random sample set selected to estimate a candidate model. Considering that our problem is severely under-determined, it could seem advisable to choose a high number of samples to make the estimate more robust. However, the probability of selecting feasible values for $\chi(\mathbf{p})$ quickly decreases with the number of selected points. If we consider $\chi(\mathbf{p})$ as a uniform discrete r.v. taking values in the set $\{r_1, \ldots, r_{\tilde{K}}\}$, we have that the probability of selecting an inlier (i.e. choosing a feasible value for $\chi(\mathbf{p})$) is:

$$w = \frac{1}{\tilde{K}}.$$

Therefore, with reference to equation (3.22), to be sure of selecting n inliers with probability p we should iterate our algorithm for a number of iterations N given by:

$$N = \frac{\log(1-p)}{\log(1-\tilde{K}^{-n})}.$$

With $\tilde{K} = 100$, $n = 3$ (i.e. the minimum number of points to fit a plane) and $p = 0.9$, we have that N is greater than $2 \cdot 10^6$. For the sake of efficiency, we will thus choose a value for N which does not guarantee the presence of an all-inlier set. Therefore, the resulting solution will possibly be sub-optimal.

As concerns the parameter t_C, the definition (3.23) somewhat loses its meaning. In our case, the true percentage of inliers should be 100%. Therefore, t_C has been considered as a tolerance value to be contented with. We have usually set it to $90 \div 95\%$ of the total number of points.

3.6 Reflectance recovery

Once we have estimated the best illumination coefficient $\hat{k}(\mathbf{p})$, recovering reflectance is almost trivial. Remembering Section 3.4, we can estimate the parameters

$$\eta_i(\mathbf{p}) = k(\mathbf{p})c_i(\mathbf{p})$$

for all points $\mathbf{p}$ and $i = 1, \ldots, M$ (M is the number of basis functions) by a simple linear least-square method. Therefore, the estimated reflectance coefficients can be computed as:

$$\hat{c}_i(\mathbf{p}) = \frac{\eta_i(\mathbf{p})}{\hat{k}(\mathbf{p})}, \quad i = 1, \ldots, M, \quad \mathbf{p} \in \mathbb{P}. \tag{3.25}$$

Note that due to the way the variation of the RANSAC algorithm has been defined (Algorithm 3), the estimated coefficients $\hat{c}_i(\mathbf{p})$ are guaranteed to satisfy condition (3.13).

3.7 PCA of the illumination spectra

In order to find a suitable basis for the illumination, we have followed the classical approach of performing Principal Component Analysis (PCA) [93] on a large dataset of illumination spectra. PCA has been extensively used in Color Science with a variety applications, such as data reduction, color estimation from mixtures or confidence interval definition [94]. As observed by Lenz [95], a common practice when dealing with color spectra (where "color" refers both to illumination and reflectance) is to compute a basis for the space of spectral distributions using the eigenvectors of the auto-correlation matrix. This is in contrast with the usual version of the algorithm, which considers the auto-covariance matrix, i.e. the auto-correlation of the spectra after the global mean has been subtracted. However, as color spectra are always non-negative functions, the former approach seems more reasonable. Moreover, Lenz has shown that if the mean of the dataset is identical to a constant multiple of the first eigenvector, both algorithms result in the same eigenfunctions [95][2].

---

[2] If this is not the case, nothing can be said about the relationship between the two bases.

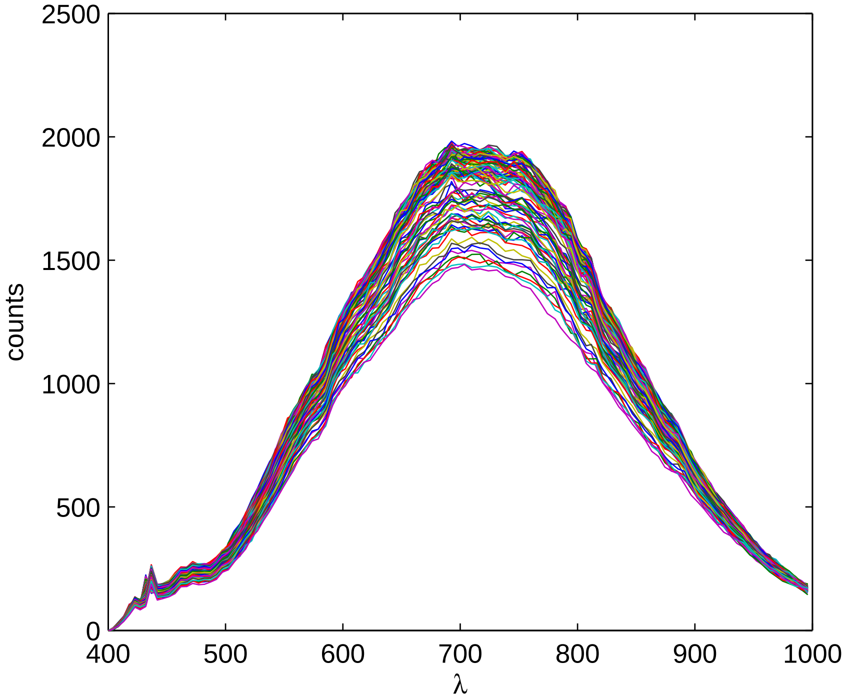


Figure 3.2: An example of illumination dataset.

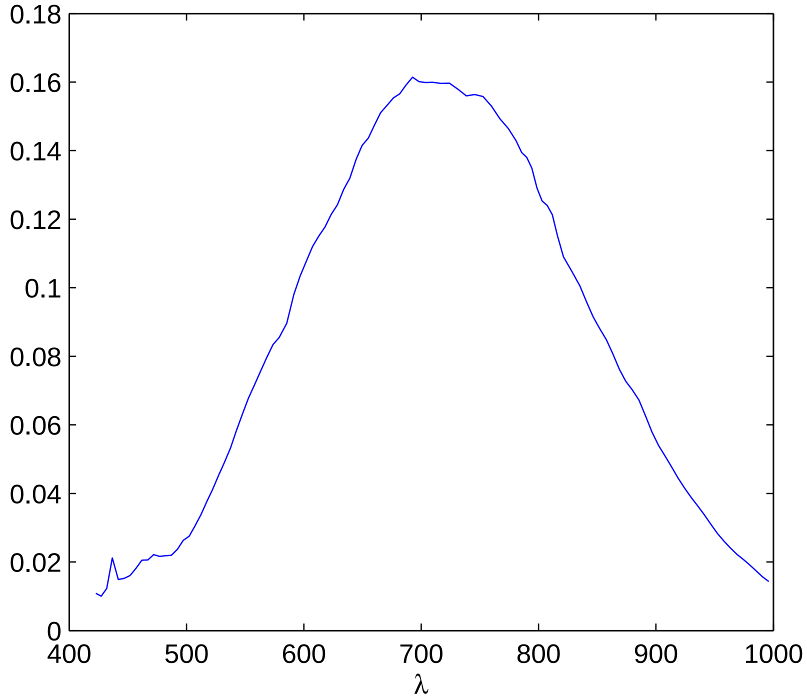


Figure 3.3: The first basis vector for the illumination.

Covariance-based PCA [96] is often formulated as the problem of finding the orthogonal transformation which projects the dataset onto a reference system where the coordinates are associated to directions of decreasing variance (the first coordinate is along the direction of maximum variance). The first principal component is the unitary vector which satisfies the maximization problem:

$$\mathbf{b}_1 = \arg\max_{\mathbf{b}} \mathrm{var}(\tilde{\mathbf{X}}\mathbf{b}), \text{ s.t. } \mathbf{b}^T\mathbf{b} = 1,$$

where $\mathbf{X} \in \mathbb{R}^{n\times m}$ is the dataset matrix, in which each row represents an observation and each column a variable, and the apex $\tilde{}$ indicates that each column has been centered around its mean. It can be easily shown that the solution to this problem is the eigenvector of the covariance matrix $\mathbf{K} = \tilde{\mathbf{X}}^T\tilde{\mathbf{X}}$ corresponding to its largest eigenvalue (greatest singular value of $\tilde{\mathbf{X}^T}$).

The subsequent principal components are the solutions of a problem of the form:

$$\mathbf{b}_n = \arg\max_{\mathbf{b}} \mathrm{var}(\tilde{\mathbf{X}}\mathbf{b}), \text{ s.t. } \mathbf{b}^T\mathbf{b} = 1,\ \mathbf{b}^T\mathbf{b}_i = 0,\ i = 1, \ldots, n-1.$$

If we substitute the covariance matrix with the correlation matrix, we can regard the first principal component as the unitary vector transformation which maximizes the squared norm of the dataset, i.e.

$$\mathbf{b'}_1 = \arg\max_{\mathbf{b}} ||\mathbf{X}\mathbf{b}||^2, \text{ s.t. } \mathbf{b}^T\mathbf{b} = 1.$$

As $||\mathbf{X}\mathbf{b}||^2 = \mathbf{b}^T\mathbf{X}^T\mathbf{X}\mathbf{b} = \mathbf{b}^T\mathbf{C}\mathbf{b}$, with $\mathbf{C}$ auto-correlation matrix, the problem has a structure identical to the previous one, and $\mathbf{b'}_1$ is the eigenvector relative to the greatest singular value of $\mathbf{X}^T$. As before, subsequent vectors can be found as

$$\mathbf{b}'_n = \arg\max_{\mathbf{b}} ||\mathbf{X}\mathbf{b}||^2, \text{ s.t. } \mathbf{b}^T\mathbf{b} = 1,\ \mathbf{b}^T\mathbf{b}'_i = 0,\ i = 1, \ldots, n-1.$$

Fig. 3.2 shows the illumination spectra of a portion of white paper acquired with our imaging spectrograph under controlled illumination. It is quite evident that all the spectra are more or less proportional to each other. This is confirmed by PCA. Given the set of singular values of the dataset matrix $(\sigma_\ell^2)_{\ell=1}^L$, where L is the original number of variables (in this case, one for each of the 112 sampled wavelengths), it is possible to evaluate the relative mean squared approximation error as:

$$\epsilon = \frac{\sum_{\ell=2}^L \sigma_\ell^2}{\sum_{\ell=1}^L \sigma_\ell^2}. \tag{3.26}$$

In this case, the error is 0.01%. This means that assumption 7 of Section 3.3 is a very close approximation of the real illumination behavior.

As for the relation between the correlation-based and covariance-based approaches, we have computed the mean vector $\mathbf{m}$ of our dataset and calculated its correlation coefficient with the first eigenvector of the correlation matrix as

$$\rho = \frac{<\mathbf{b}, \mathbf{m}>}{||\mathbf{b}||||\mathbf{m}||}, \tag{3.27}$$

where $\mathbf{b}$ is the basis vector. We have found $\rho = 1$, which means that the two vectors are exactly proportional to each other. The proportionality factor can be found as [97]:

$$\kappa = \frac{\sum_{l=1}^L b_l m_l}{\sum_{l=1}^L b_l^2}.$$

We have found $\kappa = 1.6617 \cdot 10^6$. The comparison between $\mathbf{m}$ and $\kappa\mathbf{b}$ is shown in Fig. 3.4. We can conclude that the mean vector and the first eigenvector are indeed related by a constant multiplicative factor, and therefore that in this case the correlation- and covariance-based approaches are equivalent.

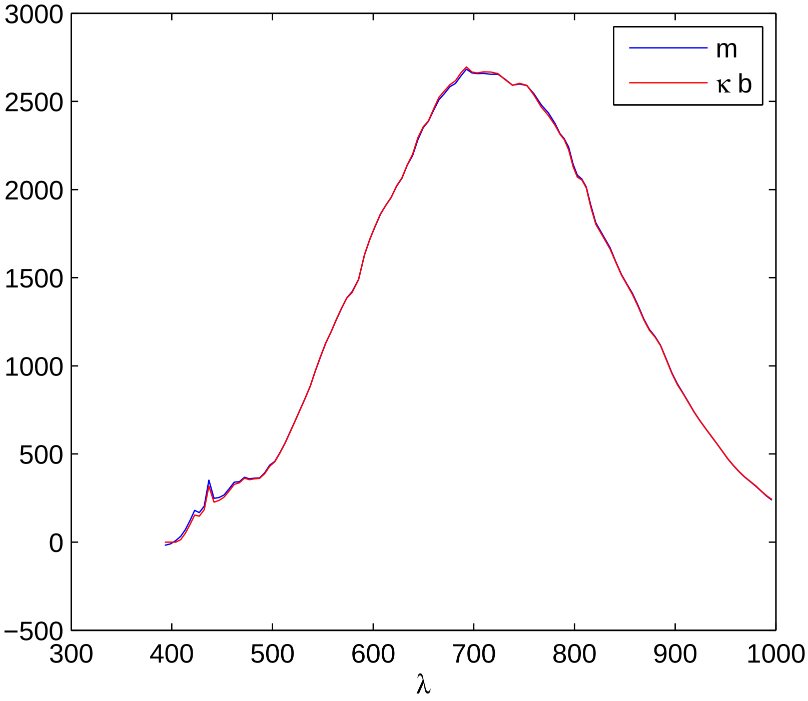


Figure 3.4: Proportionality between mean and first eigenvector of the illumination dataset.

Chapter 4

Illumination-reflectance separation on real data

We have tested our illumination-reflectance separation algorithm on four sets of data. The first represents a portion of a fresco of the Castello del Buonconsiglio in Trento (Italy). The other three have been obtained from three paintings by two contemporary artists from Padova. The paintings present a wide range of pictorial techniques and color dynamics. The acquisitions have been performed according to the procedure described in Chapter 2, using as the reference white a white panel with measured reflectance of 95%.

4.1 Case studies

4.1.1 "The dove"

The first presented case study is a portion of a fresco painted by Girolamo di Romano, known as "Romanino", in the Castello del Buonconsiglio in Trento (Italy). Romanino was born in 1485 in Brescia (Italy), a city in which one of the most important north Italian schools of Renaissance painting had developed. During his youth, he fell under the spell of Venetian painting, especially that of Giorgione and Titian, and of Milanese painting. As he matured, however, he developed a very personal style, drawing inspiration especially from the very dramatic pictorial style of German art, as demonstrated in the magnificent cycle of paintings for Cremona cathedral (1519). Romanino was a versatile artist; he painted on panels and on canvas, but he favoured the technique of fresco, the means of expression which he found most congenial. In addition, Vasari counted him among the most capable draughtsmen of his time. In 1531 he offered to decorate the Castello del Buonconsiglio for the Prince Bishop of Trento, Bernard Cles, and completed a large cycle of paintings, with secular themes, in the castle.

The acquired scene is a portion of a fresco representing a rest after the hunt, with a servant (on the right) leaning towards his master, holding a dove on his arm. We will call this acquired portion "The dove". The fresco is situated in a vault of the "Volto sotto la Loggia", one of the rooms of Magno Palazzo, and its dimensions are approximately $2\,\mathrm{m}\times 2\,\mathrm{m}$. A picture of the acquisition setup can be seen in Fig. 4.1. As mentioned in Chapter 1, the acquisition of the white signal is particularly difficult when acquiring a vault. Besides, we

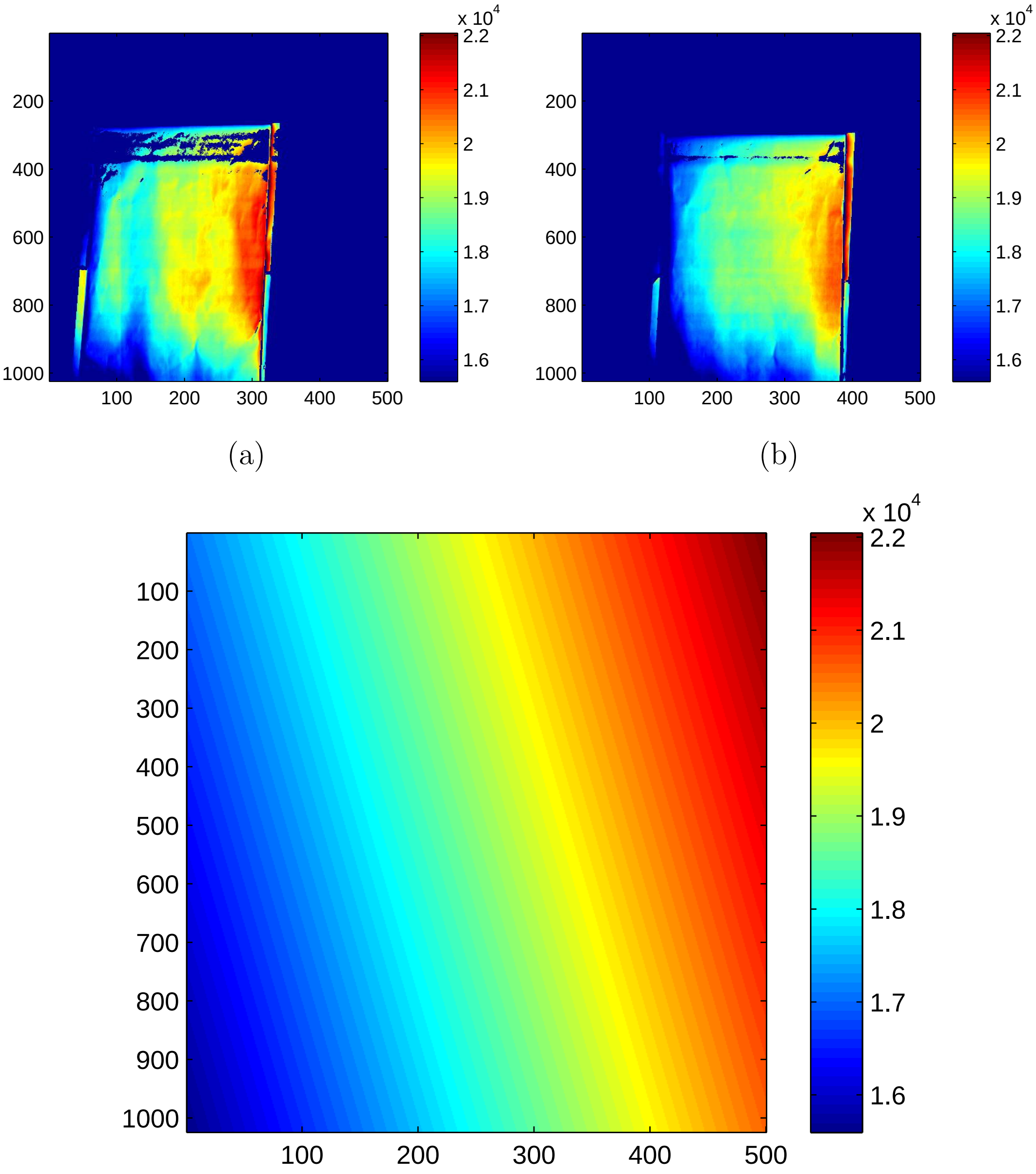
Figure 4.1: Acquisition setup in the Castello del Buonconsiglio, Trento (Italy).

did not possess a sufficiently large white panel to cover the whole scanned area. We therefore acquired the white signal in two different steps, placing the white panel in different, but partially overlapping, positions. After calculating the $k(\mathbf{p})$ coefficients corresponding to both white areas, the final reference white was obtained by fitting a plane to the obtained surfaces, and interpolating it to the whole image plane. The estimated k coefficients for the two white panel positions are shown in Fig. 4.2. It is clearly unrealistic to assume that these are the actual illumination coefficients on the left part of the vault, which lies on on a different, non-planar surface. However, as no paintings were present in that region, we gave up the acquisition of the actual white signal for the presented study, and used one single interpolated surface for the illumination coefficients. As the interpolated white signal satisfies the physical constraints of eq.(3.15) on all its support, the corresponding reflectance values are feasible, if not correct.

The (R,G,B) reconstruction of the acquired fresco can be seen in Fig. 4.3. The coefficients obtained by projecting the acquired spectral reflectance functions onto the space spanned by Parkkinen's eight basis functions are shown in Fig. 4.4. The least-square projection error is 2.7%.

4.1.2 "Tulips"

"Tulips" (oil on wood) is a painting by Matteo Massagrande. Matteo Massagrande (Padova, Italy, 1959) is a contemporary Italian painter and engraver.



(a) (b)

(c)

Figure 4.2: (a) k coefficient corresponding to the white panel in position 1; (b) k coefficient corresponding to the white panel in position 2; (c) interpolated planar surface for the coefficients $k(\mathbf{p})$.

His interests lie in the study of ancient pictorial techniques, of engraving and in the art of restoration. He frequently travels throughout as well as out of Europe, and his journeys are often occasions to develop pictorial cycles and great compositions. His first exhibition dates back to 1973. Since then, he has exhibited his artworks in several collective and personal exhibitions, and has been awarded several prizes. He currently lives in Padova, and works between his study in Padova and that in Haiós (Hungary) [98].

The (R,G,B) rendition of the painting, computed from the acquired spectral reflectance, is represented in Fig. 4.5.

The coefficients obtained by projecting the illumination data into the space

Figure 4.3: Acquired portion of “The Hunt”, painted by Romanino in the Castello del Buonconsiglio in Trento (Italy).

spanned by the basis vector $I_0(\lambda)$ defined in Section 3.7 are shown in Fig. 4.7(a). The mean relative error of the fit, defined as

$$e = \frac{1}{|\mathbb{P}|} \sum_{\mathbf{p} \in \mathbb{P}} \frac{||k(\mathbf{p}) I_0(\lambda) - I(\mathbf{p}, \lambda)||}{||I(\mathbf{p}, \lambda)||}, \tag{4.1}$$

is 1.03%. Fig. 4.6(b) shows the best plane fitted to the true illumination coefficients in the least-square sense. The mean relative error with respect to the fitted surface $I(\mathbf{p}, \lambda)$, defined as in eq.(4.1), is 2.23%. The approximation of the illumination intensity as a planar surface is therefore very good. This is due to the choice of the acquisition setup, with the multispectral camera normal to the surface and the illumination source at approximately 45° with respect to the painting surface normal, so as to minimize the reflections [44].

The coefficients of the reflectance relative to Parkkinen’s basis functions are shown in Fig. 4.7. The relative projection error is 2.82%.

4.1.3 “Peach”

We have acquired another painting by Matteo Massagrande, “Peach” (oil on wood). The (R,G,B) reconstruction of the painting can be seen in Fig. 4.8.

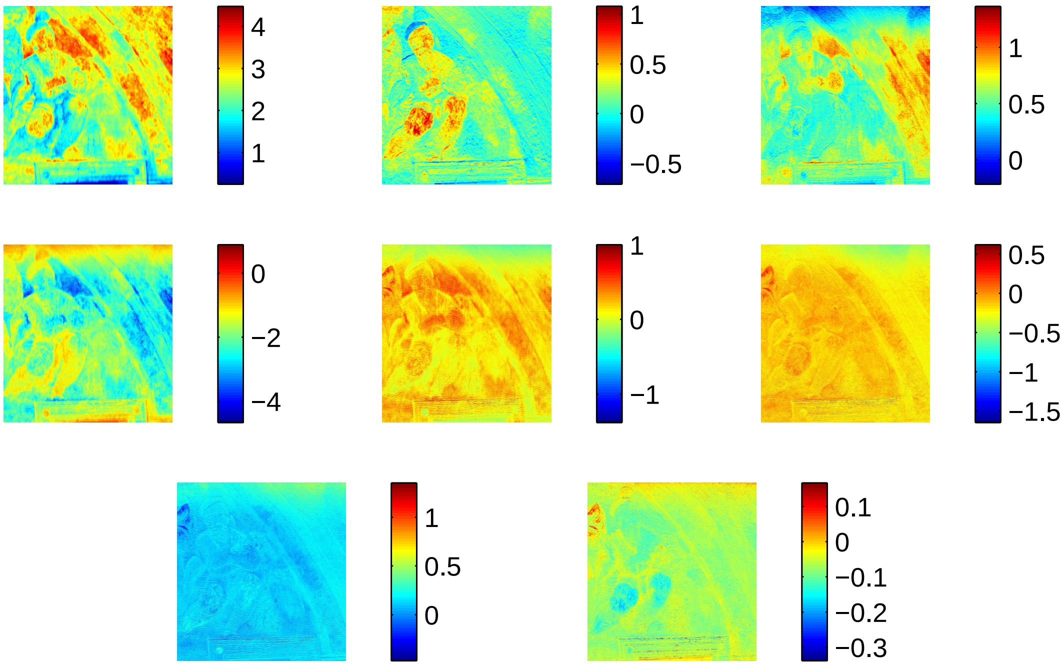


Figure 4.4: “The Dove”: reflectance coefficients for Parkkinen’s basis.

