



UNIVERSITÀ DELLA CALABRIA  
Dipartimento di MECCANICA  
DOTTORATO DI RICERCA IN INGEGNERIA MECCANICA  
CICLO XXIII (2007-2010)

---

DISEGNO E METODI DELL'INGEGNERIA INDUSTRIALE: ING-IND/15

# Digital imaging techniques for Cultural Heritage applications

By Gianfranco Bianco

A dissertation submitted in partial fulfillment of the requirements for the  
Doctoral Research degree in Mechanical Engineering

Doctoral Research Director  
*Prof. Sergio Rizzuti*

Supervisor  
*Ing. Fabio Bruno*

Candidate  
*Dr. Gianfranco Bianco*





To my family

*“Le véritable voyage de découverte ne consiste pas à chercher de nouveaux paysages mais à avoir de nouveaux yeux.” (M. Proust)*



# Abstract

This thesis deals with the use of digital imaging techniques for cultural heritage applications. Thanks to the development of digital technology during the last three decades, new tools for study and analysis of tangible objects of cultural interest have been obtained. The goals of these analyses are the protection, documentation and monitoring of historic and artistic heritage for its preservation for future generations, aiming also to facilitate the remote consultation and the dissemination of digital copies on the web. The techniques treated in this work regard the image acquisitions for both two-dimensional and three-dimensional analysis, applied in two specific areas: the analysis and virtual restoration of ancient documents using multispectral imaging techniques, and the 3D reconstruction by structured light techniques for underwater archeology. These areas require the acquisition of images that are elaborated using both image processing and signal processing techniques.

As for 3D imaging, experimental texts of a 3D scanner built for underwater applications have been conducted in turbid water conditions, using two structured-light techniques. The scanner is composed by two cameras and a projector, opportunely placed in water-proof housings. The gray-code in combination with the stereovision technique involves the projection of a number of vertical and horizontal patterns with decreasing pitch of the black/white bands. The images of an object are encoded using a binary code. This information is later used for the triangulation of the acquired points. The results show that the 3D reconstructions are acceptable also in presence of high turbidity values. Moreover, underwater 3D acquisition by fringe projection technique has been experimented. This technique employs a camera and a projector, and has the great advantage to significantly reduce the number of patterns.

About the 2D imaging, the virtual restoration of ancient document based on multispectral imaging has been analysed. Multispectral acquisition, performed using a digital camera and optical filters, allows the digitization of a scene over the wavelengths of the visible spectrum, extending the acquisition at the infrared and ultraviolet bands, which can detect details that are invisible to the naked eye or to the classic RGB acquisition. Digital analysis of ancient documents involves various steps: acquisition, registration, decorrelation and image binarization. In particular, an automated method has been applied for registration of multispectral images, that is the preliminary process needed to apply the subsequent decorrelation statistical techniques. Moreover, this registration method has been applied in several cases on two or more channels, and finally to the double-sided case (recto-verso). Image

decorrelation techniques based on blind source separation are able to extract pattern, hidden features or attenuate the degradations. Moreover, they can improve the performance of the subsequent binarization process to separate the main text in a document.

Also, the correction of geometric distortions that may affect a document surface has been studied adding 3D acquisition by structured light to the previous multispectral analysis. For this specific experimentation, a multispectral camera was coupled with a digital projector for projecting a sinusoidal fringe pattern (fringe projection technique). It is important to underline that this setup does not necessarily require the calibration of the 3D system, to the benefit of the necessary time for testing and elaboration data. The document alterations are corrected by surface parameterization, mapping the 3D model on a plane, with its texture. The mapping is done by conformal mapping algorithm.

# Sommario

La presente tesi tratta l'impiego di tecniche di imaging digitale per applicazioni nel campo dei Beni Culturali. Grazie allo sviluppo da circa tre decenni della tecnologia digitale, si sono acquisiti nuovi strumenti di indagine per lo studio e l'analisi di oggetti tangibili di carattere culturale. Gli scopi di queste investigazioni riguardano la tutela, la documentazione e il monitoraggio, per la conservazione dei beni di interesse storico e artistico per le generazioni future e per favorirne la consultazione in remoto e la diffusione di copie digitali sul web. Le tecniche studiate riguardano l'acquisizione e l'analisi di immagini sia in bidimensionale che in tridimensionale e trovano applicazione in due settori specifici: l'analisi e il restauro virtuale di documenti antichi attraverso tecniche di imaging multispettrale; il rilievo mediante tecniche di imaging 3D in luce strutturata per l'archeologia subacquea. Entrambi i campi di applicazione prevedono l'acquisizione di immagini che vengono successivamente elaborate mediante tecniche di image processing e signal processing.

Per quanto riguarda l'imaging 3D, è stata condotta la sperimentazione di uno scanner 3D progettato per il rilievo subacqueo in condizioni di acqua torbida, mediante luce strutturata. Lo scanner è dotato di due fotocamere ed un proiettore. La tecnica del gray-code in combinazione con la stereo visione prevede la proiezione di una serie di pattern a bande verticali ed orizzontali, bianche e nere con ampiezza decrescente; ciò permette la codifica della scena mediante un codice binario. Informazione che viene utilizzata successivamente per la triangolazione dei punti da acquisire. I risultati sperimentali mostrano delle acquisizioni accettabili anche per elevati valori di torbidità. Si è inoltre sperimentato la ricostruzione 3D mediante fringe projection technique, che impiega una sola fotocamera ed un proiettore, ed ha il grande vantaggio di ridurre notevolmente il numero di pattern.

L'imaging 2D è stato trattato per il restauro di documenti antichi, per mezzo di acquisizioni multispettrali. L'acquisizione multispettrale eseguita mediante opportuna camera digitale e filtri ottici permette la digitalizzazione della scena oltre le lunghezze d'onda del visibile, estendendo l'acquisizione alle bande dell'infrarosso e dell'ultravioletto, consentendo di rilevare dettagli invisibili ad occhio nudo o alla classica acquisizione RGB. Sono state affrontate pertanto le varie fasi dell'analisi digitale di documenti antichi: acquisizione, registrazione, decorrelazione e binarizzazione di immagini. In particolare, sono stati applicati metodi automatici di registrazione, fase preliminare necessaria per poter applicare le successive tecniche statistiche di decorrelazione. Altresì, sono stati applicati i vari metodi di registrazione a vari casi di due o più canali, ed infine al caso recto-verso. Le tecniche di decorrelazione basate sulla

separazione cieca delle sorgenti sono in grado di estrarre pattern, particolari nascosti o attenuare le degradazioni. Inoltre, possono migliorare le prestazioni del processo di binarizzazione successivo per separare il testo principale in un documento.

Si è studiato, inoltre, la riduzione delle distorsioni geometriche di cui può essere affetta la superficie di un documento aggiungendo l'acquisizione 3D tramite luce strutturata, alla precedente analisi multispettrale. Per questa specifica sperimentazione, alla camera multispettrale si è affiancato un proiettore digitale per la proiezione di un pattern di frange sinusoidali (fringe projection technique). E' da sottolineare che l'impiego di questo setup non richiede necessariamente la calibrazione del sistema 3D, a vantaggio del tempo impiegato per l'acquisizione ed elaborazione dei dati. Le distorsioni del documento sono corrette per mezzo di parametrizzazione della superficie, mappando il modello 3D in un piano, con relativa texture. La mappatura viene effettuata tramite algoritmo di mappa conforme.

# Acknowledgments

The research presented in this work would not have been possible without the invaluable guide and support of the following people:

- my supervisor Dr. PhD Fabio Bruno for the constant assistance and encouragement during my work;

- the Prof. Maurizio Muzzupappa and the other members of my research team (Prof M. L. Luchi and Prof. S. Rizzuti);

- my colleagues and the students with whom I constantly shared knowledge and experience;

- the researches Anna Tonazzini and Emanuele Salerno of CNR (ISTI) of Pisa who introduced me the world of virtual restoration of ancient documents;

- the head of the Department of Image Processing (UTIA) in Prague, Barbara Zitová and the deputy head Filip Šroubek for the interesting study experience during the doctoral stage;

- the company T.E.A. sas (Catanzaro) of Elena Console for the collaboration;

- the technicians of Laboratory of the Department of Mechanics at the University of Calabria;

- and finally my family for the constant moral support.





# List of Figures

I.1	The growth of peer-reviewed research publications produced by scientists in the fields of archaeology, archaeometry, and cultural heritage conservation since 1986 [Creagh & Bradley 2007]. . . . .	2
I.2	Image-based techniques scheme. . . . .	6
II.1	The interaction between the light and the underwater environment includes basically two processes: absorption and scattering. The latter can be further divided into backscatter and forward scattering. . . . .	20
II.2	At left, projection of binary patterns and vertical line coding on the image planes. At right, examples of binary patterns and code shifting. . . . .	21
II.3	3D reconstruction scheme. . . . .	22
II.4	Pinhole camera model. . . . .	24
II.5	Coordinate systems of the optical setup. . . . .	25
II.6	Samples for camera calibration (checkerboard). . . . .	26
II.7	Virtual prototype of the underwater 3D scanner. . . . .	27
II.8	Underwater 3D system . . . . .	27
II.9	Optical path of rays at water-air interface, for a flat port (top) and a dome port (bottom) [Karpel & Schechner 2004]. . . . .	28
II.10	Experimental setup for underwater 3D acquisition. . . . .	29
II.11	Calibration step, in water tank, with checkerboard . . . . .	31
II.12	Calibration panel. . . . .	31
II.13	The figures show the complete distortion model (radial plus tangential) on each pixel of the image for left and right camera in air and water. The cross indicates the centre of the image, the circle, the location of the principal point and the arrows are the optical distortions. . . . .	32
II.14	Distribution of the re-projection error in air and water. . . . .	33
II.15	Unclean point cloud: the blue points are due to scattering effects. . . . .	35

II.16	Texture mapping: underwater image (a), retouched image (b) and textured 3D model of a earthenware jar (c). . . . .	35
II.17	Distribution of the deviations between the shape acquired in water and in air (a) and comparison of the point clouds with deviations evidenced through a color map (b). . . . .	38
II.18	Optical configuration of setup in fringe projection technique.	40
II.19	Flowchart of FPT. . . . .	41
II.20	Fourier frequency spectrum of the fringe pattern. . . . .	43
II.21	One dimensional phase unwrapping. . . . .	48
II.22	Simulations scheme. . . . .	49
II.23	One-dimensional comparison between the height function $h$ ( <i>peaks</i> Matlab function) and the height reconstructed $H$ computed by WFT and WT. . . . .	50
II.24	Example of 3D reconstruction by FPT. . . . .	51
II.25	Setup . . . . .	52
II.26	Correspondence of a phase line among the projected fringe pattern (a), the acquired pattern (b) and the calculated phase-map (c). . . . .	54
II.27	Fringe patterns and 3D point clouds, in turbid water. . . . .	55
III.1	Electromagnetic spectrum. . . . .	58
III.2	Schematic view of the image acquisition process. . . . .	58
III.3	Multispectral data cube [Klein <i>et al.</i> 2008]. . . . .	60
III.4	(a) Chroma CX3 camera, filter-wheel camera (b). . . . .	61
III.5	Optical scheme. . . . .	62
III.6	Shifting of focus plane due to different refractive index of filters.	62
III.7	Multispectral imaging setup. . . . .	63
III.8	Document images acquired in different spectral bands. . . . .	64
III.9	Misalignment of two channels illustrated as overlapping in transparency. . . . .	65
III.10	Show-through: can appear in the scanned image when the paper is not completely opaque. . . . .	67
III.11	Bleed-through: due to seeping of ink from the reverse side. . . . .	67
III.12	Example of ink diffusion (spot) due to humidity. . . . .	67
III.13	Palimpsest. . . . .	68
III.14	Papyrus. . . . .	68
III.15	Registration of reference and sensed images. . . . .	70
III.16	Transformation . . . . .	75
III.17	Registration of Green and Infrared channels. The details of the document illustrated in Figure III.18 are showed as overlapped in transparency. . . . .	77

---

III.18	Multispectral images acquired from IR to UV bands to be registered. . . . .	78
III.19	Superposition of different (unknown) patterns in multispectral imaging. Channel selectivity from diversity (measured) images can be translated into pattern selectivity. The outputs are estimates of the patterns. . . . .	80
III.20	Recovering of background (b), foreground (c), and bleed-through (d) from an RGB document by ICA algorithm. . . .	81
III.21	Watermark detection: (a) infrared front view; (b) back illumination infrared view; (c) one ICA output. . . . .	82
III.22	The document images of Figure III.18 processed by SO. . . .	83
III.23	Stain attenuation by PCA. . . . .	84
III.24	Recto-verso captures. . . . .	86
III.25	Recto-verso captures. . . . .	87
III.26	Recto-verso registration: the misaligned and registered images are shown as overlaid in transparency. . . . .	88
III.27	Restoration of a pair of RGB recto and verso sides of a real document page. . . . .	90
III.28	Application of symmetric orthogonalization to extract a stamp in the RGB documents shown in Figure III.27. . . . .	91
III.29	Restoration of the pair of RGB recto and verso sides showed in Figure III.28. . . . .	92
III.30	Visual and numeric results on the comparison algorithms applied to illuminated, textual images [Sauvola & Pietaksinen 2000]. . . . .	93
III.31	Segmentation with Sauvola algorithm of a decorrelated image. . . . .	94
III.32	The decorrelation improves the segmentation process. . . . .	95
III.33	Detail of a document image captured in IR band: the image is blurred. . . . .	97
III.34	Artificial markers with three different radii introduced in an image. The out-of-focus blurring is apparent. . . . .	98
IV.1	Geometrical distortions of the pages in an old printed book. . . . .	103
IV.2	Optical configuration of . . . . .	104
IV.3	Telecentric . . . . .	105
IV.4	. . . . .	106
IV.5	3D triangle to 2D triangle mapping. . . . .	108
IV.6	Framework flowchart. . . . .	110
IV.7	Imaging system for 2D/3D data acquisition. . . . .	111
IV.8	Green channel image (a), object fringe pattern (b) and reference fringe pattern (c). . . . .	112

IV.9	Textured 3D model of a document. . . . .	113
IV.10	Triangulated 3D mesh. . . . .	113
IV.11	The developable surface is mapped on a plane (a). A comparison before (up) and after surface parameterization (down) (b). . . . .	114
IV.12	A sign is present in the original image(a). By applying Symmetric Orthogonalization to the multispectral images the sign is removed (b), to improve the binarization step (c). . . . .	114

# List of Tables

II.1	Classification of 3D acquisition techniques used in underwater environment. . . . .	16
II.2	Intrinsic parameters of cameras for calibration in clear water and air. . . . .	30
II.3	Measurements of the turbidity conditions. . . . .	34
II.4	Images of the pattern projected on various materials at different levels of turbidity. . . . .	34
II.5	3D point clouds obtained in different optical conditions. . . . .	36
II.6	Number of points per 100 pixels acquired on the various object for each turbidity level. . . . .	37
II.7	Statistical parameters (mean value $\mu$ and standard deviation $\sigma$ , expressed in mm) calculated for the various objects for each turbidity level. . . . .	37
II.8	Percentage values of the root-mean-squares error $E_{RMS}$ . . . . .	50
III.1	Results of registration with different changing of transformation. . . . .	76
III.2	Displacements obtained from the registration of the multispectral images of Figure III.17 . . . . .	76



# Contents

<b>Abstract</b>	<b>iii</b>
<b>Sommario</b>	<b>v</b>
<b>Acknowledgments</b>	<b>vii</b>
<b>List of Figures</b>	<b>ix</b>
<b>List of Tables</b>	<b>xiii</b>
<b>I Introduction</b>	<b>1</b>
I.1 Imaging . . . . .	5
I.2 Overview of imaging techniques used in the study of cultural heritage artefacts . . . . .	7
I.3 Thesis objectives . . . . .	10
I.4 Structure of the Thesis . . . . .	11
<b>II 3D imaging by structured light techniques for underwater applications</b>	<b>13</b>
II.1 Introduction . . . . .	13
II.2 Related works . . . . .	15
II.2.1 Underwater 3D optical imaging . . . . .	16
II.2.1.1 Passive techniques . . . . .	16
II.2.1.2 Active techniques . . . . .	17
II.2.1.3 Limitations of underwater imaging . . . . .	18
II.2.2 Whole-field structured light technique: pattern projection	19
II.3 Gray-code and stereo-vision techniques . . . . .	20
II.3.1 Calibration . . . . .	23
II.3.2 Accuracy . . . . .	24
II.3.3 Virtual prototype of an underwater 3D system . . . . .	25
II.3.4 Experimental setup . . . . .	26
II.3.5 Experimentation . . . . .	29
II.3.6 System calibration . . . . .	29

II.3.7	Image acquisition . . . . .	31
II.3.8	3D reconstruction . . . . .	34
II.3.9	Results . . . . .	35
II.3.10	Discussion about potential applications . . . . .	37
II.4	Fringe Projection Technique . . . . .	40
II.4.1	Fringe Analysis . . . . .	41
II.4.1.1	Phase unwrapping . . . . .	47
II.4.2	Simulations . . . . .	49
II.4.3	Phase-to-height conversion . . . . .	50
II.4.4	Calibration . . . . .	53
II.4.5	Experimentation . . . . .	53
II.5	Conclusions . . . . .	55
<b>III Multispectral imaging for ancient documents analysis</b>		<b>57</b>
III.1	Multispectral acquisition . . . . .	57
III.1.1	Multispectral imaging systems . . . . .	59
III.1.1.1	Chroma CX3 multispectral camera . . . . .	60
III.1.2	Multispectral acquisition of ancient documents . . . . .	61
III.2	Document image analysis . . . . .	64
III.2.1	Document degradations . . . . .	65
III.2.2	Enhancement methodology . . . . .	67
III.2.2.1	Related works . . . . .	68
III.2.2.2	Methodology . . . . .	69
III.2.2.3	Image registration . . . . .	69
III.2.2.4	Image decorrelation . . . . .	76
III.2.3	Recto-verso case . . . . .	85
III.2.3.1	Double-sided document registration . . . . .	85
III.2.3.2	Recto-verso decorrelation . . . . .	87
III.2.4	Image segmentation . . . . .	91
III.3	Deblurring . . . . .	94
III.3.1	Related works . . . . .	95
III.3.2	Discussion . . . . .	96
III.4	Conclusions . . . . .	99
<b>IV Combination of multispectral and 3D imaging</b>		<b>101</b>
IV.1	Introduction . . . . .	101
IV.2	Related works . . . . .	103
IV.3	3D imaging . . . . .	104
IV.4	Surface parameterization . . . . .	106
IV.5	A framework for document restoration . . . . .	109



<b>Contents</b>	<b>xvii</b>
<hr/>	
IV.6 Archiving . . . . .	112
IV.7 Conclusions . . . . .	113
<b>Conclusions</b>	<b>115</b>
<b>Bibliography</b>	<b>117</b>



## CHAPTER I

# Introduction

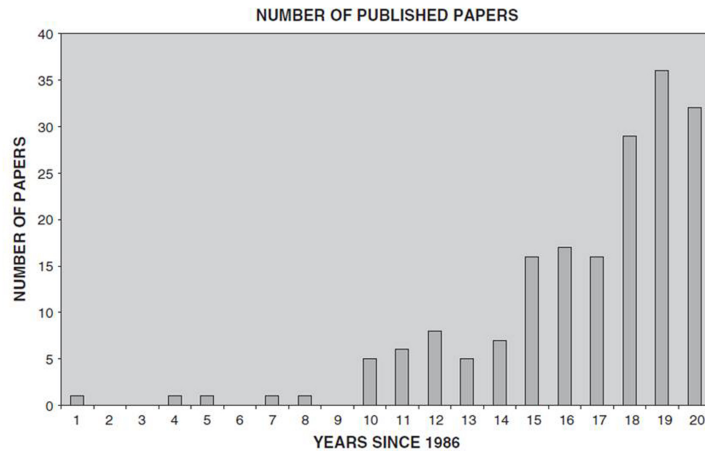
---

Cultural Heritage is a mark of the development of history and culture of humanity, which helps us to understand the past and to identify human race. It is apparent that it is essential to preserve the cultural heritage, in order to benefit the future generations and to safeguard the memory of the world. In fact, the World Heritage Convention of Unesco [Unesco 1972] encourages "*identification, protection, conservation, presentation and transmission to future generations of the cultural and natural heritage*". The same convention establishes the following as cultural heritage:

- monuments: architectural works, works of monumental sculpture and painting, elements or structures of an archaeological nature, inscriptions, cave dwellings and combinations of features, which are of outstanding universal value from the point of view of history, art or science;
- groups of buildings: groups of separate or connected buildings which, because of their architecture, their homogeneity or their place in the landscape, are of outstanding universal value from the point of view of history, art or science;
- sites: works of man or the combined works of nature and man, and areas including archaeological sites which are of outstanding universal value from the historical, aesthetic, ethnological or anthropological point of view.

In a few words a synthetic definition is: "*Every material evidence of civilisation*" (A. Guarino) [Creagh & Bradley 2007]. The tangible cultural heritage can be classified, for simplicity, in several main categories of heritage:

- movable cultural heritage (paintings, sculptures, coins, manuscripts, etc.)
- immovable cultural heritage (monuments, archaeological sites, and so on)
- underwater cultural heritage (shipwrecks, underwater ruins and cities and so on)



**Figure I.1:** The growth of peer-reviewed research publications produced by scientists in the fields of archaeology, archaeometry, and cultural heritage conservation since 1986 [Creagh & Bradley 2007].

In particular, underwater cultural heritage is fascinating due to the mystery of its underwater locations and its historical contexts. The site of a wreck or a submerged ruin is the reminder of a human tragedy. The discovery of the site of a wreck enables people to step back into the past and relive the last moments of the ship and its crew. Once out of the water and exhibited on land, objects are deprived of their context and lose part of their significance. Therefore, several recent initiatives have been undertaken to offer visitors *in situ* experiences while at the same time ensuring the conservation and protection of the original site in line with the principles of the UNESCO 2001 Convention [Unesco 2001]. The *in situ* preservation of underwater cultural heritage shall be considered as the first and preferred option before allowing or engaging in any activity directed at this heritage.

At present, there is an increasing of interest in cultural heritage issues: the peer-reviewed research publications produced by scientists has increased since 1986, as shown in Figure I.1 [Creagh & Bradley 2007]; some relevant organizations provide funding for different projects in cultural heritage preservation and documentation.

In 1997 the National Research Council of Canada and the University of Padova realized 3D digital images of the “Madonna col Bambino” of Giovanni Pisano and of two bas-reliefs of Donatello at the “Cappella degli Scrovegni” in Padova [Cortelazzo G.M. 1999, Beraldin *et al.* 1999]. However, one of the major projects in this field has been the “Michelangelo Project” [Levoy 1999], conducted in Italy by the University of Stanford, California, which acquired 3D digital images of Michelangelo’s David at the “Galleria dell’Accademia” in Florence. This statue, 7 m tall, has been digitized by two billion sampled

---

points and the analyses have revealed traces of the chisel work and the granularity of the material [Beraldin *et al.* 1998, Godin *et al.* 1999]. The VASARI, MARC projects were pioneering projects funded by the European Commission which made use of multispectral data for the acquisition and monitoring of paintings [Martinez 1991, Cupitt *et al.* 1996]. VASARI and MARC considered the multispectral acquisition of paintings in order to derive reliable colour space coordinates. A remarkable multispectral imaging project involving the CRISATEL scanner is the “Mona Lisa project” [Cotte & Dupraz 2006]. The EC project IsyReadDet developed a system for a virtual restoration and archiving of damaged manuscripts [Tonazzini *et al.* 2004a]. Google has scanned millions of books and made them available in a digital database (googlebooks) for online consultation [Google 2010]. The project Europeana was initiated by the European Commission and aims to create a collection of European digital libraries with digitised paintings, books, films and archives in a search platform [Europeana 2010]. The VENUS project has aimed to provide a virtual exploration of deep underwater archaeological sites [Chapman *et al.* 2006]. Moreover, institutions and locations of cultural significance provide an impetus for the tourist industry of a country: for many, cultural tourism contributes substantially to national economy.

Digitisation of data information is a consolidated tool to study the cultural heritage objects that may require different techniques depending on size, environment, accessibility and typology of information needed. Digital technology already plays a fundamental role in everyday life, industry and research. In the last 30 years the quick growth of personal computers, sensors, display devices has allowed a large diffusion of digital devices in all fields, due to fast acquisition, storage, manipulation and display of data. Photography has been revolutionized by CCD (Charged Coupled Device) sensor, that allows to capture images and immediately store them on memory card as digital data for subsequent elaborations by computer, for photo-retouching, printing or publication on the web. Imaging systems can employ (still or movie) cameras, video-projectors, lasers to acquire in digital format (digitisation) many types of information about objects, in two or three dimensions. Texture, shape, volume, etc. are acquired quickly with high resolution data and accuracy.

Currently, science and technology provide imaging techniques to study many aspects of the cultural heritage objects and also to explore hostile environments such as the submarine one. These techniques are non-destructive and no-contact, due to this they are very fit to investigate objects which status can be compromised by the contact, like in the case of cultural heritage artefacts. The conservation scientist needs different non-invasive techniques, for the examination, characterisation, and analysis of cultural heritage artefacts, to improve understanding of their manufacture, their evolution and/or degra-

dition during time. This understanding is necessary to give a rational basis for the restoration and conservation of objects. Generally, the scientific community needs to acquire data in heritage studies: for example, archaeologists need to map an archaeological site and to measure objects, place, flotsams or to virtually reconstruct ancient settlements; philologists are interested to decipher text in degraded manuscripts, like papyri or palimpsests, restorers are interested to discover the re-touching marks in a painting. Digital information is durable and unalterable, and thus can be used as reference for degradation monitoring and restoration of works. In particular, the artworks in an open environment are continuously and irreversibly damaged by atmospheric agents, and frequently the damages are visible too late. High accuracy digital acquisitions, at regular times, could detect very small deformations and cracks [Beraldin *et al.* 1999], allowing opportune interventions. High resolution data of the more valuable artworks could be realized for the study of the formal properties and for heritage protection. In case of damage, artworks could be restored providing a re-match with the digital model. Digital models of lower resolution could meet the issues of the museum archives, documenting a great number of pieces. Furthermore, remote study is of great interest when the artefact is located in an inaccessible site, for example at the top of monumental buildings. Standard control techniques can also benefit from the availability of digital archives: a periodic scanning can ensure a check of possible accidental or man-made alteration of the work. This context has stimulated novel “virtual restoration techniques” [Berni M. 1999]. Computer graphics tools can virtually restore damaged parts of an artwork, creating 3D models in its original aspect or reducing the interferences in a degraded manuscript. Often, ancient sculptures were originally painted with bright colours; colours different from the actual ones, suggested by historical analysis and investigation, could be experimented. Many artworks are now located in architectonic contexts very different from the original ones. Computer graphics tools can envisage the appearance of the artwork inserted in the original architectural environment.