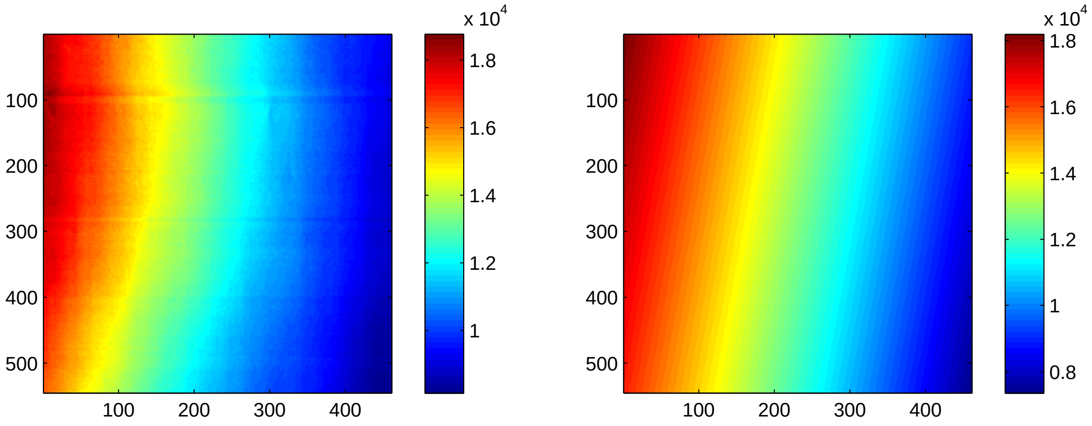
This painting was acquired using the same illumination setup as for “Tulips”. Therefore, for the analysis of the illumination characteristics we refer to Subsection 4.1.2. The fitted reflectance coefficients can be seen in Fig. 4.9. The mean fit error is 2.89%.

4.1.4 “Bull”

The last acquired painting was “Bull” (oil on canvas) by Vittorio Buzzanca, another Italian contemporary painter, which can be seen in Fig. 4.10. The real illumination coefficients and the best fitted plane can be seen in Fig. 4.11. The mean fit error is 2.17%, so we can say that also in this case the planar assumption works quite well.

The fitted reflectance coefficients for Parkkinen’s basis functions can be seen in Fig. 4.12. The relative mean fit error was found to be 2.96%.

Figure 4.5: “Tulips”, by Matteo Massagrande (oil on wood).



(a)

(b)

Figure 4.6: “Tulips”: real coefficient for the illumination (a) and best fitted plane (b).

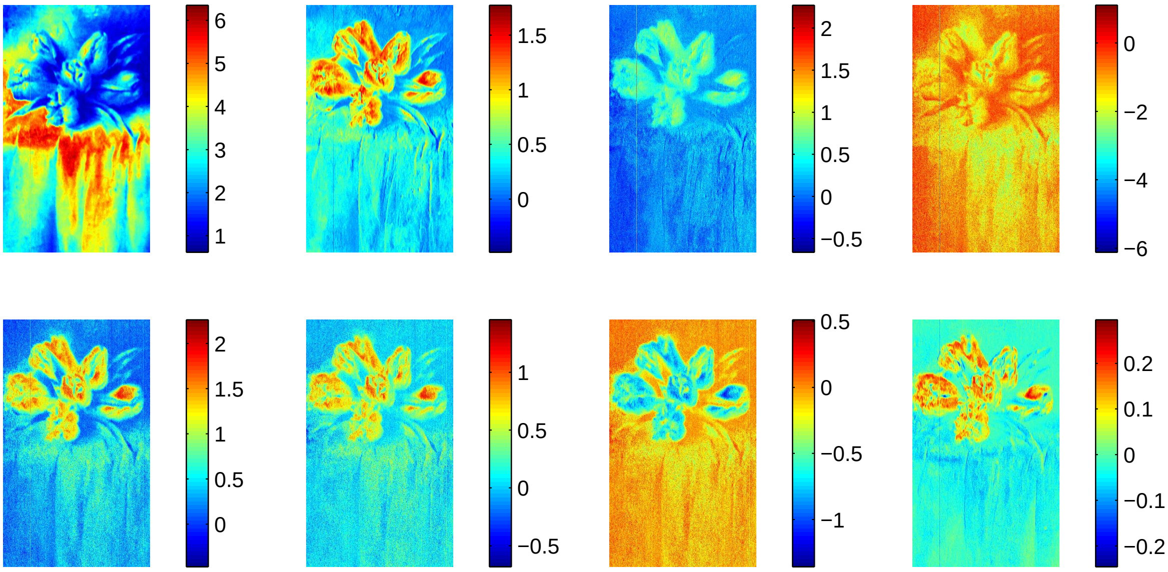


Figure 4.7: “Tulips”: reflectance coefficients for Parkkinen’s basis.

Figure 4.8: “Peach”, by Matteo Massagrande (oil on wood).

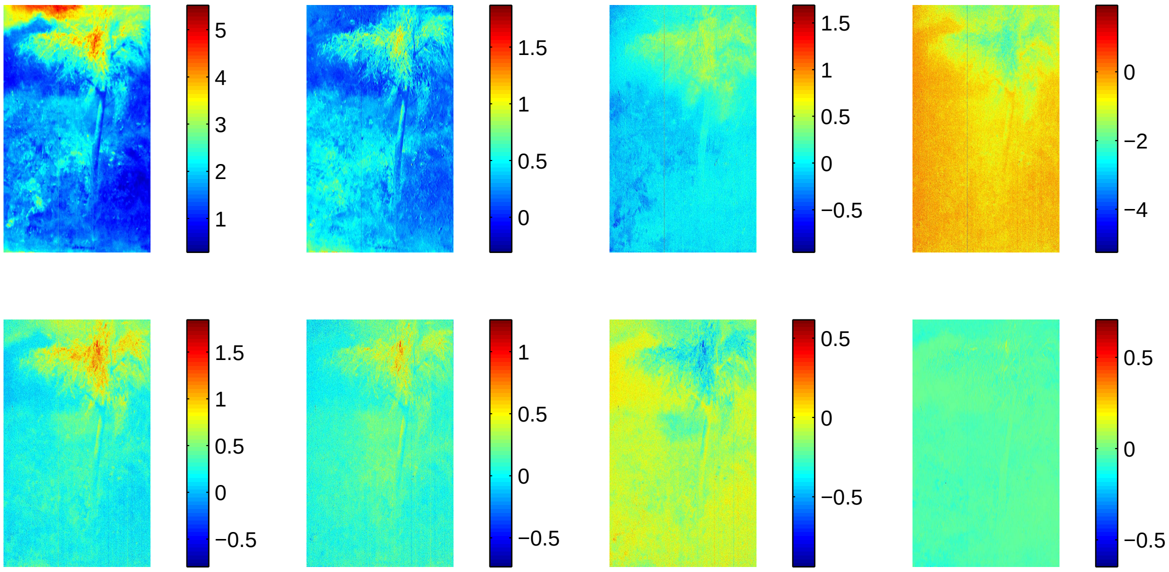
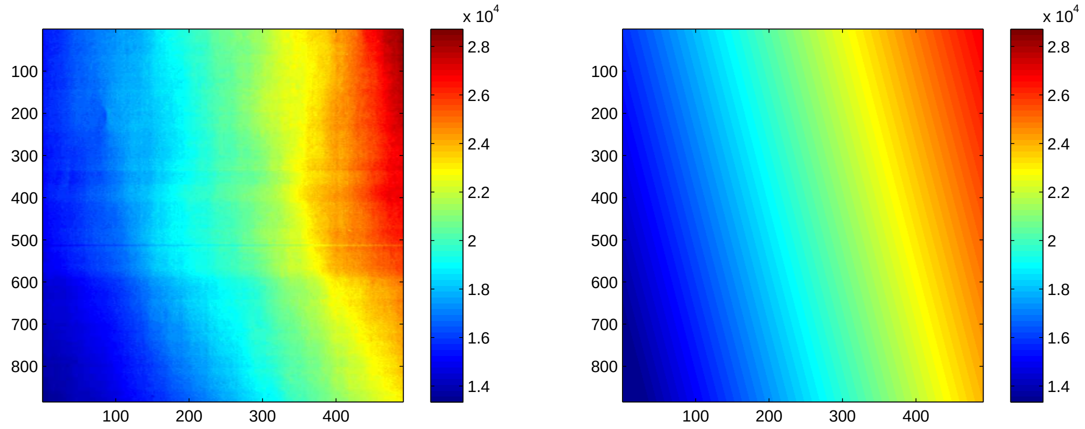


Figure 4.9: “Peach”: reflectance coefficients for Parkkinen’s basis.

Figure 4.10: “Bull”, by Vittorio Buzzanca (oil con canvas).



(a) (b)

Figure 4.11: "Bull": real coefficient for the illumination (a) and best fitted plane (b).

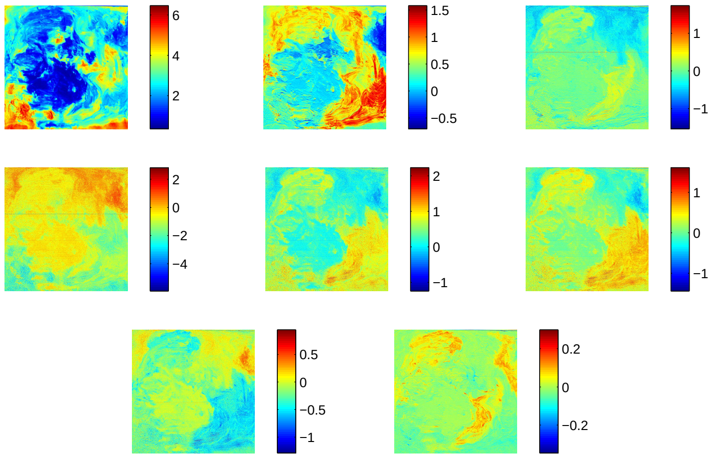


Figure 4.12: "Bull": reflectance coefficients for Parkkinen's basis.

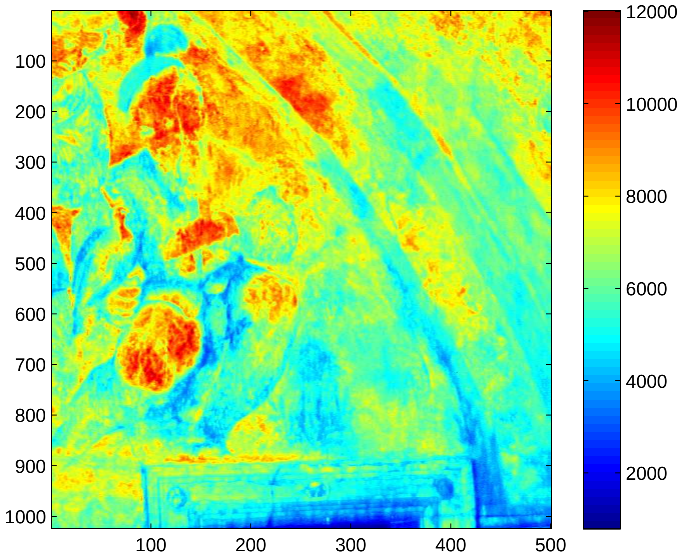


Figure 4.13: The function $\mathcal{L}(\mathbf{p})$, maximum lower bound for the illumination, for "The dove".

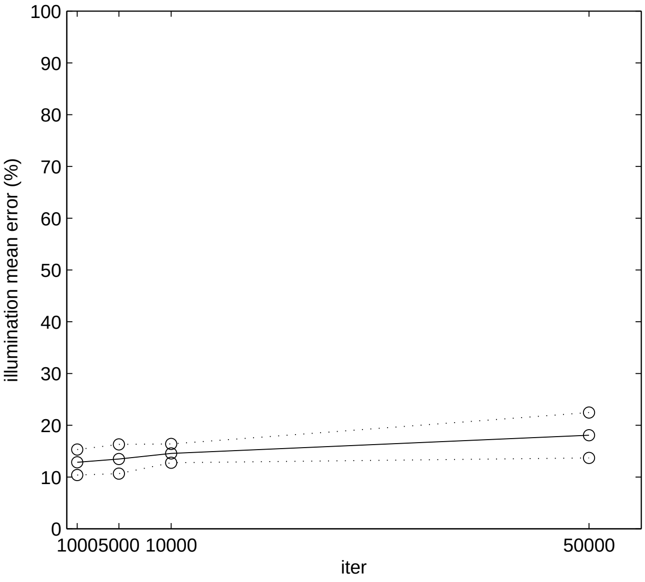


Figure 4.14: "The dove": variation of the mean error on the illumination ± its standard deviation with the number of iterations of the modified RANSAC algorithm 3.

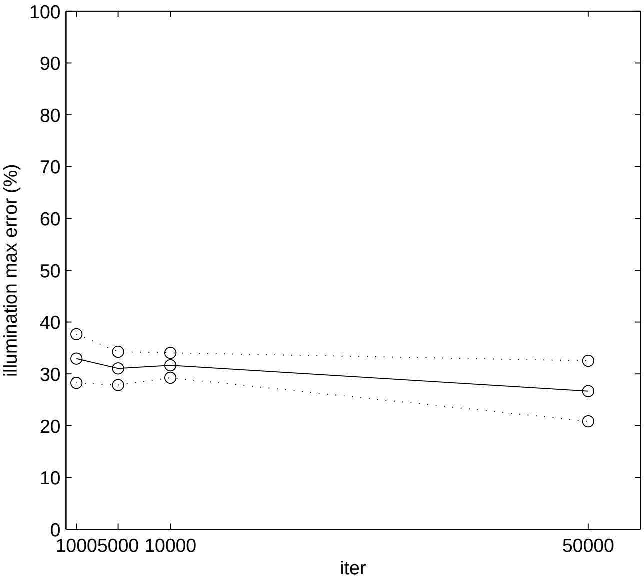


Figure 4.15: “The dove”: variation of the maximum error ± its standard deviation on the illumination with the number of iterations of the modified RANSAC algorithm 3.

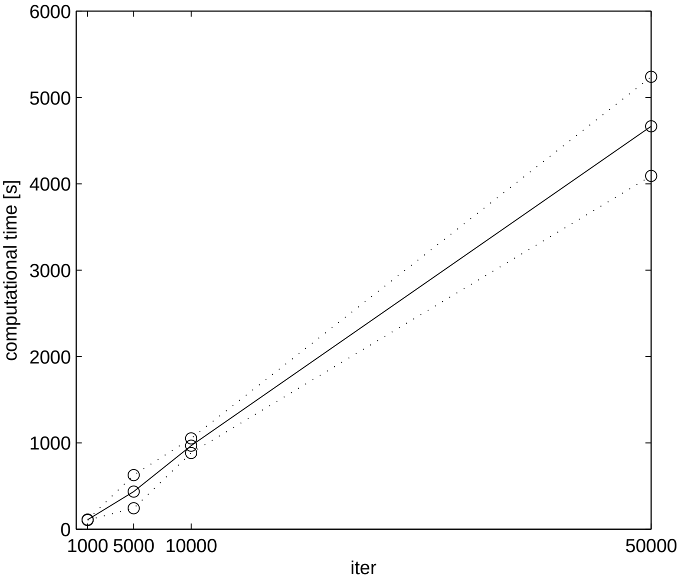


Figure 4.16: “The dove”: computational time for illumination estimation.

4.2 Illumination estimation results

4.2.1 “The dove”

As discussed in Section 3.5, we are going to estimate the illumination intensity from the function $\mathcal{L}(\mathbf{p})$, given by

$$\mathcal{L}(\mathbf{p}) = k(\mathbf{p})\chi(\mathbf{p}) = k(\mathbf{p}) \max_{\lambda} R(\mathbf{p}, \lambda).$$

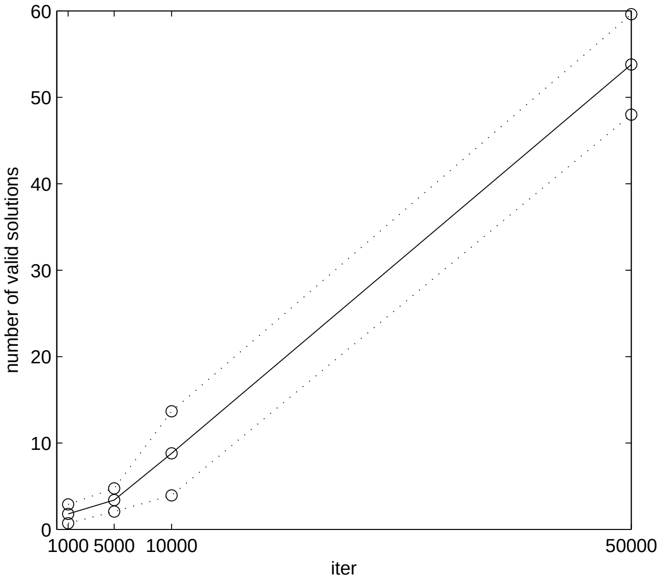


Figure 4.17: “The dove”: number of valid solutions found in the modified RANSAC algorithm 3.

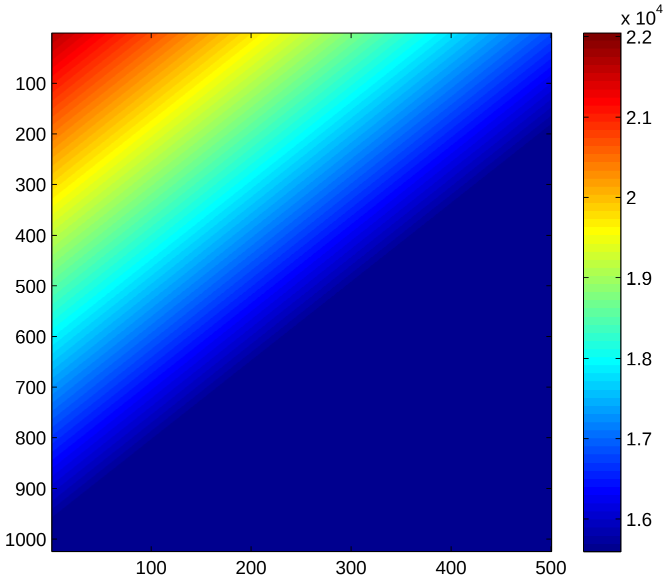


Figure 4.18: “The dove”: estimated illumination.

The image $\mathcal{L}(\mathbf{p})$ for “The dove” is represented in Fig. 4.13.

To estimate the illumination, we have followed the procedure described in Algorithm 3. Two steps of the algorithm need further discussion, namely the chosen probability distribution for $\chi(\mathbf{p})$ and the threshold t_θ. The choice for t_θ as defined in 3.24 has proved to be too stringent, meaning that in many cases the algorithm generates no valid solutions. We have therefore relaxed the definition of t_θ, fixing it to

$$t_\theta = \Delta\xi = 400.$$

As for the generation of $\hat{\chi}(\mathbf{p})$ (the maximum value of spectral reflectance at each point), we have exploited some observations regarding the specific color palette of this fresco. As can be seen from Fig. 4.3, the color dynamic range of the fresco is quite low. Pastel shades are predominant. No bright spots are present. The predominant color is beige. Supposing $\hat{\chi}(\mathbf{p})$ to assume all the values from 0.05 to 0.95, let alone with equal probability, seems therefore highly unrealistic. We have chosen to model $\hat{\chi}(\mathbf{p})$ as a Gaussian random variable with mean 0.3 and standard deviation 0.1. We have preferred this solution to that of assuming a uniform distribution for $\hat{\chi}(\mathbf{p})$ because it would have been very difficult to choose *a priori* an appropriate support for such a random variable. A Gaussian distribution is not limited to any predefined interval, so that no candidate solution is ruled out. It is worth pointing out that also unfeasible values for $\hat{\chi}(\mathbf{p})$ can possibly be generated (for example, values falling outside the [0,1] interval). However, the corresponding candidate illumination functions are simply discarded as non valid solutions by the algorithm. Choosing this procedure rather than trying to generate only valid solutions, we have privileged computational efficiency (the generation of $\hat{\chi}(\mathbf{p})$ is very fast, and additional computations are needed only if a valid solution is generated), at the expense of algorithmic efficiency (the rate of valid solutions is low).

We have subsampled $\mathcal{L}(\mathbf{p})$ by a factor of 5 along both dimensions. At each algorithm iteration, we have selected 30 points at random locations, associating them with 30 samples from the chosen Gaussian distribution $\hat{\chi}(\mathbf{p})$. After calculating the estimated illumination coefficient as

$$\hat{k}(\mathbf{p}) = \frac{\mathcal{L}_s(\mathbf{p})}{\hat{\chi}(\mathbf{p})}, \ \mathbf{p} \in \mathcal{I},$$

where $\mathcal{L}_s(\mathbf{p})$ is the subsampled version of $\mathcal{L}(\mathbf{p})$ and $\mathcal{I}$ is the set of randomly selected points, we have computed the plane parameters $\hat{\theta}$ as

$$\hat{\theta} = \arg\min_{\theta} \sum_{\mathbf{p} \in \mathcal{I}} ||\hat{k}(\mathbf{p}) - \theta^T \tilde{\mathbf{p}}||^2,$$

where $\tilde{\cdot}$ indicates homogeneous notation.

The consensus set has then been computed as

$$\mathcal{C} = \bigcup_{r_k \in \mathcal{Q}_R} \{\mathbf{p} \in \mathbb{P} \ : \ |\mathcal{L}(\mathbf{p}) - r_k \theta^T \tilde{\mathbf{p}}| < t_\theta\}, \tag{4.2}$$

with $\mathcal{Q}_R = \{0.15, 0.25, \ldots, 0.65\}$. Higher values of reflectance have not been considered in this case.

To evaluate the repeatability of our algorithm and its robustness with respect to the number of iterations used, we have repeated the estimation for five times using 1000,5000,10000 and 50000 iterations. We have defined the error as:

$$e_{ill}(\mathbf{p}) = \frac{|\hat{k}(\mathbf{p}) - k(\mathbf{p})|}{k(\mathbf{p})}, \tag{4.3}$$

for each $\mathbf{p} \in \mathbb{P}$.

Fig. 4.14 shows the average mean error, defined as

$$\mathrm{AE}_{ill}^{it} \triangleq \frac{1}{5} \sum_{rep=1}^{5} m_{e_{ill}}^{it,rep}, \ it \in \{1000, 5000, 10000, 50000\}, \tag{4.4}$$

with

$$m_{e_{ill}}^{it,rep} = \frac{1}{|\mathbb{P}|} \sum_{\mathbf{p} \in \mathbb{P}} e_{ill}^{it,rep}(\mathbf{p}),$$

and the relative standard deviation, defined as

$$\mathrm{AESTD}_{ill}^{it} = \sqrt{\frac{1}{4} \sum_{rep=1}^{5} (m_{e_{ill}}^{it,rep} - AE_{ill}^{it})^2}. \tag{4.5}$$

The global mean error, averaged over all the iterations, is 14.74 ± 2.87%.

We have also considered the maximum error, i.e. the infinity norm of the error, expressed as:

$$M_{e_{ill}} = \max_{\mathbf{p} \in \mathbb{P}} e_{ill}(\mathbf{p}). \tag{4.6}$$

The average maximum error (AME) and its standard deviation (AMESTD), defined similarly to eqs.(4.4) and (4.5), are plotted in Fig. 4.15. The global maximum error is 30.56 ± 4.04%. The maximum error standard deviation is approximately twice as large as that of the mean error. Comparing Figs. 4.14 and 4.15, it can be noted that both errors are quite stable with respect to the number of algorithm iterations. However, it is interesting to point out that as the number of iterations increases there is a slight increase in the mean illumination error, while the infinity norm of the error slightly decreases.

Fig. 4.16 shows the computational time needed by our illumination estimation algorithm. All simulations have been performed on a HP Pavillion dv4000 notebook, with a 1.83GHz Pentium M processor and 1GB RAM. Fig. 4.17 shows the number of valid solutions generated by the algorithm. The percentage wrt the number of iterations is decidedly low.

Fig. 4.18 shows the estimated illumination in the most representative case, i.e. when the mean error is closest to the global mean (10000 iterations, repetition 3).