So, many disciplines are involved in these studies to improve the performance of devices, test new algorithms for controlling acquisition and elaboration data, create database and digital archiving. Although the direct observation of images provides an initial qualitative analysis, the image analysis processing deserves particular attention, as it allows the extraction of meaningful information from digital images. These information can be extracted by signal processing techniques, for example, detecting the phase map in pattern images that is related to an object shape. Statistical techniques of image decorrelation are able to extract layers or hidden features from a collection of multispectral images. Thus a number of different techniques can be applied

for virtual restoration of ancient documents. A 3D model can be obtained from 2 or more overlapped images as done by the photogrammetry technique.

## I.1 Imaging

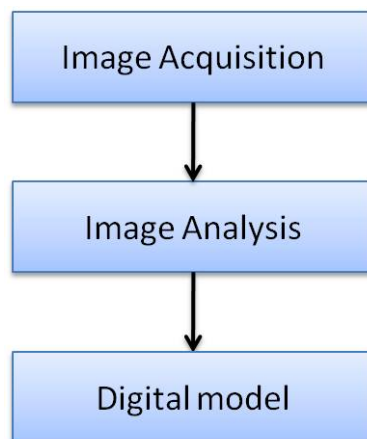
In this section a definition of imaging will be attempted, without the presumption to be exhaustive and complete, but trying to define the *concept*, and to provide any information about this matter, to underline its relevant importance in many fields and to list many different technological solutions available today. Imaging is a multidisciplinary topic, which involves different techniques, applications, methods, devices. Many fields of research and industry such as electronics, informatics, optics, mechanics and mathematics are involved in designing (hardware and software) imaging systems for different applications, like medical diagnostic, terrestrial and underwater mapping, astronomical observations, analysis of artefacts, entertainment, etc. According to the definition of Oxford Dictionary [Oxf 2003], the term imaging means: “*make a visual representation of something by scanning it with a detector or electromagnetic beam*”. In particular, imaging refers to the process of obtaining pictures or more complicated spatial representations, such as 3D computer graphics models from physical targets. Imaging also includes activities such as:

- Taking photographs using non-visible ranges of the electromagnetic spectrum.
- Extracting images from living things, through techniques such as positron emission tomography, magnetic resonance imaging, near-infrared fluorescence imaging, computed axial tomography.
- Microscopy-based image collection techniques such as dark field imaging, staining, or working with devices such as electron microscopes, atomic force microscopes or scanning tunneling microscopes.
- Writing computer programs to improve the quality of collected images, or to construct images which merge information from multiple sources. Note, though, that imaging is largely distinct from visualization.
- Making a copy of the contents of a hard disk.

At present this process involves the use of digital technologies that allow fast capture, storage, manipulation and display of data, for visualization, documentation, dissemination on the Internet; so one can use the term “digital

imaging”. Moreover, the recent quick growth of the processing and memory capacity of computers has made the acquisition, recording and remote fruition of digital images technically feasible, even by relatively inexpensive personal computers. The simplest and most popular imaging system is the digital camera (movie or still) which mounts sensors such as CCD or CMOS (Complementary Metal Oxide Semiconductor) and optics to capture images. Multispectral camera extends the acquisition range in the infrared and ultraviolet spectrum, and it is used in astronomical, remote sensing and cultural heritage applications. Infrared radiation is acquired by a thermo-camera to map the temperature distribution of an object. Imaging radar attempts to form a picture of the object as well, by mapping the electromagnetic scattering coefficient onto a two-dimensional plane. Objects with a higher coefficient are assigned a higher optical reflective index, creating an optical image.

Sonar devices deliver 3D underwater images extracted by acoustic signal. 3D models can be acquired by different techniques, that may use artificial light sources (active techniques) or environment light (passive techniques). Lasers and detectors (CCD sensor or avalanche photo-diode) are employed according to either the principles of optical triangulation or time delay. A video projector is used to project a sequence of patterns (gray-code, sinusoidal fringe) of structured light while a camera captures the relative images of the illuminated scene. Also, photogrammetry techniques create the 3D model acquiring an object with two or more cameras. Underwater imaging is seriously compromised by scattering and absorption in the medium, which decrease image contrast and attenuate light intensity [Schechner & Karpel 2004] with consequent loss of details and colour alteration. However navigation, exploration and acquisition are possible thanks to different solutions proposed to enhance the underwater visibility, i.e. polarization, spatial, time and spectrum dis-



**Figure I.2:** Image-based techniques scheme.



criminations [He & Seet 2004].

For the discussion of techniques applied in this work, it is useful to simplify the principal processes involved in the application of image-based techniques (Figure I.2). First, an image acquisition step allows to acquire 2D data as images by an apposite imaging system. Then techniques of image analysis (i.e. image processing, signal processing, etc.) elaborate the data to extract some kind of information and obtain a digital model of the target which one wants to investigate.

## **I.2 Overview of imaging techniques used in the study of cultural heritage artefacts**

An overview of the main imaging techniques used to digitize heritage objects is presented here. Certainly, conventional photography by reflection of visible light is the most used technique in scientific conservation, for documentation or inventory [Arias *et al.* 2004]. Digital photography under raking light is used by some papyrologists to help them read the incisions on ancient tablets [Anderson S. 2002]: the camera position and the tablet are kept fixed, and a number of images are taken in which the tablet is illuminated by a strongly orientated light source. It is possible also to capture images in the non-visible spectrum, from infrared to ultraviolet bands. In particular, spectral imaging refers to the acquisition of a series of digital images at different wavelength ranges. Traditionally, if the number of wavelength bands is smaller than ten, the term “multispectral imaging” is used, otherwise if the number of wavelength bands is much larger than ten, usually the term “hyperspectral imaging” is used. However, it is also possible to use spectral imaging for investigation of objects of cultural heritage [Havermans *et al.* 2003, Melessanaki *et al.* 2001, Padoan *et al.* 2008]. Spectral acquisitions have become widely employed in the non-invasive diagnosis of the conservation status of paintings and in the reconstruction of its history. For example, infrared (IR) reflectography has been used for the detection of under-drawings and “pentimenti”, while ultraviolet (UV) fluorescence can reveal restoration interventions [Pelagotti *et al.* 2006, Mansfielda *et al.* 2002]. Ancient documents affected by several degradations have been analysed in spectral modality to detect features (symbols, stamps, etc.), to improve the text legibility or to attenuate interferences [Tonazzini *et al.* 2004c]: for example, in the study of the Archimedes Palimpsest (i.e., manuscript pages that have been scraped off and re-used). The erased characters are revealed by the ultraviolet illumination [Easton *et al.* 2003]. In respect to RGB acquisition [Hill 1998] the spectral acquisition allows to measure the reflectance of

an object independently from illumination and acquisition devices, becoming the most reliable asset for faithful colour reproduction. For this reason, a number of cultural heritage institutions have carried out the creation of digital museums [Novati *et al.* 2005] digitizing their collections with multi-spectral sensors. Fluorescence lidar imaging based on laser-induced fluorescence (spectroscopic technique) has recently shown to be a helpful tool for a quick measurement of several indicators that may be used to assess the conservation status of a historical building [Weibring *et al.* 2001]. Experiments were conducted on the Parma baptistery in 1994, where a fluorescence lidar system was used for point monitoring of different lithotypes on the facade [Lognoli *et al.* 2003]. Thermal imaging or infrared thermography allows to acquire images in the infrared spectrum, measuring absolute values or changes of temperature of an object, regardless of its illumination in the visible range. Digital infrared thermography is gaining an increasing relevance in cultural heritage. In fact, several techniques based on high-resolution infrared thermography have been so far developed to provide an effective and non-destructive test to assess the integrity of historical buildings, masonry and historical masonry, frescoes, and wooden art crafts [Clark *et al.* 2003]. For example, architectural surfaces are investigated before and after a pilot cleaning intervention [Avdelidis 2004]. Thermographic image analysis using histograms was applied on the obtained thermal images with the goal of estimating the temperature distributions and identifying the problematic areas. X-ray scattering and diffraction are non-destructive methods that have been applied to a range of archaeological material, including ceramics and glazes [Fermo *et al.* 2002, Tang *et al.* 2001], textiles [Muller *et al.* 2006] and historical parchments [Kennedy *et al.* 2004] among others. In particular, X-ray diffraction is capable of describing the condition of the collagen structure within parchments. This technique has been used on valuable documents such as the Dead Sea Scrolls [Weiner *et al.* 1980]. X-ray scattering techniques are employed to investigate archaeological and experimentally altered bone [Hiller & Wess 2006]. Many other methods have also originated from medicine and are used for non-invasive examination of objects of art. In addition to standard radiography, 3D imaging coupled with computed tomography is used in examination of various art objects successfully. However, the offered resolution is not sufficient for the examination of paintings. Among other methods, there are X-ray fluorescence [Woll *et al.* 2005], neutron-induced autoradiography [Taylor *et al.* 1975], high-energy proton-induced X-ray emission [Griesser *et al.* 2000]. Unfortunately, these methods lack in-depth resolution: although they permit the identification of certain components of the object, it is not possible to determine to which thin layer of the painting they belong. Optical Coherence Tomography utilises

infrared light for non-invasive structure examination of weakly absorbing objects, and has been under consideration for the examining of objects of art [Kwiatkowska *et al.* 2008]. Raman spectroscopy, improved with recent advances in microscopy as well as with implementation of non-conventional techniques, has become a very useful molecular diagnostic technique in the cultural heritage field [Howell & Tasnim 2005]. In fact, Raman microscopy is nowadays present in the Conservation and Restoration departments of the most renowned Museums and Libraries all around the world, where it is used as a non-destructive technique for (even in situ) diagnostics of different materials, ranging from inorganic pigments to biomaterials, employed in artifacts as manuscripts, paintings, textiles, ceramics, glasses, sculptures, stone monuments, and even mummies or underwater cannonballs, where degradation estimates can also be assessed [Perardi *et al.* 2003]. Recently, a relevant development has been the optical 3D imaging techniques for cultural heritage applications. Generally, 3D optical techniques can be classified into passive and active, depending on the use of the light source [Remondino & El-Hakim 2006a]. 3D active techniques use artificial light sources to illuminate the scene to be acquired by a sensor (usually a CCD sensor) and are based on two main different principles: triangulation and time delay. In the first case the geometry of the object is reconstructed by means of a triangle configuration measurement devices, knowing the direction of emission patterns and the location of source and sensor. As sources of emission patterns, both laser sources and source of white light (digital projector) can be used. A number of different shapes of light patterns (structured light) can be used, e.g., spot, stripes, coded patterns. Triangulation-based laser scanners provide high accuracy in the 3D measurements, essential for 3D documentation of very detailed object as sculptures, bas-reliefs and other complex small objects. For example, a structured-light scanner and a laser scanner were experimented on the Minerva of Arezzo to realize a digital model of the statue [Rocchini *et al.* 2001b]. Shadow moiré method and fringe projection profilometry, based on the pattern projection technique, have been used to study the 3D profile of the wood support of the “Mona Lisa” [F. *et al.* 2007]. The second category of active techniques is based on time delay. In particular, 3D reconstruction is based on the measurement of the time of flight of a light pulse sent towards the object. Assuming a constant speed light, the distance from the device to the object can be calculated. Time delay active sensors are instead employed for large objects, i.e. architectures, rooms, walls and archaeological sites as “Chiostro degli Olivetani” and “Grotta della Poesia” in Lecce (Italy) [Fontana *et al.* 2004], respectively. These sensors work on longer distances (i.e. 2 - 1000 m), but are less accurate than triangulation systems. Passive techniques use the environment light for the acquisition of 3D models, working on 2D images collected from different

view angles [Remondino & El-Hakim 2006a] as well. This technique operates on the same principle as the human vision system. When a person looks at a point, the distance to the point is determined by comparing its apparent shift in position between the two eyes. Then, these techniques use 2D image measurements to recover 3D object information through a mathematical model. Generally, at least two images are required, even if it is possible to generate 3D data using a single image. The main advantages of such an approach is the simple instrumentation needed, usually two digital cameras, and the speed of acquisition. These peculiarities are important when one deals with complex sites or large architectures, because many parts of the model could be occluded by some obstructing objects, as well as narrow paths [Rizzi *et al.* 2006]. 3D reconstruction of the great Buddha of Bamiyan (Afghanistan) has been obtained by photogrammetric techniques [Grun *et al.* 2004]. Also photogrammetry has been employed to survey the underwater archaeological site of Pianosa (Tuscany, Italy) [Drap *et al.* 2007].

Finally, it is important to point out that there are many combinations and integrations between 2D and 3D techniques for purposes of texture mapping and conservation status monitoring. A 3D multispectral scanner has been proposed by Mansouri *et al.* [Mansouri *et al.* 2007] for the acquisition of small objects such as pottery or small statues. It couples an LCD projector with a camera equipped with a filter-wheel. This system was used for visualization purposes, in order to acquire faithfully textured 3D models of a jug and some small statues, and to relight them with various illuminants in a virtual museum application. Cabrelles *et al.* [Navarro *et al.* 2009] proposed a methodology to accurately and exhaustively record a tomb, Djin Block No. 9, located in the archaeological monumental park of Petra (Jordan), by means of terrestrial laser scanning, close range photogrammetry and thermal imagery.

### I.3 Thesis objectives

The main objective of this thesis is the study and the application of optical imaging techniques for Cultural Heritage purposes, aiming to satisfy the following needs:

- Testing whole-field techniques for 3D underwater acquisition;
- Suggesting components of an experimental setup for acquisitions in underwater environment,
- Understanding the performances of a 3D underwater system;
- Finding a sequence of tools to restore ancient documents;

- Using automatic methods in the document restoration process to reduce the operator interventions;
- Identifying limitations and problems of a multispectral imaging system based on a filter-wheel, suggesting a solution for these issues;

## I.4 Structure of the Thesis

The Introduction chapter presents the background and a review about the digital imaging techniques.

3D imaging by structured light techniques used in underwater applications is discussed in Chapter II. Two whole-field techniques i.e. gray-code and fringe projection are tested in water tank to acquire medium-size objects. The experimental tests has been conducted in different turbidity conditions.

Chapter III deals with multispectral acquisition and image processing techniques applied to virtual restoration of ancient documents affected by degradations. Each step of the restoration process, i.e. registration, decorrelation and binarization, is discussed. Moreover, some typical problems of a multispectral imaging system based on a filter-wheel are presented.

Chapter IV introduces a combination of multispectral and 3D imaging to study and correct the geometrical distortions that may be present in a historical document. A framework which exploits 3D and 2D data is proposed.

The final Conclusions chapter summarizes considerations about the results of 2D and 3D imaging techniques obtained through different applications.



# 3D imaging by structured light techniques for underwater applications

---

*Two 3D structured-light techniques adapted for underwater applications are described, both based on pattern projection: the Gray-Code technique and the Fringe Projection Technique. The former uses an optical configuration consisting of a projector and a stereo pair of digital cameras, while in the latter a projector and a digital camera are employed. The results of the 3D reconstructions at different turbidity levels are shown.*

## II.1 Introduction

There are several 3D reconstruction techniques used in underwater applications such as inspection, mapping, monitoring, scientific investigation, etc. [Kocak *et al.* 2008, Kocak & Caimi 2005]. 3D range sensor systems are usually installed on ships, submarines and underwater robots like Autonomous Underwater Vehicles (AUVs) or Remotely Operated Vehicles (ROVs) [Yuh 2000] to provide many quantitative measurements: object size, area, volume, and shape. These data are used both for manipulative tasks and for on-line/off-line analyses and investigations.

Among the various available solutions, the optical techniques provide higher accuracy and resolution compared to those based on radio and ultrasonic sources that cannot be adequately focused [Amann *et al.* 2001]. The applications of the optical techniques are spreading in many fields such as 3D morphometric measurement in marine biology [Bythell *et al.* 2001, Cocito *et al.* 2003], study of underwater fauna [Bräger & Chong 1999], seafloor mapping [Gracias & Santos-Victor 2000], archeological site and shipwreck surveys [Canciani *et al.* 2003] and inspection of sea bottom structures [Foresti 2001].

The state of the art, analyzed in the next section, presents two kinds of optical techniques for 3D reconstruction:

- *Active techniques* are based on the projection of structured light patterns that scan the surfaces with a spot or a stripe of light and are employed both in long and close range acquisitions. These techniques determine the distance between the device and the illuminated object according either to the time of flight of the projected light, or to the triangulation principle based on the known relative positions of the light source and an optical sensor. A classic example of the latter is found in laser scanning, which makes use of a coherent light stripe to illuminate the surface of the object acquired with an imaging system.
- *Passive techniques* are based on multiple views of the scene, working on the whole Field-Of-View (FOV) of the optical sensor. They are typically used in close and medium range acquisitions. Photogrammetry is widely used for underwater 3D acquisition, because it requires just one or two digital cameras to acquire a number of images taken from different viewpoints with natural or artificial illumination.

Underwater conditions are characterized by non-uniform lighting and poor visibility due to scattering and absorption in the medium [Schechner & Karpel 2004]. Both active and passive techniques suffer from scattering and absorption in the medium, which decrease image contrast and attenuate light intensity. To discriminate the principal component of degradation, i.e. the backscattering, different solutions have been proposed based on polarization, spatial, time and spectrum discriminations [He & Seet 2004]. Passive systems use sunlight or other light sources (i.e. spotlights, lamps) just to ensure an efficient lighting of the scene [Allais *et al.* 2007], while structured light sources (i.e. lasers), used by active systems, allow imaging at greater ranges and higher contrast reducing the backscattering [Gupta *et al.* 2008, Tetlow & Spours 1999, Jaffe & Dunn 1988]. On the other hand, passive systems (i.e. stereo systems) offer several advantages: lower cost of the devices, acquisition of the surface color information (i.e.: texture), and the ability to acquire data from the whole FOV of the image sensor.

Whole-field active techniques based on light patterns emitted by a projector are often used with very good results in air, but the difficulties related to the use of a projector in submarine environments have probably discouraged, until now, the experimentation of active whole-field techniques for underwater 3D acquisition. In fact, they have never been tested in water because it is apparent that these techniques are seriously affected by water turbidity. Moreover, the scattering effects are increased by the wider water volume illuminated by the light source.

Although underwater imaging is strongly influenced by backscattering, we think that the whole-field active technique may produce interesting results



in close range 3D acquisition. In fact, there is an increasing need to have more precision (in the order of millimeters) in some application fields like, for example, monitoring the growth of coral reefs, or the reconstruction of underwater archaeological pieces, that in most cases are not recovered from the seabed. Objects that have dimensions of less than one meter may be effectively acquired using an active whole-field technique because the equipment can stay close to the objects, thus reducing the backscattering effect generated by the particles suspended in water.

The experimentation presented in this chapter aims to evaluate the feasibility and the efficiency of whole-field structured light 3D imaging in turbid water, considering a particularly efficient technique, based on the combination of stereo-photogrammetry and gray-code patterns projection used to resolve the correspondence in stereo vision. This technique has been evaluated by conducting a series of experiments addressed to estimate the effect of turbidity on performance. Performance was expressed in terms of number of 3D points calculated and distribution of the measurement error on these points.

Therefore, an experimental setup composed of a video projector and two cameras in stereo vision configuration has been realized. The optical devices are mounted in waterproof housings, and fixed on an aluminum frame. A number of patterns (black and white bands) are projected on the object and synchronously acquired by the two cameras. The active stereo system solves the correspondence of the photogrammetry, establishing the matching between the two images. Image elaboration computes the 3D model by applying the raw texture. The experimentation has been done in water tank, varying both the turbidity conditions and the objects acquired, in order to test the reaction of materials with different diffusive reflection properties.

## II.2 Related works

In this section a quick review of the 3D optical techniques usually employed in the submarine field is presented, focusing attention on the operating principles and the limitations of underwater imaging. For a more complete review of underwater imaging techniques see also [Kocak & Caimi 2005, Kocak *et al.* 2008, Klepsvik *et al.* 1990, Jaffe *et al.* 2001, Shortis *et al.* 2007]. At the end of this section the whole-field structured light methods for 3D reconstruction that have been applied in the experiments are described. A comprehensive review of 3D acquisition techniques in air can be found in [Chen *et al.* 2000, Blais 2004, Remondino & El-hakim 2006b].

	ACTIVE		PASSIVE
	Time delay	Triangulation	
Global	<i>holography</i>		<i>multi-views</i>
Local	<i>time-of-flight</i>	<i>laser scanning</i>	

**Table II.1:** Classification of 3D acquisition techniques used in underwater environment.

## II.2.1 Underwater 3D optical imaging

In general, as for terrestrial systems, underwater 3D optical techniques can be classified into passive and active, depending on the use of the light source. Moreover, one can distinguish between global (whole-field) techniques that simultaneously measure the depth of all the points in the FOV of the sensor, and local techniques that measure only one point or a single line at a time [Desmangles 2003]. Table II.1 presents a classification of the 3D imaging techniques based on these two criteria that have been employed in underwater environment.

### II.2.1.1 Passive techniques

These techniques use ambient light or artificial light (lamps and spotlights, for example) to adequately illuminate the scene and make the objects visible. Essentially, passive techniques are based on multiple acquisitions, taken by different viewpoints or under different lighting conditions, with both still and video cameras, to produce multiple overlapping images of the scene.

The easiest way to obtain the depth of a scene is the stereopsis (or binocular vision), that is commonly known as *visual perception process* which gives the sensation of depth from two slightly different views of the scene. The observation of this principle has led to the development of photogrammetric techniques [Mikhail *et al.* 2001] based on the triangulation of corresponding points on the two views, by applying the epipolar constraints and computing the fundamental matrix to estimate the relative camera positions [Ma *et al.* 2003]. This requires a solution to the so-called *correspondence problem* [Scharstein & Szeliski 2001], i.e. matching common points extracted, for example, by a SIFT operator [Lowe 2004]. Stereo systems may be installed on underwater robots [Brandou *et al.* 2007] or used by scuba divers, and are widely employed in many applications like seabed mapping with 3D mosaicking [Leone *et al.* 2006], *in situ* estimation of the length of marine organisms [Harvey & Shortis 1995], and reconstruction of complex 3D structures in aquatic environment [Jenkin *et al.* 2007]. Monocular systems use a single

camera to take multiple overlapping shots of the scene, in order to survey archaeological sites [Drap *et al.* 2007], to measure three-dimensional biological objects [Cocito *et al.* 2003], and for oceanographic surveys of coral reefs [Pizarro *et al.* 2004].

The structure from motion technique uses an image sequence acquired by a moving camera [Tomasi & Kanade 1992]. It involves the extraction and tracking of a sparse set of features in an image sequence, and the estimation of their 3D positions using multiple views [Saito *et al.* 1995]. A bundle adjustment technique can be adopted to improve the robustness of the reconstruction [Triggs *et al.* 2000]. Photometric stereo methods involve the use of shading patterns in multiple images (three at least) acquired from a fixed viewpoint under different lighting conditions [Negahdaripour *et al.* 2002].

The multiple-camera system, with overlapping FOVs, allows for a much wider FOV for panoramic views, photo-mosaics, 3D motion estimation and positioning [Firoozfam & Negahdaripour 2003, Negahdaripour *et al.* 2001]. The configuration with three synchronized still cameras is a good compromise between accuracy and encumbrance, so it can be a good solution for underwater surveys conducted by a scuba diver [Menci-Software 2010]. With respect to a standard stereo system, the use of the third camera can improve the quality of the acquisition, reducing the errors due to the mismatching and the occlusions, because it makes the epipolar constraint more robust by means of the trifocal tensor [Hartley & Zisserman 2000]. Quadranocular stereo (two pairs) videos are used for 3D reconstruction of reefs and benthic structures [Zhang & Negahdaripour 2003].

### II.2.1.2 Active techniques

In general, the 3D underwater active systems make use of coherent light (laser) for a better light propagation in scattering medium, and for the capability to make beams highly collimated. These characteristics make the laser-based technique suitable for long range acquisition [Jaffe 1990]. 3D active optical techniques are based either on the triangulation principle or on time delay (time-of-flight and interference) of the laser light [Remondino & El-hakim 2006b].

In the first case, a light source projects structured light (sheet, narrow beam) on the object and an image sensor acquires the light reflected from another location; the 3D points are acquired by knowing the relative position of two devices (i.e.: source and receiver). The time-of-flight or LIDAR (LIght Detection And Ranging) techniques operate on the same principles used by conventional RADAR (RAdio Detection And Ranging) and SONAR (Sound Navigation And Ranging) systems by transmitting short pulses of light and

detecting the light reflected from objects within the receiver FOV. Range information is obtained by measuring the time delay between the transmitted and received light pulses. Finally, holography provides the three-dimensional coordinates by the interference between two coherent waves propagating from the object and a reference wave: the geometrical shape of an object is related to the whole-field measure of the optical phase.

Underwater techniques based on triangulation and time-of-flight adopt a local approach to scan the entire target of interest, using a sheet light or a narrow beam. The single sheet projection is useful in several applications like underwater navigation [Dalglish *et al.* 2004], pipe inspection [Rives & Borrelly 1997] and mapping of underwater archaeological sites [Roman *et al.* 2010]. Using the forward movement of the AUV, a stripe of light can provide the map of a seabed [Jaffe & Dunn 1988]. Multiple stripes are used to scan large FOVs [Jalkio *et al.* 1985], but backscattering would be superimposed over some of the stripes. A two-axis scanning system allows to perform a point-to-point scan on the object from a stationary platform [Tetlow & Allwood 1994, Tetlow & Allwood 1995]. This technique is used in marine biology, plankton study [Hobson & Watson 1999] and underwater holography (in-line or off-line) because it allows small-scale 3D reconstruction with high accuracy [Hobson & Watson 2002].

### II.2.1.3 Limitations of underwater imaging

Underwater imaging is seriously compromised by scattering and absorption in the medium, which decrease image contrast and attenuate light intensity [Harsdorf *et al.* 1999] with consequent loss of details and colour alteration. The light intensity decays exponentially in function of the distance from the source, and depends on the attenuation coefficient the sum of absorption and scattering coefficients [Jaffe 1990]. In particular, the total attenuation coefficient  $c$  is the decay constant associated with the removal of light intensity per unit distance:

$$E(r) = E(p)e^{-cr} \quad (\text{II.1})$$

Here  $r$  is the distance and  $E$  is the irradiance at positions  $p$  and  $r$ . The constant  $c$  can be further decomposed as a sum of two quantities, the absorption and scattering coefficients, so that

$$c = a + b \quad (\text{II.2})$$

and

$$E(r) = E(p)e^{-ar} e^{-br} \quad (\text{II.3})$$

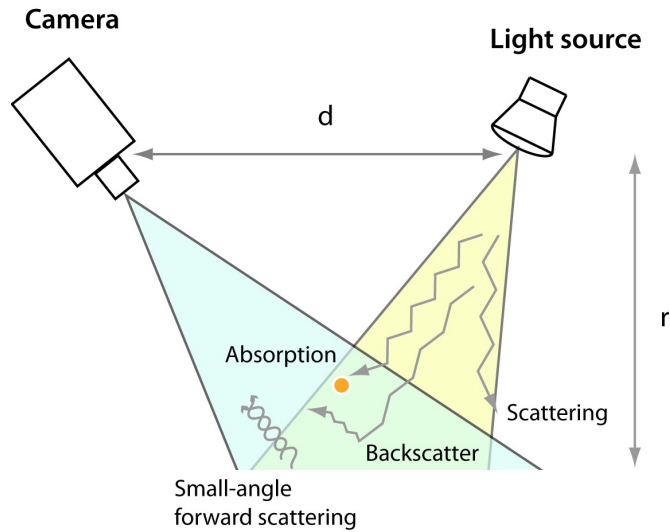
The light absorption, i.e. a conversion in other forms of energy, is due to the interaction with the matter. Scattering, which is the divergence from a straight-line path, is composed by two factors: forward scattering (small angular deviation) and background scattering (deviation of about 180 degrees); the latter is the dominant contribution to image degradation (Figure II.1).

In order to enhance underwater imaging, one can use four hardware solutions: polarization, spatial, time and spectrum discriminations [He & Seet 2004]. Polarization discrimination [Lewis *et al.* 1999] uses the polarization difference between the backscattered light and the light reflected by the object. Schechner and Karpel [Schechner & Karpel 2004] have exploited this property to develop an algorithm that improves the image contrast using a couple of images taken through a polarizer at different orientations. The spatial discrimination reduces the degrading effect by increasing the separation between the light source and the sensor in order to minimize the backscattering volume due to the overlap of the device FOVs [Narasimhan *et al.* 2008]. Laser Line Scan (volume scanning or synchronous scanning) methods, based on triangulation, involve the optical scanning of a narrow instantaneous FOV of the receiver in a synchronous mode, with a highly collimated laser source over a wide swath [Dalglish *et al.* 2006]. The time discrimination rejects the backscattered signal with a gated receiver that confines the instantaneous scattering of a small volume of water in close proximity to the target [Tan *et al.* 2005]. The distance is calculated by measuring the time took by a coherent light beam to be emitted, reflected by the object and acquired by a time-gated sensor, in order to discriminate the noise component. This approach is employed by time-of-flight systems. Finally, the effects of light absorption can be reduced by selecting the wavelength of the light source in the green/blue range (550-470 nm) spectrum, where attenuation is the lowest.

### II.2.2 Whole-field structured light technique: pattern projection

The whole-field structured light techniques are widely used in air, both in industrial and cultural heritage applications [Heredia-Ortiz & Patterson 2003, Rocchini *et al.* 2001a, Godin *et al.* 2002], and allow a high-accuracy 3D shape reconstruction of medium/small scale objects. These techniques are based on the projection of white light patterns. Different patterns as gray-code bands, sinusoidal fringes, grids, etc. are projected on the surface and the scene is taken by a camera placed in a known position to calculate the triangulation [Salvi *et al.* 2004].

Some of these techniques, based on the moiré phenomenon, compare the interference between two line patterns - the reference (undeformed) and the



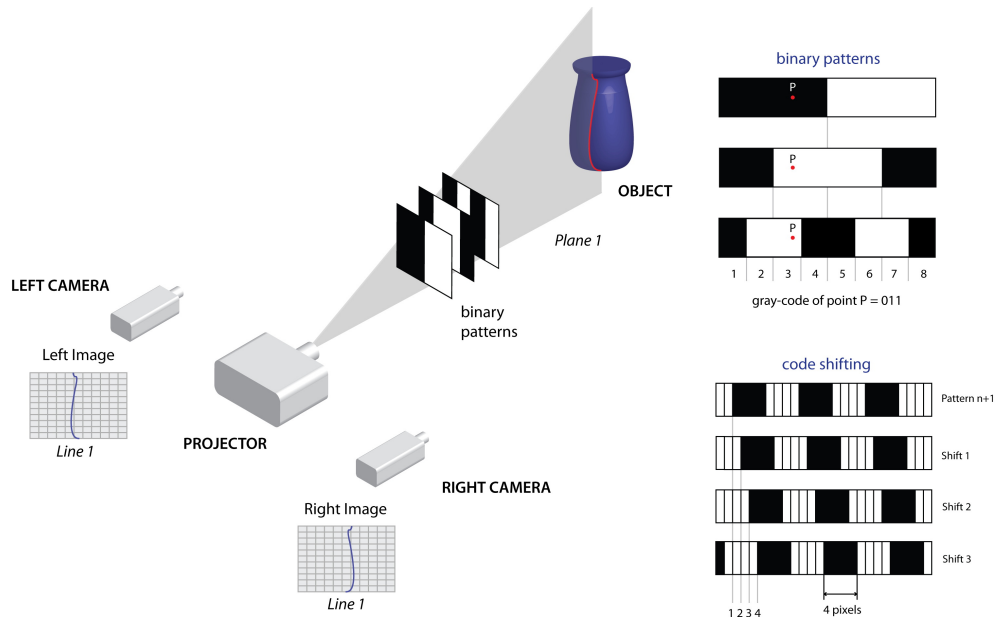
**Figure II.1:** The interaction between the light and the underwater environment includes basically two processes: absorption and scattering. The latter can be further divided into backscatter and forward scattering.

deformed - to obtain the height map of the object [Creath & Wyant 1992]. The simpler systems employ a single camera and a digital video projector [Sansoni *et al.* 1997], but systems with a single projector and two cameras [Barone *et al.* 2003] are also widely used. In this latter case, the structured light is used only to solve the correspondence problem of the stereo pair. The success of these systems is due to the good balance between accuracy and costs, as it is possible to build them using low-cost commercial devices.

### II.3 Gray-code and stereo-vision techniques

The 3D reconstruction technique employed in this application is based on the combination of stereo-photogrammetry and patterns projection. The use of structured light, and in particular the gray-code [Posdamer & Altschuler 1982, Krattenthaler *et al.* 1993], makes the matching between correspondent points in the stereo pairs more effective [Barone & Razionale 2004]. In contrast to traditional passive approaches this technique does not rely on images with consistent textures, because each point on the object surface is precisely identified thanks to a double binary code, as better explained in next section.

As previously mentioned in the introduction, the experimental setup uses two cameras for stereo vision and a video projector as a source of structured light to generate a set of vertically and horizontally striped binary light pat-



**Figure II.2:** At left, projection of binary patterns and vertical line coding on the image planes. At right, examples of binary patterns and code shifting.

terns of varying widths. (Figure II.2). This particular setup calculates the 3D points using the positions of the two cameras determined during the calibration phase. The projector is only used to establish the correspondences and it is not involved in the triangulation, so it is not required to calibrate its optics. For the same reason, conventional projectors, commercially available at low cost, may be used without compromising the accuracy of measurement.

The gray-code is a technique that helps to solve the correspondence among the points in the stereo pairs. The object is illuminated by a set of  $n$  temporally encoded patterns of black/white bands with width progressively halved, and  $n$  images are captured, respectively. A binary code (0, 1 with  $n$  bit) is assigned to each point in the camera frame, and the values 0 and 1 are associated to intensity levels, i.e. 0=black and 1=white. This procedure allows to codify  $2^n-1$  lines defined as crossing zones between white and black bands.

Figure II.2 illustrates an example of three binary patterns with an associated 3-bit code that allows to distinguish 8 stripes. It is also shown how the point P assumes the '011' code since it is obscured in the first pattern and illuminated in the others. By projecting both horizontally and vertically striped coded patterns, a double code is assigned to the intersection points through the horizontal and vertical lines. This binary code identifies uniquely and automatically each point of the object surface, taking very easy to establish the point correspondences between the left and the right camera even when the surface does not have any texture.



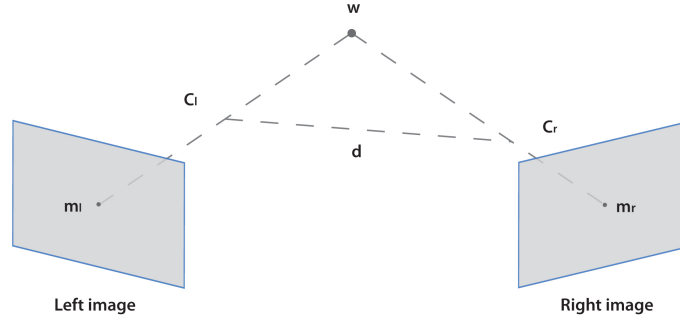


Figure II.3: 3D reconstruction scheme.

Coded light approach allows for high resolutions, which are obtained by increasing the number  $n$  of projected binary patterns. The maximum resolution depends on the projector resolution and, usually, finer bands (1 pixel width) are included in the coding process. However, the use of finer bands can lead to a reduction of accuracy in the line extraction process (i.e. images with finer bands are harder to process). To overcome this limitation, the used methodology does not rely on the finest bands, but on larger ones (4 pixels width), and shifts them in steps of 1 pixel for a total of  $N=4$  pattern positions. This procedure yields the maximum resolution without loss of accuracy.

A set of 8 vertical and 8 horizontal patterns (8-bit code) are used for gray-code. Other  $4 + 4$  patterns are projected for vertical and horizontal code shifting, respectively, with a bandwidth of four pixels (Figure II.2). The current projector resolution is  $800 \times 600$  points that allow us to codify  $799$  lines  $\times$   $599$  lines =  $478.601$  points.

Typically, crossing zones in acquired images do not present an abrupt variation of gray levels, but a relatively smooth intensity gradient, depending on environment light conditions. The process to determine crossing lines on the camera image can be improved by projecting more images (the negative images), created by inverting the light intensities of the initial patterns (a total of 48 patterns). The images created by subtracting positive and negative images provide a better scene coding and sub-pixel points extraction.

Given a pair of conjugate points  $m_l$  and  $m_r$ , the corresponding 3D  $w$  point is calculated by a triangulation procedure based on the extrinsic and intrinsic parameters obtained from system calibration (Figure II.3), explained in next section. The point correspondence is determined without using geometrical constraints like intersections between projected planes and optical rays or epipolar geometry [Barone *et al.* 2003].



### II.3.1 Calibration

A calibration procedure is adopted to calculate the so-called *intrinsic* and *extrinsic* parameters of the cameras. In particular, a camera is described by a pinhole model, with intrinsic parameters (including focal length, principle point, pixel skew factor, pixel size and the optical distortions), and extrinsic parameters (including rotation and translation from a world coordinate system to a camera coordinate system). Figure II.4 shows a typical diagram of a pinhole camera model, where  $P$  is an arbitrary point with coordinates  $X_w, Y_w, Z_w$  and  $X_c, Y_c, Z_c$  in the world coordinate system ( $O_w; X_w, Y_w, Z_w$ ) and camera coordinate system ( $O_c; X_c, Y_c, Z_c$ ), respectively. The relationship between these two coordinate systems is given by:

$$\mathbf{X}_c = R\mathbf{X}_w + \mathbf{T} \quad (\text{II.4})$$

where  $\mathbf{X}_c$  is the vector of camera coordinate system,  $R$  the rotation matrix,  $\mathbf{X}_w$  the vector of the world coordinate system and  $\mathbf{T}$  the translation vector.  $R$  and  $\mathbf{T}$  represent the extrinsic parameters to be calculated. The projection of point  $P$ , in the image coordinate system ( $O; x_u, y_v$ ) is  $(x_u, y_v)$ . Its projection on the image sensor is related to the camera coordinate system as follows, based on a projective model:

$$x_u = f \frac{X_c}{Z_c} \quad y_v = f \frac{Y_c}{Z_c} \quad (\text{II.5})$$

$f$  is the focal length. In real cases, the lens of the camera involves optical distortions which must be taken into account in the calibration model. So the ideal image coordinates  $(x_u, y_u)$  are distorted in the real image coordinates  $(x_d, y_d)$ :

$$x_d = x_u + k_x(x_u, y_u) \quad y_d = y_u + k_y(x_u, y_u) \quad (\text{II.6})$$

where  $k_x$  and  $k_y$  are the distortion coefficients. There are two different distortion components: radial and tangential, the latter is orthogonal to the former. The radial distortion changes with focal length, while the tangential distortion depends on the axial alignment of the optics in the lens. Finally, the real coordinates are transformed in pixel coordinates  $(u, v)$  by:

$$u = s \frac{x_d}{d_x} + u_o \quad v = \frac{y_d}{d_y} + v_o \quad (\text{II.7})$$

with  $s$  is a scale factor,  $(u_o, v_o)$  are the pixel coordinates of the principal point, i.e. the intersection of optical axis with the sensor image plane,  $(d_x, d_y)$  are the cell sizes of the sensor array. The parameters of each camera are obtained by correlating the coordinates of known markers located on a calibrating sample with the corresponding coordinates on the image plane.

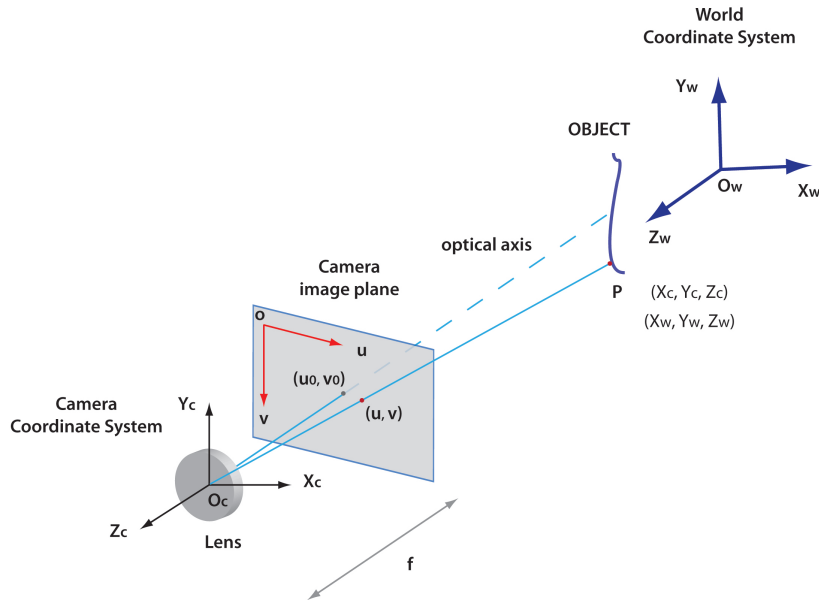
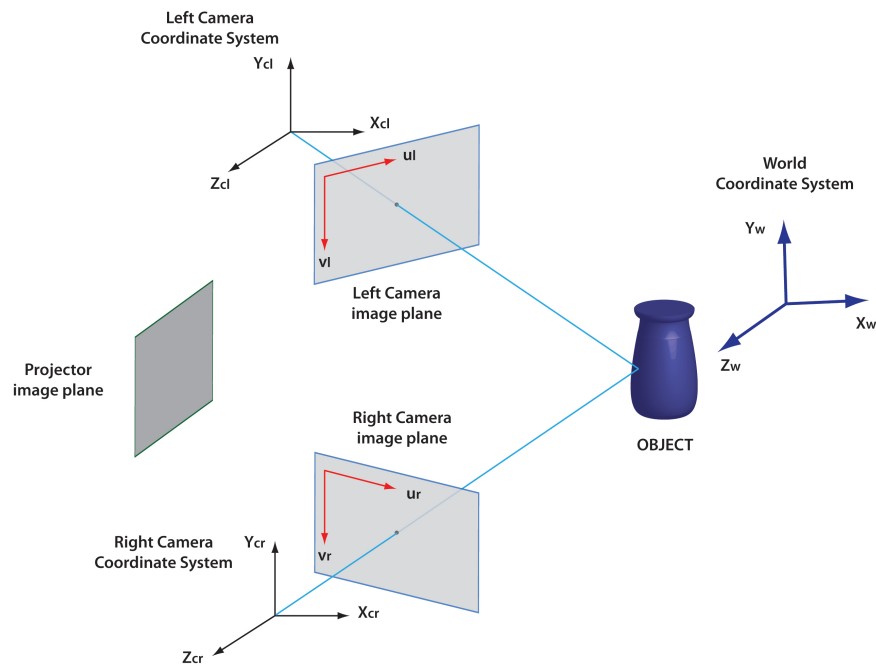


Figure II.4: Pinhole camera model.

The procedure consists of acquiring several images with the cameras of the calibration sample (checkerboard, Figure II.6) placed in different positions. For each position, an iterative process solves the correlation equation and provides, separately, the intrinsic and extrinsic parameters of each camera. Calibration consists of an overall optimization problem solving, that minimizes re-projection error between camera planes. The optimized parameters are both intrinsic and extrinsic of each camera and extrinsic of the stereo pair. The stereo calibration is based on correspondence between the points of the checkerboard in the two camera images. It serves to determinate the relative position between the two cameras. So, by the extrinsic parameters of each camera we can obtain the extrinsic parameters of stereo system.

### II.3.2 Accuracy

Experimental analyses have been conducted to validate the system in different configurations out of the water [Barone & Razonale 2004]. A statistic approach has been used to plan and execute the experimental tests. A rectified planar surface and a cylindrical surface (radius=125 mm) have been scanned by the optical system. The maximum error and the standard deviation have been estimated by comparing the point clouds obtained by single scans with the corresponding surfaces of best fit. The analysis has provided the following indications about the accuracy: a maximum deviation of 0.1 mm and a standard deviation of 0.05 mm with a working distance of 800 mm; a



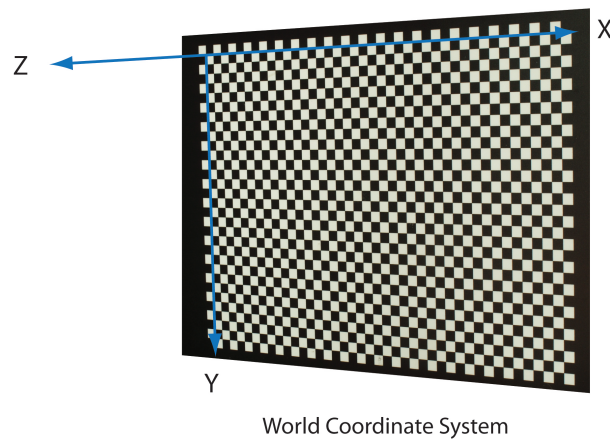
**Figure II.5:** Coordinate systems of the optical setup.

measured maximum deviation of 0.4 mm, and a standard deviation of 0.2 mm with a working distance of 1200 mm.

### II.3.3 Virtual prototype of an underwater 3D system

The preliminary phase to build the underwater 3D system is the design of a virtual prototype using CAD software, which creates a 3D model to study the optical configuration and functionalities of devices, and to calculate the encumbrance. Figure II.7 shows the 3D model obtained with Pro Engineer software. Specifications of 3D scanner are listed as follows:

- the scanner must be portable so it can be moved and positioned *in situ* by a scuba diver;
- regulation of the devices;
- small encumbrance taking into account the size of the optical devices;
- the lens of projector and cameras must ensure a FOV to acquire medium-size objects (about 60 cm x 60 cm) with a working distance of 100 cm;
- remote control for automatic operations of projection patterns and acquisition images;



**Figure II.6:** Samples for camera calibration (checkerboard).

- the scanner must be power independent for at least 1 hour

### II.3.4 Experimental setup

The setup used in the experimentation is composed of a commercial projector and two cameras (protected by acrylic watertight), as shown in Figure II.8. The digital cameras are two Nikon D200s with a CCD (Charge Coupled Device) sensor size of 23.6x15.8 mm and resolution of 3872x2592 pixels, equipped with an AF-Nikkor 20 mm lens. The projector is a Mitsubishi PK20, characterized by a very small size (123x48x97 mm) and an acceptable resolution of 800x600 pixels and a brightness of 25 lumens. The projector mounts a DLP (Digital Light Processing) matrix and a LED lamp, which ensures lower power consumption (max 37 W) if compared to the traditional lamp-coupled color wheel usually used in DLP projectors. The presence of an SD (Secure Digital) card slot allows the projection of JPEG and Bitmap photos stored on the card. This function is very useful, because the patterns needed for the active technique are loaded on the SD card and the projector can work without any connection to a PC.

The cameras and the projector are mounted in waterproof housings and fixed on an aluminum frame. The spherical port of the camera housing (dome port) significantly reduces the refractive effect of light at water-air interface [Edge 2006] preserving the angle of view of the lens. Moreover, optical distortions and chromatic aberrations decrease if the dome's center is placed close to the nodal point of the lens. The dome port creates a virtual image in front of the housing and the camera captures this virtual image. For a distance of about 1 m the distance of the virtual image is at 186 mm



Figure II.7: Virtual prototype of the underwater 3D scanner.

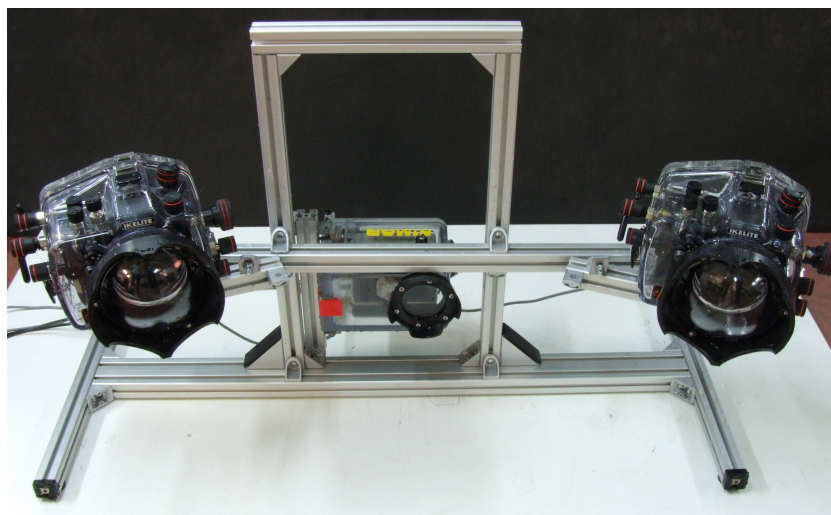
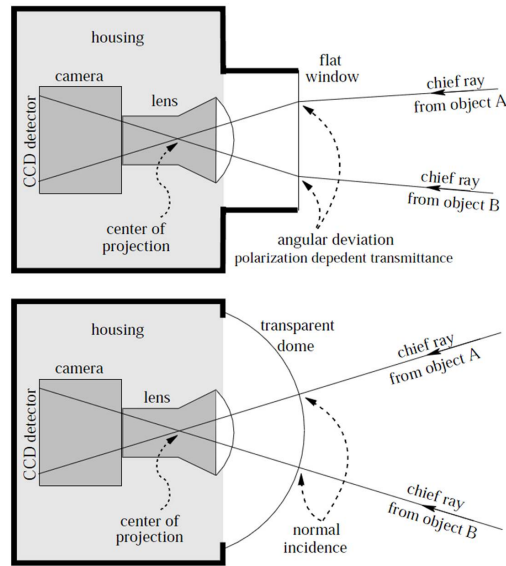


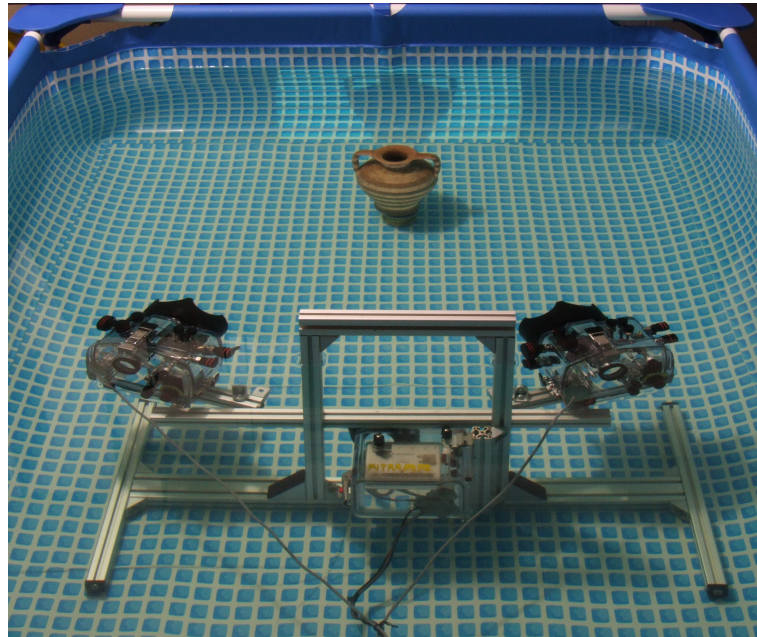
Figure II.8: Underwater 3D system



**Figure II.9:** Optical path of rays at water-air interface, for a flat port (top) and a dome port (bottom) [Karpel & Schechner 2004].

[Karpel & Schechner 2004]. The housing for the projector has been built by modifying a commercial product, in order to customize the position of the commands and allow for a remote control cable connection. As the projector housing mounts a flat port, the refraction at water-air interface causes a shrinking of the FOV. In fact, in air it projects a 76 cm diagonal image from about one meter, but in underwater conditions the FOV is reduced of 75%.

The design of the frame had to take into account the differences between the FOVs of projector and cameras. In fact, the camera FOV is symmetric with respect to the optical axis, while the projector generates an oblique off-axis projection. To solve this issue the cameras are placed on a single bar, while the projector is fixed independently in a lower position. Each 3D acquisition requires the projection of 50 patterns synchronized with the image capturing done by the cameras. This sequence of pattern projection and image capturing has to be as automatic as possible, because it is essential to avoid any movement or vibration, which would result in a less precise acquisition. Moreover, the automation of the process reduces the time needed for each 3D acquisition. For these reasons we have designed a remote control system, placed out of the water tank, for the cameras and the projector. In particular, we have designed a simple wired remote controller that allows us to advance in the sequence of patterns stored in the memory card. The cameras are connected to a PC with a USB cable and are controlled through the Nikon Camera Control Software. The remote control of the cameras allows us to change the settings (exposure time, aperture, focusing) in an interactive way



**Figure II.10:** Experimental setup for underwater 3D acquisition.

and to transfer the images on the PC, thus verifying acquisitions in real-time. The images are acquired in gray-scale because colors are not needed to calculate the binary code. Moreover, the various algorithms that we have implemented work only with gray-scale images that are easier to be processed and occupy less memory (8 bit per pixel) than color images (24 bit per pixel). The images are stored in the camera memory card using the RAW format that avoids any loss of information due to the JPEG compression. The images are directly transferred on the PC and converted in raster images (bitmap) for the subsequent image elaboration step without moving the cameras out of the housing.