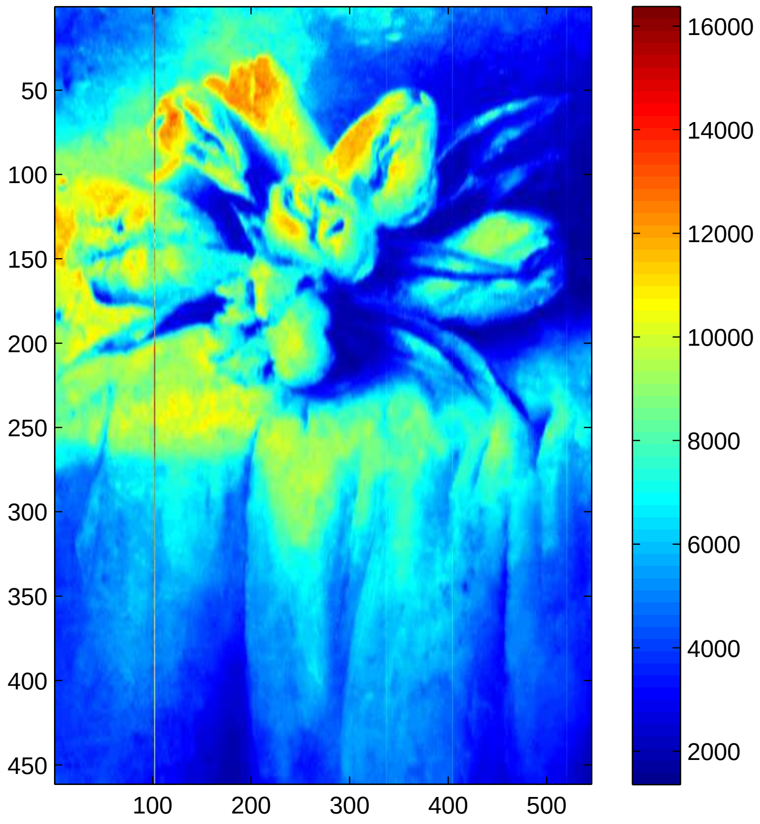


Figure 4.19: $\mathcal{L}(\mathbf{p})$ for the painting "Tulips".

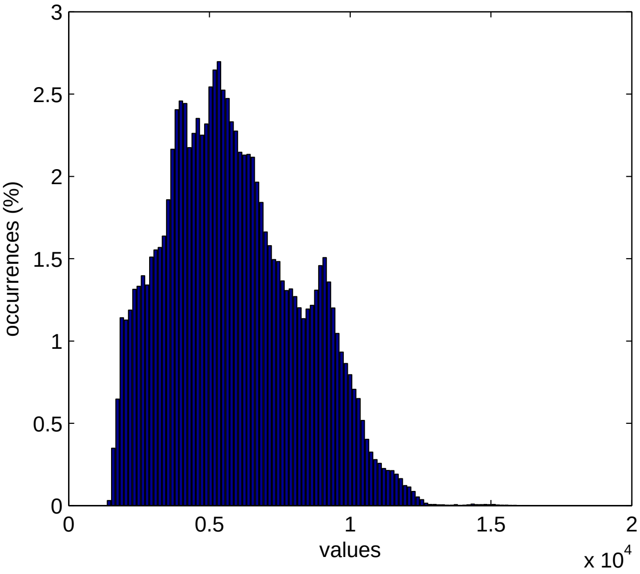


Figure 4.20: "Tulips": $\mathcal{L}(\mathbf{p})$ histogram.

4.2.2 "Tulips"

The image $\mathcal{L}(\mathbf{p})$ for the painting "Tulips" is represented in Fig. 4.19. The red stripe on the left in the painting is the result of some measurement error, probably due to the protective glass in front of the painting.

Looking at Fig. 4.5, it is easy to note that this painting is characterized by a high color dynamic. White areas cover most of the painting, but also a rather dark spot is present (top-right corner). The tulips are of a very full pink. While looking for a suitable distribution for $\hat{\chi}(\mathbf{p})$, we have moved from the observation that not all the values of $\mathcal{L}(\mathbf{p})$ can be assumed to have been generated by any value of $\chi(\mathbf{p})$ with equal probability. This issue has been raised in the literature before, see for example [69], as concerns illumination and color gamut. Here we have assumed that high reflectance values are more likely for high $\chi(\mathbf{p})$ values, and that a low reflectance value can be expected where $\chi(\mathbf{p})$ is low. This assumption is certainly not always satisfied in practice, especially for images with a low reflectance dynamic such as "The dove". However, it has been useful in practice after suitable developments.

The assumption that high $\mathcal{L}(\mathbf{p})$ values are likely to be generated by high reflectance values naturally leads to a segmentation of $\mathcal{L}(\mathbf{p})$ in an appropriate number of regions. However, the decision on how to perform such a segmentation might not be immediate. Looking at the histogram in Fig. 4.20, it is possible to see that the distribution of $\mathcal{L}(\mathbf{p})$ is far from being uniform. Assuming that the distribution of $\mathcal{L}(\mathbf{p})$ were equal to that of $k(\mathbf{p})$ would be equivalent to supposing $\max_\lambda R(\mathbf{p}, \lambda)$ constant over the image, which we certainly cannot do. Therefore, uniform quantization does not seem the best choice. Another option would be that of using μ-law quantization [92], a technique employed in digital telecommunication systems to reduce the dynamic of a signal in order to improve its signal-to-quantization noise ratio. However, we have found the results extremely sensitive to the choice of the quantization parameter μ. We have therefore preferred a third option, that is to perform histogram equalization [99] before segmenting $\mathcal{L}(\mathbf{p})$.

Let us indicate as $\mathcal{L}_s(\mathbf{p})$ the subsampled and equalized version of $\mathcal{L}(\mathbf{p})$. We have segmented $\mathcal{L}_s(\mathbf{p})$ into five regions. The centers of the clusters have been chosen as

$$C_i = \min_{\mathbf{p}} \mathcal{L}(\mathbf{p}) + i\frac{\Delta'}{2},$$

with $i = 1, \ldots, 5$ and $\Delta' = \frac{\max_{\mathbf{p}} \mathcal{L}(\mathbf{p}) - \min_{\mathbf{p}} \mathcal{L}(\mathbf{p})}{5}$. $\mathcal{L}_s(\mathbf{p})$ and the resulting segmentation are represented in Fig. 4.21.

At each algorithm iteration, we have randomly chosen 10 points from each region, and generated values for $\hat{\chi}(\mathbf{p})$ according to a Gaussian distribution with standard deviation 0.1 and mean μ chosen in $\mathcal{A} = \{0.1, 0.2, 0.4, 0.6, 0.75\}$, according to the region index (i.e., choosing $\mu = 0.1$ for $i = 1$ and $\mu = 0.75$ for $i = 5$). The choice of a Gaussian distribution is motivated by similar reasons than those outlined in Subsection 4.2.1. As for this particular case, note that the generated values for χ are not bounded to any pre-defined interval, but are only generated with higher probability around their mean. Also note again that $\hat{\chi}$ may assume non-valid values (for example, it might get negative). If that be the case, the corresponding plane is simply not selected as a valid one.

The definition of the consensus set is the same as in eq.(4.2).

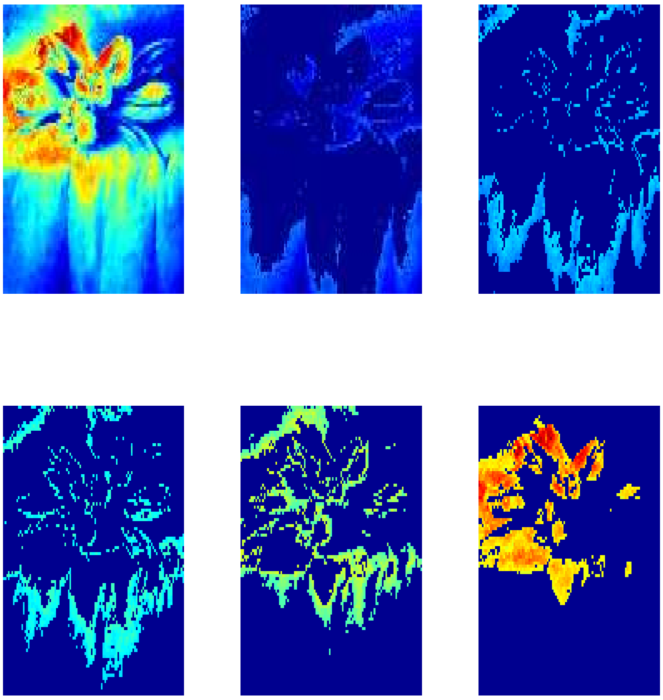


Figure 4.21: “Tulips”: $\mathcal{L}_s(\mathbf{p})$ (top left) and its five segments.

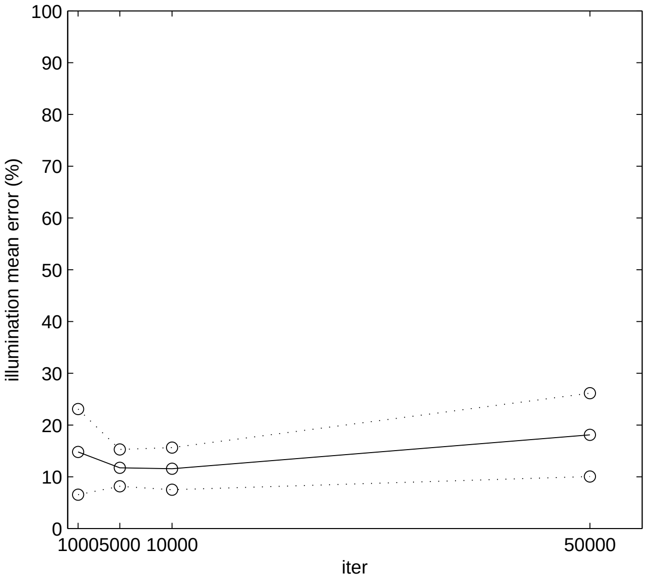


Figure 4.22: “Tulips”: variation of the mean error on the illumination ± its standard deviation with the number of iterations of the modified RANSAC algorithm 3.

The average mean and maximum error on the illumination, together with their standard deviations, have been defined as in Subsection 4.2.1. Note that, although

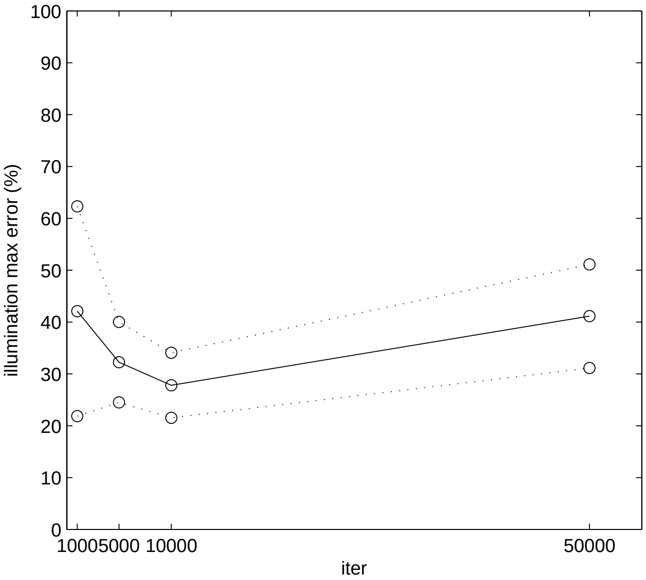


Figure 4.23: “Tulips”: variation of the maximum error ± its standard deviation on the illumination with the number of iterations of the modified RANSAC algorithm 3.

in this case the true illumination was available, we have defined the estimation error with respect to the linear space approximation for the illumination, and not to the measured illumination spectra. As noted by Ho *et al.* [67], this is more correct, as our algorithm rests upon the finite-dimensional linear space approximation assumption.

The illumination estimation results can be seen in Figs. 4.22 and 4.23. The global mean error is 14.04% ± 3.08%, while the global maximum error is 35.81% ± 6.94%. It can be seen that the infinity norm of the illumination estimation error is a little less stable than its mean. Besides, both the mean and maximum errors are slightly higher with 50000 iterations than they are when using the 5000-10000 iterations. The maximum error standard deviation is again approximately twice as large as the mean error standard deviation.

Fig. 4.24 shows the computational time needed by our algorithm, while Fig. 4.25 shows the number of valid solutions generated by the algorithm. The percentage wrt the number of iterations is again rather low.

Fig. 4.26 shows the estimated illumination in the most representative case, i.e. when the mean error is closest to the global mean (10000 iterations, repetition 3).

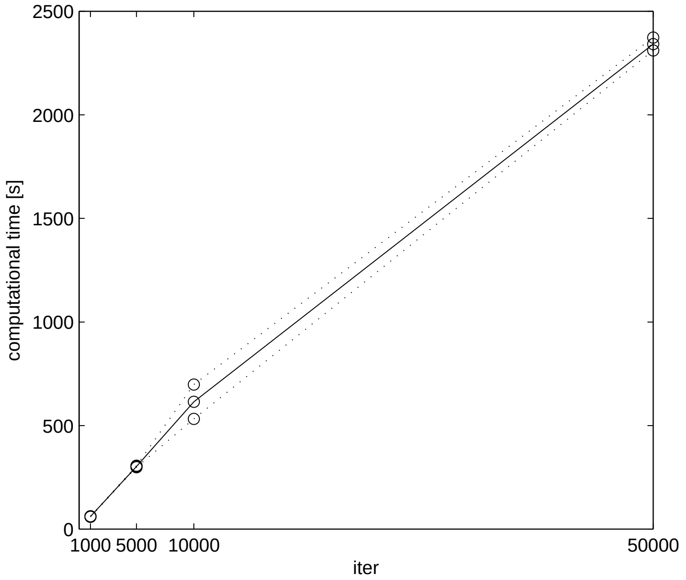


Figure 4.24: "Tulips": computational time for illumination estimation.

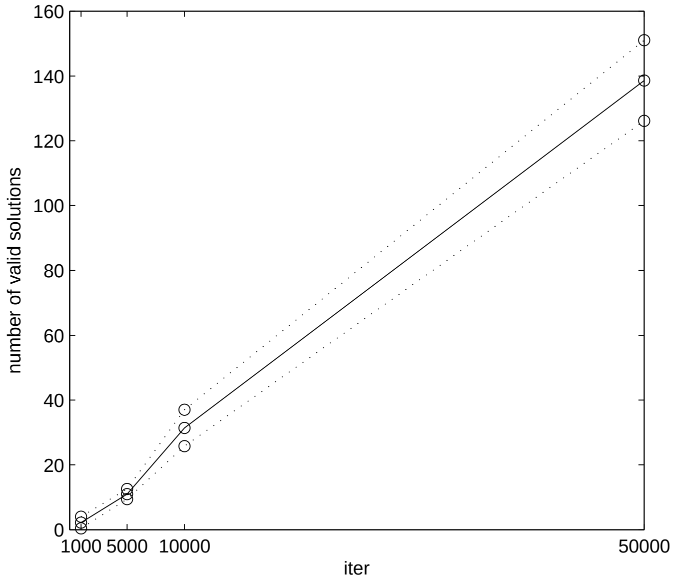


Figure 4.25: "Tulips": number of valid solutions found in the modified RANSAC algorithm 3.

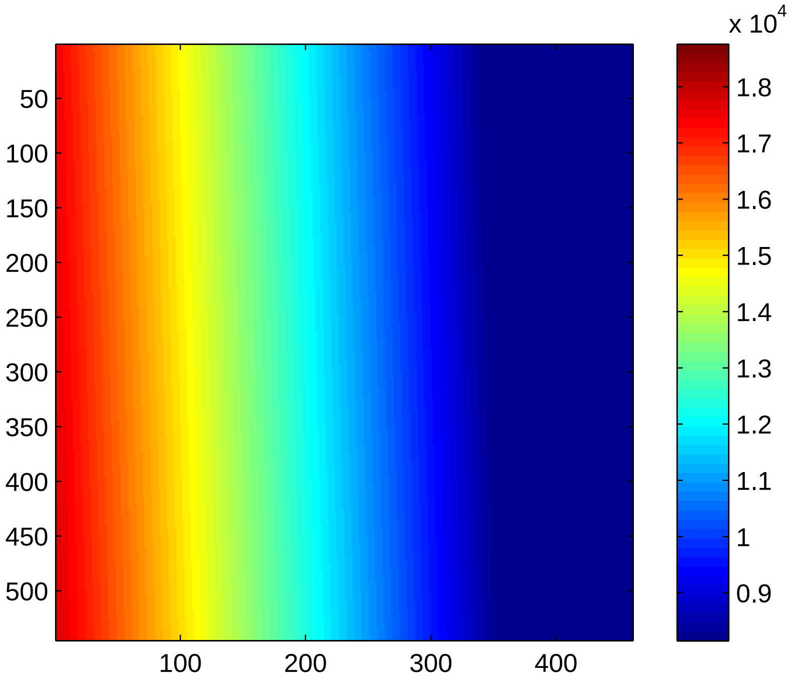


Figure 4.26: “Tulips”: estimated illumination.

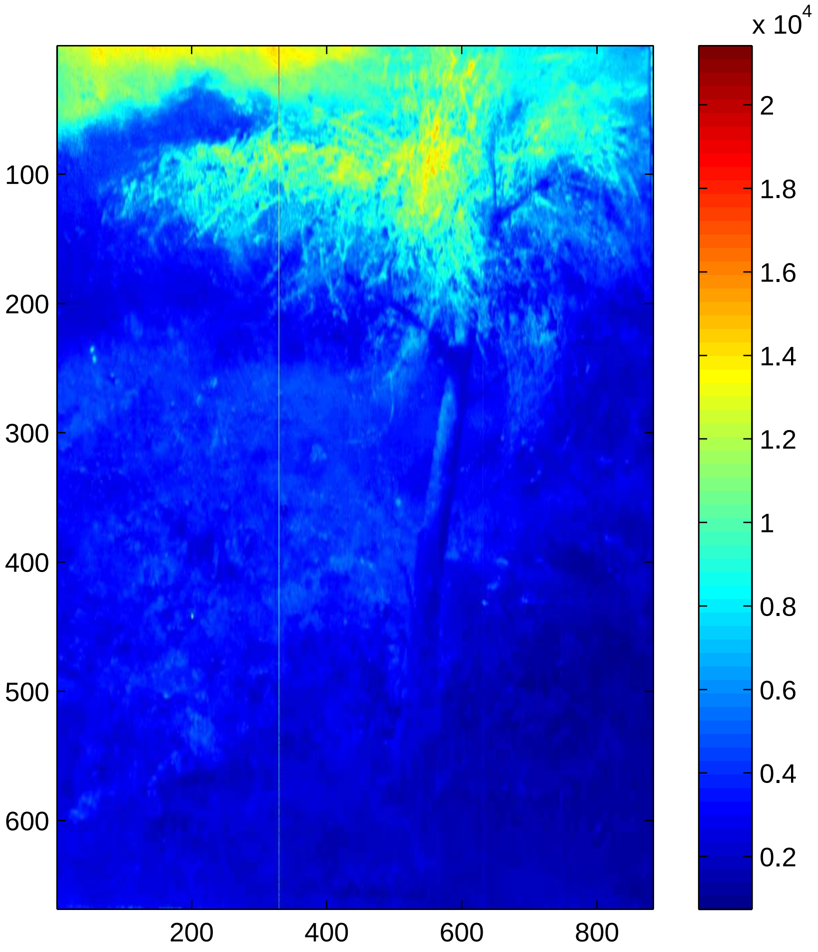


Figure 4.27: $\mathcal{L}(\mathbf{p})$ for the painting "Peach".

4.2.3 "Peach"

The image $\mathcal{L}(\mathbf{p})$ for the painting "Peach" can be seen in Fig. 4.27. For this painting the same observations made for "Tulips" hold: the color dynamic is very high, with white highlights and dark regions. We have therefore used the modified version of the algorithm described in Subsection 4.2.2, comprising histogram equalization followed by a segmentation of $\mathcal{L}(\mathbf{p})$ into five regions. The histogram is represented in Fig. 4.28. It can be noted that also in this case the histogram is markedly non-uniform. The five segmented regions can be seen in Fig. 4.29.

The illumination estimation results can be seen in Figs. 4.30 and 4.31. In this case, 1000 iterations were too few to obtain at least two valid solutions, therefore that datum has been omitted. The mean error on the illumination is stable wrt the number of iterations, while the maximum error shows a peak for 5000 iterations. The global mean illumination error is 18.29 ± 3.97 %, while the maximum error is 41.65 ± 12.57%. The maximum error on the illumination is larger in this case than for the previous ones. The maximum error standard deviation is in this case three times larger than the mean error standard deviation, probably because of the small value the latter assumes for 5000 iterations.

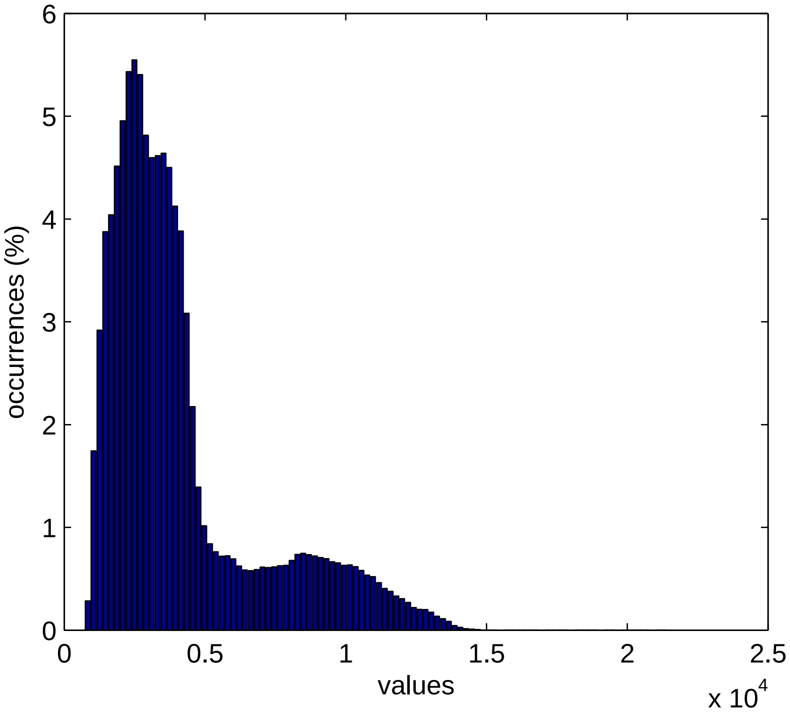


Figure 4.28: "Peach": $\mathcal{L}(\mathbf{p})$ histogram.

The simulations for the painting "Peach" were carried out on different computers, therefore no data regarding the algorithm computational time are presented. The dependence of the number of valid solutions on the number of iterations is shown in Fig. 4.32.

The estimated illumination in the most representative case (50000 iterations, repetition 2), defined as in Subsection 4.2.1, is shown in Fig. 4.33.

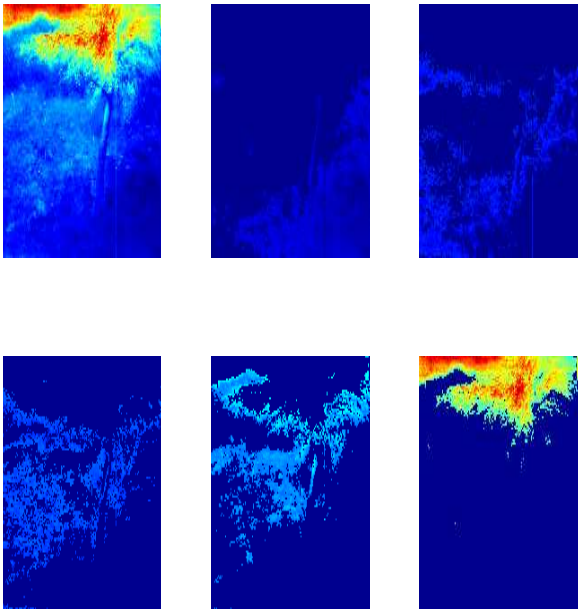


Figure 4.29: “Peach”: $\mathcal{L}_s(\mathbf{p})$ (top left) and its five segments.

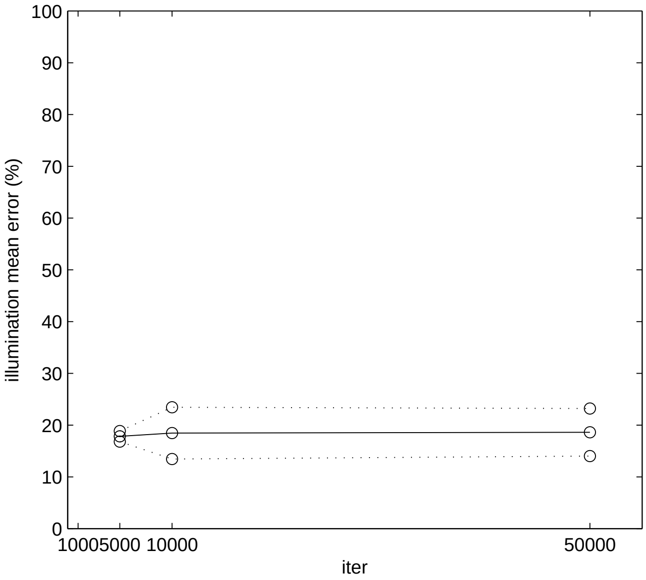


Figure 4.30: “Peach”: variation of the mean error on the illumination ± its standard deviation with the number of iterations of the modified RANSAC algorithm 3.