### II.3.5 Experimentation

The experimental tests have been conducted in a water tank (Figure II.10) initially filled with clear water. The laboratory tests have allowed to control the turbidity of the water by dissolving different quantities of solid particles. The tests have been conducted with three different objects, ranging from 20 to 30 cm.

### II.3.6 System calibration

Calibration is an essential step to ensure high accuracy for the adopted 3D reconstruction technique. In particular we had to verify that the model adopted



		Air		Water	
		Left Camera	Right Camera	Left Camera	Right Camera
Focal length (mm)		20.822	20.920	22.185	22.281
Principal point (mm)	$p_x$	12.050	12.077	11.989	12.098
	$p_y$	8.193	7.920	8.257	7.978

**Table II.2:** Intrinsic parameters of cameras for calibration in clear water and air.

in air calibration is also able to work properly in underwater conditions. The calibration has been carried out using an 8x8 mm checkerboard (32x42 squares), mounted on an aluminium frame. The calibration has been done at a distance of about 1 m, acquiring on each camera, 17 images for the 17 different panel poses (Figure II.11,12). The refractive indices of air, dome port and water, are not considered in our linear calibration model. In this work have been supposed that the calibration procedure is able to take into account all the optical variables related to optical setup [Turner 1993, Schewe *et al.* 1996]. Following this approach, it has been used Bouguet’s Camera Calibration Toolbox [Bouguet 2010], to calibrate the stereo system, as was done in previous works [Allais *et al.* 2007, Liu 2001].

Then, the results of calibration carried out in air with those obtained in clear water are compared. The intrinsic parameters (focal length and principal point) for both left and right camera are presented in Table II.2, and the distortions are shown in Figure II.13. The apparent focal length in water is higher than one in air, because the double refraction effect at air-acrylic and acrylic-water interfaces involves a magnification of objects in underwater acquisitions. As explained in a previous section, the use of dome ports for the camera housing should reduce the refraction effect, but our measures on the focal length evidence that it is increased by about 7%.

The principal point  $(p_x, p_y)$ , however, is not affected by any significant change. In Figure II.11 the distributions of complete optical distortions (radial plus tangential) are shown, obtained in air and in clear water, where each arrow represents the effective displacement of a pixel induced by the lens distortion over the image. As it is apparent, the maximum values are located at the image borders and the distortions obtained in water are greater than those obtained in air. These distortions are corrected by the calibration algorithm; however the borders are usually not processed, because the images are cropped in order to select only the central area where the object is present. In order to evaluate the effectiveness of the underwater calibration, we considered the re-projection error for each corner point of the checkerboard, defined as the distance between the points of the actual image, and those computed from the



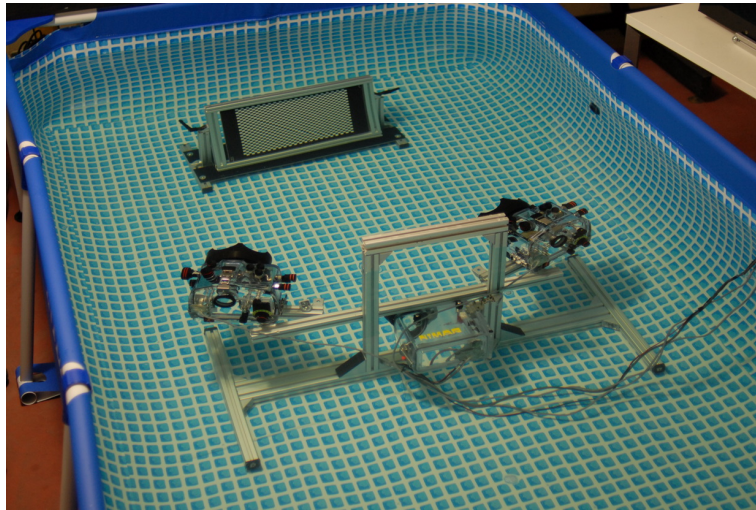


Figure II.11: Calibration step, in water tank, with checkerboard

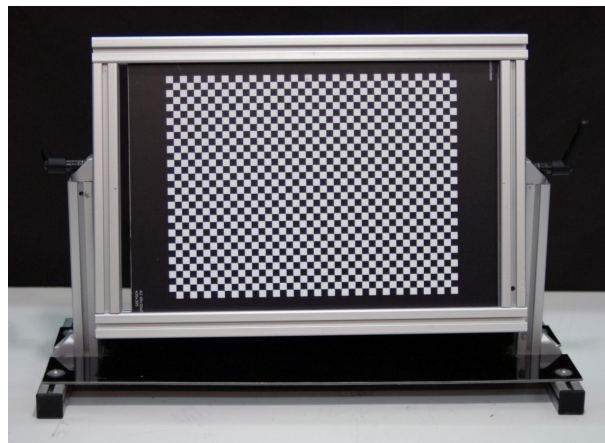
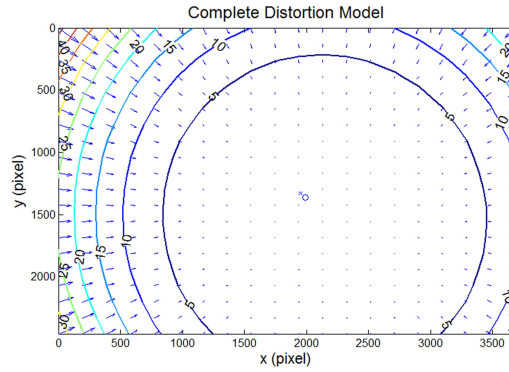


Figure II.12: Calibration panel.

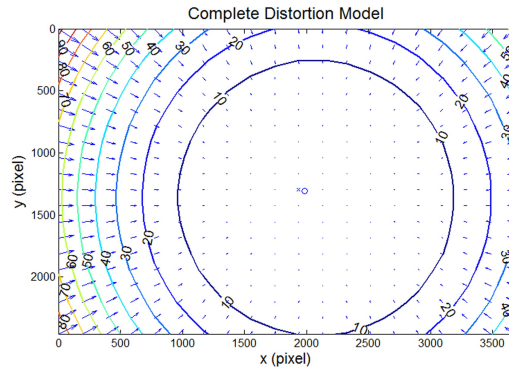
back projection of the ideal checkerboard [Bouguet 2010]. As shown in Figure II.14, the error dispersion in water is comparable with the one obtained in air.

### II.3.7 Image acquisition

The tests have been conducted in a water tank (3x2x0.7 m, 4500 liters of water), in different light conditions, each time varying the concentration of clay dissolved in water. Particles of clay, with a size ranging from a few micrometres to a millimetre, remain suspended in water for a time sufficient to conduct the test, thus creating scattering conditions similar to the real marine environment. Three differently reflective objects have been acquired: shiny, opaque and dark objects (Obj1, Obj2, Obj3, respectively), in air, in



(a) Air



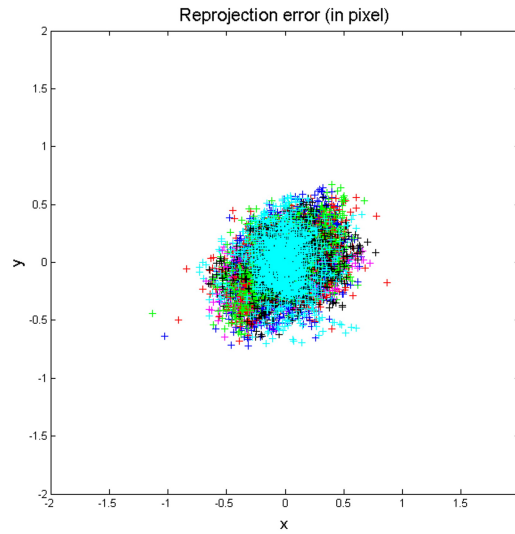
(b) Water

**Figure II.13:** The figures show the complete distortion model (radial plus tangential) on each pixel of the image for left and right camera in air and water. The cross indicates the centre of the image, the circle, the location of the principal point and the arrows are the optical distortions.

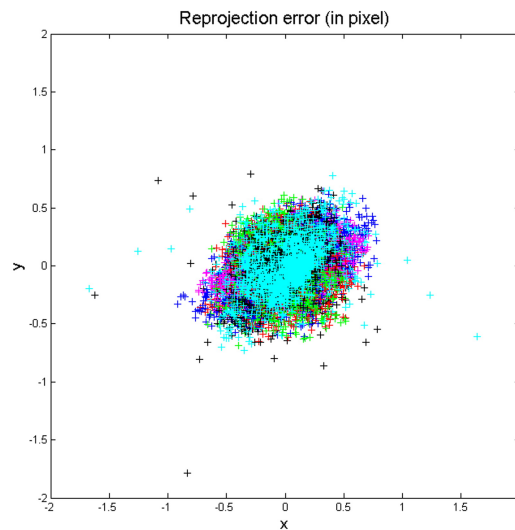
clear and in turbid water with four clay concentration levels: low (2.2 mg/l), medium (6.6 mg/l), high (15.6 mg/l), and very high (22.2 mg/l), dissolving 10, 30, 70 and 100g, respectively.

Table II.4 shows the images taken, while projecting a gray-code pattern, for each object and for three turbidity levels (low, high and very high). The images present a gradual loss of quality, visible in the attenuation of contrast between black and white bands. It is necessary point out that, at the moment, the images have not been processed with any particular algorithms to remove the noise and improve the contrast.

To define the turbidity conditions, the illuminance, the absorbance and the Total Solids Suspended (TSS) have been measured . The illuminance has been measured by projecting a white pattern on a luxmeter LSI-Lastem placed at a distance of 1m from the scanner. In air the projector is able to



(a) Air



(b) Water

**Figure II.14:** Distribution of the re-projection error in air and water.

provide an illuminance of about 90 lx at a distance of 1 m. The absorbance and the TSS are measured with a UV-VIS spectrophotometer produced by Scan Messtechnik GmbH. For very small particles TSS may be considered equal to NTU (Nephelometric Turbidity Units) which are another common measurement unit.

Table II.3 reports all the values of the measured quantities that are useful in order to replicate the experimental conditions. In fact the clay dissolved

Turbidity level	Clay dissolved (mg/l)	TSS (mg/l)	Absorbance	Illuminance (lx)
Clear water	0	63.5	0.712	80.4
Low turbidity	2.2	75.6	3.523	80.3
Medium turbidity	6.6	96.1	3.952	96.1
High turbidity	15.6	128.0	5.052	47.2
Very high turbidity	22.2	148.7	6.070	23.5

Table II.3: Measurements of the turbidity conditions.

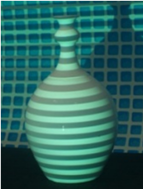
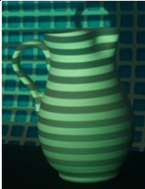
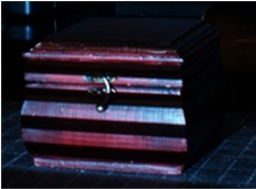





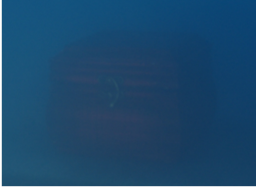
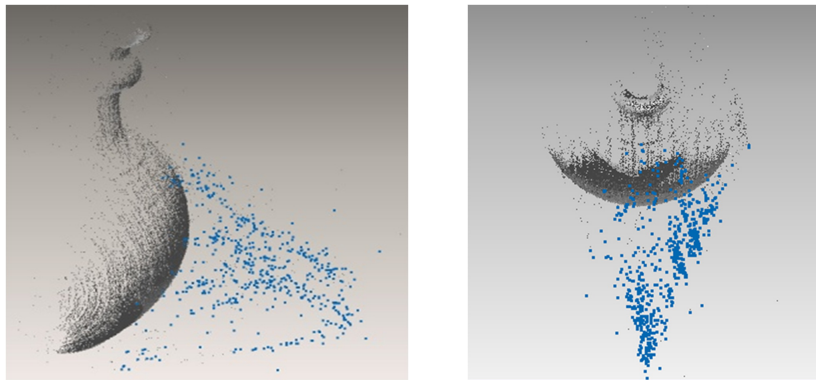
	Obj1	Obj2	Obj3
Clear water			
High turbidity			
Very high turbidity			

Table II.4: Images of the pattern projected on various materials at different levels of turbidity.

might precipitate or aggregate thus altering the turbidity conditions between two different experiments or also during the same experiment, if it requires too much time.

### II.3.8 3D reconstruction

Subsequently, the acquired images have been processed to obtain the point clouds of the objects. The raw point clouds acquired with the higher level of turbidity are strongly affected by the scattering that generates many *outlier points* within the volume of water illuminated by the projector (Figure II.15). All the point clouds are filtered and cleaned with Rapidform software that is also used to generate the polygonal surface of the 3D model (Table 5).



**Figure II.15:** Unclean point cloud: the blue points are due to scattering effects.


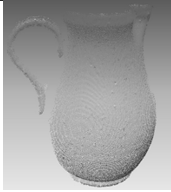
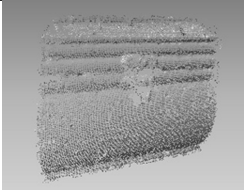

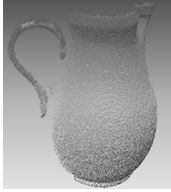
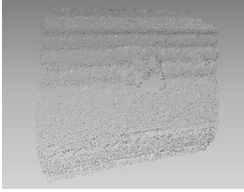
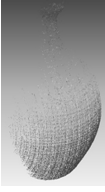
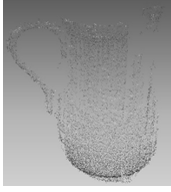
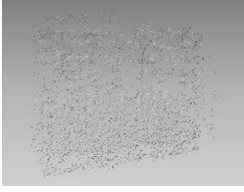

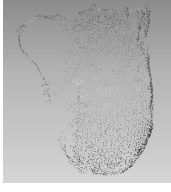
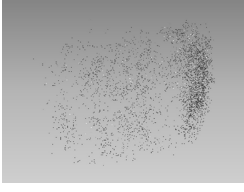


**Figure II.16:** Texture mapping: underwater image (a), retouched image (b) and textured 3D model of a earthenware jar (c).

In Figure II.16, a 3D model of a textured earthenware jar is represented. As the pictures taken underwater present irregularities in the spectrum as red channel loss, irregular illumination and missing tones [Yamashita *et al.* 2007], we retouched the pictures through channel mixing, saturation and levels correction, then we applied the texture on the 3D model.

### II.3.9 Results

The main results that are considered to evaluate the performance of the 3D technique are the number of acquired points and the geometrical error. The following consideration must be done for which concerns the interpretation of the results: a large quantity of acquired points does not necessarily correspond to a better acquisition, because it is necessary to check how the point cloud fits the shape of the real model: a high number may be the result of scattering. Therefore, for accuracy evaluation of the 3D reconstructions, the point

	Obj1	Obj2	Obj3
Air			
Clear water			
High turbidity			
Very high turbidity			

**Table II.5:** 3D point clouds obtained in different optical conditions.

clouds in water have been compared with those acquired, in air, by qualified equipment (Scanprobe by Scansystems) for each object. In order to compare the three objects, we have used as a reference parameter the number of points  $N_{100}$  acquired each 100 pixels. This value is calculated as follows:

$$N_{100} = \frac{n_{3D}}{n_{px}} * 100 \quad (\text{II.8})$$

where:  $n_{3D}$  is the number of 3D points, and  $n_{px}$  the number of pixels in the object image. As we can see in Table II.6, the increase of turbidity causes a drastic decrease of the number of points per 100 pixels. The highest value, recorded in clear water, is of 14.67 points for the opaque object, while the lowest is 1.23 points, recorded while sampling a dark object in heavily turbid water. We have experienced a strong reduction in the number of acquired points when the turbidity level changed from low to medium (from 10.91 to 1.44 points) while sampling the dark object. This is probably due to the fact that the acquisition of dark objects is, in general, strongly influenced by illumination and contrast conditions, so a little variation of these conditions



Turbidity level	Obj1	Obj2	Obj3
Clear water	13.42	14.67	10.54
Low turbidity	9.47	10.45	10.91
Medium turbidity	9.19	10.20	1.44
High turbidity	3.54	3.13	1.99
Very high turbidity	1.38	2.10	1.23

**Table II.6:** Number of points per 100 pixels acquired on the various object for each turbidity level.

Turbidity level	Obj1		Obj2		Obj3	
	$\mu_1$	$\sigma_1$	$\mu_2$	$\sigma_2$	$\mu_3$	$\sigma_3$
Clear water	0.0036	0.4294	-0.0033	0.2708	0.0044	0.3268
Low turbidity	-0.0208	0.34695	0.0131	0.2622	0.0650	0.5273
Medium turbidity	0.0245	0.3317	-0.0119	0.3503	0.0009	0.5475
High turbidity	-0.0137	0.3339	-0.0662	1.1385	0.0409	1.0014
Very high turbidity	-0.378	1.4091	-0.1491	1.4161	-	-

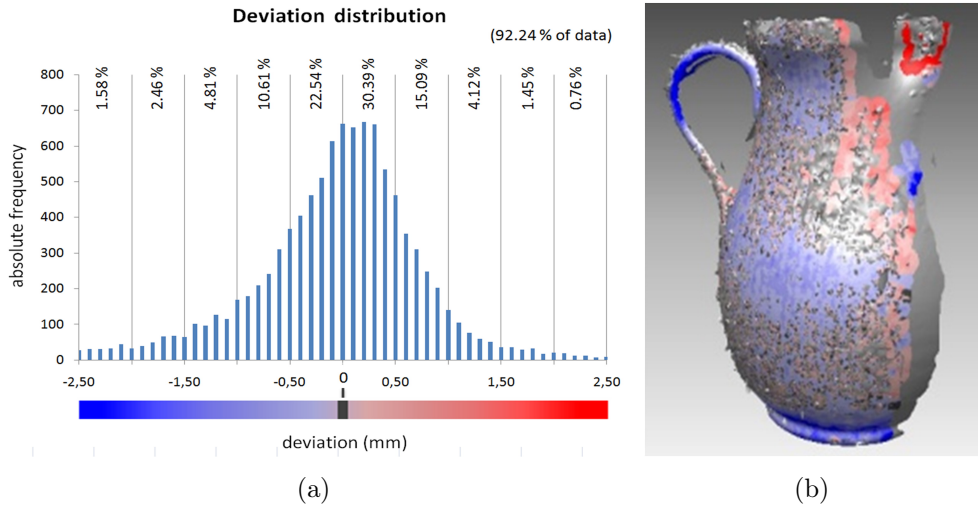
**Table II.7:** Statistical parameters (mean value  $\mu$  and standard deviation  $\sigma$ , expressed in mm) calculated for the various objects for each turbidity level.

can drastically worsen the results of the reconstruction. Further analysis concerning the geometrical error has been conducted with Rapidform software that provides qualitative visualization of discards, as shown in Figure II.17. The software computes the mean value  $\mu$  and the standard deviation  $\sigma$  (Table II.7): in presence of heavy turbidity, the values of  $\sigma$  are high for every acquired object. At parity of environmental conditions, the material properties in terms of reflectivity become the discriminating value: the dark coloring of object Obj3 prevented acquisition of any significant result, as the point cloud was highly dispersed, but it is also apparent that the deviation values remain quite stable in presence of low and medium turbidity.

### II.3.10 Discussion about potential applications

The main contributions of this work can be summarized as follow:

1. A simple, low-cost and easily customizable experimental setup that may be used to test the effectiveness of the various active whole-field techniques in underwater conditions.
2. The description of an experimental approach to evaluate the performances of these systems with different water turbidity levels.



**Figure II.17:** Distribution of the deviations between the shape acquired in water and in air (a) and comparison of the point clouds with deviations evidenced through a color map (b).

3. The results of the first tests made with a whole-field technique based on stereo-photogrammetry and gray-code patterns projection.

The results of 3D reconstructions demonstrate that 3D underwater acquisition of small scale objects can be effectively made using a whole-field active system, also in presence of turbid water. As stated in Section II.2 (see Table II.1), in literature there are no studies about the use of active whole-field techniques in underwater environment. So this work could be considered as a first exploration of the potentialities and the feasibility of such an application. At the moment, the attention has been focused on the testing phase, trying to determine the influence of the turbidity on the accuracy of the measurements. So the imaging system, described before, has been developed taking in mind that its main application was the laboratory study in water tank. All the issues related to the use in real conditions have been deliberately ignored because they require an engineering effort that is not currently in the scope of our researches.

The experience acquired during these first experiments allow us to point out some considerations about the solutions adopted for the experimental setup:

- The use of the dome port for the camera significantly reduces the refraction effects, making the same calibration fit for both air and water environment;
- The availability of the new mini projectors with LED lamps allow us to



place the projector inside a waterproof housing without having problems related to the heat dissipation or the power consumption;

- The setup can be easily adapted to be used with different 3D imaging techniques (i.e.: stereo-photogrammetry, fringe projection, gray-code, etc.) in order to quickly perform comparative experimentations.

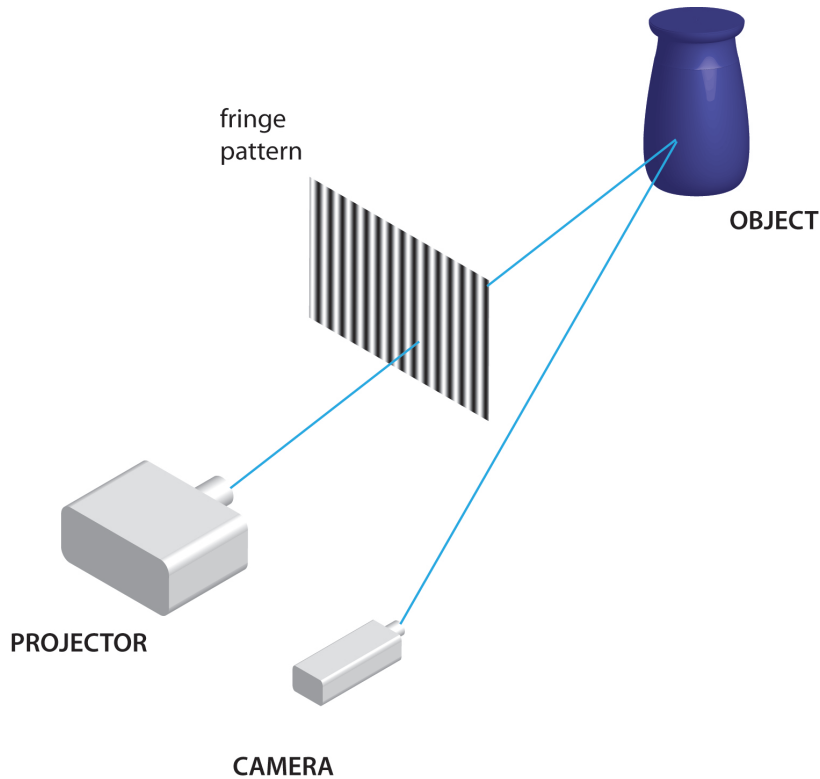
Moreover, the problems have been clearly identified, that currently limit the use of such a system in a marine environment, trying to suggest some possible solutions. The main limitation is related to the use of the multi-frame process (patterns projection and capture), that at present requires a long time for the acquisition. This problem could be considered negligible in some applications. As an example, the system could be fixed on a frame placed on the seabed for the relief of small objects in archaeological sites or in the study of the growth of coral reefs, where a very high precision is required. Another possible application could be in pipeline inspections [Rives & Borrelly 1997, Tetlow & Spours 1999] where a ROV may anchor the scanner to the pipeline (as previously done for gamma ray imaging [Roche *et al.* 2006]).

Anyway, in general, the long acquisition time is a relevant problem that limits the use of the system in real conditions. For this reason it is interesting to experiment different 3D imaging techniques that significantly reduces the number of projected patterns. Obviously, there is accuracy loss when using less patterns, but this could allow us to install the device on a ROV if it may ensure a sufficient stability during the acquisition.

In particular, in the next section the application of the Fringe Projection Technique is proposed [Bianco *et al.* 2010]. This technique works also with a single pattern and ensures a very fast acquisition, as well as the capture of a single frame, that may allow the use of the system both by ROV and scuba divers. The single pattern technique combined with a video camera may also be applied for close-range bathymetric mapping. Moreover, laser beams could be used to produce a grid of points, as well as interfering beams to expose a pattern of interference fringes. Both these solutions could reduce the scattering effects, but they make the setup more complex and costly.

## II.4 Fringe Projection Technique

The Fringe Projection Technique (FPT) is a whole-field technique based on the moiré phenomenon. The basis of moiré techniques is the comparison of two states of the same system of lines: the reference or undeformed system, and the modified or deformed system. Sinusoidal fringe patterns are projected on the surface of an object, and a camera records the images relative to these projections from another direction. These images are analyzed to evaluate the phase-map of the object which contains the height information, through signal processing techniques that provide the wrapped phase-map of the object. Then it is necessary to use phase unwrapping to obtain continuous phase distribution. One or more patterns can be used to extract the phase. The geometrical setup of the system is based on the principle of optical triangulation (Figure II.18). By calibrating the system, one can calculate the height map from the extracted phase distribution. Figure II.19 illustrates the various steps of the 3D reconstruction by FPT: fringe pattern projection, image acquisition, image processing, fringe analysis, calibration and 3D reconstruction. Moreover, software implemented in Matlab environment has been developed for 3D shape measurement. In general, a sinusoidal fringe pattern along a

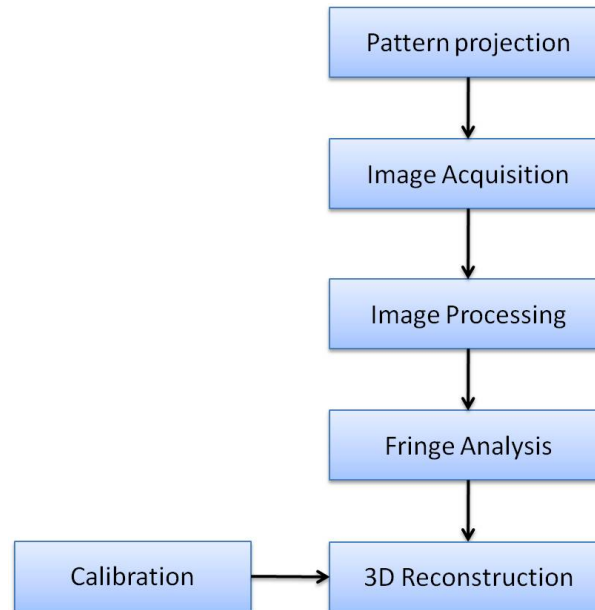


**Figure II.18:** Optical configuration of setup in fringe projection technique.

vertical line can be represented by:

$$g(x, y) = a(x, y) + b(x, y)\cos(2f_0x + \varphi(x, y)) \quad (\text{II.9})$$

where  $g(x, y)$  is the intensity value at the point  $(x, y)$ ;  $a(x, y)$  and  $b(x, y)$  represents irradiance variations arising from the non-uniform light reflection by a test object, in case of pattern image is recorded;  $a(x, y)$  is due to the texture variations on the object surface and can be estimated by a low-pass filtering process. Otherwise, in case of the pattern is projected,  $a$  and  $b$  are constant and equal to 0.5;  $f_0$  is the fundamental frequency of the sinusoidal pattern and  $\varphi(x, y)$  is the phase value which contains the information about the object profile: it stays constant if the pattern is projected or acquired from a reference plane. Then, one or more fringe patterns are projected by a digital projector on the object and acquired by a camera. Image processing techniques (resizing, filtering, cropping, creating mask, etc.) are applied in order to reduce noise, select the area of the image interested by the fringes where next techniques can be compute the phase, as described in Section II.4.2.



**Figure II.19:** Flowchart of FPT.

### II.4.1 Fringe Analysis

Fringe analysis is the process of extracting quantitative measurement data from fringe patterns [Surrel 2000]. It usually consists of phase detection and

phase unwrapping. The phase detection is the calculation of the fringes phase from the recorded intensity patterns. The intensity pattern encodes the physical quantity which is measured and this intensity varies as the cosine of a phase which is most often directly proportional to that physical quantity. Fringe processing is the extraction of the phase field from one or many intensity fields which have been acquired. The phase is obtained within the range  $[-\pi; +\pi]$ , i.e. the so-called wrapped phase-map, by different signal processing techniques that can be distinguished in *spatial* and *temporal*, if one or more patterns are used, respectively. Phase unwrapping technique is then applied to obtain a continuous phase distribution (unwrapped phase-map).

The spatial techniques extract the phase-map using one pattern. In this work, four different techniques are implemented and tested based on the signal processing techniques: Fourier Transform (FT), Windowed Fourier Transform (WFT), Wavelet Transform (WT) and Spatial Phase Demodulation (SPD).

- **Fourier Transform**

The FPT that uses the Fourier Transform to process the image pattern is called Fourier Transform Profilometry [Takeda & Mutoh 1983]. The Eq. (II.5) can be rewritten as

$$g(x, y) = a(x, y) + c(x, y)e^{j2\pi f_0 x} + c^*(x, y)e^{-j2\pi f_0 x} \quad (\text{II.10})$$

where

$$c(x, y) = \frac{1}{2}b(x, y)e^{j\varphi(x, y)} \quad (\text{II.11})$$

\* denotes complex conjugate. Taking the Fourier Transform of Eq. (II.10), for the variable  $x$ , we have

$$G(f, y) = A(x, y) + C(f - f_0, y) + C^*(f - f_0, y) \quad (\text{II.12})$$

where capital letters denote the Fourier spectrum and  $f$  is the spatial frequency in the  $x$  direction. Since the spatial variations of  $a(x, y)$ ,  $b(x, y)$  are slow compared with the spatial frequency  $f_0$ . One can make use of either of the two spectra centered around the carrier, say  $C(f - f_0, y)$  and  $C(f + f_0, y)$ , and translate it by  $f_0$  on the frequency axis towards the origin to obtain  $C(f, y)$  (see Figure II.20). The unwanted background variations  $a(x, y)$  are thus filtered out in this stage. The inverse Fourier transform of  $C(f, y)$  is computed to obtain  $c(x, y)$ . The phase of  $c(x, y)$  is obtained as:

$$\varphi(x, y) = \Im \log[c(x, y)] \quad (\text{II.13})$$

The phase map for the whole image is formed by repeating this process for each horizontal line in the image. First, the Fourier Transform of the fringe pattern image is calculated, then a band-pass filter is applied to select the phase information around the carrier  $f_0$ . The wrapped phase-map is obtained by calculating the Inverse Fourier Transform of this filtered signal (just the image with the carrier).

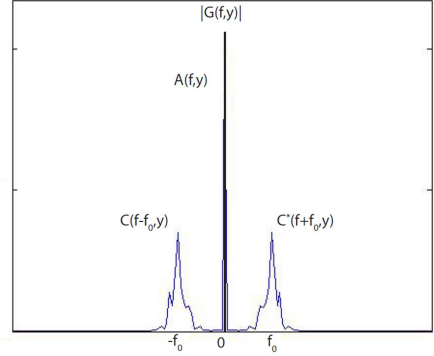


Figure II.20: Fourier frequency spectrum of the fringe pattern.

#### • Windowed Fourier Transform

The Windowed Fourier Transform of the signal  $g(x, y)$  (Eq. II.9) can be written as

$$W^F(u, \xi) = \int_{+\infty}^{-\infty} g(x, y)w(x - u)e^{-j\xi x} dx \quad (\text{II.14})$$

and the inverse is

$$g(x, y) = \frac{1}{2\pi} \int_{+\infty}^{-\infty} W^F(u, \xi)w(x - u)e^{j\xi x} dx \quad (\text{II.15})$$

where  $W^F(u, \xi)$  denotes the WFT spectrum;  $w(x)$  is a window which can be chosen as a Gaussian function

$$w(x) = \exp(-x^2/2\sigma^2) \quad (\text{II.16})$$

where the parameter  $\sigma$  controls the extension of  $w(x)$ . From Fourier Transform we know what frequencies appear from the spectrum, but we do not know where they occur in the signal. In contrast, using WFT, we obtain  $W^F(u, \xi)$  from which we know not only the spectrum components, but also where a component appears in the time domain. As a WFT is performed over a local area, determined by the extension of  $w(x)$ , a signal in one position will not affect the signal in another position in spectral analysis, provided that their distance is larger than

the effective radius of the Gaussian window. Moreover the spectrum of the signal in a local area is expected to be simpler than the spectrum of the whole-field signal. More effective operation of the spectrum is hence possible. The spectrum of the WFT can be processed in two ways:

1. Threshold the spectrum and set the spectral components with low amplitude to zero. This approach assumes that the spectrum for noise is widely distributed with low coefficients. As the noise is eliminated, a high-quality signal can be reconstructed.
2. Find the peak of the spectrum. The frequency corresponding to the peak is considered the local frequency or instantaneous frequency. This approach assumes that the spectrum of the signal in a local area is extremely simple and that it consists of only one frequency.

Based on the WFT, two different approaches [Kemaio 2004] for fringe demodulation are thus possible and are called the Windowed Fourier Filtering (WFF) method and the Windowed Fourier Ridges (WFR) method, respectively. The first approach filters the fringe patterns in the windowed Fourier domain, and the second approach searches for the best match between the fringe pattern and the computer-generated windowed exponential elements.

This work makes use of the WFF, described as follows. We can rewrite the Eq. (II.14)

$$W^F(u, \xi) = e^{-j\xi u} [g(u, y) * h(u, \xi)] \quad (\text{II.17})$$

with

$$h(u, \xi) = w(u)e^{-j\xi u} \quad (\text{II.18})$$

then the filtering process is applied, given a threshold value  $T$

$$\bar{W}^F(u, \xi) = \begin{cases} W^F(u, \xi), & \text{if } W^F(u, \xi) \geq T \\ 0, & \text{if } W^F(u, \xi) < T \end{cases} \quad (\text{II.19})$$

by computing the inverse transform

$$\bar{g}(x, y) = \int_{+\infty}^{-\infty} \int_{+\infty}^{-\infty} \bar{W}^F(u, \xi) w(x - u) e^{j\xi x} du d\xi \quad (\text{II.20})$$

or using the convolution

$$\bar{g}(x, y) = \int_{+\infty}^{-\infty} \int_{+\infty}^{-\infty} e^{j\xi x} [W^F(u, \xi) * h(u, \xi)] du d\xi \quad (\text{II.21})$$

and finally the phase distribution is:

$$\varphi(x, y) = \arctan \frac{\Im(\bar{g}(x, y))}{\Re(\bar{g}(x, y))} \quad (\text{II.22})$$

- **Wavelet Transform**

Wavelets offer a powerful method to quantify how energy is spatially distributed at multiple frequencies. The one dimensional wavelet transform of the fringe signal is defined by

$$W(a, b) = \frac{1}{\sqrt{a}} \int h^* \frac{x-b}{a} g(x, y) dx \quad (\text{II.23})$$

Each wavelet is obtained by scaling a mother wavelet  $h(x)$  by  $a > 0$  and translating it by  $b$ . The choice of an appropriate wavelet for a given application is an important practical question. A requirement for phase recovery is the use of complex wavelets. In this study, the Morlet wavelet was used to deal with phase recovery, because it is reputed to provide a better localization in both spatial and frequency domains [Dursun *et al.* 2004]. The Morlet wavelet, which is a plane wave modulated by a Gaussian, is defined as

$$h(x) = \pi^{\frac{1}{4}} e^{jcx} e^{-\frac{x^2}{2}} \quad (\text{II.24})$$

where  $c$  is a fixed spatial frequency. The Fourier Transform of  $h(x)$  is given by

$$H(k) = \frac{\sqrt{2\pi}}{\pi^{\frac{1}{4}}} e^{(-\frac{(k-c)^2}{2})} \quad (\text{II.25})$$

So for  $h^* \frac{x-b}{a}$  the Fourier Transform is

$$F(h^*(\frac{x-b}{a})) = H'(k) = a \frac{\sqrt{2\pi}}{\pi^{\frac{1}{4}}} e^{(-\frac{(k-\frac{c}{a})^2}{2a^2})} e^{jbk} \quad (\text{II.26})$$

Multiplying  $g(x, y)$  by  $h^*(\frac{x-b}{a})$  is equivalent to convolving with  $H'(k)$  in the frequency domain. Now  $H'(k)$  is a modulated Gaussian centered at  $\frac{c}{a}$ . It gives maximum response at the peak in  $G(f; y)$ . The center of the Gaussian is controlled by the magnitude of the dilation  $a$ . Since  $a$  is finite, we eliminate the DC. Also the phase of the result is unaffected by this convolution since the  $e^{jbk}$  term just amounts to a shift in the space domain. So the phase corresponding to the maximum response gives us the phase of the original fringe with the carrier, and the amplitudes of WT coefficients are defined as

$$A(a, b) = \sqrt{(\Im(W(a, b)))^2 + (\Re(W(a, b)))^2} \quad (\text{II.27})$$

$$a_m = \max_a(A(a, b)) \quad (\text{II.28})$$

So computing the 1D complex Morlet wavelet transform of the signal only in the  $x$  direction with  $y$  fixed, we obtain the phase of WT coefficients by taking the inverse tangent of the quotient of the imaginary part and dividing it by the real part:

$$\Phi(a_m, b) = \arctan\left(\frac{\Im(W(a_m, b))}{\Re(W(a_m, b))}\right) \quad (\text{II.29})$$

So the desired phase distribution  $\varphi(x, y)$  which contains the information about the object topographical variations can be obtained from the phase  $\Phi(a, b)$  at ridge points [Zhong & Weng 2004]. The maximum of the WT amplitudes at every position is defined as the ridge of the WT coefficient. Phase distribution of the entire image can be obtained by repeating this process for each value of  $y$ .

- **Spatial Phase Demodulation**

In respect to Fourier Transform technique, this method filters the signal in spatial domain [Lin & Lin 1999, Chen *et al.* 1996]. First, we multiply Eq. (II.9) by  $\cos(2\pi f_0 x)$  and obtain the following new equation:

$$\begin{aligned} g(x, y) \cos(2\pi f_0 x) &= a(x, y) \cos(2\pi f_0 x) + \\ & b(x, y) \cos(2\pi f_0 x + \varphi(x, y)) \cos(2\pi f_0 x) = \\ & a(x, y) \cos(2\pi f_0 x) \frac{1}{2} b(x, y) \cos(4\pi f_0 x + \varphi(x, y)) + \\ & \frac{1}{2} b(x, y) \cos(\varphi(x, y)) \quad (\text{II.30}) \end{aligned}$$

There is only one low frequency term in the above equation, i.e.  $\frac{1}{2} b(x, y) \cos(\varphi(x, y))$  (because the first two terms own the  $f_0$  values. Eq. (II.31) can be further generated after processing Eq. (II.30) by a low-pass filter:

$$g_1(x, y) = \frac{1}{2} b(x, y) \cos(\varphi(x, y)) \quad (\text{II.31})$$

Similarly, we multiply Eq. (II.9) by  $\sin(2\pi f_0 x)$  and obtain the following new equation:

$$\begin{aligned} g(x, y) \sin(2\pi f_0 x) &= a(x, y) \sin(2\pi f_0 x) + \\ & b(x, y) \cos(2\pi f_0 x + \varphi(x, y)) \sin(2\pi f_0 x) \\ & = a(x, y) \sin(2\pi f_0 x) \frac{1}{2} b(x, y) \sin(4\pi f_0 x + \varphi(x, y)) + \\ & \frac{1}{2} b(x, y) \sin(\varphi(x, y)) \quad (\text{II.32}) \end{aligned}$$



and there is only a low frequency term in Eq. (II.32) i.e.  $\frac{1}{2}\sin(\varphi(x, y))$ . Similarly, we process Eq. (II.32) by a low-pass filter and obtain the following equation:

$$g_2(x, y) = -\frac{1}{2}b(x, y)\sin(\varphi(x, y)) \quad (\text{II.33})$$

We divide Eq. (II.31) by Eq. (II.32) and obtain the following equation:

$$\frac{g_1(x, y)}{g_2(x, y)} = -\frac{\frac{1}{2}b(x, y)\sin(\varphi(x, y))}{\frac{1}{2}b(x, y)\cos(\varphi(x, y))} = -\tan(\varphi(x, y)) \quad (\text{II.34})$$

So, we can obtain

$$\varphi(x, y) = \frac{g_1(x, y)}{g_2(x, y)} \quad (\text{II.35})$$

where  $\varphi(x, y)$  represents the phase wrapped distribution.

**Phase stepping** or phase shifting is a temporal technique that uses 3 or more fringe patterns shifted by known quantities to extract the phase map. The phase stepping methods can use different phase shift: in this work four patterns are projected [Quan *et al.* 2001], with a shift of  $\pi/2$ . In this case we can represent the fringe patterns by the equation

$$g_n(x, y) = a(x, y) + b(x, y)\cos(2f_0x + \varphi(x, y) + \frac{\pi}{2}n) \quad (\text{II.36})$$

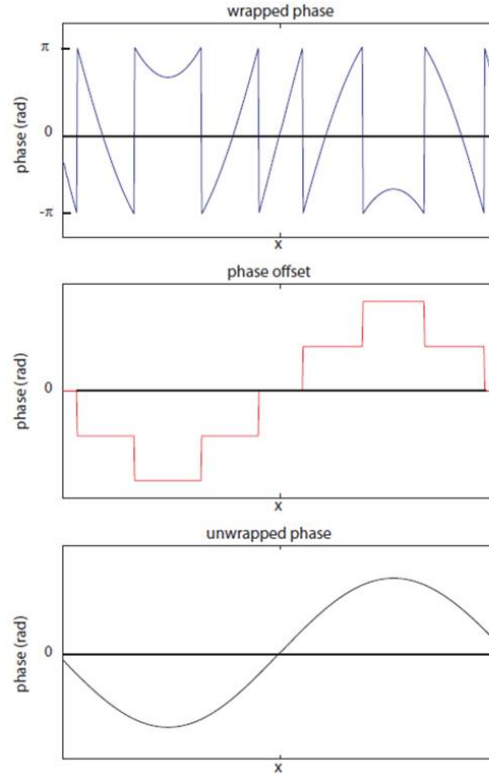
where  $n = 0, 1, 2, 3$ . The phase map (discontinuous) is simply computed by difference of images:

$$\varphi(x, y) = -\arctan\left(\frac{g_1(x, y) - g_3(x, y)}{g_0(x, y) - g_2(x, y)}\right) \quad (\text{II.37})$$

to faithfully map the real profile. Note that very often the phase is only meaningful up to an additive constant. There is no requirement on the fringe pattern, and the calculated phase at a given pixel depends only on the intensities recorded at the same pixel, unless spatial smoothing is used. Obviously, dynamic events cannot be studied as the phase-stepping takes some time.

#### II.4.1.1 Phase unwrapping

The phase values calculated by signal processing techniques are located between  $(-\pi/2, \pi/2)$  and are a wrapped value [Ghiglia & Pritt 1998]. This means that if the actual phase value is less than  $-\pi/2$ , it will be increased by  $\pi/2$ , and if the actual phase value is greater than  $\pi/2$ , it will be decreased by  $\pi/2$ . The increasing and decreasing operations are executed repeatedly until the



**Figure II.21:** One dimensional phase unwrapping.

final phase value is located between  $(-\pi/2, \pi/2)$ . So, the values calculated by Eqs.(II.13,22,29,35,37) are not the required phase values. Let the unwrapped phase value of the object at point  $(x, y)$  be denoted by  $\varphi(x, y)$ . The relation between  $\varphi_w(x, y)$  (wrapped phase) and  $\varphi(x, y)$  can be expressed by:

$$\varphi(x, y) = n\pi + \varphi_w(x, y) \quad (\text{II.38})$$

where  $n$  is a constant. To recover  $\varphi(x, y)$  from  $\varphi_w$  the unwrapped approaches can be adopted. An illustrative example showing the relation between the wrapped and unwrapped values is shown in Figure II.21, in one dimensional case: the continuous phase line is obtained by properly adding a phase offset. Several methods have been proposed in literature, based on different approaches, such as: basic residue algorithms [Feng *et al.* 2007], branch-cut [Nan & Dazheng 2004, Bioucas-Dias & Valadao 2005], max-flow/min-cut [Dazheng *et al.* 2006] and estimation of local frequency [Baldi *et al.* 2002]. In this work a phase unwrapping algorithm based on *minimum weighted discontinuity* [Flynn 1997] is used. The algorithm is robust also in presence of noise.

### II.4.2 Simulations

Preliminary simulations allow to test the implemented algorithms for 3D reconstruction by FPT. The scheme followed to obtain 3D shape with the different fringe analysis is illustrated in Figure II.22. The procedure is inverse in respect to the real case: a phase map (theoretical) is computed from a height function  $h(i, j)$ , then the fringe patterns are obtained and used to calculate the phase distribution by fringe analysis techniques; finally the height is reconstructed  $H(i, j)$ . A root-mean-square (RMS) error can be calculated to quantify the deviation of the 3D reconstructions:

$$E_{RMS} = \sqrt{\frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N [H(i, j) - h(i, j)]^2} \quad (\text{II.39})$$

In particular, to simulate the underwater conditions, a noise component is added to fringe patterns and the fringe analysis techniques are used for 3D shape reconstruction. The percentage values of  $E_{RMS}$  are reported in Table II.8. The WT and WFT are less sensitive to the noise and provide the minimum values of  $E_{RMS}$ . In Figure II.23 a comparison along a line between the height function and the height reconstructed with WT and WFT techniques is illustrated. An example of 3D shape reconstruction obtained in simulation is reported in Figure II.24.

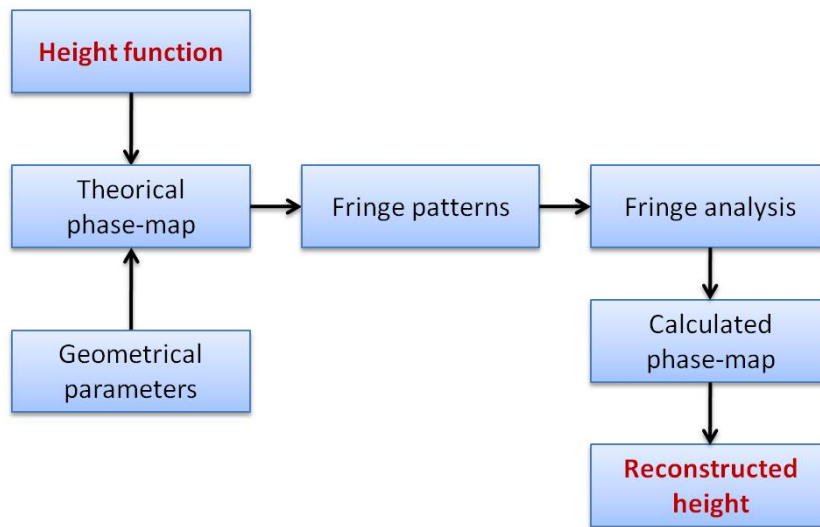
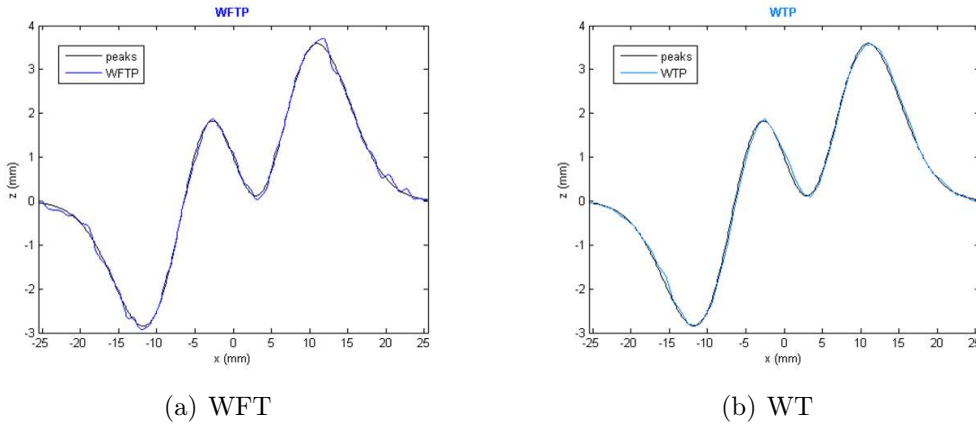


Figure II.22: Simulations scheme.

Fringe Analysis Technique	$E_{RMS}$ %
Fourier Transform	1.115
Windowed Fourier Transform	0.871
Wavelet Transform	0.784
Spatial Phase Demodulation	1.021
Phase Stepping	1.286

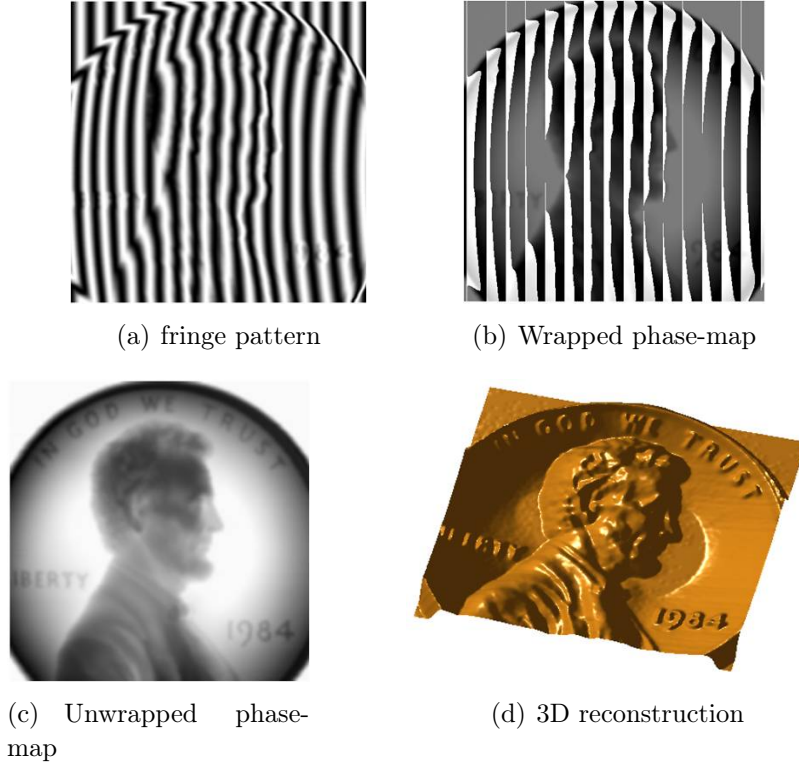
**Table II.8:** Percentage values of the root-mean-squares error  $E_{RMS}$ .



**Figure II.23:** One-dimensional comparison between the height function  $h$  (*peaks* Matlab function) and the height reconstructed  $H$  computed by WFT and WT.