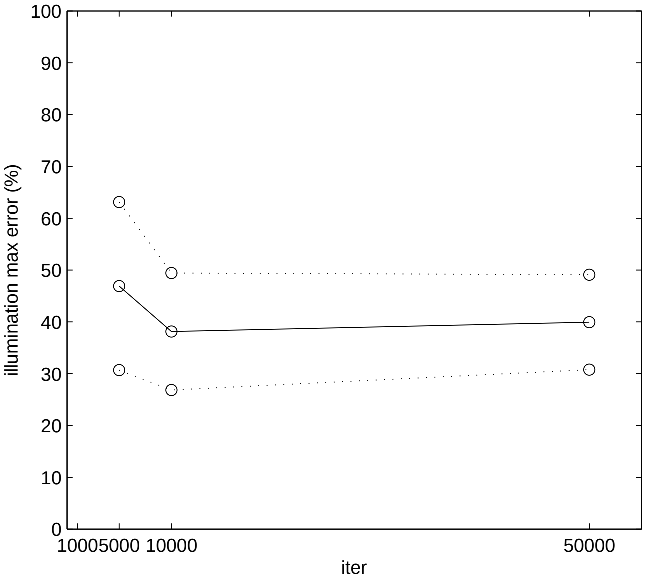


Figure 4.31: "Peach": variation of the maximum error ± its standard deviation on the illumination with the number of iterations of the modified RANSAC algorithm 3.

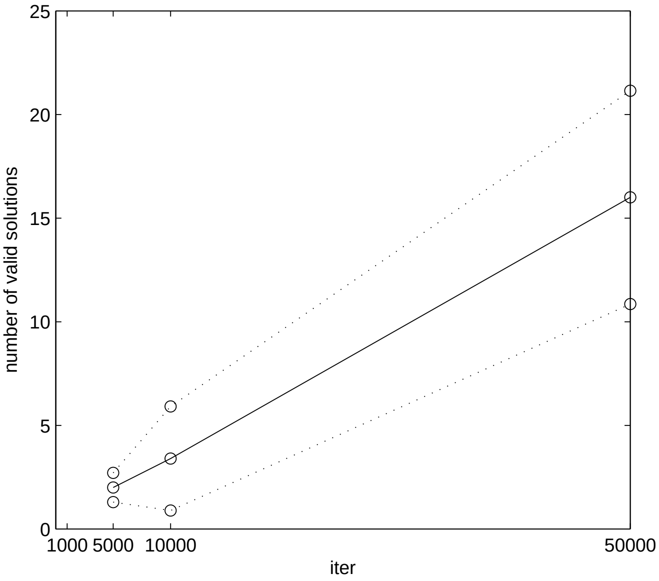


Figure 4.32: "Peach": number of valid solutions found in the modified RANSAC algorithm 3.

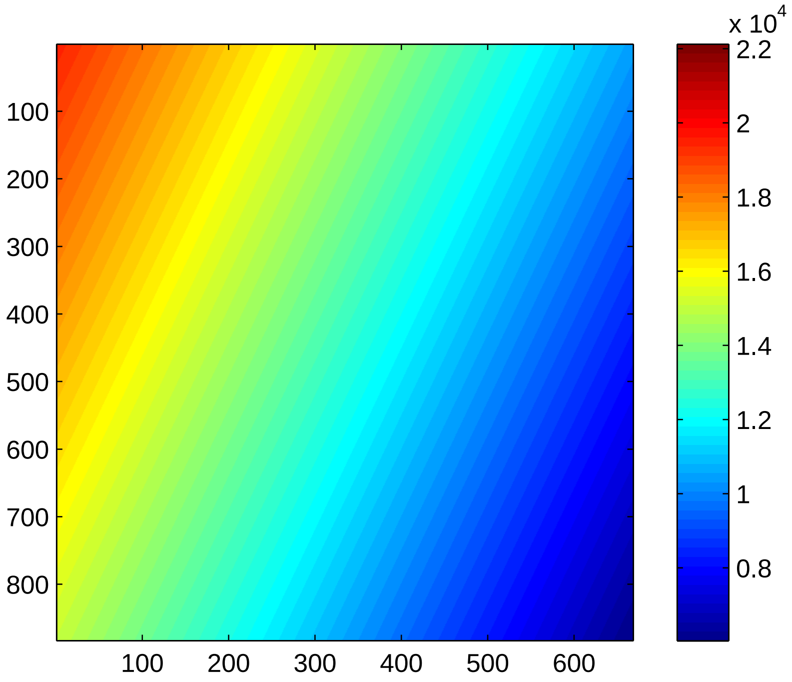


Figure 4.33: “Peach”: estimated illumination.

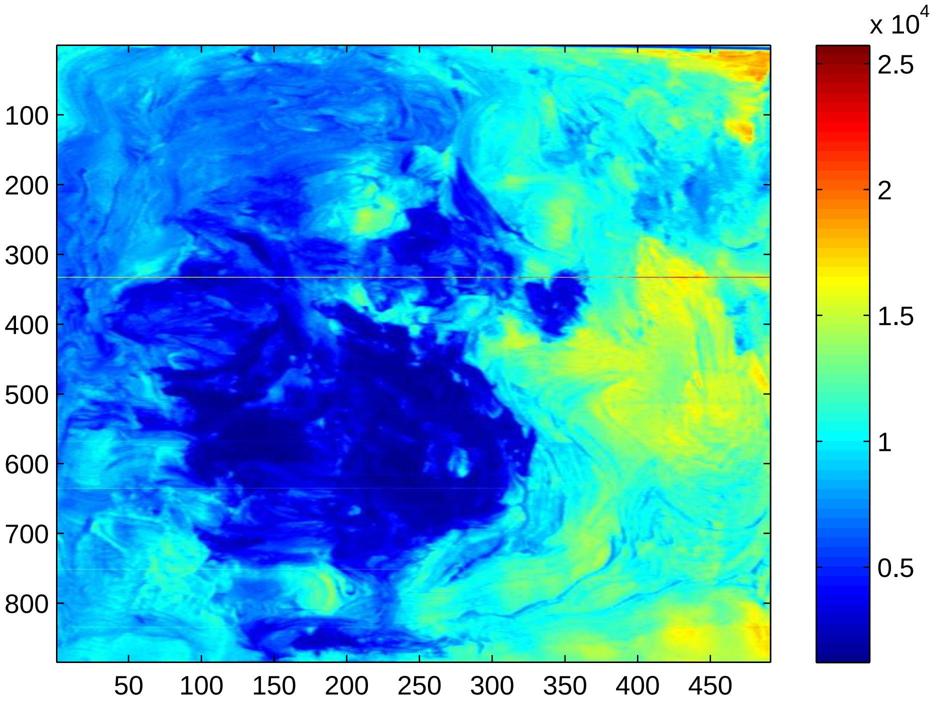


Figure 4.34: $\mathcal{L}(\mathbf{p})$ for the painting “Bull”.

4.2.4 “Bull”

Fig. 4.34 shows $\mathcal{L}(\mathbf{p})$ for the painting “Bull”. Looking at Fig. 4.10, one can see that the same considerations done for “Tulips” and “Peach” apply. Therefore, we have deployed the procedure described in Subsection 4.2.2. The histogram of $\mathcal{L}(\mathbf{p})$ is plotted in Fig. 4.35. Once again, the histogram is not uniform, therefore we have performed histogram equalization before segmentation. The segmentation output can be seen in Fig. 4.36.

The mean error on the illumination estimations is represented in Fig. 4.37. The maximum illumination error is shown in Fig. 4.38. Apart from the value for 5000 iterations, no great variation can be detected wrt the number of iterations. The global mean error is $11.6 \pm 4.11\%$, while the global infinity norm of the error is 35.23 ± 13.81. Again, the maximum error standard deviation is approximately three times as large as that of the mean error.

The computational time needed by the algorithm is shown in Fig. 4.39, while the number of valid solutions is represented in Fig. 4.40. Fig. 4.41 shows the estimated illumination in the most representative case (10000 iterations, repetition 1).

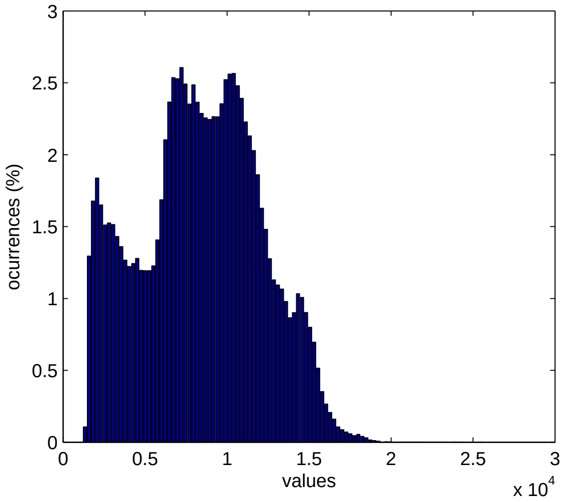


Figure 4.35: “Bull”: $\mathcal{L}(\mathbf{p})$ histogram.

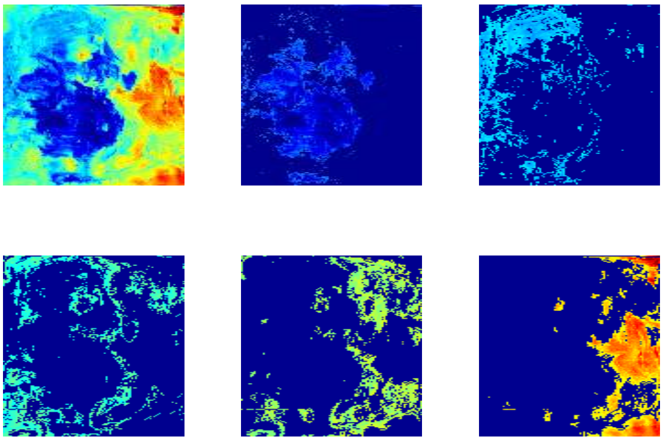


Figure 4.36: “Bull”: $\mathcal{L}_s(\mathbf{p})$ (top left) and its five segments.

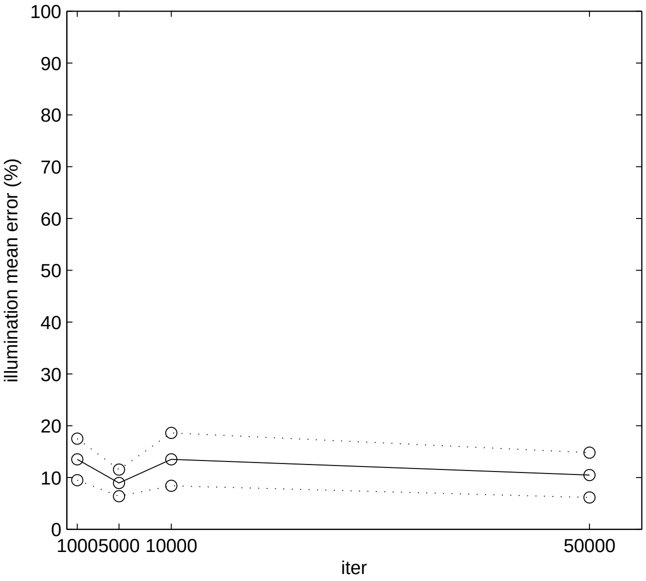


Figure 4.37: “Bull”: variation of the mean error on the illumination ± its standard deviation with the number of iterations of the modified RANSAC algorithm 3.

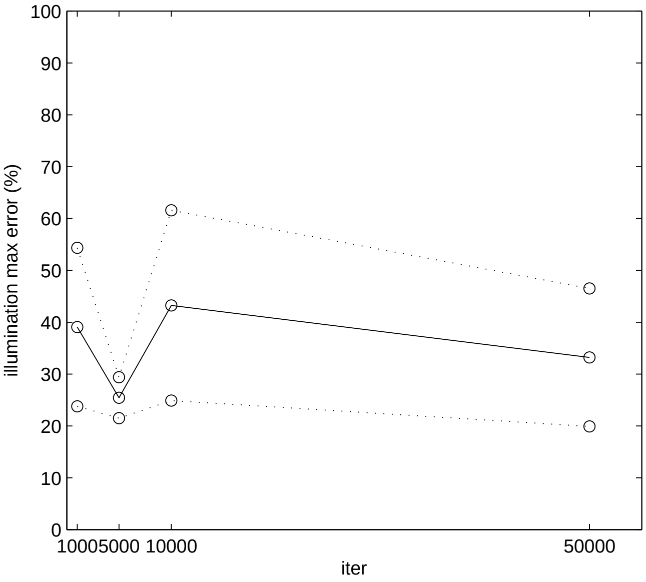


Figure 4.38: “Bull”: variation of the maximum error ± its standard deviation on the illumination with the number of iterations of the modified RANSAC algorithm 3.

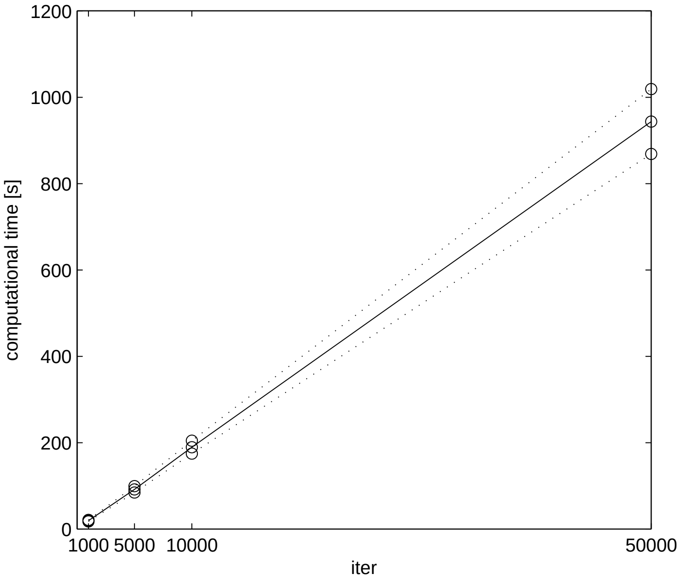


Figure 4.39: “Bull”: computational time for illumination estimation.

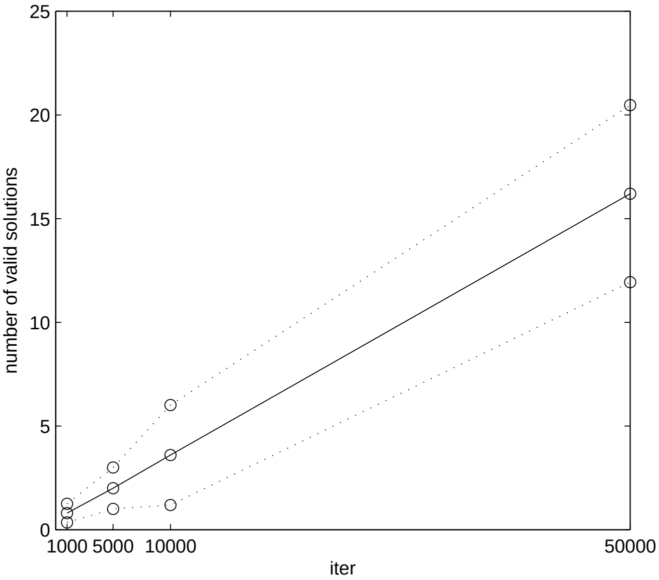


Figure 4.40: “Bull”: number of valid solutions found in the modified RANSAC algorithm 3.

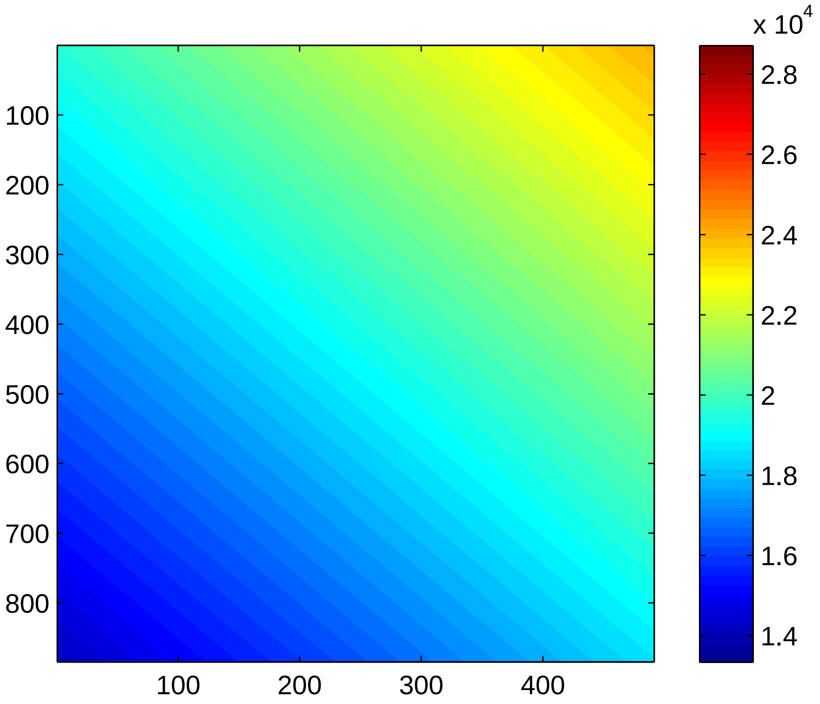


Figure 4.41: “Bull”: estimated illumination.

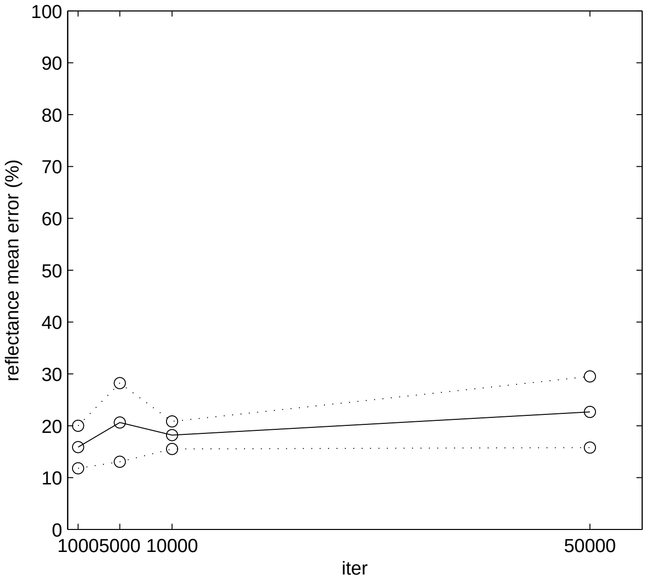


Figure 4.42: "The dove": variation of the mean error on the reflectance with the number of iterations of the modified RANSAC algorithm 3.

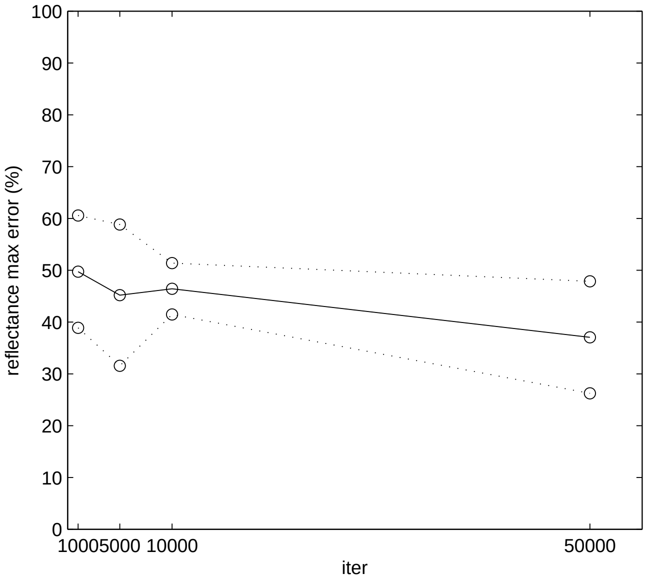


Figure 4.43: "The dove": variation of the maximum error on the reflectance with the number of iterations of the modified RANSAC algorithm 3.

4.3 Reflectance estimation results

4.3.1 "The dove"

The results for reflectance estimation relative to the experimental trial described in Subsection 4.2.1 are shown in Figs. 4.42 and 4.43. We recall that, once the illumination has been estimated, the coefficients of the reflectance are simply found by dividing the coefficients $\eta_i(\mathbf{p})$ by $k(\mathbf{p})$, as described in Section 3.6. We

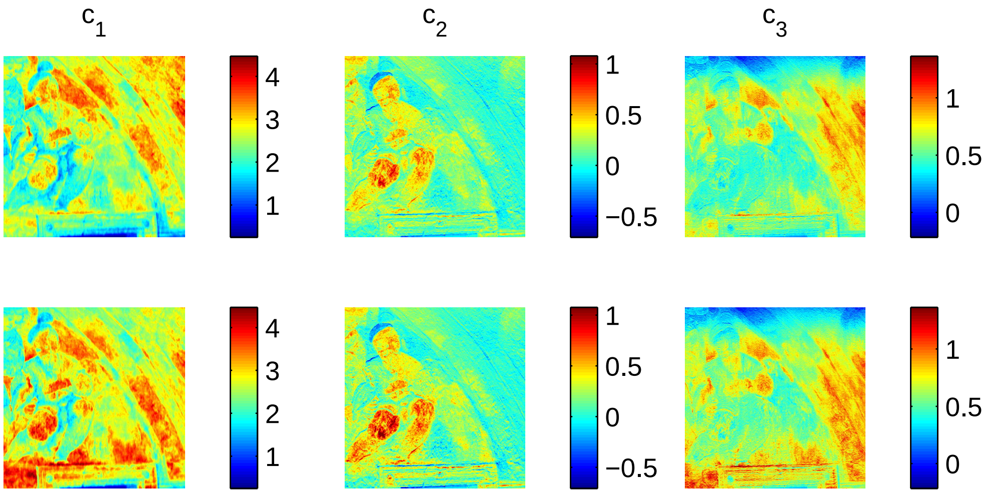


Figure 4.44: “The dove”: true (top) and estimated (bottom) reflectance coefficients 1,2,3.

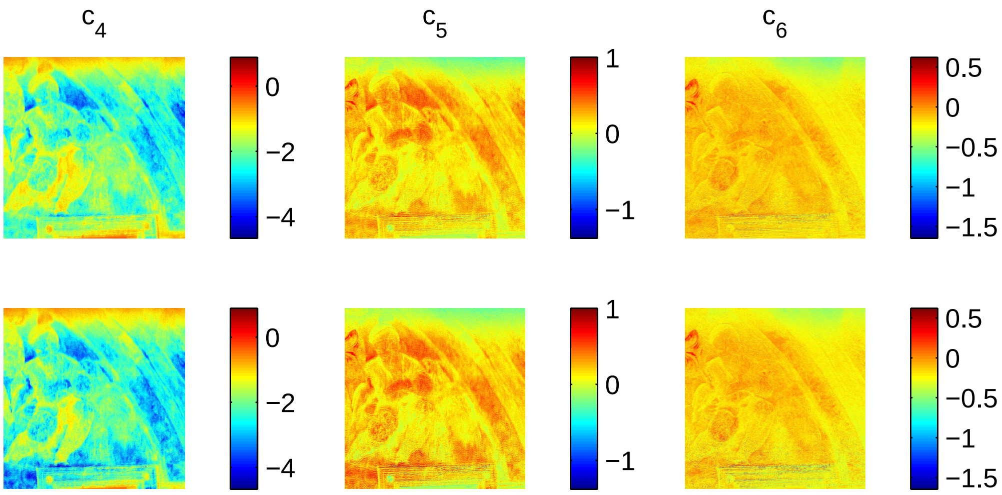


Figure 4.45: “The dove”: true (top) and estimated (bottom) reflectance coefficients 4,5,6.

have defined the reflectance estimation error as:

$$e_{ref}(\mathbf{p}) = \frac{|| \sum_{i=1}^{M} c_i(\mathbf{p}) r_i(\lambda) - \sum_{i=1}^{M} \hat{c}_i(\mathbf{p}) r_i(\lambda) ||}{|| \sum_{i=1}^{M} c_i(\mathbf{p}) r_i(\lambda) ||}. \tag{4.7}$$

The average mean reflectance estimation error has been defined as

$$\mathrm{AE}_{ref}^{it} \triangleq \frac{1}{5} \sum_{rep} m_{e_{ref}}^{it,rep}, \ \ it \in \{1000, 5000, 10000, 50000\}, \tag{4.8}$$

with

$$m_{e_{ref}}^{it,rep} = \frac{1}{|\mathbb{P}|} \sum_{\mathbf{p} \in \mathbb{P}} e_{ref}^{it,rep}(\mathbf{p}).$$

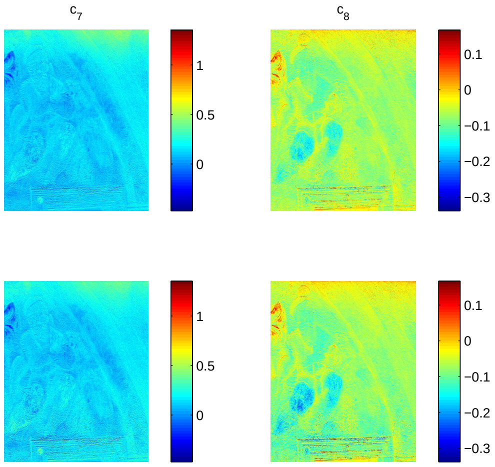


Figure 4.46: "The dove": true (top) and estimated (bottom) reflectance coefficients 7,8.

The average error standard deviation has been found as

$$\mathrm{AESTD}_{ref} = \sqrt{\frac{1}{4}\sum_{rep}(m_{e_{ref}}^{it,rep} - AE_{ref}^{it})^2}. \tag{4.9}$$

As for the illumination, we have also considered the infinity norm of the reflectance estimation error, defined by

$$M_{e_{ref}} = \max_{\mathbf{p}\in\mathbb{P}} e_{ref}(\mathbf{p}). \tag{4.10}$$

Comparing Figs. 4.42 and 4.43 with Figs. 4.14 and 4.15, one can see that corresponding errors exhibit the same behavior, but that such behavior is amplified for the reflectance error. It is worth pointing out that the propagation of the relative error on $k(\mathbf{p})$ to that on the reflectance functions cannot be easily modeled. Looking at eq.(3.25), it is easy to see that

$$\frac{dc_i(\mathbf{p})}{c_i(\mathbf{p})} = -\frac{dk(\mathbf{p})}{k(\mathbf{p})}. \tag{4.11}$$

In other words, the relative error on each reflectance coefficient $c_i(\mathbf{p})$ is equal (in absolute value) to that on the illumination. However, looking at eq.(4.7), one gets that

$$e_{ref}(\mathbf{p}) = \frac{||\sum_{i=1}^{8} dc_i(\mathbf{p})r_i(\lambda)||}{||\sum_{i=1}^{8} c_i(\mathbf{p})r_i(\lambda)||}. \tag{4.12}$$

The relative error on the reflectance thus involves the linear combination of the coefficients and the coefficient errors multiplied by the reflectance basis functions. The resulting amplification factor is therefore highly variable, and not *a priori* computable.

The global mean error on reflectance estimation is 19.33 ± 5.67%. The overall maximum error is 44.59 ± 10.55%. Also for reflectance, the standard deviation of the infinity norm of the error is twice as large as that on the mean error. The

(a)

(b)

Figure 4.47: “The dove”: (R,G,B) representation from true (a) and estimated (b) reflectance data.

comparison between the true and estimated reflectance coefficients in the most representative case (it = 1000, repetition 3) is represented in Figs. 4.44,4.45 and 4.46. Fig. 4.47 shows the perceptive comparison between the true reconstructed (R,G,B) image and that obtained from the estimated reflectance functions, i.e.

$$\hat{R}(\mathbf{p}, \lambda) = \sum_{i=1}^{8} \hat{c}_i(\mathbf{p}) r_i(\lambda). \tag{4.13}$$

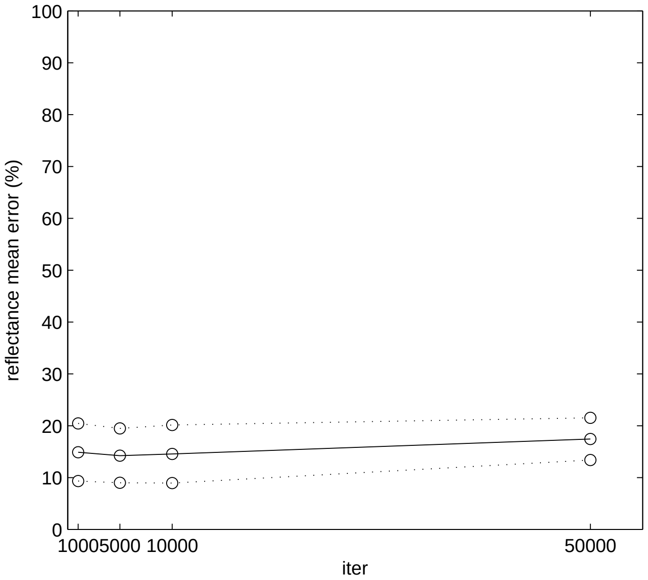


Figure 4.48: “Tulips”: variation of the mean error on the reflectance with the number of iterations of the modified RANSAC algorithm 3.

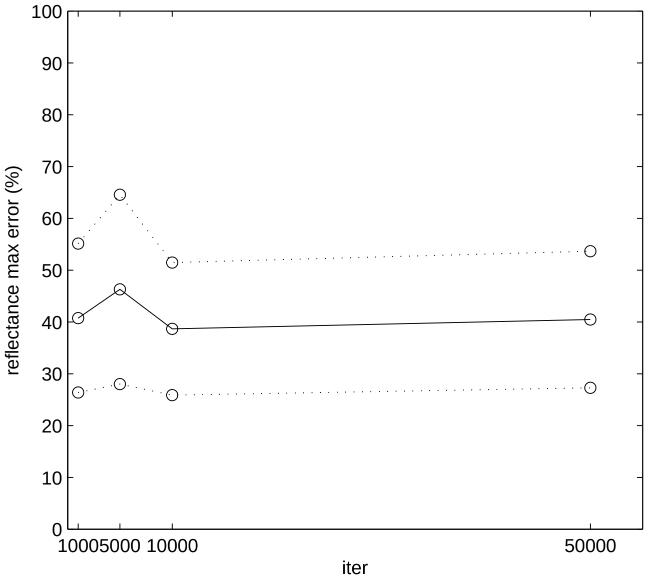


Figure 4.49: “Tulips”: variation of the maximum error on the reflectance with the number of iterations of the modified RANSAC algorithm 3.

4.3.2 “Tulips”

The results for reflectance estimation relative to the experimental trial described in Subsection 4.2.2 are shown in Figs. 4.48 and 4.49. We have defined the reflectance estimation error as in Subsection 4.3.1.

Note that in this case no evident resemblance between the illumination and reflectance errors is detectable. The mean error on reflectance is almost constant, while the infinity norm of the error presents a peak for 5000 iterations. The

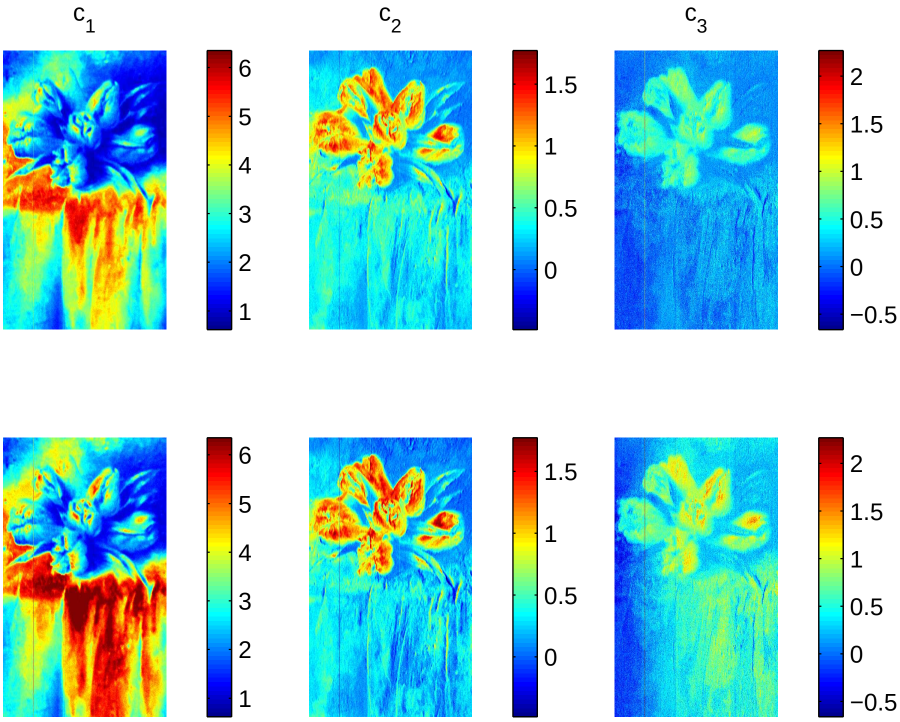


Figure 4.50: “Tulips”: true (top) and estimated (bottom) reflectance coefficients 1,2,3.

global mean error is 15.27% ± 4.91%, while the global maximum error is 41.54% ± 13.9%. The comparison between the true and estimated reflectance coefficients in the most representative case (it = 5000, repetition 2) is represented in Figs. 4.50,4.51 and 4.52. Fig. 4.53 shows the perceptive comparison between real and estimated (R,G,B) renditions of the painting.

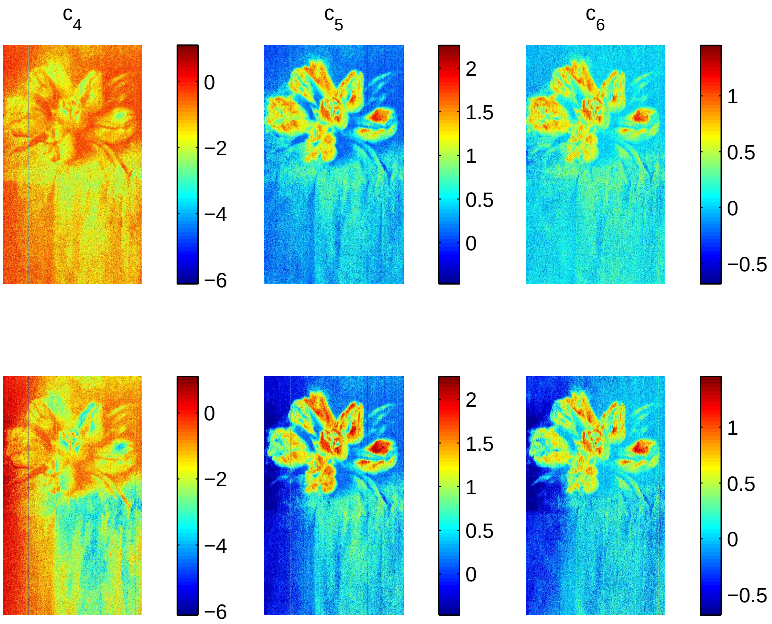


Figure 4.51: “Tulips”: true (top) and estimated (bottom) reflectance coefficients 4,5,6.

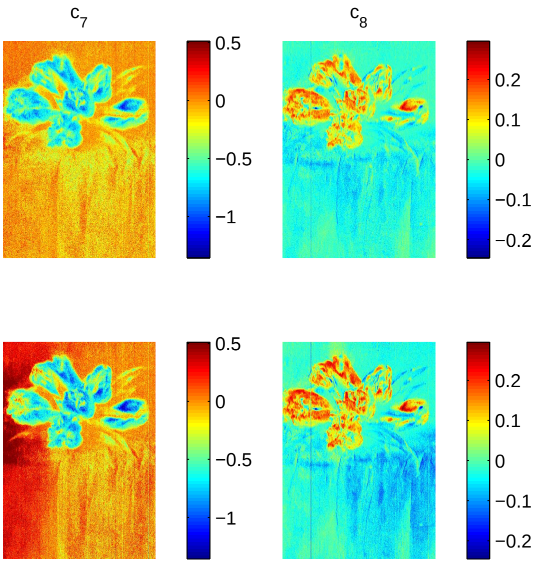


Figure 4.52: “Tulips”: true (top) and estimated (bottom) reflectance coefficients 7,8.

(a)

(b)

Figure 4.53: “Tulips”: (R,G,B) representation from true (a) and estimated (b) reflectance data.

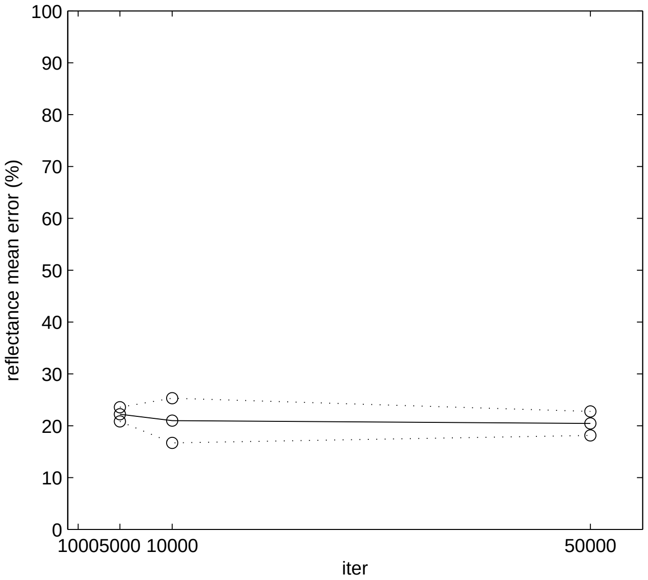


Figure 4.54: "Peach": variation of the mean error on the reflectance with the number of iterations of the modified RANSAC algorithm 3.

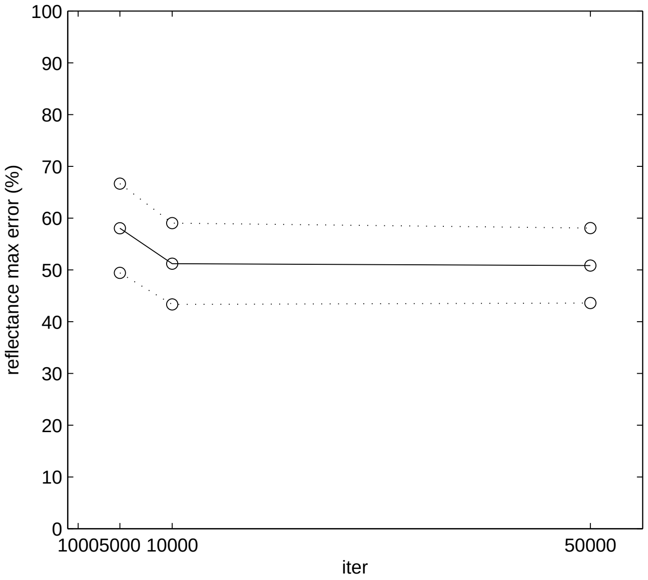


Figure 4.55: "Peach": variation of the maximum error on the reflectance with the number of iterations of the modified RANSAC algorithm 3.

4.3.3 "Peach"

The reflectance estimation results relative to the experiment described in Subsection 4.2.3 are shown in Figs. 4.30 and 4.31. In this case, the behavior of the reflectance errors is similar to that of the errors on the illumination. The global mean error is 21.21 ± 2.93%. As for the illumination, the global maximum error on reflectance is quite high: 53.35 ± 7.92%. The estimated coefficients in the most representative case (it = 10000, repetition 1) can be seen in Figs. 4.56-

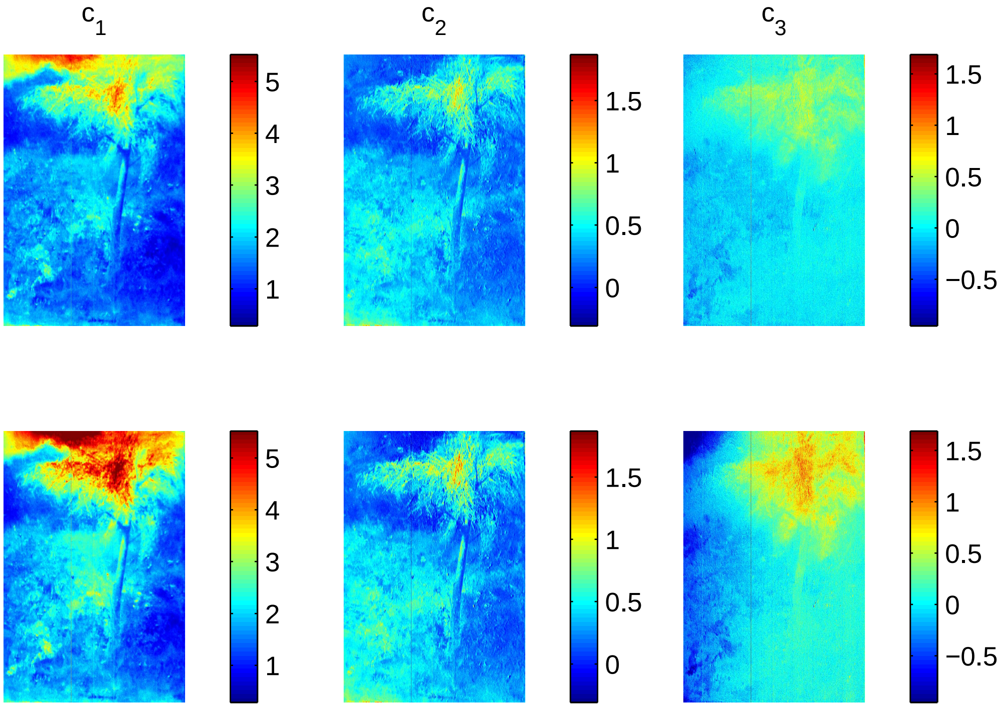


Figure 4.56: “Peach”: true (top) and estimated (bottom) reflectance coefficients 1,2,3.

4.58. Fig. 4.59 shows the perceptive comparison between the true and estimated reconstructed images.

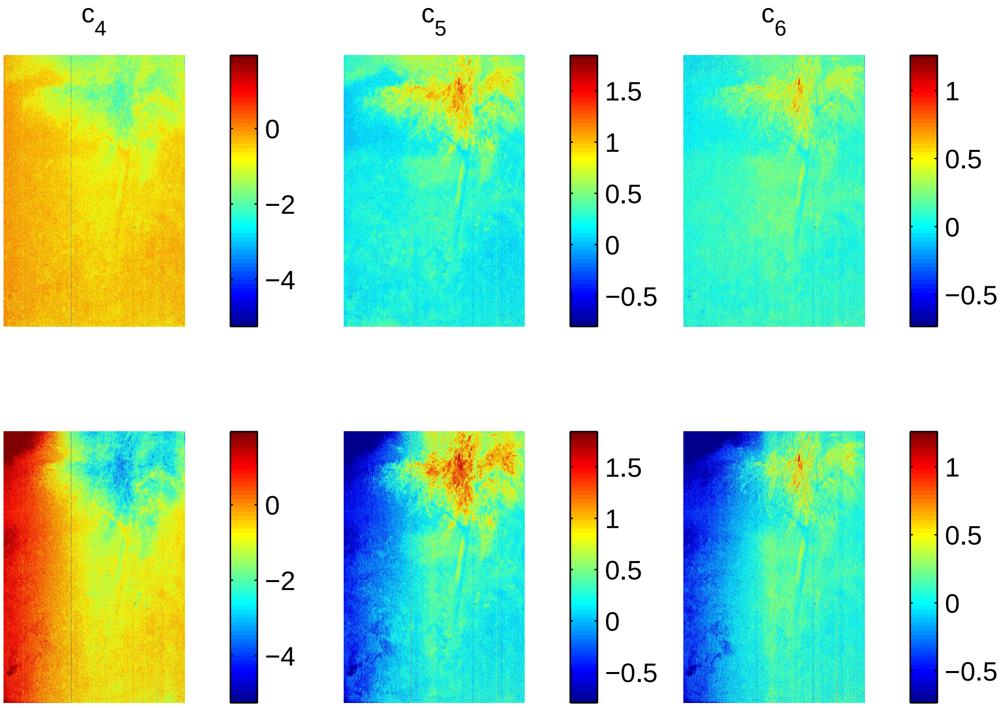


Figure 4.57: “Peach”: true (top) and estimated (bottom) reflectance coefficients 4,5,6.

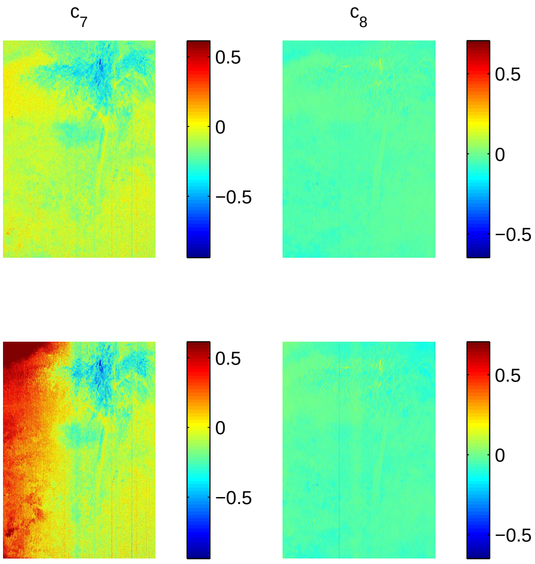


Figure 4.58: “Peach”: true (top) and estimated (bottom) reflectance coefficients 7,8.

(a)

(b)

Figure 4.59: “Peach”: (R,G,B) representation from true (a) and estimated (b) reflectance data.

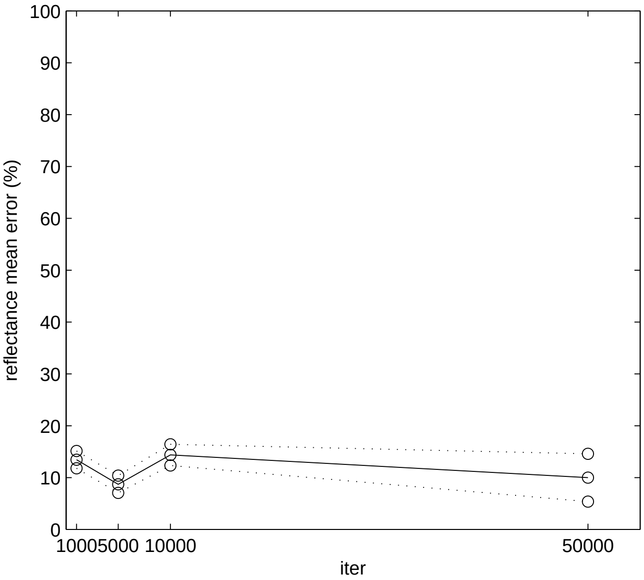


Figure 4.60: “Bull”: variation of the mean error on the reflectance with the number of iterations of the modified RANSAC algorithm 3.

4.3.4 “Bull”

The mean and maximum errors on reflectance are shown in Figs. 4.60 and 4.61. Also in this case, the reflectance error behavior is similar to that of the illumination error. The global mean error is 11.63 ± 2.78%, while the global maximum error is 36.23 ± 7.77%. Figs. 4.62, 4.63, 4.64 show the estimated coefficients in the most representative experiment (it = 1000, repetition 1). The corresponding (R,G,B) rendition of the estimated image is compared in Fig. 4.65 with the true reconstructed image.

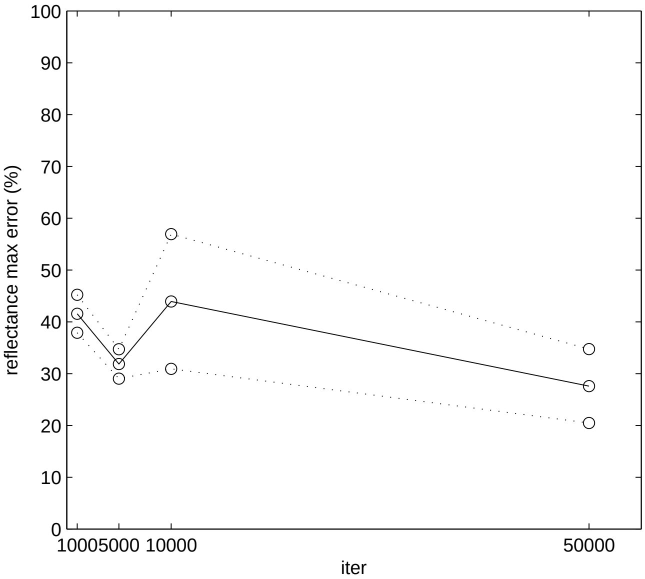


Figure 4.61: “Bull”: variation of the maximum error on the reflectance with the number of iterations of the modified RANSAC algorithm 3.