### II.4.3 Phase-to-height conversion

A sinusoidal fringe pattern is projected on the object, with varying brightness profile along the horizontal direction  $x_p$  and uniform brightness profile along the vertical direction  $y_p$ . As shown in Figure II.25, equiphase lines in the projected sinusoidal grating produce a set of equiphase planes in the image space of the projector. When the line of sight of a pixel sequentially pierces these equiphase planes, monotonically changing phase values are observed at the pixel location. For phase value  $\varphi$ , the corresponding depth is given by the  $z$  coordinate of the point in which the line of sight of the pixel intersects with the plane of constant phase  $\varphi$ . Each pixel  $\mathbf{x}_c = [x_c \ y_c]^T$  in the camera frame is the projection of a point  $P$  in the scene that lies on the optical ray  $(O_c, \mathbf{x}_c)$ . Therefore, once its associated projector coordinate  $x_p$  is identified, the point  $P$  may be localized by intersecting the projector plane  $\alpha$  with the corresponding optical ray. The 3D shape is reconstructed by triangulation. The correspondence between image coordinates and projector coordinates are established exploiting the phase map computed by fringe analysis techniques. Note that the phase map calculated by the unwrapping algorithm is a map



**Figure II.24:** Example of 3D reconstruction by FPT.

relative, in the sense that the zero phase is arbitrary, so if one wants to obtain the absolute phase map, it is necessary to project a centerline on the object. The phase of points along this line is fixed to zero. Let  $\mathbf{X}_c = [X_c \ Y_c \ Z_c]^T$  and  $\mathbf{X}_p = [X_p \ Y_p \ Z_p]^T$  be the 3D position coordinate vectors of  $P$  in the camera and the projector reference frame, respectively. Denote  $\mathbf{x}_c = [x_c \ y_c \ 1]^T$  and  $\mathbf{x}_p = [x_p \ y_p \ 1]^T$  the respective homogeneous coordinates of the projections of  $P$  onto the camera and projector image planes. The following expressions relate image coordinates to 3D coordinates:

$$\mathbf{x}_c = \begin{bmatrix} x_c \\ y_c \\ 1 \end{bmatrix} = \frac{1}{Z_c} \mathbf{X}_c \quad (\text{II.40})$$

$$\mathbf{x}_p = \begin{bmatrix} x_p \\ y_p \\ 1 \end{bmatrix} = \frac{1}{Z_p} \mathbf{X}_p \quad (\text{II.41})$$

The relative positions of the optical devices are known from calibration. In other words, compute the rigid transformation equation that leads the 3D coordinates  $\mathbf{X}_c$  of any point  $P$  in the camera reference frame to its coordinates

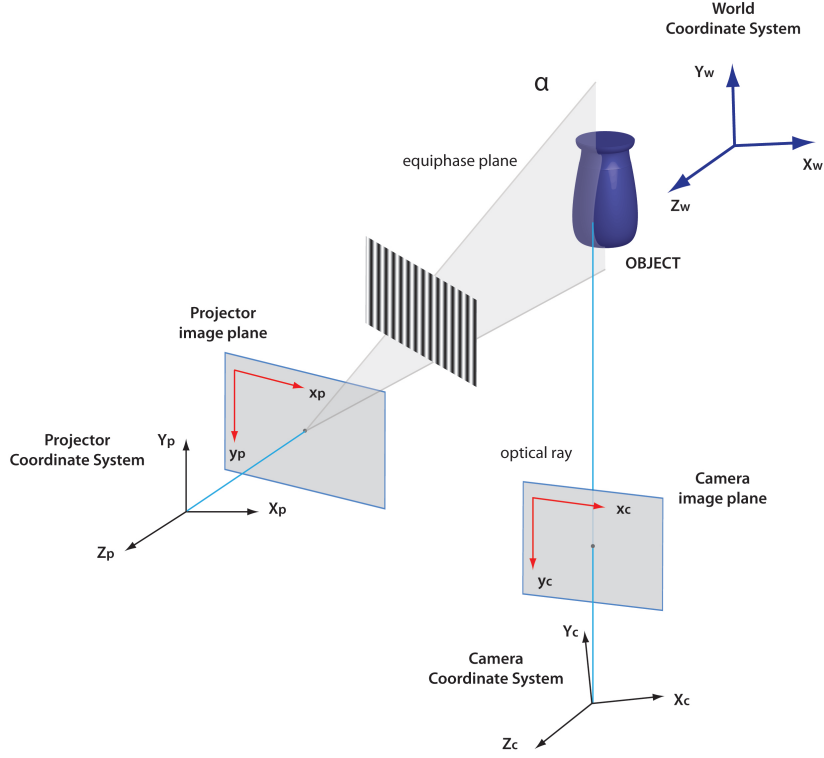


Figure II.25: Setup

in the projector reference frame  $\mathbf{X}_p$ :

$$\mathbf{X}_p = R\mathbf{X}_c + T \quad (\text{II.42})$$

where  $R$  and  $T$  are the rotation matrix and the translation vector that define the rigid motion between projector and camera, and they are calculated through the calibration process. Substituting Eq. (II.40) and (II.41) into Eq. (II.42):

$$Z_p \begin{bmatrix} x_p \\ y_p \\ 1 \end{bmatrix} = Z_c R \mathbf{x}_c + T \quad (\text{II.43})$$

Since the projector coordinate  $y_p$  is unknown, we have:

$$Z_p \begin{bmatrix} x_p \\ 1 \end{bmatrix} = Z_c R_{[1,3]} \mathbf{x}_c + T_{[1,3]} \quad (\text{II.44})$$

where  $R_{[1,3]}$  is the  $2 \times 3$  matrix containing the first and third row of the rotation matrix  $R$ , and  $T_{[1,3]} = [T_1 \ T_3]^T$  is a 2-vector containing the first and third coordinates of the translation vector  $T$ . From the previous equation, solving for the unknowns  $Z_c$  and  $Z_p$ :

$$\begin{bmatrix} Z_c \\ Z_p \end{bmatrix} = \mathbf{M}^{-1} T_{[1,3]} \quad (\text{II.45})$$

where  $\mathbf{M}$  is the following  $2 \times 2$  matrix:

$$\mathbf{M} = \begin{bmatrix} -R_{[1,3]}\mathbf{x}_c & \begin{bmatrix} x_p \\ 1 \end{bmatrix} \end{bmatrix} \quad (\text{II.46})$$

A closed form expression for the depth in the camera reference frame  $Z_c$  may then be derived:

$$Z_c = \frac{T_1 - x_p T_3}{\langle -R_{[1]} + x_p R_{[3]}, \mathbf{x}_c \rangle} \quad (\text{II.47})$$

where  $R_{[1]} = [R_{11} \ R_{12} \ R_{13}]^T = [R_{31} \ R_{32} \ R_{33}]^T$  are the first and third rows vectors of  $R$ . Notice that the denominator of Eq. (II.47) is a scalar product. The final expression for  $\mathbf{X}_c$  is:

$$\mathbf{X}_c = Z_c \mathbf{x}_c = \frac{T_1 - x_p T_3}{\langle -R_{[1]} + x_p R_{[3]}, \mathbf{x}_c \rangle} \quad (\text{II.48})$$

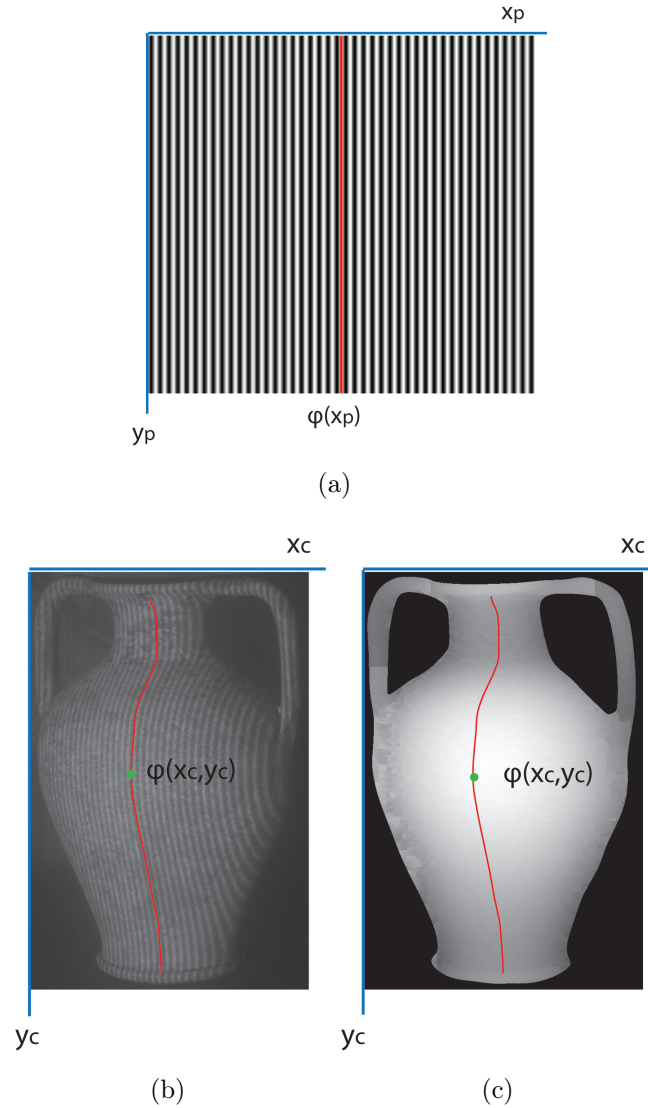
#### II.4.4 Calibration

The extrinsic parameters are computed by the calibration procedure to be used in Eq. (II.48). Also the intrinsic parameters are obtained to take into account the optical distortions. The procedure consists of correlating the coordinates of known markers located on a checkerboard with the corresponding coordinates on the image planes of camera and projector. Several images of the calibration panel are acquired from different positions for camera calibration, as described in Section II.3.6, while a checkerboard is projected on the same panel for each pose and captured in order to calibrate the projector.

The Toolbox of Bouguet is used to calibrate the camera [Bouguet 2010], and the projector is calibrated as an inverse camera. Once the calibration for each device has been obtained, the stereo calibration provides the relative position between the devices and fixes a reference coordinate system which will be used to compute the point cloud of the object.

#### II.4.5 Experimentation

The setup used in the experimentation is described in Section II.3.4: it features only one Nikon D200 digital camera and the DLP projector Mitsubishi PK20. The experimentation has been conducted in water tank, in different turbidity conditions. The calibration has been performed in clear water. A checkerboard with gray squares is used to detect the corners of the projected checkerboard in the images captured, for each different pose of the panel. A centerline is projected on the object to compute the absolute phase distribution. Then, one image of fringe pattern is captured.

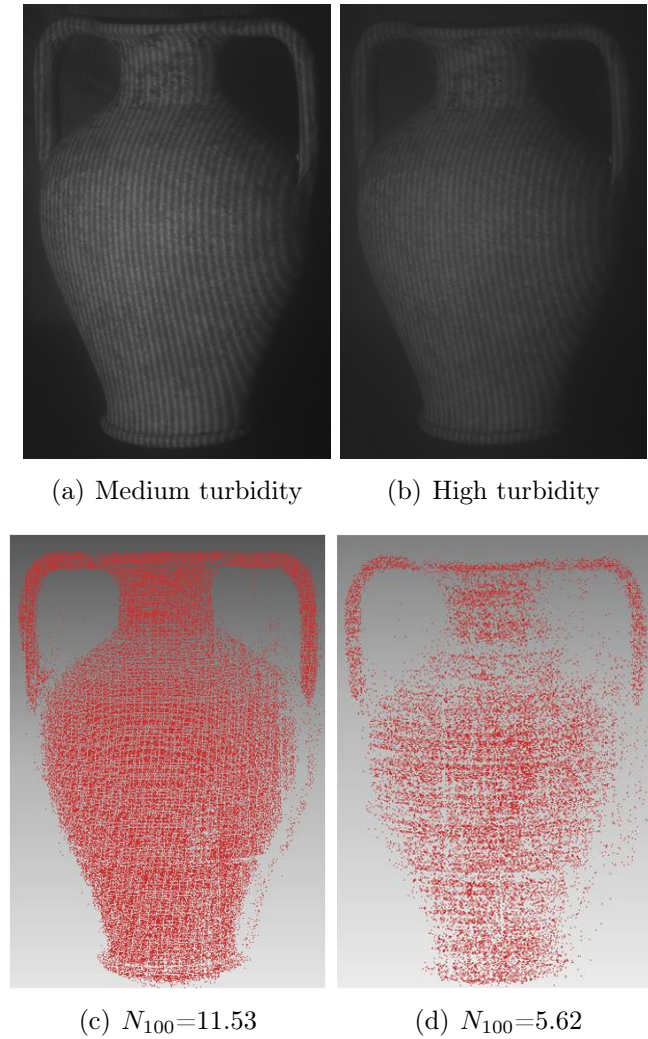


**Figure II.26:** Correspondence of a phase line among the projected fringe pattern (a), the acquired pattern (b) and the calculated phase-map (c).

Figure II.26 shows the correspondence among a phase value  $\varphi(x_p)$  in the projector image plane (which identify a vertical line), the phase in the camera image plane and the value in the calculated phase-map. For each  $\varphi(x_p)$ , there is a corresponding range of phase values along a line. By intersecting the optical ray at a point  $P$  on the object with phase  $\varphi(x_c, y_c)$  and the equiphase plane  $\varphi(x_p)$ , the 3D point cloud is obtained.

3D models are also acquired in turbidity as shown in Figure II.27. The number of 3D points for 100 pixels  $N_{100}$  in medium and high turbidity are 11.53 and 5.62, respectively. As it is apparent, the number of point clouds





**Figure II.27:** Fringe patterns and 3D point clouds, in turbid water.

decreases according to turbidity, but it can be still considered acceptable. In fact, these results show that the FPT improves the 3D acquisition in underwater environment with respect to the gray-code technique combined with the stereovision. Moreover, the obtained point clouds present less *outlier points*, so the noise due to scattering is attenuated.

## II.5 Conclusions

The possibility of acquiring 3D models of objects in an underwater environment by whole-field structured light technique has been investigated in this chapter. A first technique based on gray-code patterns and stereovision has

been tested. To achieve this goal, an underwater 3D system has been built. The designed 3D optical system is compact, low-cost, easy to build, based on commodity hardware, and has a battery life of one hour thanks to its low power consumption. At the moment, it has been designed only for laboratory testing, but it could be easily adapted for use both by scuba divers or to be installed on a ROV, AUV or submarine. The tests were conducted in a water tank with various concentrations of clay, in order to simulate a typical turbid environment like a seafloor.

The light conditions of the experiments have been monitored using a spectrophotometer and a luxmeter. The experiments have demonstrated the ability of the adopted technique to create a point cloud of the object, as long as the cameras are able to take sufficiently contrasted images of the patterns projected on the object. The accuracy of the acquisition has been determined by comparing the point clouds acquired in water, with different levels of turbidity, with the reference models acquired in air. Acceptable results are obtained in presence of low and medium turbidity; the technique managed to work also with high turbidity, while scattering did not permit acquisition of a point cloud in presence of extreme turbidity. Another interesting result is the comparison of the equipment performance in scanning different materials. The obtained point cloud can be used to create a 3D model, which can be textured with a slight photo retouch. The obtained results can be considered satisfactory, and are encouraging for future development of the equipment.

The main drawback of this technique is the long time of acquisition, due to the high number (50) of projected patterns. So, another whole-field technique has been used: the Fringe Projection Technique, that involves only one pattern to be projected and, of course, to be acquired, obtaining a drastic reduction of the acquisition time. Also in this second case, the experimental results are acceptable in turbidity conditions and comparable with the previous ones.

The main challenge for future development is the application in a real underwater environment. The optimizations may concern a quantitative study about the effect of polarization in backscattering; the evaluation of the benefits obtainable by varying the wavelength of the light source; the application of a mathematical model able to describe and reduce the degradation of the image induced by the scattering; a comparison of different whole-field techniques (e.g.: photogrammetry).

# Multispectral imaging for ancient documents analysis

---

*Historical documents often undergo various changes over time that alter their original state and reduce their legibility. Multispectral imaging offers a tool to analyse and elaborate the degradations which can affect these documents. The goal of virtual restoration is to improve the document legibility for both human and automatic operators, to extract patterns and enhance colour reproduction. The processing of document images is basically composed by the following steps: registration of images misaligned during the capture, decorrelation by statistical techniques to separate layers that are present in the document, and segmentation to isolate patterns.*

## III.1 Multispectral acquisition

Multispectral acquisition consists in taking pictures using more than one spectral component of electromagnetic radiation from the same region of an object. The portion of the electromagnetic spectrum (Figure III.1) that may be acquired extends from infrared to ultraviolet wavelengths and is shared in different channels. The channels are well-defined spectrally and often calibrated radiometrically. Generally speaking, multispectral acquisition implies that the number of acquired channels is smaller than ten. The conventional RGB capture is a simple case of multispectral acquisition of three channels in the visible bands, i.e. Red, Green, and Blue. Usually RGB imaging captures the three channels simultaneously: multispectral imaging instead captures images at different times. In fact, RGB imaging systems use colour sensors that are composed, for example, by a monochrome sensor with a Bayer mask, allowing to capture a colour scene with one shot, while multispectral imaging systems employ a monochrome sensor and a different optical pass-band filter for each channel that can be displayed as a gray-scale representation.

Much of the process described above was not practicable before the digital imaging revolution, and reflects the dual nature of multispectral imaging, as a source of both spectral data to be analyzed and images to be visualized. The data acquired may be radiometric (recording brightness or intensity in a broad,

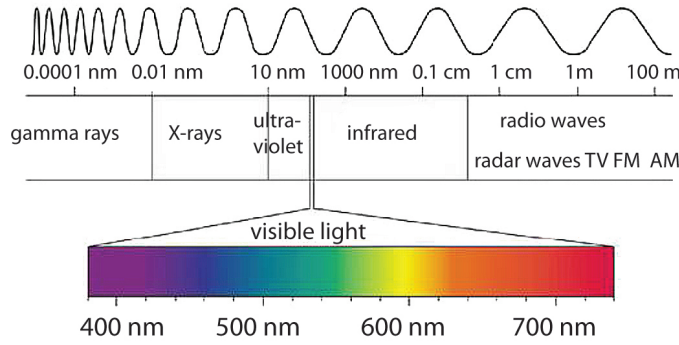


Figure III.1: Electromagnetic spectrum.

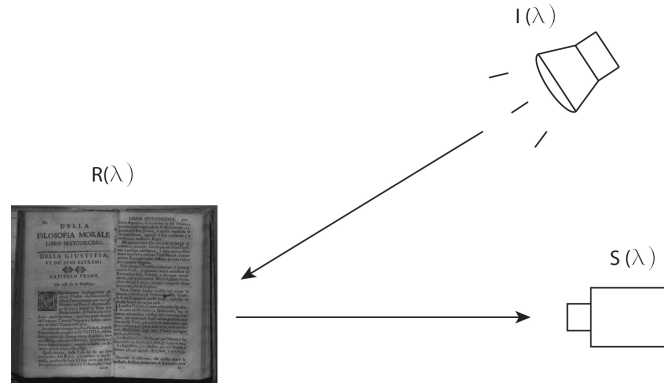


Figure III.2: Schematic view of the image acquisition process.

defined pass-band) or may consist of many narrow pass-bands, recording the spectral energy distribution at higher resolution. These data are in addition to the texture, geometry, and context that would normally be expected from images. The interpretation of a multispectral image requires an understanding of the characteristics of the filter-detector combination used to obtain the data, so that the spectral signature of the scene can be recovered.

The acquisition process of an image depends on the light emitted from the source of illumination, that is reflected by the object and acquired by a sensing element [Hardeberg *et al.* 1999]. The main components involved in the image acquisition process are illustrated in Figure III.2. Supposing a linear optoelectronic transfer function of the acquisition system, the camera response to an image pixel is then equal to:

$$c = \int_{\lambda_{min}}^{\lambda_{max}} l_R(\lambda)r(\lambda)o(\lambda)\phi_k(\lambda)s(\lambda), d\lambda \quad (III.1)$$

where:  $l_R(\lambda)$  denotes the spectral radiance of the illuminant,  $r(\lambda)$  the spectral reflectance of the object surface imaged in a pixel,  $o(\lambda)$  the spectral transmit-

tance of the optical systems in front of the detector array,  $\phi_k(\lambda)$  the spectral transmittance of an optical filter, and  $s(\lambda)$  the spectral sensitivity of the digital sensor array. For example, given a wavelength  $\lambda$  in the visible spectrum, the amount of light emitted at this wavelength is expressed through the power of the radiation source  $l_R(\lambda)$ , instead the light that is reflected from the object is  $l_R(\lambda)r(\lambda)$  that is the intrinsic colour of the object, in other words the contribution to the visual perception of object colour. This quantity is transmitted through the lens and the filters, and digitized by the sensor. A sensor CCD or CMOS is a digital device that reacts to light. The sensitivity function  $s(\lambda)$  indicates the sensor responsiveness to radiation at a given wavelength, and depends on *quantum efficiency*, i.e. the percentage of quanta that is converted to electrons.

So the imaging depends very much on the spectral characteristics of the light source illuminating the object and on the properties of the instrument itself. For example, the value of a pixel in a spectral image depends on the light irradiance at the corresponding object point, the reflectance of the object at this wavelength, and the sensitivity, gain and exposure time of the camera. The same object recorded with a different light source or with different camera settings will result in different spectral images. Nevertheless, such spectral images can be very useful, for example to show hidden features which might not be revealed in conventional photographs. However, a numeric (quantitative) evaluation of such differences is typically not very useful, because the resulting values are comparable only within the same spectral image or even only within parts of it. Combining information from several qualitative spectral images, as well as comparing images recorded at different times or even with different instruments, is practically impossible. In quantitative multispectral imaging the value of any pixel of a calibrated image represents a precise measurement of the portion of light that is reflected from the corresponding location on the object in a particular wavelength band. The resulting multispectral cube is a stack of such calibrated spectral reflectance images, which contains one image for each wavelength band (Figure III.3). For a given pixel coordinate, the sequence of pixel values in the images of the stack corresponds to an entire reflectance spectrum of the document at this location.

### III.1.1 Multispectral imaging systems

Multispectral imaging systems are widely used for cultural heritage investigations. CRISATEL High Resolution Multispectral System, by LUMIERE Technologies, is a camera designed for multispectral imaging of paintings. It is equipped with a 12 megapixel sensor mounted on a moving frame that slides horizontally to reach the resolution of 12.000x20.000 pixels. This camera has

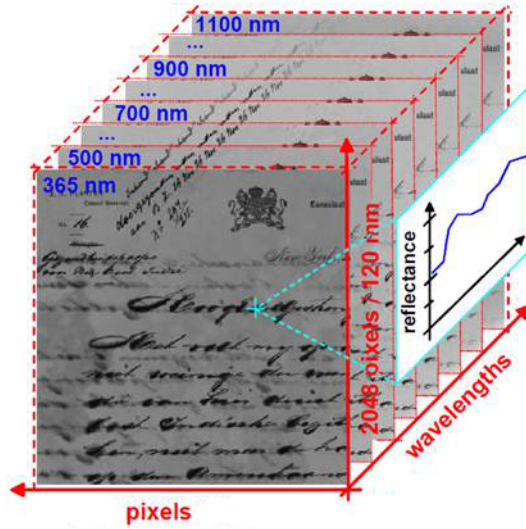
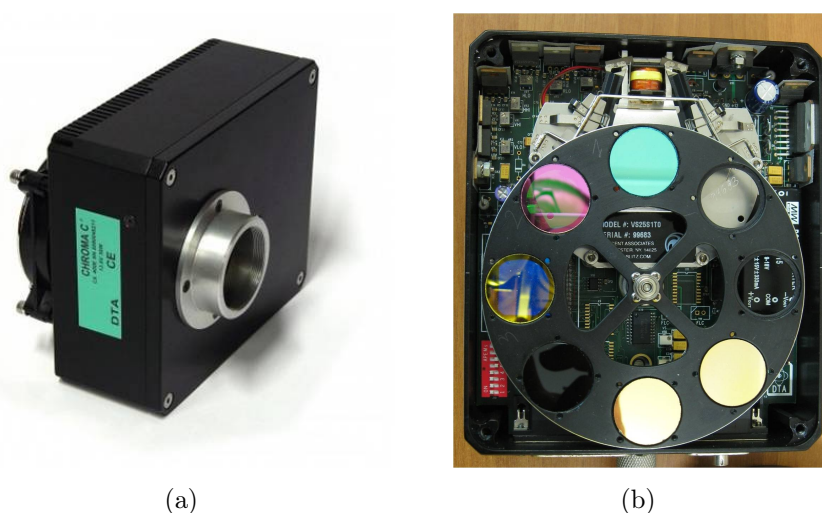


Figure III.3: Multispectral data cube [Klein *et al.* 2008].

a motorized filter-wheel, on which 13 interference filters are installed: ten covering the visible band and the other three covering near-infrared wavelengths [Ribés *et al.* 2005]. A camera used in the registration and enhancement of ancient manuscripts images [Lettner *et al.* 2008] is the Hamamatsu C9300-124. This compact device is equipped with a 10 megapixel grayscale CCD with a spectral response of 330-1000 nm. Because of its small size, the camera does not incorporate an internal filter-wheel, that must be placed outside, in front of the lens. A spectral imaging system (EurekaVision) system comprising a 39-Mpixel monochrome camera, LED-based narrowband illumination, and acquisition/control software has been designed for investigations of cultural heritage objects [Easton *et al.* 2003, Eur 2010]. The EurekaVision system captures high resolution images over 12 or more spectral bands from the near UV to the near IR. The spectral bands are created not by using band-pass filters to filter reflected light, but by using narrow-band LED illumination. A computer controllable hyper-spectral imaging apparatus, capable of acquiring spectral images of 5 nm bandwidth and with 3 nm tuning step in the spectral range between 380-1000 nm, has been developed in [Rapantzikos & Balas 2005a]. This device was selected as the instrument of choice for the Codex Sinaiticus Digitization Project conducted by the British Library in London [Cod 2010].

### III.1.1.1 Chroma CX3 multispectral camera

The experimentation makes use of the D.T.A. Chroma CX3 C1600E, shown in Figure III.4. It is a multispectral camera equipped with a 1.6 megapixel grayscale sensor (Kodak KAF-1602E) cooled by a Peltier cell that proved to



**Figure III.4:** (a) Chroma CX3 camera, filter-wheel camera (b).

be reliable even with very long exposure times, and Nikkor 35 mm lens. The camera is connected to the PC by a FDL-PCI interface card and powered at 13.8 V. It is characterized by a high sensitivity to light, as evidenced by its spectral response, that extends from 350 to 1100 nm. The peculiarity that makes the CX3 particularly suitable for multispectral acquisition in virtual restoration is the 8-positions motorized filter-wheel (the model at our disposal has just 6 of these positions actually occupied by the filters). Three filters correspond to RGB channels, while the other five slots can be customized according to the user's needs. In particular the slots are filled with three interference filters to perform acquisitions in the near-infrared spectral region (in the range from 740 nm to 1100 nm). A band-pass filter near the red (720-880 nm) has been chosen, a long-wave filter that allows transmission of radiation with a wavelength greater than 850 nm and a high-pass filter (from 740 nm). Figure III.5 shows the optical scheme of the camera. Moreover, an IR cut-off filter is used to block the infrared radiation in the acquisition of visible bands (Red, Green, Blue). The image at ultraviolet (UV) wavelengths is obtained by fluorescence of UV light, in the 350- 400 nm range.

### III.1.2 Multispectral acquisition of ancient documents

The multispectral images acquired with Chroma CX3 can be affected by blur and geometrical misalignment due to the different filters and the manual focus settings. For each filter shifting, it is necessary to put the image in-focus by manually rotating the ring of the lens, since the different refractive indices of the optical filters alter the FOV (Field of View) and the focal plane. The



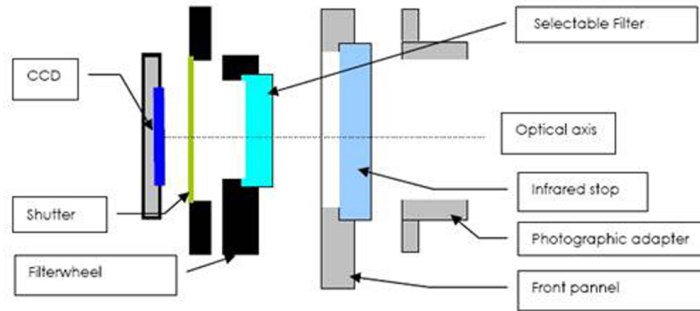


Figure III.5: Optical scheme.

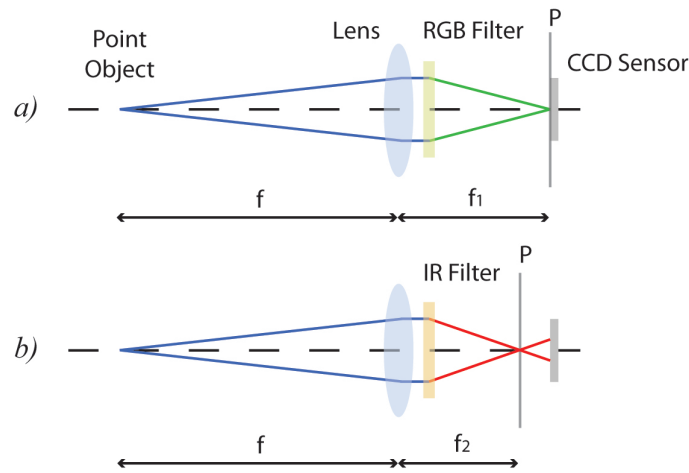


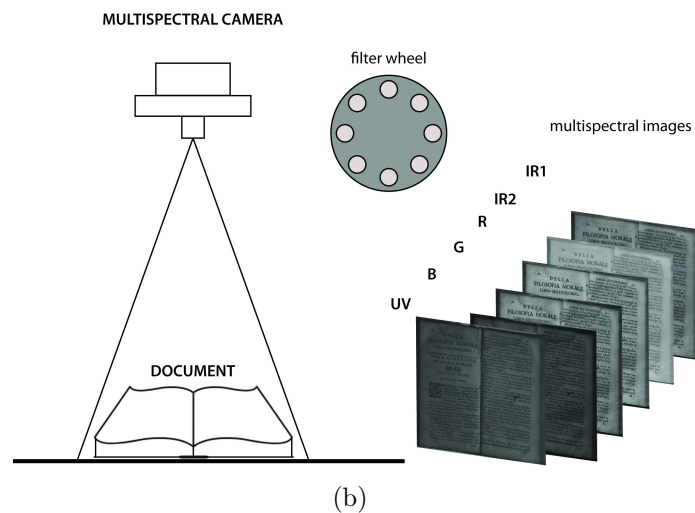
Figure III.6: Shifting of focus plane due to different refractive index of filters.

focusing, manually operated, consists in capturing images and subsequently adjusting the lens, thus the image quality can be compromised by an accidental error. The variation of optical path and refractive indices, caused by filters switching, affects the Chroma CX3, as well as every single sensor camera equipped with a filter wheel. In Figure III.6 this problem is shown: in *a*), with a RGB filter, the focal plane  $P$  coincides with sensor surface and we get a in-focus image; in *b*), with an IR filter, at the same target distance  $f$ , the focal plane  $P$  is at  $f_1 - f_2$  distance from the CCD surface and we get a blurred image [Mansouri & Marzani 2005]. Although multi-CCD cameras do not present this problem (the filter in this case is laminated directly on the sensor area, allowing to correct optical misalignment) [Brauers *et al.* 2008], the ability to easily modify the spectral response of the camera by quickly replacing the filters, combined with the lower cost of single CCD equipment, make the filter-wheel cameras the most used solution in virtual restoration. The experimental





(a)



(b)

**Figure III.7:** Multispectral imaging setup.

setup is illustrated in Figure III.7 and an example of document acquisition in the infrared, visible and ultraviolet bands is shown in Figure III.8. The information selectivity across the channels is apparent: in the infrared bands the foreground text disappears and hidden stamps are visible, whereas in the ultraviolet band blur and contrast increase. Moreover, the misalignment between two channels (G and UV), due to accidental displacements of the camera during the manual focus setting operation, is illustrated in Figure III.9.

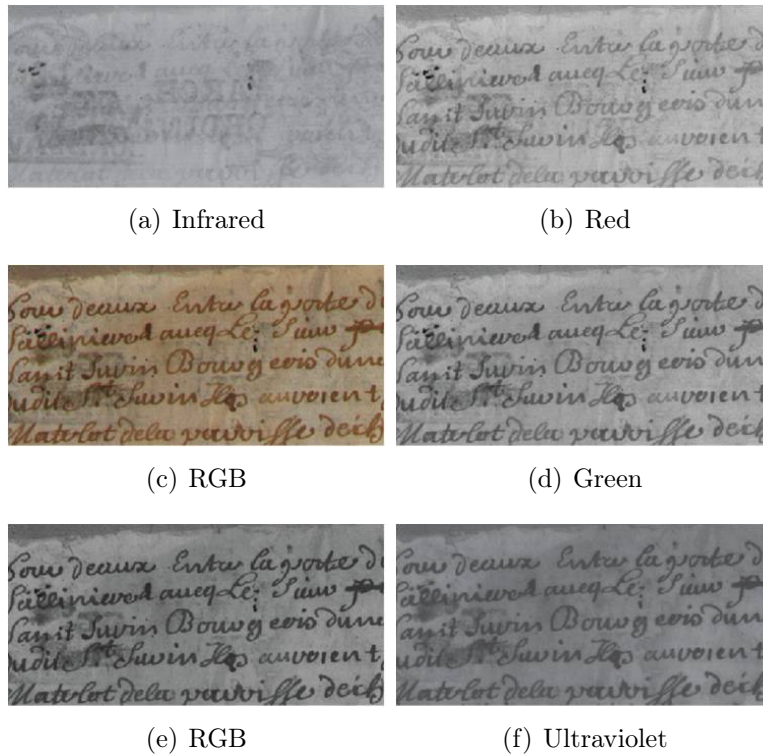
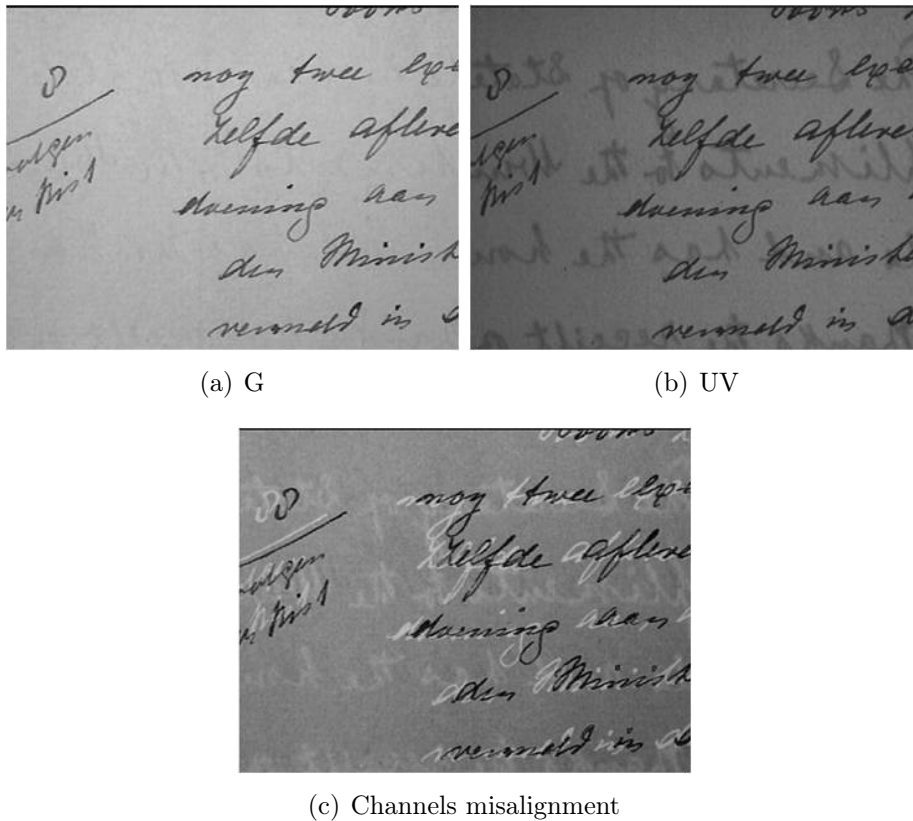


Figure III.8: Document images acquired in different spectral bands.

## III.2 Document image analysis

In the late 1980's the advent of fast computers, large computer memories, and inexpensive scanners fostered an increasing interest in document image analysis. With many paper documents being sent and received via fax machines and being stored digitally in large document databases, the interest in doing more with these images than simply view and print them had already grown. Just as humans extract information from these images, research was performed and commercial systems to read text on a page were built, in order to find fields on a form, to locate lines and symbols on a diagram, etc.

At present, a virtual restoration of such documents is of great interest for preservation, digital archiving, study and dissemination on the Internet. The goal of virtual restoration is to obtain digital data that can be used for subsequent image processing tasks, such as patterns extraction (stamps, text, etc.), improvement of readability and aesthetics, Optical Character Recognition, or to bring the document back to its original appearance.



**Figure III.9:** Misalignment of two channels illustrated as overlapping in transparency.

### III.2.1 Document degradations

The digital image quality of old documents depends greatly on the quality of the original documents that are often affected by several types of degradations. Old documents, supported by fragile materials, are easily affected by bad environmental conditions as the time elapses. Manipulation, humidity and poor storage for many years affect heritage documents and make them difficult to read. The deterioration level increases with excessive manipulations and consultations of the ancient document. Moreover, digitizing techniques inevitably add up, during the scanning process, some degradation to the quality of the images taken for the documents. The actual state of the readability of the digital copy of the degraded document is very bad. The term *degradation* refers to the following definition suggested by Henry S. Baird [Baird 2000]: “By degradation (or defects), we mean every sort of less than ideal properties of real document images”. Document image degradations could be divided in different sets according to their origin. We can distinguish for example defects due to the time effect and defects due to digitization. By time effect defects,

we refer to all defects related to bad environmental conditions, mainly the humidity, caused by a poor storage during many years. Once the old document is scanned, the previous defects become also a part of the digitized image document. During the scanning process, other type of degradations could be added. In [Drira 2006] an adequate typology of document image degradations is proposed. The proposed typology is made according to the further treatment that will be applied in the context of virtual document image restoration; this typology is thus intended to image processing.

- *Noisy background degradations*

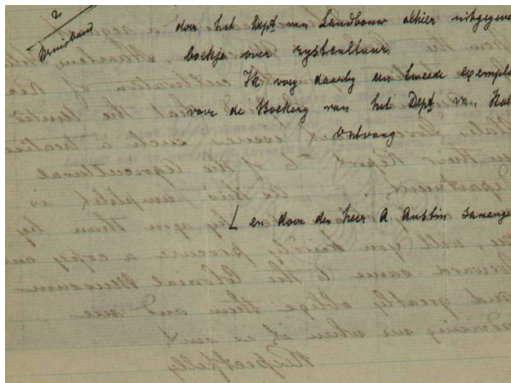
The legibility of digital document images is often compromised by the presence of artefacts in the background. This can derive from many kinds of degradations: spots due to the humidity and the digitization, marks resulting from the ink that goes through the paper generally called *show-through* (Figure III.10) (transparency of text through the page) or *bleed-through* (Figure III.11) (seeping of ink from the reverse side), interference strokes effects, strokes of pen and underlines.

- *Noisy foreground degradations*

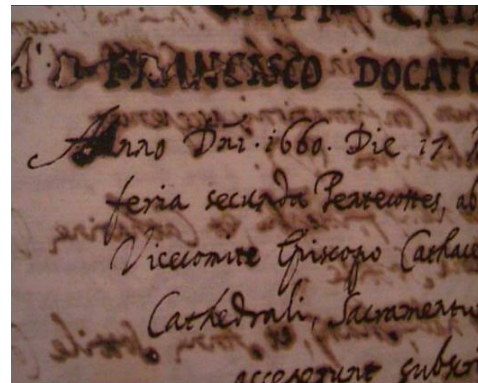
These degradations can lead to broken or touching characters, which are among the major causes of OCR errors. Noisy foreground degradations could contribute to a degraded foreground. This can occur especially if some degradation (i.e. spots, Figure III.12) interferes with text characters. In this case, each pixel in the spot contains both information about the real data and noise. Age effects can also affect the ink components. Many chemical effects can occur leading to ink disappearance and some gaps can even appear in the document image, causing significant loss of data affecting the document's content. Once the document is digitized, these degradations become a part of the document. Gaps, for example, create regions with homogeneous gray level. It is important to observe that semitransparent blotches could preserve part of the original information whereas gaps imply a complete loss of the information.

- *Global degradations*

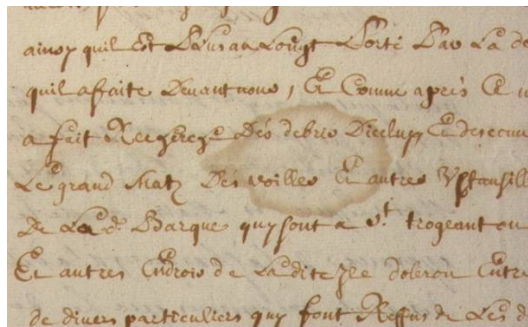
This class of degradations refers to degradations affecting the entire document: for example, the geometrical degradation, which is a common phenomenon in document analysis resulting from the scanning of thick documents. The paper surface of a thick book is usually curved during the scanning. The presence of this curvature leads to warped words appearing around the book spine area and to a non-uniform illumination.



**Figure III.10:** Show-through: can appear in the scanned image when the paper is not completely opaque.



**Figure III.11:** Bleed-through: due to seeping of ink from the reverse side.



**Figure III.12:** Example of ink diffusion (spot) due to humidity.

Time effect is another source of degradation which can degrade original colours. This issue is addressed in Chapter IV.

### III.2.2 Enhancement methodology

The goal of virtual restoration is to obtain digital data that can be used for subsequent image processing tasks, such as patterns extraction (stamps, text, etc.), improvement of readability and aesthetics, Optical Character Recognition, or to bring the document back to its original appearance. Recently, enhancement techniques have become a tool for scientific analysis and documentation of old degraded manuscripts, overwritten (Figure III.13) or disintegrating (Figure III.14) texts [Easton *et al.* 2003]. The main advantage of the analysis in different wavelengths extending the visible light is the additional information that the human eye cannot see. In particular, statistical techniques of decorrelation work on images captured in multispectral modality.



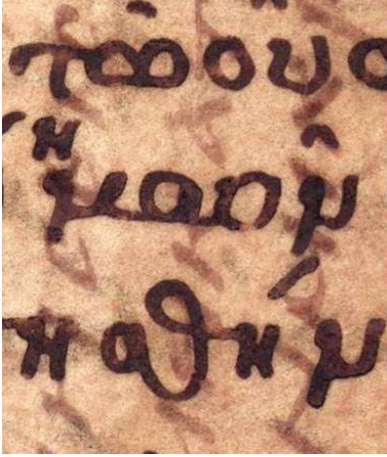


Figure III.13: Palimpsest.



Figure III.14: Papyrus.

### III.2.2.1 Related works

Among the several studies conducted in image analysis of historical documents, some of the most pertinent will be presented in this section [Lettner *et al.* 2007]. Two prominent representatives are the Archimedes Palimpsests [Easton *et al.* 2003] and the Herculaneum papyri [Ware *et al.* 2000]. Easton *et al.* were the first to capture and enhance the erased writing of the famous Archimedes palimpsest by multispectral methods [Easton *et al.* 2003]. The system they propose is modeled on the VASARI illumination system developed at the National Gallery of London [Martinez *et al.* 2002]. In this project it turned out that the adoption of spectral imaging produces higher and better readability of the texts than conventional thresholding methods. Very good results in the enhancement of spectral images of palimpsests and other latent texts have also been achieved by Italian company FOTOSCIENTIFICA Re.co.rd. [FOTOSCIENTIFICA 2010] which provided for instance the pictures for the EC project Rinascimento Virtuale, devoted to the decipherment of Greek palimpsest manuscripts [Virtuale 2010]. The EC project IsyReaDeT developed a system for a virtual restoration and archiving of damaged manuscripts, using a multi-spectral imaging camera, advanced image enhancement and document management software [Tonazzini *et al.* 2004a]. For the readability enhancement, Easton *et al.* have used an unconstrained least squares algorithm for spectral unmixing and produced normalized and non-negative fraction maps of text and parchment [Easton *et al.* 2003]. The combination of these fraction maps can be used to highlight different classes, e.g. the underlying text and the overwriting. Rapantzikos and Balas separated the overwritten from the underwritten text-layer by means of the Principal Component Analysis (PCA) and

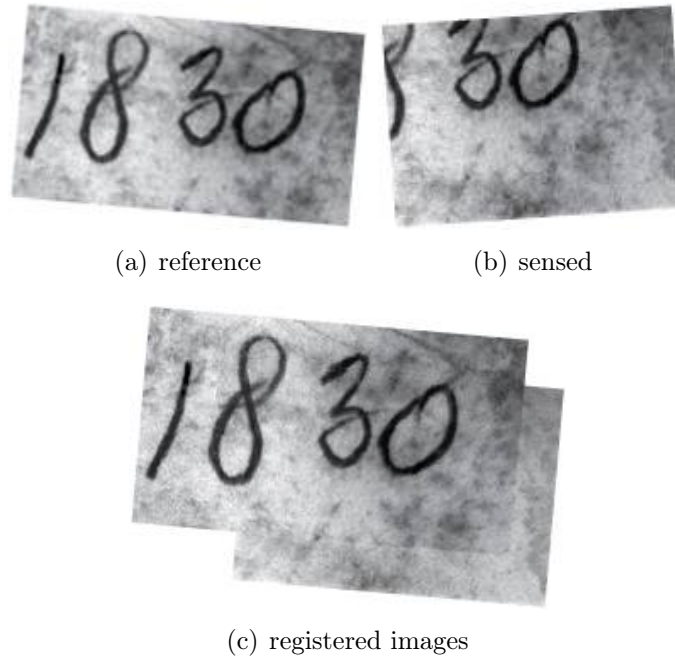
a linear spectral mixture analysis [Rapantzikos & Balas 2005b]. Promising results have also been obtained by Tonazzini *et al.* who applied an Independent Component Analysis (ICA) to the spectral components at different bands to separate bleed-through and show-through texts as well as palimpsests [Tonazzini *et al.* 2004a]. The major advantage of this method is that no models are required (Blind Source Separation). For the extraction of *hidden information* of multispectral images, methods adapted from remote sensing applications, e.g. PCA, are also appreciable to emphasize information hidden in individual spectral bands. The generation of binary images out of ancient manuscripts is discussed by Bar-Yosef [Yosef *et al.* 2004]. Since the behavior of distinct inks is different in multispectral bands, the result depends on the method used for enhancing the digital images. Some studies on historical documents cover not only the enhancement of the readability, but also the more detailed analyses of the text contents. For instance, Manmatha and Rothfeder [Manmatha & Rothfeder 2005] as well as Feng and Manmatha [Feng & Manmatha 2005] investigate several methods for the segmentation of words in historical documents, focusing on the whole word in order to avoid character segmentation, which is difficult to perform on degraded documents. Features for word spotting are also described by Rath and Manmatha [Rath & Manmatha 2003] and Leydier *et al.* [Leydier *et al.* 2005]. Text extraction from gray scale historical document images is investigated by Shi *et al.* [Shi *et al.* 2005], who use an adaptive local connectivity map to extract text from background. Blobs containing several instances of degraded characters are analyzed by Tonazzini *et al.* [Tonazzini *et al.* 2004c], but this study is already based on printed characters and not on medieval written sources.

### III.2.2.2 Methodology

The proposed procedure to analyse the degradations and improve the legibility of a degraded document is composed mainly of three steps: registration by Fourier-based technique, decorrelation through statistical techniques and finally, image segmentation by local technique.

### III.2.2.3 Image registration

Image registration is the process of overlaying two or more images of the same scene taken at different times, from different viewpoints, and/or by different sensors. It geometrically aligns two images: the *reference* and *sensed* images (Figure III.15). The present differences between images are due to different imaging conditions. Image registration is a crucial step in all image analysis tasks in which the final information is gained from the combination of various



**Figure III.15:** Registration of reference and sensed images.

data sources like in image fusion, change detection, and multichannel image restoration.

In general, the registration applications can be divided into four main groups according to the nature of the image acquisition process [Zitová & Flusser 2003]:

1. Different viewpoints (*multiview analysis*)

Images of the same scene are acquired from different viewpoints. The aim is to gain a larger 2D view or a 3D representation of the scanned scene.

2. Different times (*multitemporal analysis*)

Images of the same scene are acquired at different times, often on regular basis, and possibly under different conditions. The aim is to find and evaluate changes in the scene which appeared between the consecutive image acquisitions.

3. Different sensors (*multimodal analysis*)

Images of the same scene are acquired by different sensors. The aim is to integrate the information obtained from different source streams to gain more complex and detailed scene representation.



#### 4. Scene to model registration

Images of a scene and a model of the scene are registered. The model can be a computer representation of the scene, for instance maps or digital elevation models (DEM) in GIS, another scene with similar content (another patient), “average” specimen, etc. The aim is to localize the acquired image in the scene/model and/or to compare them.

Due to the diversity of the images to be registered, and to various types of degradations, it is impossible to design a universal method applicable to all registration tasks. Every method should take into account not only the assumed type of geometric deformation between the images but also radiometric deformations and noise corruption, registration accuracy and application-dependent data characteristics. Nevertheless, the majority of the registration methods consists of the following four steps:

- *Feature detection*

Salient and distinctive objects (closed-boundary regions, edges, contours, line intersections, corners, etc.) are manually or (preferably) automatically detected. For further processing, these features can be represented by their point representatives (centers of gravity, line endings, distinctive points), which are called control points (CPs) in literature.

- *Feature matching*

In this step, the correspondence between the features detected in the sensed image and those detected in the reference image is established. Various feature descriptors and similarity measures, along with spatial relationships among the features, are used for that purpose.

- *Transform model estimation*

The type and parameters of the so-called mapping functions, aligning the sensed image with the reference image, are estimated. The parameters of the mapping functions are computed by means of the established feature correspondence.

- *Image resampling and transformation*

The sensed image is transformed by means of the mapping functions. Image values in non-integer coordinates are computed by the appropriate interpolation technique.

The implementation of each registration step has its typical problems. First, it is necessary to decide what kind of features is appropriate for the

given task. The features should be distinctive objects, which are frequently spread over the images and which are easily detectable. The feature sets detected both in the reference and the sensed images must have enough common elements, even in situations when the images do not cover exactly the same scene or when there are object occlusions and/or other unexpected changes. The detection methods should have a good localization accuracy and should not be sensitive to the assumed image degradation.

In the feature matching step, problems caused by an incorrect feature detection or by image degradations can arise. Physically corresponding features can be dissimilar due to the different imaging conditions and/or to the different spectral sensitivity of the sensors. The choice of the feature description and the similarity measure has to consider these factors. The feature descriptors should be invariant to the assumed degradations. Simultaneously, they have to be discriminable enough to be able to distinguish among different features and sufficiently stable as well, and they have not to be influenced by slight unexpected feature variations and noise. The matching algorithm in the space of invariants should be robust and efficient. Single features without corresponding counterparts in the other image should not affect its performance.

The type of the mapping functions should be chosen according to the *a priori* known information about the acquisition process and the expected image degradations. If no *a priori* information is available, the model should be flexible and general enough to handle all the possible degradations which might appear. The accuracy of the feature detection method, the reliability of feature correspondence estimation, and the acceptable approximation error need to be considered too.

Finally, the choice of the appropriate type of resampling technique depends on the trade-off between the demanded accuracy of the interpolation and the computational complexity. The nearest-neighbor or bilinear interpolation are sufficient in most cases; however, some applications require more precise methods. Because of its importance in various application areas and its complicated nature, image registration has been the topic of much recent research.

### **An image registration method based on Fourier-Mellin Transform**

Two main approaches to feature understanding have been established:

1. *Area-based methods*

Put emphasis rather on the feature matching step than on their detection. No features are detected in these approaches, so the first step of image registration is omitted.

### 2. Feature-based methods

Use of extracted features to estimate the registration parameters. Significant regions (forests, lakes, fields), lines (region boundaries, coastlines, roads, rivers) or points (region corners, line intersections, points on curves with high curvature) are understood as features here. They should be distinct, spread all over the image and efficiently detectable in both images.

The use of feature-based methods is recommended if the images contain enough distinctive and easily detectable objects (applications in remote sensing and computer vision). On the other hand, medical images are not so rich in such details, and thus area-based methods are usually employed here. The applicability of area-based and feature-based methods for images with various contrast and sharpness is analyzed in [Rezaie & Srinath 1984]. Recently, registration methods using simultaneously both area-based and feature-based approaches have started to appear [P. Hellier 2000].

Due to the diversity of multispectral images in which the context information changes from infrared to ultraviolet channel, in this work an automatic registration method based on the Fourier-Mellin transform has been applied. The whole images are used to estimate the registration parameters and match them automatically. To relate each pixel to a precise location, one of the data images is taken as a reference, and all the others are aligned to create a data cube where all the information associated to one pixel is also associated to a fixed spatial location. Supposing that the document page undergoes rigid transformations in its plane and taking into account the changing of scale due to the different FOVs of the optical filters, one can assume that a similarity transformation is sufficient to model the misalignment. Once a reference image has been fixed, the other images are registered by computing the translation, rotation and scale parameters. This method applies the Fourier Shift Theorem [Bracewell 2000] to recover the translation, and a phase correlation in the log-polar domain to obtain the rotation and scale parameters [Reddy & Chatterji 1996]. Once all the transformations are computed, the images are cropped to let them depict the same area and have the same size.

Fourier methods differ from other registration strategies because they search for the optimal match according to information in the frequency domain. In this correspondence, an extension of the phase correlation technique is applied for automatic image registration, which is characterized by its insensitivity to translation, rotation, scaling and noise, and by its low computational cost as well.

Let  $i_1$  and  $i_2$  are two images that differ only by a displacement i.e.  $(x_0, y_0)$ ,

$$i_2(x, y) = i_1(x - x_0, y - y_0) \quad (\text{III.2})$$

Their corresponding Fourier transforms  $I_1$  and  $I_2$  will be related by:

$$I_2(\xi, \eta) = e^{-j2\pi(\xi x_0 + \eta y_0)} I_1(\xi, \eta) \quad (\text{III.3})$$

The cross-power spectrum of the two images is defined as:

$$\frac{I_1(\xi, \eta) I_2^*(\xi, \eta)}{|I_1(\xi, \eta) I_2(\xi, \eta)|} = e^{j2\pi(\xi x_0 + \eta y_0)} \quad (\text{III.4})$$

where  $I_2^*$  is the complex conjugate of  $I_2$ . The shift theorem guarantees that the phase of the cross-power spectrum is equivalent to the phase difference between the images. By taking inverse Fourier transform of the representation in the frequency domain, we will have a function that is an impulse; that is, it is approximately zero everywhere except at the displacement that is needed to optimally register the two images. If  $i_2(x, y)$  is a translated and rotated replica of  $i_1(x, y)$  with translation  $(x_0, y_0)$  and rotation  $\theta$ , then

$$i_2(x, y) = i_1(x \cos \theta + y \sin \theta - x_0, -x \sin \theta + y \cos \theta - y_0) \quad (\text{III.5})$$

According to the Fourier translation property and the Fourier rotation property, transforms of  $i_1$  and  $i_2$  are related by

$$I_2(\xi, \eta) = e^{-j2\pi(\xi x_0 + \eta y_0)} I_1(\xi \cos \theta + \eta \sin \theta, -\xi \sin \theta + \eta \cos \theta) \quad (\text{III.6})$$

Let  $A_1$  and  $A_2$  be the magnitudes of  $I_1$  and  $I_2$ . Therefore, from Eq. (III.6) we have

$$A_2(\xi, \eta) = A_1(\xi \cos \theta + \eta \sin \theta, -\xi \sin \theta + \eta \cos \theta) \quad (\text{III.7})$$

then a rotation of angle  $\theta$  in the spatial domain causes the Fourier representation to be rotated through the same angle. If  $i_1$  is scaled replica of  $i_2$  with scale factor  $k$ , according to the Fourier scale property, the Fourier transforms of  $i_1$  and  $i_2$  are related by

$$I_2(\xi, \eta) = \frac{1}{|k|} I_1\left(\frac{\xi}{k}, \frac{\eta}{k}\right) \quad (\text{III.8})$$

Scaling in the spatial domain causes an inverse scaling in the frequency domain.