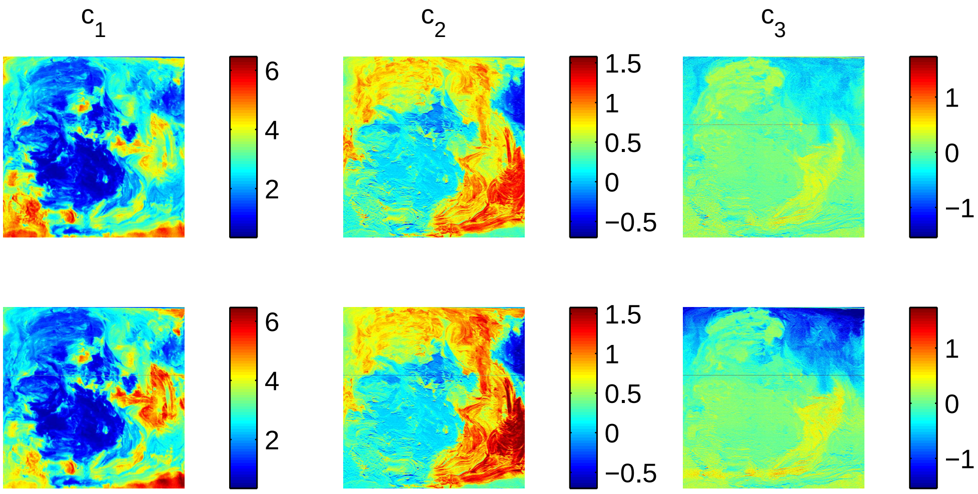


Figure 4.62: “Bull”: true (top) and estimated (bottom) reflectance coefficients 1,2,3.

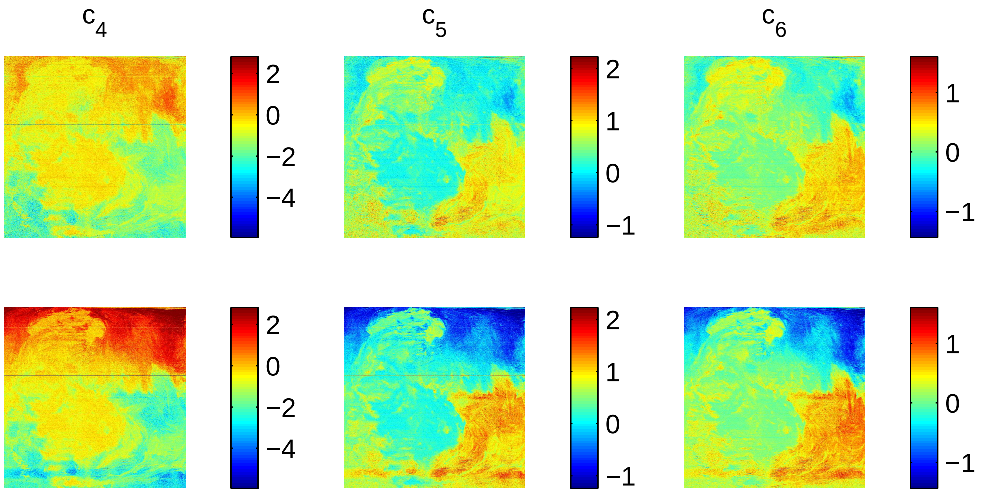


Figure 4.63: “Bull”: true (top) and estimated (bottom) reflectance coefficients 4,5,6.

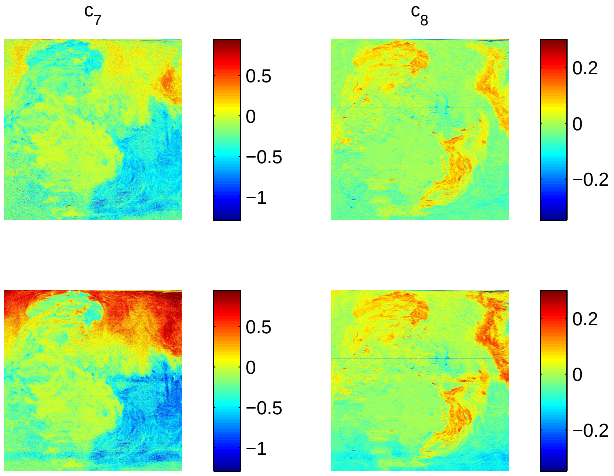


Figure 4.64: “Bull”: true (top) and estimated (bottom) reflectance coefficients 7,8.

(a)

(b)

Figure 4.65: “Bull”: (R,G,B) representation from true (a) and estimated (b) reflectance data.

	$AE_{ill}(\%)$	$AESTD_{ill}(\%)$	$M_{e_{ill}}(\%)$	$STD_{M_{e_{ill}}}(\%)$
The dove	14.74	2.87	30.56	4.04
Tulips	14.04	3.08	35.81	6.94
Peach	18.29	3.97	41.65	12.57
Bull	11.6	4.11	35.23	13.81
Mean	14.67	3.51	35.81	9.34

Table 4.1: Summary of results on illumination estimation.

	$AE_{ref}(\%)$	$AESTD_{ref}(\%)$	$M_{e_{ref}}(\%)$	$STD_{M_{e_{ref}}}(\%)$
The dove	19.33	5.67	44.59	10.55
Tulips	15.27	4.91	41.54	13.9
Peach	21.21	2.93	53.35	7.92
Bull	11.63	2.78	36.23	7.77
Mean	16.86	4.07	43.93	10.04

Table 4.2: Summary of results on reflectance estimation.

4.4 Conclusions

We have presented a statistical algorithm to separate the contributions of spectral reflectance and illumination from a multispectral color signal. The algorithm rests upon the finite linear space hypothesis for both reflectance and illumination, i.e. assumes that both signals can be expressed as linear combinations of a finite number of wavelength-dependent basis functions through spatially-dependent coefficients. Physical considerations related to the specific illumination setup allow to reduce the number of illumination basis functions to one. The illumination is therefore separable into the product of a spatially varying coefficient $k(\mathbf{p})$ and a known spectrally dependent basis function. Solving this formulation of the problem requires the estimation of the scale factor (intensity) of the illumination, a parameter which is left unestimated in the illumination-reflectance separation algorithms present in the literature. Estimating this scale factor requires solving an ill-conditioned problem, as the illumination and reflectance coefficients are multiplied into the color signal.

The algorithm consists of two steps. The first is that of illumination estimation. Based on the physical constraints that the illumination is subject to, we have derived a relation linking the maximum (over wavelength) of the color signal with the product of the illumination coefficient and the maximum (over wavelength) of the spectral reflectance, at each spatial location. The illumination estimation method is based on the observation that the illumination is a slowly-varying function of spatial coordinates, in opposition to reflectance. Therefore, the algorithm fits a smooth surface (a plane, in this implementation) to the measured signal $\mathcal{L}(\mathbf{p})$ through a modification of the RANSAC algorithm. In this context, the illumination coefficient $k(\mathbf{p})$ is regarded as the useful signal to be estimated, whereas the maximum reflectance at $\mathbf{p}$ is seen as a multiplicative random noise, with values ranging in the [0,1] interval, which corrupts $k(\mathbf{p})$, yielding the observed $\mathcal{L}(\mathbf{p})$. The model of the underlying statistics for the multiplicative noise is part of the solution to the problem, and is of crucial importance. The random noise has been assumed to have independent, identically distributed samples. This hypothesis does not necessarily hold, as reflectance values are highly correlated in space, but it allows for great generality and avoids the need of making further assumptions on the sample correlation. The RANSAC algorithm has been modified according to this formulation, changing the definitions of the thresholds and of the consensus set. All candidate illuminations are checked for consistency with the physical constraints, so that only feasible solutions are considered.

Once the illumination has been estimated, the reflectance coefficients are derived from the linear coefficients $\eta_i(\mathbf{p})$, found as the solution of a linear system. The reflectance coefficients are simply computed as the ratio between $\eta_i(\mathbf{p})$ and $k(\mathbf{p})$.

The presented algorithm proposes a solution to an otherwise ill-posed problem, formulating it as a parameter estimation problem. All the assumptions made for the illumination are physically based, and have been previously suggested in the literature. The assumption about the illumination coefficients lying on a plane is not crucial, and can be easily generalized to other smooth surfaces (i.e. quadrics). The estimated reflectance and illumination functions satisfy their physical constraints, without the need to resort to constrained minimization techniques. A further development should certainly pertain the estimation of the modified RANSAC algorithm parameters (specifically, the values assumed by the multiplicative random noise) directly from the multispectral image. All the assumptions made in the presented case studies are reasonable and do not rest on the knowledge of the true reflectance functions, but they are nonetheless highly subjective. As the algorithm performance has proved to be quite sensitive to the choice of the parameters, automating their selection would certainly lead to more repeatable results.

The algorithm has been tested on four case studies, representing artworks of different pictorial techniques, color characteristics and dimensions. Tables 4.1 and 4.2 summarize the results. The mean relative error on both illumination and reflectance is lower than 20%, and is only 2% higher for the latter. On the contrary, the infinity norm of the reflectance error is 10% higher than that on the illumination, reaching almost 44%. This suggests that the infinity norm of the error is more sensitive to the effects of error propagation, and therefore that it represents the most significant performance evaluation parameter to consider in further research.

Chapter 5

Multispectral camera calibration for texturing uncalibrated 3D data

This chapter presents a method for registering textures acquired with the multispectral scanner described in Chapter 1 onto 3D data acquired by other systems than the Shape&Color laser scanner. What makes the problem worth addressing is that due to its structure (a linear camera mounted on a rotating stage), the multispectral camera cannot be approximated with the pinhole model. Indeed, the image plane of the traditional pinhole camera is transformed into a cylinder. Moreover, the camera cannot even be classified as a central panoramic camera, as the center of rotation does not coincide with the optical center. A camera model which can be applied to our multispectral camera is the general *cylindrical panoramic camera*, which has been extensively considered in the literature. In this chapter, we will review this model with reference to our multispectral camera, and consider its application to texture mapping of multispectral images onto general, uncalibrated 3D data.

5.1 Multispectral camera model

Fig. 5.1, replicated here from Chapter 1 for the sake of readability, represents the Shape&Color scanner architecture. Neglecting the spectral dimension, for example considering one wavelength at a time, or the (R,G,B) rendition of an acquired multispectral image, our multispectral camera behaves as a line-camera rotating by means of a rotating stage. The rotating stage center does not coincide with the optical center of the camera lens. Moreover, we cannot assume the optical axis of the camera to be aligned with any radial direction of the stage. These features lead to the classification of our multispectral camera as a *cylindrical panoramic camera*. We will thus briefly review the characteristics of such model and describe its parameters.

For a thorough review and classification of general cylindrical panorama systems the reader is referred to [100]. The structure of the camera model in the (x, z) plane is represented in Fig. 5.2. The axes of the camera model are indicated with the subscript $_O$. The actual camera is a pinhole line-camera moving on a rotation platform centered at O, with radius ρ. Therefore, our camera model has

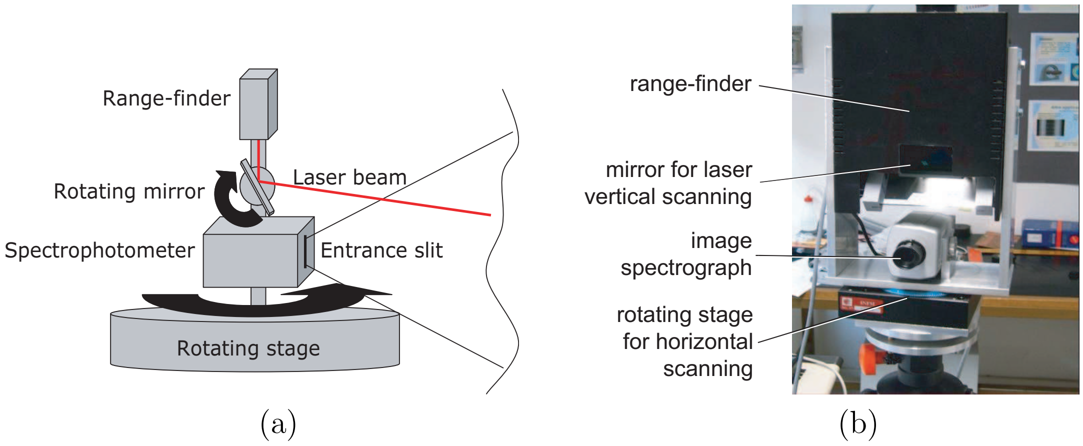


Figure 5.1: Schematic (a) and picture (b) of the instrument.

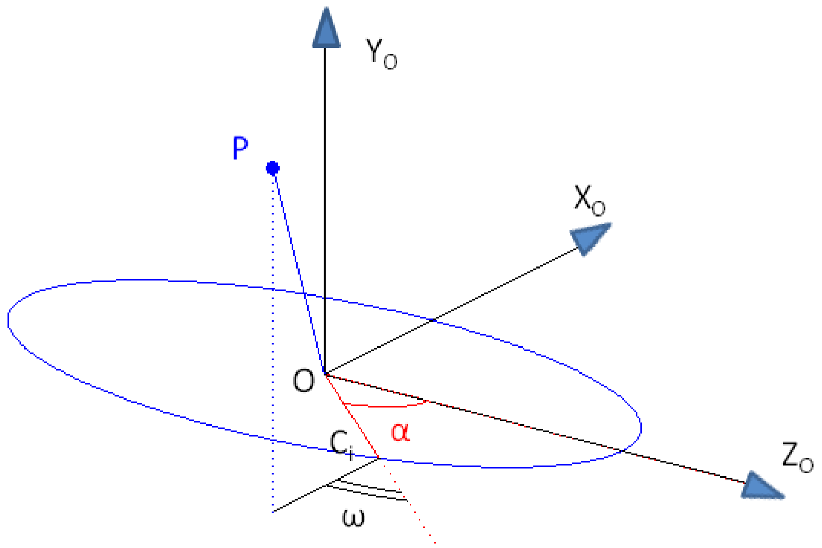


Figure 5.2: Model of the Multispectral camera in the (x, z) plane.

multiple projection centers and a cylindrical image plane. The multiple camera optical centers $\mathbf{C}_i$ lie on a circle with radius ρ, representing their distance from the rotating stage center. The angle between the z axis and the vector $\overrightarrow{OC}$ has been denoted by α. With reference to our multispectral camera, α is a multiple of the chosen angular step of the rotating stage. As previously mentioned, we cannot assume the slit camera axis to be aligned with the radial direction of the rotating platform, so a parameter ω must be introduced to represent the constant rotation angle between this axis and the radial direction in the (x, z) plane.

In the following, we will often refer to the central camera reference system and the multispectral line-camera reference system. The former, whose axes are denoted with the subscript $_O$, is defined as in Fig. 5.2. The latter is centered at C, with the y axes parallel to Y_O, and the z axis aligned with the camera axis (see Fig. 5.3). The coordinates expressed in this reference system will be denoted by the subscript $_C$. Image formation follows the traditional pinhole model along the Y_C axis. The image formation model in the central camera coordinate system

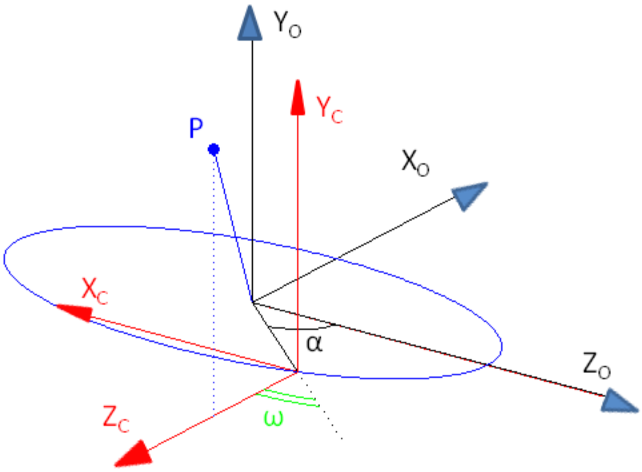


Figure 5.3: Central and camera-centered reference systems.

is more complicated for both the y and x coordinates, and will be discussed in Section 5.3.

5.2 Related work

The camera model described in Section 5.1 has been extensively considered in the literature [27, 28, 100–103]. In the following, we will review the most relevant issues with respect to the presented work.

5.2.1 Intrinsic parameter calibration

For "intrinsic parameters", we mean the effective focal length of the camera f, its principal row v_0, the off-axis radius ρ and the tilt ω. In [100], the authors correctly point out that ρ and ω are extrinsic parameters wrt the line-camera. However, as we will consider as a camera the whole system (rotating stage + line-camera), we will regard ρ and ω as intrinsic parameters of the central camera. The focal length f and the principal point v_0 can be calibrated independently using one of the classical calibration algorithms for pinhole cameras, such as [104–106]. In [100], the authors propose an alternative method to estimate the two parameters. Their algorithm uses only one image row at a time, and estimates the intrinsic parameters with a least-square approach from at least five correspondences.

As for ρ and ω, the authors extensively discuss three different possible calibration methods [100–102]: a point-based approach, a correspondence-based approach, and a line-based approach. The first approach is particularly relevant for our work, as it involves defining the projection model for the cylindrical camera. The method consists in estimating ρ and ω through a classical computer vision technique, that is, minimizing the least-square norm of the reprojection error [1]. Considering the projections of reference 3D points implies including the 6 unknown rototranslation parameters (3 for the rotation and 3 for the translation)

---

[1] The *reprojection error* is defined as the distance between the actual 2D correspondence points and the projection of the 3D correspondence points onto the image plane.

between world-centered and camera-centered reference systems into the objective function expression. The total number of unknown parameters, considering also a scale factor, is therefore nine[2]. The authors observe that the highly non-linear nature of the projection operator (see Section 5.3) makes the objective function rather complicated, and the minimization strongly dependent on the chosen initial parameters. However, they do not further investigate the minimization procedure. Their observations can be extended to the case of texture mapping. Although the intrinsic parameters ρ and ω are in that case considered pre-calibrated, the rototranslation parameters are included in non-linear functions through the projection operator. Without any prior knowledge of the extrinsic parameters, the minimization of the reprojection error is therefore a very difficult task. The main contribution of this chapter consists in indicating a way to compute a reliable initial point for the minimization procedure, so as to reduce the risk of falling into a local minimum.

The above-described approach to the estimation of ρ and ω has been ultimately discarded by the authors due to its high error sensitivity (exponential growth with noise level). The second considered approach to the problem is that of estimating ρ and ω from the selection of corresponding points in two panoramas. This method has the advantage of avoiding the necessity to include the rototranslation parameters into the minimization. Moreover, the objective function is linear in the (four) parameters, so that no initial value is needed. The drawbacks are that only the ratio between ρ_1 and ρ_2 (the subscript identify the two different panoramas) can be estimated, and that the error still grows exponentially with the input noise level. It is also worth pointing out that from our point of view, this method would not be applicable, as we do not have the chance of modifying ρ.

The last approach to the problem of estimating ρ and ω exploits the parallel-line based approach, which consists in exploiting geometrical features of lines on known calibration objects (e.g. length or distances) to derive constraints for the camera parameters. The authors explore two different kinds of constraints, one requiring the knowledge of the lengths and distances between at least three line pairs, the second exploiting orthogonal linear line triplets, that is triplets of lines defining two orthogonal planes. Both approaches allow the derivation of a single linear equation (for each line pair or line triplet, respectively) linking the considered 3D geometric feature to the desired camera parameters. The intrinsic paramters ρ and ω can be easily computed from the estimated equation coefficients, and the authors prove that the error sensitivity wrt the input noise level follows a linear relation.

5.2.2 Cylindrical panoramas for texture mapping

Although its main application area has been that of stereo reconstruction from cylindrical panorama pairs [27, 28, 100, 103], the cylindrical camera model has also been considered in relation to the problem of texturing 3D laser data with cylindrical panoramas [28, 103]. The specific investigated problem has been that

---

[2] In [100–102], the authors erroneously claim that the total number of parameters is 14, considering 12 parameters for the rototranslation. In fact, only 6 parameters are sufficient to define a rototranslation matrix, while an additional parameter is the scale factor.

of mapping cylindrical panorama textures onto 3D data obtained with a phase-difference laser scanner whose functioning principle is very similar to that of a cylindrical camera. The laser scanner acquires range data along a single line, and a whole range image is obtained by performing a rotation of the laser in the plane orthogonal to the scanning direction. The proposed algorithm starts by calibrating the extrinsic parameters of the cylindrical camera and of the laser scanner (possibly using multiple viewpoints) wrt a reference world coordinate system using a photometric approach. This technique exploits the cylindrical nature of the camera and laser data. Moreover, it requires the possibility to adjust the tilt angle ω to different values. Namely, the extrinsic parameter calibration procedure is derived for $\omega = 0$. After the positions of the camera centers have been computed, the authors suggest that the value of ω can be changed, and re-calibrated using one of the approaches described in Subsection 5.2.1. Once the range data and the camera centers have been transformed into a single reference system, the correspondences between 3D points and 2D cylindrical panorama coordinates can be easily recovered.

This approach cannot be extended to our case. First of all, we are considering 3D models obtained with generic 3D range scanners and already integrated. Most of all, the extrinsic parameters estimation procedure proposed in [28,103] requires fixing $\omega = 0$, which we cannot do (obviously unless this was the value returned by the intrinsic parameters estimation procedure). Therefore, we will consider a more general formulation of the extrinsic parameter estimation problem, that is, that of the minimization of the reprojection error.

5.3 Image formation model

5.3.1 (y, z) plane

As previously mentioned, the line-camera follows the 1D pinhole model along the y axis. We thus have that a point $P'_C = (Y_C, Z_C)^T$, with coordinates expressed in the camera-centered reference system, is projected onto the pixel coordinate y given by:

$$y = f\frac{Y_C}{Z_C} + v_0, \tag{5.1}$$

where f is the effective focal length of the camera in pixel and v_0 its principal row [50] expressed in pixel. Eq.(5.2) can also be formulated in homogeneous coordinates as:

$$\begin{pmatrix} y \\ 1 \end{pmatrix} = \mathbf{K}[\mathbf{0}\ \mathbf{I}_{2\times 2}\ \mathbf{0}] \begin{pmatrix} X_C \\ Y_C \\ Z_C \\ 1 \end{pmatrix}, \tag{5.2}$$

with

$$\mathbf{K} = \begin{bmatrix} f & v_0 \\ 0 & 1 \end{bmatrix}.$$

We now have to find the relationship between the two reference systems (O, X_o, Y_o, Z_o) and (C, X_C, Y_C, Z_C). It is quite easy to see that the latter can

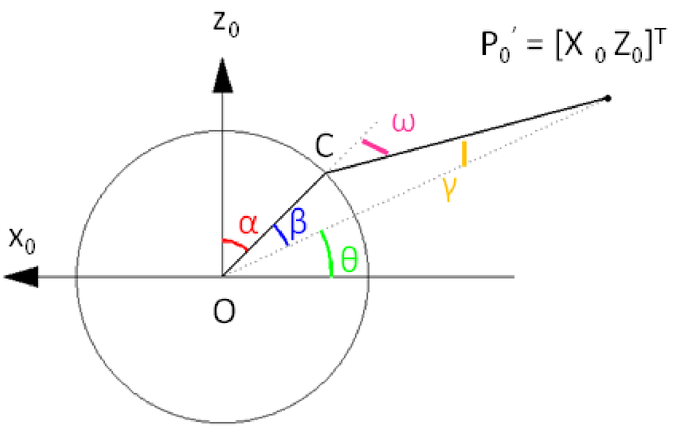


Figure 5.4: Image formation in the (x, z) plane.

be obtained by performing a rotation of $\omega + \alpha$ around the Y_o axis, followed by a translation of $(\rho \sin\omega, 0, -\rho\cos\omega)^T$. In other words,

$$\begin{pmatrix} X_C \\ Y_C \\ Z_C \\ 1 \end{pmatrix} = \begin{bmatrix} \mathbf{Q}(\alpha) & \mathbf{d} \\ \mathbf{0} & 1 \end{bmatrix} \begin{pmatrix} X_O \\ Y_O \\ Z_O \\ 1 \end{pmatrix}, \tag{5.3}$$

where

$$\mathbf{Q}(\alpha) = \begin{bmatrix} \cos(\omega+\alpha) & 0 & -\sin(\omega+\alpha) \\ 0 & 1 & 0 \\ \sin(\omega+\alpha) & 0 & \cos(\omega+\alpha) \end{bmatrix}$$

and

$$\mathbf{d} = \begin{pmatrix} \rho\sin\omega \\ 0 \\ -\rho\cos\omega \end{pmatrix}.$$

Substituting this relation into eq.(5.2), we obtain that:

$$y = \frac{fY_O}{X_O\sin(\alpha+\omega) + Z_O\cos(\alpha+\omega) - \rho\cos\omega} + v_0. \tag{5.4}$$

5.3.2 (x, z) plane

Modeling image formation in the (x, z) plane requires to determine the relationship between points $P_O = (X_O\ Z_O)^T$ and α. Fig. 5.4 illustrates the meaning of the parameters considered in the following.

We can start by noticing that

$$\alpha = \frac{\pi}{2} - \beta - \theta. \tag{5.5}$$

We can then derive β as

$$\beta = \pi - (\pi - \omega) - \gamma = \omega - \gamma, \tag{5.6}$$

where γ can in turn be computed applying the sine law to the triangle $\triangle(\mathbf{O}, \mathbf{C}, \mathbf{P'_O})$:

$$\frac{\rho}{\sin\gamma} = \frac{||P'_o||}{\sin(\pi - \omega)} = \frac{||P'_o||}{\sin\omega} \Rightarrow \gamma = \arcsin\left(\frac{\rho}{||P'_o||}\sin\omega\right). \tag{5.7}$$

Lastly, θ can be easily derived as

$$\theta = \arctan\left(\frac{Z_0}{X_0}\right). \tag{5.8}$$

Merging eqs.(5.5),(5.6),(5.7),(5.8), we finally get

$$\alpha = \frac{\pi}{2} - \omega + \gamma - \theta = \frac{\pi}{2} - \omega + \arcsin\left(\frac{\rho}{||P'_o||}\sin\omega\right) - \arctan\left(\frac{Z_0}{X_0}\right). \tag{5.9}$$

The variables α and x are related by the very simple equation (with this choice of the reference system)

$$\alpha = -x\Delta_\alpha, \tag{5.10}$$

where Δ_α is a constant and known angular step .

5.4 Estimation of the extrinsic parameters

The problem of estimating the rigid transformation between world-centered and camera-centered systems can be formulated as follows. Given a set of correspondences

$$\begin{pmatrix} X_w \\ Y_w \\ Z_w \end{pmatrix} \leftrightarrow \begin{pmatrix} \alpha \\ y \end{pmatrix},$$

we want to recover the roto-translation aligning the world reference system to the central camera reference system. The two coordinate systems are related by:

$$\tilde{\mathbf{P}}_{\mathbf{O}} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix} \tilde{\mathbf{P}_w}, \tag{5.11}$$

with

$$\mathbf{R} = \begin{bmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ R_{31} & R_{32} & R_{33} \end{bmatrix} = \begin{bmatrix} \mathbf{r}_1^T \\ \mathbf{r}_2^T \\ \mathbf{r}_3^T \end{bmatrix} \in SO(3)$$

and

$$\mathbf{t} = \begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix}.$$

The problem of estimating the 6 parameters of the rotation and the translation can be solved in a classical way by minimizing the reprojection error between actual and estimated projections [50]. If we call

$$\mathbf{m}_i \triangleq \begin{pmatrix} \alpha_i \\ y_i \end{pmatrix}$$

the actual coordinates of the correspondence points on the image, and

$$\hat{\mathbf{m}}_i(\mathbf{p}) \triangleq \left(\begin{array}{c} \hat{\alpha}_i \\ \hat{y}_i \end{array} \right) = \mathcal{P}(\mathbf{P}_{w,i}),$$

the projections of the correspondence 3D points ($\mathcal{P}(\cdot)$ indicates the projection operator), we want to estimate the parameter vector $\mathbf{p}$ as the solution to the minimization problem

$$\hat{\mathbf{p}} = \arg\min_{\mathbf{p}} \sum_{i=1}^{n} \left\| \mathbf{m}_i - \hat{\mathbf{m}}_i(\mathbf{p}) \right\|^2 , \tag{5.12}$$

with n the number of correspondences.

The minimization of eq.(5.12) can be performed using nonlinear minimization methods such as the Levenberg-Marquardt optimization algorithm. As observed in [102], the nonlinear nature of the projection operator makes such objective function quite complicated. The authors also observe that as a consequence, the estimation accuracy highly depends on the chosen initial parameters. In this chapter, we would like to outline a method for choosing suitable initial parameters for the nonlinear minimization problem. To do so, we will assume that the chosen set of correspondences is free from projection errors, and will estimate the rototranslation parameters between world and camera-centered reference systems from the available 3D-2D point pairs. This approach is similar to that followed, for example, in [104].

In order to estimate the roto-translation parameters, we would like to exploit as much as possible the projection equation along the y axis, since it is linear in the parameters. Merging Eqs.(5.2),(5.3) and (5.11), we get:

$$\left(\begin{array}{c} \mathrm{y} \\ 1 \end{array} \right) = \lambda \mathbf{K} \left[\begin{array}{cccc} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array} \right] \left[\begin{array}{cc} \mathbf{Q}(\alpha) & \mathbf{d} \\ \mathbf{0} & 1 \end{array} \right] \left[\begin{array}{cc} \mathbf{R} & \mathbf{t} \\ \mathbf{0} & 1 \end{array} \right] \tilde{\mathbf{P}}_w, \tag{5.13}$$

where λ is the unknown scale factor.

Expanding eq.(5.13), we find that

$$\left(\begin{array}{c} \mathrm{y} \\ 1 \end{array} \right) = \lambda \mathbf{K}\mathbf{A} = \lambda \mathbf{K} \left[\begin{array}{c} \mathbf{a}_1^T \\ \mathbf{a}_2^T(\alpha) \end{array} \right] \tilde{\mathbf{P}}_w, \tag{5.14}$$

with

$$\mathbf{a}_1 = \left[\begin{array}{c} R_{21} \\ R_{22} \\ R_{23} \\ t_y \end{array} \right], \tag{5.15}$$

$$\mathbf{a}_2(\alpha) = \left[\begin{array}{c} R_{11}\sin(\omega+\alpha) + R_{31}\cos(\omega+\alpha) \\ R_{12}\sin(\omega+\alpha) + R_{32}\cos(\omega+\alpha) \\ R_{13}\sin(\omega+\alpha) + R_{33}\cos(\omega+\alpha) \\ t_x\sin(\omega+\alpha) + t_z\cos(\omega+\alpha) - \rho\cos\omega \end{array} \right]. \tag{5.16}$$

$\mathbf{K} \in \mathbb{R}^2$ can be inverted and the right term rewritten as:

$$\begin{aligned} \mathbf{K}^{-1} \left(\begin{array}{c} \mathrm{y} \\ 1 \end{array} \right) &= \lambda \left[\begin{array}{cc} \tilde{\mathbf{P}}_w^T & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_w^T \end{array} \right] [\mathbf{a}_1^T \; \mathbf{a}_2^T(\alpha)]^T = \\ &= \lambda \tilde{\mathbf{M}} \mathbf{a}(\alpha) \end{aligned} \tag{5.17}$$

with $\mathbf{a} \in \mathbb{R}^{8\times1}$, $\tilde{\mathbf{M}} \in \mathbb{R}^{2\times8}$. The vector $\mathbf{a}$ has 6 degrees of freedom (3 for the rotation and 3 for the translation). One more is given by the scale.

The extrinsic parameters of the camera can be recovered in two steps. The first step consists in estimating $\mathbf{a}_1$ by considering the first row of the system of equations represented by eq.(5.17), which can be rewritten as:

$$v_1 = \lambda \mathbf{P}_w^T \mathbf{a}_1, \tag{5.18}$$

where $\mathbf{v} \triangleq \mathbf{K}^{-1} \begin{pmatrix} \mathrm{y} \\ 1 \end{pmatrix}$ and v_1 is its first element. The four elements of $\mathbf{b} \triangleq \lambda \mathbf{a}_1$ can be estimated in the least-square sense by stacking at least four such equations for different correspondence points. Once $\mathbf{b}$ has been estimated, we can recover a first slot of extrinsic parameters by considering the following equalities:

$$\begin{aligned} b_1 &= \lambda R_{21} \\ b_2 &= \lambda R_{22} \\ b_3 &= \lambda R_{23} \\ b_4 &= \lambda t_y. \end{aligned} \tag{5.19}$$

λ and $\mathbf{r}_2$ can be easily derived by $\{b_1, b_2, b_3\}$ using the fact that $||\mathbf{r}_2|| = 1$:

$$\lambda = \sqrt{b_1^2 + b_2^2 + b_3^2} \tag{5.20}$$

$$\mathbf{r_2} = \frac{1}{\lambda} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}. \tag{5.21}$$

The translation component along the y axis can also be easily computed as:

$$t_y = \frac{b_4}{\lambda}. \tag{5.22}$$

Looking at eq.(5.16), we can see that a similar approach cannot be used to estimate $\mathbf{a}_2(\alpha)$, which depends upon the point itself. However, as $\mathbf{a}_2(\alpha)$ is a linear combination of the estrinsic parameters, the second row of eq.(5.17) can be rearranged as

$$\begin{aligned} v_2 &= \lambda \mathbf{P}_w^T \left(\mathbf{M}' \begin{bmatrix} R_{11} \\ R_{12} \\ R_{13} \\ R_{31} \\ R_{32} \\ R_{33} \\ t_x \\ t_z \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ -\rho\cos(\alpha) \end{bmatrix} \right) \\ &= \lambda \mathbf{P}_w^T \left(\mathbf{M}' \mathbf{a}_2' + \begin{bmatrix} 0 \\ 0 \\ 0 \\ -\rho\cos\omega \end{bmatrix} \right). \end{aligned} \tag{5.23}$$

Figure 5.5: Snapshot of "The captain".

where v_2 is the second element of vector $\mathbf{v}$ and

$$\mathbf{M}' \triangleq \begin{bmatrix} \cos(\alpha+\omega) & 0 & 0 & \sin(\alpha+\omega) & 0 & 0 & 0 & 0 \\ 0 & \cos(\alpha+\omega) & 0 & 0 & \sin(\alpha+\omega) & 0 & 0 & 0 \\ 0 & 0 & \cos(\alpha+\omega) & 0 & 0 & \sin(\alpha+\omega) & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \cos(\alpha+\omega) & \sin(\alpha+\omega) \end{bmatrix}$$

Developing eq.(5.23), we obtain

$$\frac{v_2}{\lambda} = \mathbf{P}_w^T \mathbf{M}' \mathbf{a}_2' - \rho \cos\omega, \tag{5.24}$$

and lastly

$$v_2' \triangleq \frac{v_2}{\lambda} + \rho \cos\omega = \mathbf{P}_w^T \mathbf{M}' \mathbf{a}_2'. \tag{5.25}$$

As v_2' and $\mathbf{M}'$ are completely determined by the 3D-2D correspondences, the intrinsic parameters and the so-far estimated extrinsic parameters, the vector $\mathbf{a}_2'$ can be estimated using at least eight correspondences, and the remaining extrinsic parameters can be recovered. The estimated rotation matrix $\mathbf{R}$ will not be, in general, belonging to $SO(3)$. The best approximating rotation matrix has to be extracted from it using, for example, the method cited in [104, Appendix C], where it is proved that the matrix $\mathbf{R}' \in SO(3)$ which is closest to $\mathbf{R}$ according to Frobenius norm can be computed as $\mathbf{R}' = \mathbf{U}\mathbf{V}^T$, where $\mathbf{R} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^T$ is the singular value decomposition of $\mathbf{R}$.

5.5 Results

Our initial point estimation algorithm has been tested on synthetic data obtained from the 3D model of a real object. The 3D model we have utilized is that of a small statue called "The captain" (a snapshot can be seen in Fig. 5.5), which was digitized using a structured light range camera. We have projected a subsample of its points according to two different cylindrical projections, whose intrinsic

	f [pixel]	v_0 [pixel]	ρ [mm]	ω [deg]	Δ_α [rad]
$\mathcal{P}_1$	3000	1024	100	0	10^{-4}
$\mathcal{P}_2$	3000	1024	100	30	$5 \cdot 10^{-4}$

Table 5.1: Cylindrical projection parameters used in the experimental trial.

	$\mathcal{P}_1$			$\mathcal{P}_2$			
3σ [pixel]	RT$_1$	RT$_2$	RT$_3$	RT$_4$	RT$_5$	RT$_6$	mean
0.5	0.0002	0.0006	0.0010	0.0007	0.0003	0.0009	0.0006
1	0.0002	0.0009	0.0010	0.0014	0.0008	0.0010	0.0009
1.5	0.0005	0.0003	0.0008	0.0018	0.0011	0.0002	0.0008
2	0.0014	0.0010	0.0012	0.0012	0.0021	0.0006	0.0013
3	0.0017	0.0012	0.0013	0.0008	0.0018	0.0046	0.0019
4	0.0002	0.0012	0.0015	0.0035	0.0029	0.0070	0.0027
5	0.0026	0.0011	0.0032	0.0075	0.0010	0.0012	0.0027

Table 5.2: Estimation results for the rotation matrix.

parameters are summarized in Table 5.1. We have subsequently applied three different rototranslations for each projection operator to the sampled 3D points, and corrupted the 2D projection points with Gaussian noise of increasing standard deviation. In conformity with [100], we have chosen $3\sigma \in \{0.5, 1, 1.5, 2, 3, 4, 5\}$ pixels. We have then tried to estimate the extrinsic parameters using the procedure described in Section 5.4, that is, estimating the initial point for the minimization from the rototraslated 3D and noisy 2D correspondence points, and then using the Levenberg-Marquardt minimization algorithm to find the optimal parameters. Each time, we have used 100 correspondence points randomly chosen in the point set. To evaluate the estimation results for the translation vector $\mathbf{t}$, we have used the Euclidean vector norm, defining the relative error on $\mathbf{t}$, $e_\mathbf{t}$, as

$$e_\mathbf{t} = \frac{||\mathbf{t} - \hat{\mathbf{t}}||}{||\mathbf{t}||}, \tag{5.26}$$

where $\hat{\cdot}$ denotes an estimated parameter. As for the rotation, we have used the second criteria proposed in [52], that is, we have considered the error defined as

$$e_\mathbf{R} = ||\mathbf{I_{3\times 3}} - \mathbf{R}\hat{\mathbf{R}}^T||, \tag{5.27}$$

where $\mathbf{I_{3\times 3}}$ is the 3×3 identity matrix, and the norm $||\mathbf{M}||$ is computed as the largest singular value of $\mathbf{M}$.

The estimation results are summarized in Tables 5.2 and 5.3. The mean estimation results are also shown in Figs. 5.6 and 5.7. In [100, 102], the authors have claimed that the error sensitivity in the estimate of ρ and ω using the point-based approach is exponential. Although the relative error on the translation vector can be said to have a somewhat exponential behavior, there is a reasonable point for defining it linear as well. The error on rotation is decidedly linear. More extensive experimental results could clarify this point. As a general comment, the estimation results are good even when the input noise level is high. The mean error on translation is below 4%, while the maximum relative error is 11.3%; the mean error on rotation is below $3\cdot 10^{-3}$, its maximum value being $7.5\cdot 10^{-3}$.

	$\mathcal{P}_1$			$\mathcal{P}_2$			
3σ [pixel]	RT_1	RT_2	RT_3	RT_4	RT_5	RT_6	mean
0.5	0.3102	0.5635	1.3443	1.0848	0.2205	0.9123	0.7393
1	0.2658	0.7944	1.2617	1.8811	0.5085	0.7525	0.9107
1.5	0.8724	0.3671	0.9598	2.6822	0.6934	0.4856	1.0101
2	1.9488	1.0117	1.4843	0.8365	1.3775	0.6650	1.2206
3	2.5015	1.1089	1.8280	1.4579	1.1571	3.4476	1.9168
4	0.5042	1.0314	1.8733	4.5939	1.9197	6.2587	2.6968
5	4.0707	0.8808	4.0973	11.3066	0.7841	1.9986	3.8564

Table 5.3: Estimation results for the translation vector (%).

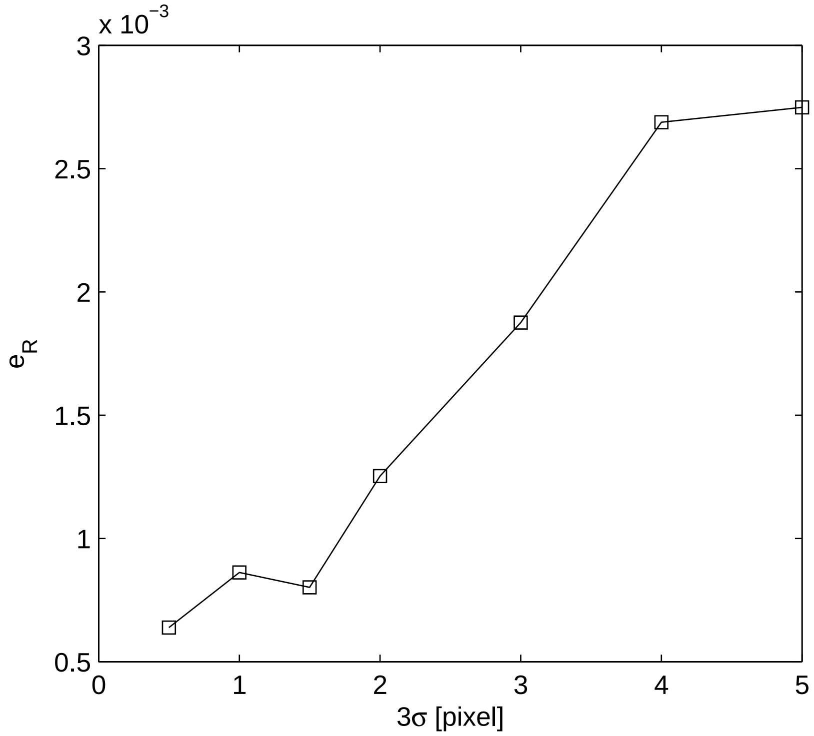


Figure 5.6: Mean error on the rotation matrix as a function of noise level.

5.6 Conclusions

We have considered the problem of computing the extrinsic parameters of our multispectral camera from a set of known 3D-2D correspondences, with the aim of texturing uncalibrated 3D models with the acquired multispectral images. The Shape&Color multispectral camera can be regarded as a cylindrical panoramic camera, a camera model which has been extensively studied in the literature [27, 28, 100–103]. Previous work has concerned the formulation of the projection function, the estimation of the intrinsic parameters, and the definition of the epipolar geometry between pairs of panorama images, with the aim of performing stereo panorama reconstruction. An application to texture mapping onto laser 3D data has also been considered by other authors [28, 103], but the proposed extrinsic parameter estimation procedure is very specific to the considered application, and cannot be applied in our case.

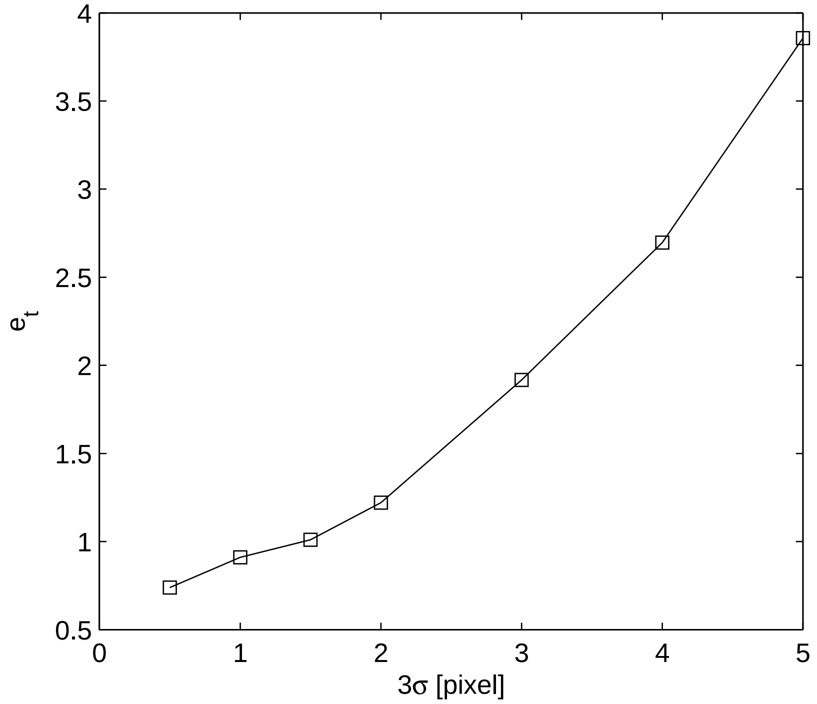


Figure 5.7: Mean error on the translation vector (%) as a function of noise level.

The extrinsic parameter estimation problem has been formulated in a classical way as a reprojection error minimization problem. Due to the nonlinear nature of the projection operator and the high dimensionality of the problem, the resulting objective function can be expected to exhibit a very complicated behavior. As a consequence, the minimization results are strongly dependent on the chosen initial values for the parameters, and the risk of being trapped in a local minimum is high. We have proposed a way of computing a reliable initial point for the minimization procedure, in order to lower such risk. Our method follows a classical approach, that is, that of computing the rototraslation parameters in the ideal case, as if the 3D and 2D correspondence points were not affected by measurement errors. The problem has been solved using uniquely the linear projection operator along the y axis. The extrinsic rototranslation parameter estimation has been performed in two steps via standard linear algebra computations. The proposed approach has been tested on synthetic 3D-2D correspondences corrupted by Gaussian random noise of increasing standard deviation. The estimation results are good. An increase of the estimation error with the input noise level has been detected. The error behavior is nearly linear in the considered interval, but more data are needed to extend this consideration to broader ranges.

As a last observation, we would like to point out that the proposed initial parameter estimation method could be useful even when dealing with the problem addressed in [102], that is, estimating the intrinsic parameters of the circular camera (ρ and ω) from 3D-2D correspondences. While no plausible initial values for the rototranslation parameters can in general be figured out, reasonable initial choices for the intrinsic parameters can easily be obtained by experimental measurements on the central camera system. Therefore, the presented methodology can be leveraged to subsequently derive the corresponding initial rototranslation estimation. Further work will concern the comparison of the circular camera parameter estimation using the proposed extrinsic parameter estimation method with those presented in [100–102].

Conclusions

We have presented three problems concerning the use of a multispectral imaging spectrograph in applications of cultural heritage. The multispectral camera is part of an instrument developed within the project "Shape&Color" (CARIPARO, 2003-2005), coupling the spectrograph with a 3D laser scanner. Although the issues we have addressed arose from the characteristics of this specific instrument, they can be regarded as general problems concerning multispectral imaging, and are therefore of broader interest.

First of all, we have assessed the performance of our spectral camera in the measurement of spectral reflectance in the 400-830 nm interval by acquiring a set of colored calibrated tiles under different illumination conditions. We have used a metallic iodide lamp, the metallic iodide lamp together with a halogen lamp, the metallic iodide and halogen lamps one after the other, and the metallic iodide and incandescence lamp in a sequence. In the last two cases, we have obtained spectral reflectances as a juxtaposition of the reflectance measured with the metallic iodide lamp from 400 to 600 nm and that acquired with the other lamp from 600 to 830 nm. To evaluate the system performance, we have defined the error as a function of wavelength and used a metrological procedure to infer the uncertainty of the computed error from the statistics of the measured variables. To describe the reflectance measurement performance, we have used the average error (AE) and the average error standard deviation (AESTD), calculated for each illumination setup and averaged over the eight-tile set. The best results have been obtained with the metallic iodide and incandescence lamps used in a sequence. In this case, the absolute AE is less than 0.02 over the whole spectrum, and the AESTD less than 0.01 between 420 and 830 nm and less than 0.08 between 400 and 420 nm. The proposed methodology can be generalized to quantify the effects of other influence quantities (e.g. the surface material) onto the accuracy of reflectance measurements, and can also be extended to other types of multispectral sensors.

We have then presented a statistical algorithm to separate the contributions of spectral reflectance and illumination from a multispectral color signal. The algorithm rests upon the finite linear space hypothesis for both reflectance and illumination, i.e. assumes that both signals can be expressed as linear combinations of a finite number of wavelength-dependent basis functions through spatially-dependent coefficients. Physical considerations related to the specific illumination setup allow to reduce the number of illumination basis functions to one. The illumination is therefore separable into the product of a spatially varying coefficient $k(\mathbf{p})$ and a known spectrally dependent basis function. Solving this formulation of the problem requires the estimation of the scale factor (intensity) of the illumination, a parameter which is left unestimated in the illumination-reflectance separation algorithms present in the literature. Estimating this scale factor requires solving an ill-conditioned problem, as the illumination and reflectance coefficients are multiplied into the color signal.

The algorithm consists of two steps. The first is that of illumination estimation. Based on the physical constraints that the illumination is subject to, we have derived a relation linking the maximum (over wavelength) of the color signal with the product of the illumination coefficient and the maximum (over wavelength) of the spectral reflectance, at each spatial location. The illumination estimation method is based on the observation that the illumination is a slowly-varying function of spatial coordinates, in opposition to reflectance. Therefore, the algorithm fits a smooth surface (a plane, in this implementation) to the measured signal $\mathcal{L}(\mathbf{p})$ through a modification of the RANSAC algorithm. In this context, the illumination coefficient $k(\mathbf{p})$ is regarded as the useful signal to be estimated, whereas the maximum reflectance at $\mathbf{p}$ is seen as a multiplicative random noise, with values ranging in the [0,1] interval, which corrupts $k(\mathbf{p})$, yielding the observed $\mathcal{L}(\mathbf{p})$. The model of the underlying statistics for the multiplicative noise is part of the solution to the problem, and is of crucial importance. The random noise has been assumed to have independent, identically distributed samples. This hypothesis does not necessarily hold, as reflectance values are highly correlated in space, but it allows for great generality and avoids the need of making further assumptions on the sample correlation. The RANSAC algorithm has been modified according to this formulation, changing the definitions of the thresholds and of the consensus set. All candidate illuminations are checked for consistency with the physical constraints, so that only feasible solutions are considered. Once the illumination has been estimated, the reflectance coefficients are derived from the linear coefficients $\eta_i(\mathbf{p})$, found as the solution of a linear system. The reflectance coefficients are simply computed as the ratio between $\eta_i(\mathbf{p})$ and $k(\mathbf{p})$.

The algorithm has been tested on four case studies, representing artworks of different pictorial techniques, color characteristics and dimensions. The mean relative error on both illumination and reflectance is lower than 20%, and is only 2% higher for the latter. On the contrary, the infinity norm of the reflectance error is 10% higher than that on the illumination, reaching almost 44%. This suggests that the infinity norm of the error is more sensitive to the effects of error propagation, and therefore that it represents the most significant performance evaluation parameter to consider in further research.

Lastly, we have considered the problem of computing the extrinsic parameters of our multispectral camera from a set of known 3D-2D correspondences, with the aim of texturing uncalibrated 3D models with the acquired multispectral images. The Shape&Color multispectral camera can be regarded as a cylindrical panoramic camera, a camera model which has been extensively studied in the literature. Previous work has concerned the formulation of the projection function, the estimation of the intrinsic parameters, and the definition of the epipolar geometry between pairs of panorama images, with the aim of performing stereo panorama reconstruction. An application to texture mapping onto laser 3D data has also been considered by other authors, but the proposed extrinsic parameter estimation procedure is very specific to the considered application, and cannot be applied in our case.

The extrinsic parameter estimation problem has been formulated in a classical way as a reprojection error minimization problem. Due to the nonlinear nature of the projection operator and the high dimensionality of the problem, the resulting objective function can be expected to exhibit a very complicated behavior. As a consequence, the minimization results are strongly dependent on the chosen initial values for the parameters, and the risk of being trapped in a local minimum is high. We have proposed a way of computing a reliable initial point for the minimization procedure, in order to lower such risk. Our method follows a classical approach, that is, that of computing the rototraslation parameters in the ideal case, as if the 3D and 2D correspondence points were not affected by measurement errors. The problem has been solved using uniquely the linear projection operator along the y axis. The extrinsic rototranslation parameter estimation has been performed in two steps via standard linear algebra computations. The proposed approach has been tested on synthetic 3D-2D correspondences corrupted by Gaussian random noise of increasing standard deviation. The estimation results are good. An increase of the estimation error with the input noise level has been detected. The error behavior is nearly linear in the considered interval, but more data are needed to extend this consideration to broader ranges.

Bibliography

[1] M. Barni, A. Pelagotti, and A. Piva, “Image processing for the analysis and conservation of paintings: opportunities and challenges,” *Signal Processing Magazine*, vol. 22, no. 5, pp. 141–144, Sep. 2005.

[2] B. Hill, “Multispectral color technology: A way towards high definition color image scanning and encoding,” in *Electronic Imaging*, 1998, pp. 2–13.

[3] G. Novati, P. Pellegri, and R. Schettini, “An affordable multispectral imaging system for the digital museum,” *International Journal on Digital Libraries*, vol. 5, no. 3, pp. 167–178, 2005.

[4] K. Martinez, J. Cupitt, D. Saunders, and R. Pillay, “Ten years of art imaging research,” *Proceedings of the IEEE*, vol. 90, no. 1, pp. 28–41, 2002.

[5] K. Martinez and A. Hamber, “Towards a colorimetric digital image archive for the visual arts,” in *Proceedings of the IEEE*, vol. 1073, Jan. 1989, pp. 114–121.

[6] K. Martinez, “High resolution digital imaging of paintings: The vasari project,” *Microcomputers for Information Management*, vol. 8, no. 4, pp. 277–283, 1991.

[7] K. Martinez, J. Cupitt, and D. Saunders, “High resolution colorimetric imaging of paintings,” in *Proceedings of SPIE - Cameras, Scanners, and Image Acquisition Systems*, vol. 1901, Feb. 1993, pp. 25–36.

[8] J. Farrell, J. Cupitt, D. Saunders, and B. Wandell, “Estimating spectral reflectances of digital artwork,” in *Proceedings of the International Symposium on Multispectral Imaging and Color Reproduction for Digital Archival*, 1999, pp. 58–64.

[9] J. Cupitt, K. Martinez, and D. Saunders, “Methodology for art reproduction in colour: the MARC project,” *Computers and the History of Art Journal*, vol. 6, no. 2, pp. 1–20, 1996.

[10] F. Schmitt, H. Brettel, and J. Hardeberg, “Multispectral imaging development at ENST,” *Displaying and Imaging*, vol. 8, pp. 261–268, Jun. 2000.

[11] H. Maître, F. Schmitt, and J.-P. Crettez, “High quality imaging in museum: from theory to practice,” in *Proceedings of SPIE - Very High Resolution and Quality Imaging II*, vol. 3025, pp. 30–39.

[12] H. Maître, F. Schmitt, and C. Lahanier, "15 years of image processing and the fine arts," in *Proceedings of the IEEE International Conference on Image Processing (ICIP01)*, vol. 1, Thessaloniki, Greece, Oct. 2001, pp. 557–561.

[13] F. Schmitt, "High quality digital color images," in *Proceedings of the International Conference on High Technology, (CHIBA'96)*, 1996, pp. 55–62.

[14] H. Maître, F. Schmitt, J. Crettez, Y. Wu, and J. Y. Hardeberg, "Spectrophotometric image analysis of fine art paintings," in *Proceedings of the Fourth Colour Imaging Conference, IS&T - SID*, Scottsdale, Arizona, Nov. 1996, pp. 50–53.

[15] J. Hardeberg, F. Schmitt, H. Brettel, J.-P. Crettez, and H. Maître, "Multispectral image acquisition and simulation of illuminant changes," *Colour Imaging: Vision and Technology*, pp. 145–164, 1999.

[16] J. Hardeberg, F. Schmitt, and H. Brettel, "Multispectral image capture using a tunable filter," in *Proceedings of SPIE - Color Imaging: Device Independent Color, Color Hardcopy and Graphic Arts V*, vol. 3963, 2000, pp. 77–88.

[17] J. Hardeberg, "Multispectral color image acquisition," in *Proceedings of NORSIG-2001*, Trondheim, Norway, Oct. 2001, pp. 77–82.

[18] ——, *Acquisition and Reproduction of Color Images - Colorimetric and Multispectral Approaches.* Dissertation.com, USA, 2001.

[19] Y. Yokoyama, T. N, H. Haneishi, Y. Miyake, J. Hayashi, and M. Saito, "A new color management system based on human perception and its application to recording and reproduction of art paintings," in *Proc. of the Fifth Color Imaging Conference: Color Science, Systems, and Applications*, pp. 169–172.

[20] H. Haneishi, T. Hasegawa, A. Hosoi, Y. Yokoyama, N. Tsumura, and Y. Miyake, "System design for accurately estimating spectral reflectance of art paintings," *Applied Optics*, vol. 39, no. 35, pp. 6621–6632, 2000.

[21] H. Haneishi, T. Hasegawa, N. Tsumura, and Y. Miyake, "Design of color filters for recording artworks," in *Proceedings of the IS&T 50th Annual Conference*, 1997, pp. 369–372.

[22] J. Y. Hardeberg, F. J. Schmitt, and H. Brettel, "Multispectral color image capture using a liquid crystal tunable filter," *Optical Engineering*, vol. 41, pp. 2532–2548, Oct. 2002.

[23] G. Antonioli, F. Fermi, C. Oleari, and R. Reverberi, "Spectrophotometric scanner for imaging of paintings and other works of art," in *Proceedings of CGIV 2004: The Second European Conference on Colour Graphics, Imaging and Vision*, 2004, pp. 219–224.

[24] N. Brusco, S. Capeleto, M. Fedel, A. Paviotti, L. Poletto, G. M. Cortelazzo, and G. Tondello, "A system for 3D modeling frescoed historical buildings with multispectral texture information," *Machine Vision and Applications*, vol. 17, pp. 373–393, Jun. 2006.

[25] A. Mansouri, A. Lathuiliére, F. S. Marzani, Y. Voisin, and P. Gouton, "Toward a 3D multispectral scanner: an application to multimedia," *IEEE MultiMedia*, vol. 14, no. 1, pp. 40–47, Jan 2007.

[26] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, Jun. 1981.

[27] F. Huang, S. K. Wei, R. Klette, G. Gimelfarb, R. Reulke, M. Scheele, and K. Scheibe, "Cylindrical panoramic cameras - from basic design to applications," in *Image and Video Computing New Zealand*, Auckland, New Zealand, 2002, pp. 101–106.

[28] R. Klette and K. Scheibe, "Combinations of range data and panoramic images - new opportunities in 3D scene modeling," in *International Conference on Geometry, Morphology, and Computational Imaging Computer Graphics, Imaging and Vision: New Trends*, 2005, pp. 3–10.

[29] M. Caldon, M. Fedel, L. Poletto, N. Brusco, G. Cortelazzo, and A. Paviotti, "Simultaneous acquisition of 3D shape and color texture from large frescoed areas," in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, vol. 6618, Jul. 2007.

[30] [Online]. Available: RIEGL+LMS-Z420ihttp://www.riegl.com/

[31] [Online]. Available: CYRAX2500http://www.cyra.com/news/cyrax2500.html

[32] P. Dias, V. Sequeira, F. Vaz, and J. G. M. Goncalves, "Fusion of intensity and range data for improved 3D models," in *International Conference on Image Processing (ICIP01)*. Thessaloniki, Greece: IEEE Press, 2001.

[33] ——, "Registration and fusion of intensity and range data for 3D modelling of real world scenes," in *Fourth International Conference on 3-D Digital Imaging and Modeling (3DIM03)*. Banff, Canada: IEEE Press, 2003.

[34] G. Bostrom, M. Fiocco, D. Puig, A. Rossini, J. G. M. Goncalves, and V. Sequeira, "Acquisition, modelling and rendering of very large urban environments," in *Second International Symposium on 3D Data Processing, Visualization and Transmission (3DPVT04)*. Thessaloniki, Greece: IEEE Press, 2004.

[35] Eyetronics: 3d scanning services and hardware. [Online]. Available: http://www.eyetronics.com/

[36] Cyberware: 3d scanning systems and software. [Online]. Available: http://www.cyberware.com/

[37] Konica minolta: 3d scanning systems. [Online]. Available: http://www.konicaminolta-3d.com/

[38] Cam2. [Online]. Available: http://www.faro.com/default.aspx?ct=it

[39] T. Bosch and M. Lescure, *Selected Papers on Laser Distance Measurement.* Bellingham, WA: SPIE Milestone Series, MS115, 1995.

[40] R. G. Dorsch, G. Hausler, and J. M. Hermann, "Laser triangulation: Foundamental uncertainty in distance measurement," *Applied optics*, vol. 33, pp. 1306–1314, 1994.

[41] S. Donati, *Electro-optical Instrumentation.* Prentice Hall, 2004.

[42] J. P. Mohen, M. Menu, and B. Mottin, *Mona Lisa: inside the painting.* Harry N. Abrams Inc., 2006.

[43] A. Ribes, F. Schmitt, R. Pillay, and C. Lahanier, "Calibration and spectral reconstruction for CRISATEL: an art painting multispectral system," *Journal of Imaging Science and Technology*, vol. 49, no. 6, pp. 563–573, 2005.

[44] G. S. E. Ciliberto, Ed., *Modern Analytical Methods in Art and Archeology.* New York, USA: John Wiley & Sons, 2000.

[45] S. Tominaga, "Spectral imaging by a multichannel camera," *Journal of Electronic Imaging*, vol. 8, no. 4, pp. 332–341, 1999.

[46] F. H. Imai, M. R. Rosen, and R. S. Berns, "Multi-spectral imaging of van Gogh's self-portrait at the National Gallery of Art, Washington, D.C." in *Proceedings of the IS&T PICS Conference*, Springfield, Virginia, Oct. 2001, pp. 185–189.

[47] H. Sugiura, T. Kuno, N. Watanabe, N. Matoba, J. Hayashi, and Y. Miyake, "Development of a multispectral camera system," in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, ser. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, M. M. Blouke, N. Sampat, G. M. Williams, and T. Yeh, Eds., vol. 3965, May 2000, pp. 331–339.

[48] W. L. Wolfe, *Introduction to imaging spectrometers.* SPIE Tutorial Text Vol. 25, 1997.

[49] P. Palojärvi, "Integrated electronic and optoelectronic circuits and devices for pulsed time-of-flight laser rangefinding," Master's thesis, Oulu University, Oulu, FI, 2003.

[50] R. I. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd ed. Cambridge University Press, ISBN: 0521540518, 2004.

[51] N. Brusco, A. Giorgi, M. Andreetto, and G. Cortelazzo, "3D registration by textured spin-images," in *Proceedings of The 5th International Conference on 3D Digital Imaging and Modeling (3DIM05).* Banff, Alberta, Canada: IEEE Press, June 2005, pp. 592–599.

[52] L. Lucchese, G. Doretto, and G. M. Cortelazzo, "A frequency domain technique for 3-D view registration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 11, pp. 1468–1484, November 2002.

[53] K. Pulli, "Surface Reconstruction and Display from Range and Color Data," Ph.D. dissertation, University of Washington, 1997.

[54] N. Brusco, P. Zanuttigh, G. M. Cortelazzo, and D. Taubman, "Distortion-sensitive synthesis of texture and geometry in interactive 3D visualization," in *Third international Symposium on 3D Data Processing, Visualization and Transmission (3DPVT06)*, Chapel Hill, N Carolina, USA, 2006.

[55] M. Wheeler, Y. Sato, and K. Ikeuchi, "Consensus surfaces for modeling 3D objects from multiple range images," in *Proceedings of the Sixth International Conference on Computer Vision.* IEEE, 1998, pp. 917–924.

[56] S. F. El-Hakim, "Three-dimensional modeling of complex environments," in *Proceedings of SPIE*, vol. 4309, Dec. 2000, pp. 162–173.

[57] G. Wyszecki and W. S. Stiles, *Color Science - Concepts and Methods, Quantitative Data and Formulae.* Wiley, 1982.

[58] (2006) Labsphere, Inc. [Online]. Available: http://www.labsphere.com/

[59] S. B. Howell, *Handbook of CCD astronomy.* Cambridge, UK: Cambridge University Press, 2006.

[60] *Guide to the expression of uncertainty in measurement*, ISO Std. 17.020, 2007.

[61] F. E. Nicodemus, J. C. Richmond, J. J. Hsia, I. W. Ginsberg, and T. Limperis, "Geometrical considerations and nomenclature for reflectance," Final Report National Bureau of Standards, Washington, DC. Inst. for Basic Standards., Tech. Rep., Oct. 1977.

[62] R. Ramamoorthi and P. Hanrahan, "A signal-processing framework for inverse rendering," in *SIGGRAPH 2001, Computer Graphics Proceedings*, E. Fiume, Ed. ACM Press / ACM SIGGRAPH, 2001, pp. 117–128. [Online]. Available: citeseer.ist.psu.edu/ramamoorthi01signalprocessing.html

[63] P. Debevec, C. Tchou, A. Gardner, T. Hawkins, C. Poullis, J. Stumpfel, A. Jones, N. Yun, P. Einarsson, T. Lundgren, M. Fajardo, and P. Martinez, "Estimating surface reflectance properties of a complex scene under captured natural illumination," USC ICT, MA, Tech. Rep. ICT-TR-06.2004, Dec. 2004.

[64] K. J. Dana, B. van Ginneken, S. K. Nayar, and J. J. Koenderink, "Reflectance and texture of real-world surfaces," *ACM Transactions on Graphics*, vol. 18, pp. 1–34, Jan. 1999.

[65] C. Schaaf. (2006, Jan.) BRDF and albedo research at Boston University. [Online]. Available: http://www-modis.bu.edu/brdf/

[66] D. H. Brainard and B. A. Wandell, “Analysis of the Retinex theory of color vision,” *Journal of the Optical Society of America A*, vol. 3, pp. 1651–1661, Oct. 1986.

[67] 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, pp. 1651–1661, Oct. 1990.

[68] E. H. Land and J. J. McCann, “Lightness and Retinex theory,” *Journal of the Optical Society of America (1917-1983)*, vol. 61, pp. 1–11, Jan. 1971.

[69] D. A. Forsyth, “A novel algorithm for color constancy,” *International Journal of Computer Vision*, vol. 5, pp. 5–36, Jan. 1990.

[70] E. Angelopoulou, S. Lee, and R. Bajcsy, “Spectral gradient: A material descriptor invariant to geometry and incident illumination,” in *Proceedings of the Seventh International Conference on Computer Vision*, Kerkyra, Greece, Sep. 1999, pp. 861–867.

[71] J. M. Geusebroek, R. Boomgaard, S. Smeulders, and H. Geert, “Color invariance,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, pp. 1338–1350, Dec. 2001.

[72] J. Geusebroek, R. Boomgaard, S. Smeulders, and T. Gevers, “A physical basis for color constancy,” in *Proceeding of The First European Conference on Colour in Graphics, Image and Vision*, Poitiers, France, Apr. 2002, pp. 3–6.

[73] P.-R. Chang and T.-H. Hsieh, “Constrained nonlinear optimization approaches to color-signal separation,” *IEEE Transactions on Image Processing*, vol. 4, pp. 81–94, Jan. 1995.

[74] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, “Optimization by simulated annealing,” *Science*, vol. 220, pp. 671–680, 1983.

[75] A. Ikari, R. Kawakami, R. T. Tan, and K. Ikeuchi, “Separating illumination and surface spectral from multiple color signals,” in *Digitally Archiving Cultural Objects*, K. Ikeuchi and D. Miyazaki, Eds. New York. NY: Springer US, 2008, ch. 15, pp. 297–321.

[76] M. Tsukada and Y. Ohta, “An approach to color constancy using multiple images,” *Proceedings of the Third International Conference on Computer Vision*, pp. 385–389, Dec 1990.

[77] S. Tominaga, “Surface identification using the dichromatic reflection model,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 7, pp. 658–670, 1991.

[78] B. K. P. Horn, “Determining Lightness from an Image,” *Computer Grahics and Image Processing*, vol. 3, pp. 277–299, Dec. 1974.

[79] G. D. Finlayson, S. D. Hordley, and I. Tastl, "Gamut constrained illuminant estimation," *International Journal of Computer Vision*, vol. 67, no. 1, pp. 93–109, 2006.

[80] G. J. Klinker, S. A. Shafer, and T. Kanade, "A physical approach to color image understanding," *International Journal of Computer Vision*, vol. 4, pp. 7–38, 1990.

[81] E. Land, "The Retinex theory of color vision," *Scientific American*, vol. 237, pp. 108–129, 1977.

[82] R. Lenz, P. Meer, and M. Hauta-Kasari, "Spectral-based illumination estimation and color correction," *Color Research & Application*, vol. 24, pp. 98–111, Mar. 1999.

[83] L. Fellin, G. Forcolini, and P. Palladino, *Manuale di illuminotecnica*, 2nd ed. Milan, Italy: Tecniche Nuove, 1999.

[84] D. B. Judd, D. L. Macadam, and G. Wyszecki, "Spectral distribution of typical daylight as a function of correlated color temperature." *Journal of the Optical Society of America*, vol. 54, pp. 1031–1040, Aug. 1964.

[85] J. Cohen, "Dependency of the spectral reflectance curves of Munsell color chips," *Psychonomic Science 1*, vol. 1, pp. 369–370, 1964.

[86] L. T. Maloney, "Computational approaches to color constancy," Ph.D. dissertation, Stanford Univ., Stanford, CA, 1984.

[87] ——, "Evaluation of linear models of surface spectral reflectance with a small number of parameters." *Journal of the Optical Society of America A*, vol. 3, pp. 1673–1683, Oct. 1986.

[88] J. P. S. Parkkinen, J. Hallikainen, and T. Jaaskelainen, "Characteristic spectra of Munsell colors," *Journal of the Optical Society of America A*, vol. 6, pp. 318–322, Feb. 1989.

[89] J. L. Dannemiller, "Spectral reflectance of natural objects: how many basis functions are necessary?" *Journal of the Optical Society of America A*, vol. 9, pp. 507–515, Apr. 1992.

[90] G. A. Mian. (2006) Elaborazione numerica dei segnali. [Online]. Available: http://lttm.dei.unipd.it/nuovo/teaching/ens/index.html

[91] Munsell Color, Ed., *Munsell Book of Color - Matte Finish Collection*, Baltimore, Md., 1976.

[92] N. Benvenuto, R. Corvaja, T. Erseghe, and N. Laurenti, *Communication Systems: Fundamentals and Design Methods.* Chichester, UK: John Wiley & Sons, 2006.

[93] K. V. Mardia, J. T. Kent, and J. M. Bibby, *Multivariate Analysis.* London, UK: Academic Press, 1979.

[94] D. Y. Tzeng and R. S. Berns, “A review of principal component analysis and its applications to color technology,” *Color Research & Application*, vol. 30, pp. 84–89, 2005.

[95] R. Lenz, “Spaces of spectral distributions and their natural geometry,” in *Proceedings of The First European Conference on Color in Graphics, Imaging and Vision (CGIV02)*, Poitiers, France, Apr. 2002, pp. 249–254.

[96] C. R. Rao, “The Use and Interpretation of Principal Component Analysis in Applied Research,” *SHANKHYA (Series A)*, vol. 26, pp. 328–358, 1964.

[97] S. M. Kay, *Fundamentals of Statistical Signal Processing, Vol. I - Estimation Theory.* Upper Saddle River, NJ: Prentice Hall, 1993.

[98] Galleria Stefano Forni: Matteo Massagrande. [Online]. Available: http://www.galleriastefanoforni.com/?page_id=176

[99] R. C. Gonzalez and R. E. Woods, *Digital Image Processing.* Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc., 2001.

[100] F. Huang, S. K. Wei, and R. Klette, *Calibration of Line-based Panoramic Cameras.* Springer Netherlands, 2006, pp. 55–84.

[101] ——, “Calibration of panoramic cameras using 3D scene information,” in *Geometry, Morphology, and Computational Imaging*, Springer, Berlin, 2003, pp. 335–345.

[102] F. Huang, S.-K. Wei, and R. Klette, “Comparative studies of line-based panoramic camera calibration,” in *IEEE Workshop on Omnidirectional Vision (Omnivis03)*, 2003.

[103] R. Klette, G. Gimelfarb, S. K. Wei, F. Huang, K. Scheibe, M. Scheele, A. Borner, and R. Reulke, “A review on research and applications of cylindrical panoramas,” CITR, Auckland University, New Zealand, Tech. Rep. 123, Jan. 2003.

[104] Z. Zhang, “A flexible new technique for camera calibration,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 11, pp. 1330–1334, 2000.

[105] T. Y. Tsai, “An efficient and accurate camera calibration technique for 3D machine vision,” in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, Miami Beach, FL, 1986, pp. 364–374.

[106] O. Faugeras, *Three-Dimensional Computer Vision.* Cambridge, MA: MIT Press, 1993.