### *Invariance*

The translation property of the Fourier transform shows that spatial shifts affect only the phase representation of an image. This leads to the well known result that the Discrete Fourier transform magnitude is a circular translation invariant. The basic translation invariants may be converted to



Transformation	Displacement	Estimation
Translation (px)	(10,25) (25,70)	(10,25) (25,70)
Rotation (degree)	1° 2°	1.14 2.23
Scaling (%)	5 10	4.9 9.4

**Table III.1:** Results of registration with different changing of transformation.

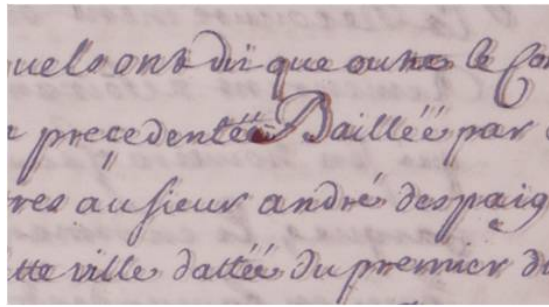
Channels	Translation (px)	Rotation (degree)	Scale (%)
IR	(1,4)	0	3
R	(1,0)	0	0
G	(1,0)	0	0
B	(1,0)	0	0
UV	(-6,-4)	0.4	1

**Table III.2:** Displacements obtained from the registration of the multispectral images of Figure III.17

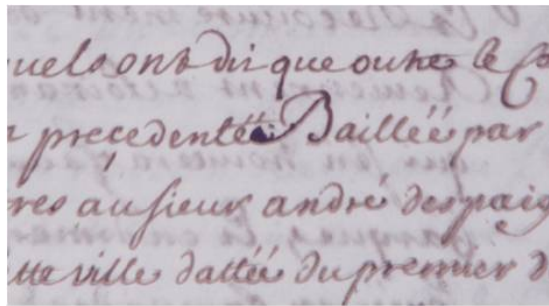
estimated values of these parameters are reported in Table III.1. The alignment of the images are perfect in the case of translation. The rotation and the scaling present small percentage errors. The case of multispectral images has been analysed, choosing as reference a channel with good contrast (usually the green). Figure III.17 illustrates an example of multispectral images to be registered with Fourier-Mellin algorithm. The misalignments between Green and Infrared channels are apparent and also the alignment after registration. The relative displacements calculated to align these images are listed in Table III.2. Moreover, it has to be noted that the time to process these five images with resolution of  $1236 \times 940$  is about 37.4 sec.

#### III.2.2.4 Image decorrelation

Removing interferences and degradations is not trivial, since their intensity is often close to the one of the main text. On the basis of a single-side scan, the techniques attempted to reduce the interference are based on thresholding or segmentation classification. The challenging problem of double-sided (or recto-verso) scan will be discussed in the next section. From a colour scan, thresholding can only be used effectively within multi-resolution analysis with adaptive binarization, or using directional wavelets [Nishida & Suzuki 2003, Wang *et al.* 2003]. Other proposals include colour cluster segmentation via an



(a) Unregistered images



(b) Registered images

**Figure III.17:** Registration of Green and Infrared channels. The details of the document illustrated in Figure III.18 are shown as overlapped in transparency.

adaptation of the k-means algorithm [Drira *et al.* 2006, Leydier *et al.* 2004], and a single-scan grayscale classification based on a double binary Markov random field [Wolf 2006]. While these methods certainly perform better than simple thresholding, they do not completely succeed in discriminating unambiguously the foreground text from interferences. Especially for gray-level images, techniques that exploit the two sides of a document promise to perform better [Dubois & Pathak 2001, Sharma 2001]. These techniques rely on segmentation to identify the show-through areas, and then suitably inpaint estimated pure background areas [Dano 2003, Knox 1998, Wang & Tan 2001].

In [Sharma 2001] a nonlinear model is introduced for show-through, and a compensation using adaptive filters is proposed. A nonlinear diffusion approach is proposed in [Cheriet & Moghaddam 2008] to model and remove these degradations.

All these approaches, however, pose two kinds of problems. Firstly, all the data maps must be spatially registered before processing. This is not trivial, since the different colour channels have often different focus conditions and, above all, recto and verso normally suffer from topographical differences. Secondly, since the above mentioned solutions remove all the structured background from the foreground text, other than erasing the strokes coming

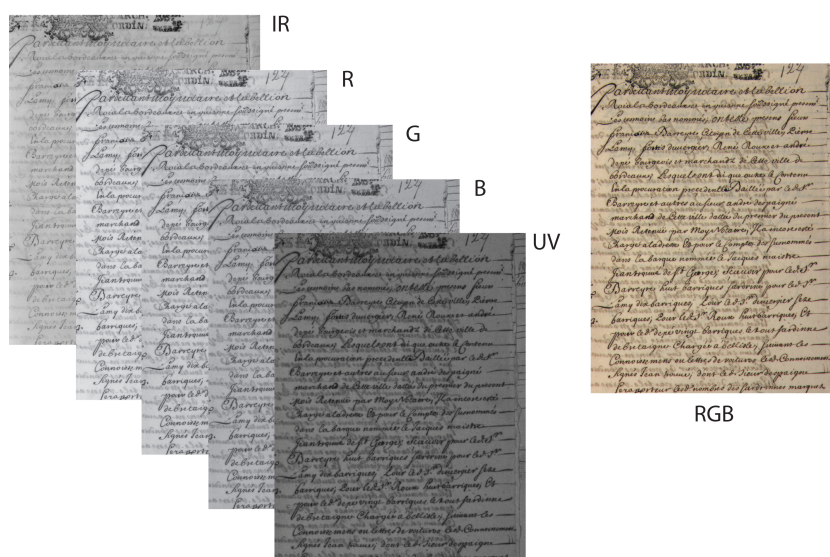


Figure III.18: Multispectral images acquired from IR to UV bands to be registered.

from the reverse side, they can also remove other patterns belonging to the front side (e.g. stamps). This may be undesirable, when these patterns are signs of the document history and authenticity. A specific research line focuses on applying *Blind Source Separation* (BSS) algorithms, viewing the foreground and the interference as overlapping individual patterns in the document, and relying on multiple observations for their separation. This approach has been adopted in this work for document analysis of degraded documents using the information spread across the multispectral channels.

### Data model

Multispectral acquisitions can be modeled as unknown mixtures of all the patterns that overlap in the document. First the image formation and capture process is modeled, assuming that the document is made of a mixture of distinct *patterns*, through some physical process that we can hypothesize but is actually unknown (Figure III.19). Not all the patterns may be immediately visible or identifiable in the document. An imaging device captures a number of different *channels*. At this point, some of the patterns we are interested in may be visible and distinguishable in some channel, or one may be able to extract them by selecting and composing different channels. Another option is to hypothesize some mutual statistical relationship of the patterns and process all of part of the data cube by some statistical technique. In this way, some of the outputs of the processing can reproduce some of the individual patterns that form the data image. Thus



a suitable number of selective acquisitions can produce a selectivity in the patterns that may be revealed.

In [Tonazzini *et al.* 2004a] a linear instantaneous mixture model for the data is proposed, as well as an independent component analysis strategy to analyze multispectral single-side scans with several overlapping information layers. In the linear model, it is assumed that an  $N$ -channel multispectral scan produces an  $N$ -vector  $x(t)$  at each pixel  $t$ . Similarly, it is assumed that  $M$  source patterns, represented by an  $M$ -vector  $s(t)$ , are superimposed to form the appearance of the document. Since we consider documents containing homogeneous texts or drawings, it can be also reasonably assumed that the reflectance of each undegraded source is almost uniform, and denote by  $a_{ij}$  the mean reflectance index for the  $j$ -th source at the  $i$ -th wavelength. Thus, the individual source functions  $s_i(t)$ ,  $i = 1, 2, \dots, M$ , denote the quantity of the  $M$  patterns that concur to form the colour at point  $t$ . It is unlikely that this simple model is able to account for the true mixing process, especially where two or more patterns overlap and the mixing becomes nonlinear [Sharma 2001]. To keep a linear model, we should at least assume convolutive mixtures with channel-specific kernels:

$$x_i(t) = h_i(t) * \sum_{j=1}^M a_{ij} s_{ij}(t) + n_i(t) \quad (\text{III.13})$$

$$t = 1, 2, \dots, T$$

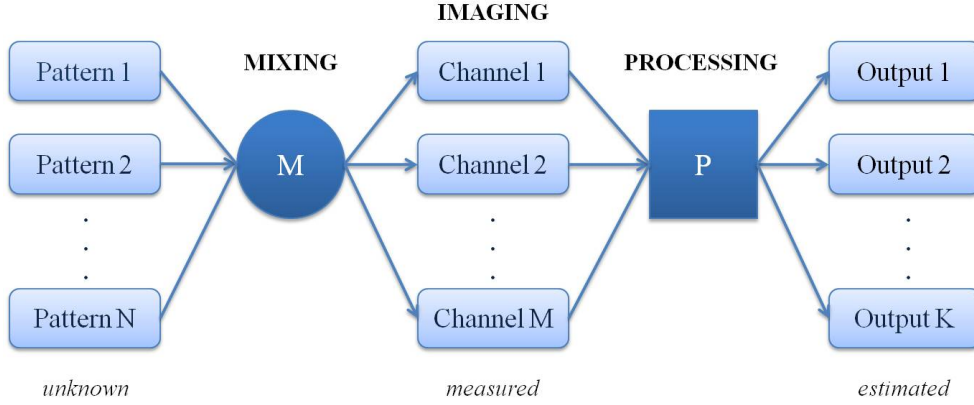
$$i = 1, 2, \dots, N$$

where  $x_i(t)$ , and  $s_i(t)$ , are elements of  $x(t)$  and  $s(t)$ , respectively, and  $n_i(t)$  is the  $i$ -th channel noise at pixel  $t$ . Function  $h_i(t)$  is the blur kernel affecting channel  $i$ , and  $*$  denotes convolution. The  $N \times M$  mixing coefficients  $a_{ij}$  must be estimated along with the functions  $s_i(t)$  to solve the problem. Therefore, this becomes a problem of blind source separation from noisy convolutive mixtures, where the kernels  $h_i(t)$  are often known.

A synthetic form of the Eq. (III.13) is given by:

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) \quad (\text{III.14})$$

When no additional assumption is made, the problem expressed by the Eq. (III.14) is clearly underdetermined, since any nonsingular choice for  $A$  can give an estimate of  $\mathbf{s}(t)$  that accounts for the evidence  $\mathbf{x}(t)$ . Even if no specific information is available, statistical assumptions can often be made on the sources. In particular, it can be assumed that the sources are mutually independent. If this assumption is justified, both  $A$  and  $\mathbf{s}$  can be estimated from  $\mathbf{x}$  (ICA approach [Hyvärinen *et al.* 2001]). If the prior distribution for



**Figure III.19:** Superposition of different (unknown) patterns in multispectral imaging. Channel selectivity from diversity (measured) images can be translated into pattern selectivity. The outputs are estimates of the patterns.

each source is known, independence is equivalent to assume a factorized form for the joint prior distribution of  $\mathbf{s}$ :

$$P(\mathbf{s}(t)) = \prod_{i=1}^N P_i(s_i(t)) \quad \forall t \quad (\text{III.15})$$

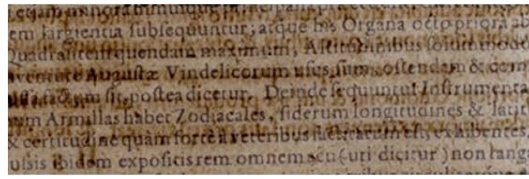
The separation problem can be formulated as the maximization of Eq. (III.15) subject to the constraint  $\mathbf{x} = \mathbf{A}\mathbf{s}$ . This is equivalent to the search for a  $W$ :

$$W = (w_1, w_2, \dots, w_N)^T \quad (\text{III.16})$$

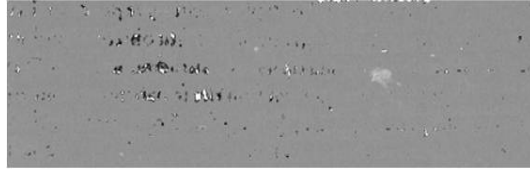
such that, when applied to the data  $\mathbf{x} = (x_1, x_2, \dots, x_N)$ , produces the set of vectors  $\mathbf{w}_i^T \mathbf{x}$  that are maximally independent, and whose distributions are given by the  $P_i$ . By taking the logarithm of Eq. (III.15) the problem solved by ICA algorithms is then:

$$\widehat{W} = \arg \max_W \sum_t \sum_i \log P_i(\mathbf{w}_i^T \mathbf{x}(t) + T \log \|\det(W)\|) \quad (\text{III.17})$$

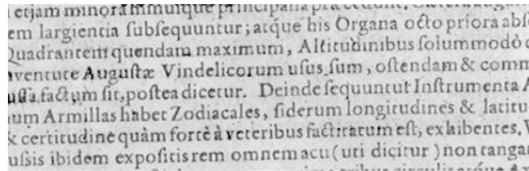
Matrix  $\widehat{W}$  is an estimate of  $A^{-1}$  up to arbitrary scale factors and permutations on the columns. Hence, each vector  $\widehat{\mathbf{s}}_i = \widehat{\mathbf{s}}_i^T \mathbf{x}$  is one of the original source vectors up to a scale factor. Besides independence, a necessary extra condition to make separation possible is that the all sources, except at most one, must be non-Gaussian. To enforce non-Gaussianity, generic super-Gaussian or sub-Gaussian distributions can be used as priors for the sources. These have proven to give very good estimates for the mixing matrix and for the sources as well, regardless of the true source distributions which, on the other hand, are usually unknown. There is no apparent physical reason for our original



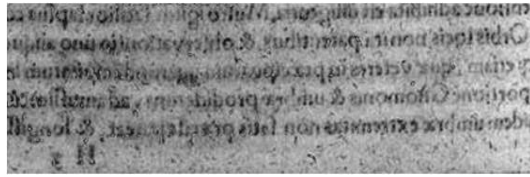
(a) RGB document



(b) background



(c) foreground

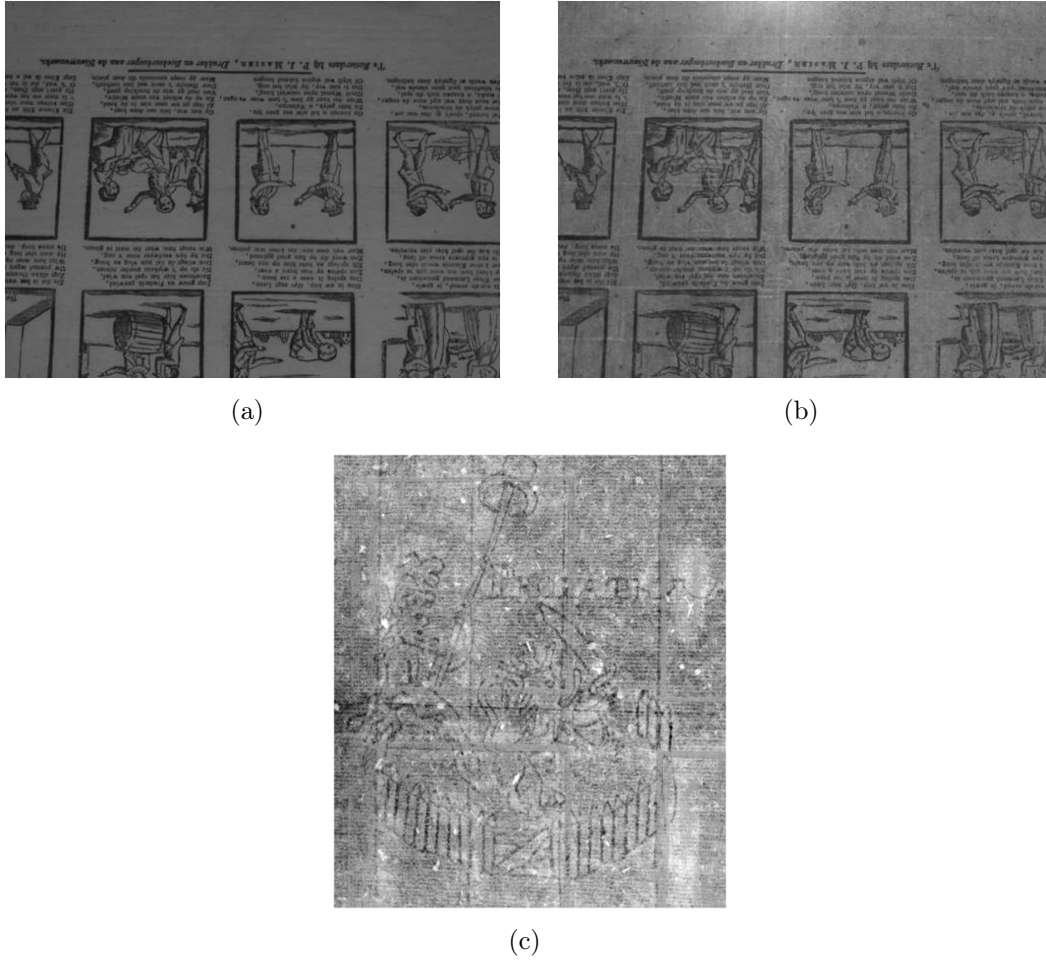


(d) bleed-through

**Figure III.20:** Recovering of background (b), foreground (c), and bleed-through (d) from an RGB document by ICA algorithm.

sources to be mutually independent, so, even if the data model (Eq. (III.14)) was correct, the ICA principle might not be able to separate the different classes. However, it is intuitively clear that one can try to maximize the information content in each component of the data vector by decorrelating the observed image channels. To avoid cumbersome notation, and without loss of generality, it is assumed to have zero-mean data vectors. Seeking for a linear transformation  $\mathbf{y}(t) = W\mathbf{x}(t)$  such that  $\langle y_i y_j \rangle = 0, \forall i, j = 1, \dots, M, i \neq j$ , where  $W$  is generally an  $M \times N$  matrix and the notation  $\langle \cdot \rangle$  means expectation. In other words, the components of the transformed data vector  $\mathbf{y}$  are orthogonal. It is clear that this operation is not unique since that, given an orthonormal basis of a subspace, any rigid rotation of it still yields an orthonormal basis of the same subspace.

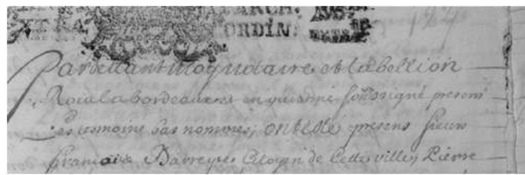
It is well known that linear data processing can help to restore colour text



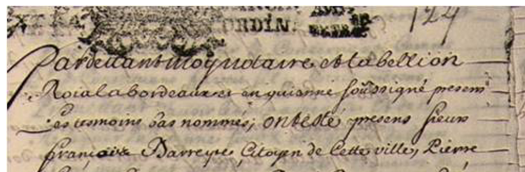
**Figure III.21:** Watermark detection: (a) infrared front view; (b) back illumination infrared view; (c) one ICA output.

images, although the linear model is not fully justified. In [Ohta *et al.* 1980] the authors compare the effect of many fixed linear colour transformations on the performance of a recursive segmentation algorithm. They argue that the linear transformation that results in maximum-variance components is the most effective. From this they derive a fixed transformation that, for a large class of images, approximates the Karhunen-Loeve transformation, which is known to give orthogonal output vectors: one of these components has the maximum variance. This approach is also called principal component analysis (PCA), and one of its purposes is to find the most useful among a number of variables [Cichocki & Amari 2002]. The data covariance matrix is the  $N \times N$  matrix:

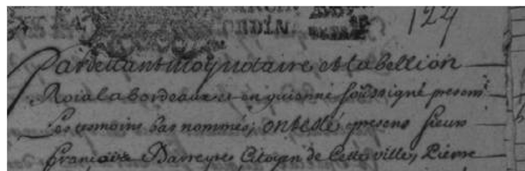
$$R_{xx} = \langle \mathbf{x}\mathbf{x}^T \rangle \approx \frac{1}{T} \sum_{t=1}^T \mathbf{x}(t)\mathbf{x}^T(t) \quad (\text{III.18})$$



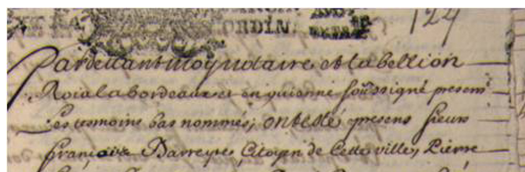
(a) IR



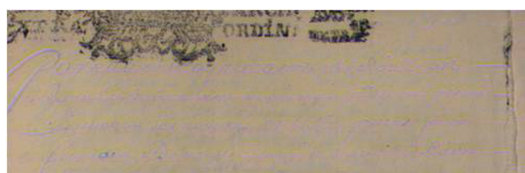
(b) RGB



(c) UV



(d) Attenuation of bleed-through



(e) Extraction of a symbol

**Figure III.22:** The document images of Figure III.18 processed by SO.

Since the data are normally correlated, matrix  $R_{xx}$  will be non-diagonal. The covariance matrix of vector  $\mathbf{y}$  is:

$$R_{yy} = \langle W \mathbf{x} \mathbf{x}^T W^T \rangle = W R_{xx} W^T \quad (\text{III.19})$$

To obtain orthogonal  $\mathbf{y}$ ,  $R_{yy}$  should be diagonal. Let perform the eigenvalue decomposition of matrix  $R_{xx}$  and call  $V_x$  the matrix of the eigenvectors of  $R_{xx}$ , and  $\Lambda_x$  the diagonal matrix of its eigenvalues, in decreasing order. Now, it is



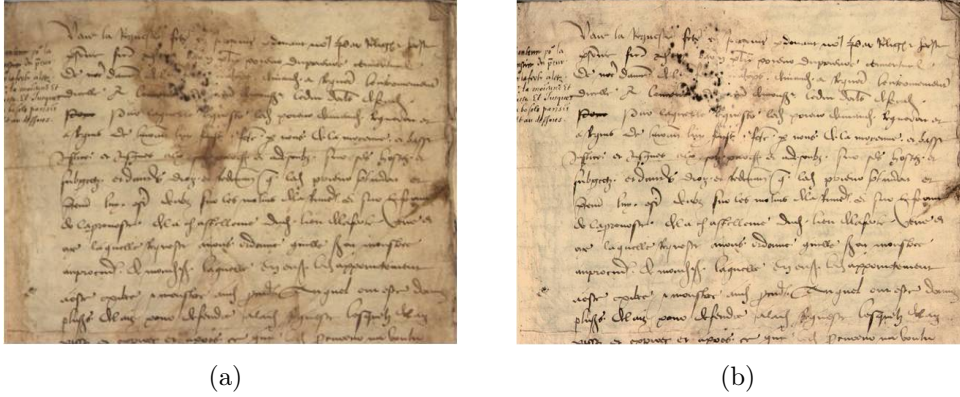


Figure III.23: Stain attenuation by PCA.

easy to verify that all of the following choices for  $W$  yield a diagonal  $R_{yy}$ :

$$W_{PCA} = V_x^T \tag{III.20}$$

$$W_{SO} = V_x \Lambda_x^{-\frac{1}{2}} V_x^T \tag{III.21}$$

Matrix  $W_{PCA}$  produces a set of vectors  $y_i(t)$  that are orthogonal to each other and whose Euclidean norms are equal to the eigenvalues of the data covariance matrix. This is what PCA does [Cichocki & Amari 2002]. By using matrix  $W_{SO}$  the Symmetric Orthogonalization (SO) approach is adopted [Tonazzini *et al.* 2004b]. In [Cichocki & Amari 2002] it is observed that the application of matrix  $W_{SO}$  is equivalent to ICA when matrix  $A$  is symmetric. In general, ICA applies a further rotation to the output vectors, based on higher-order statistics.

The ICA, PCA, and SO algorithms are tested to process multispectral images of ancient documents, preliminarily registered. So,  $\widehat{W}$ ,  $W_{PCA}$ ,  $W_{SO}$  are computed to obtain the “estimated” patterns. The aim is to emphasize the hidden features of the document, attenuate the degradations and improve the document legibility. For each study case are showed the best results obtained with one of the three techniques. Moreover, it has to be noted that the number of outputs computed by decorrelation is the same of the number of inputs (multispectral images), and the selection (subjective) of output/s where particular patterns are recovered is conducted by the human operator.

An simple case of decorrelation by ICA is illustrated in Figure III.20. The outputs obtained from an RGB document are: the background, the foreground (main text) and the bleed-through. In Figure III.21, it is reported another example where a paper watermark pattern is detected and extracted. In this case, the document is constituted of only two classes: the foreground pattern, with drawings and text, and the background pattern with the watermark, so

that only two views are needed. Two infrared acquisitions are used: the first taken under front illumination, the second taken under back illumination. In this case a good extraction is achieved by using all the three methods proposed. The SO has been applied to process the multispectral images shown in Figure III.18, obtaining the attenuation of bleed-through and the extraction of a symbol (Figure III.22). Another good result from PCA is in the attenuation of stains: in Figure III.23 the RGB images are shown. Note that this document is much more legible after decorrelation.

These experiments demonstrate the validity of blind source separation techniques for enhancing and separating the various features that appear overlapped in several ancient documents. The multispectral acquisition allows to extract hidden information. No conclusions can be instead drawn about the superiority of one method over the others for all documents. It is possible only to say that, when the main goal is to enhance partially hidden features, at least one of the three methods proposed always succeeded in reaching the goal in all experiments, in a simple and fast way.

### III.2.3 Recto-verso case

This section deals with the challenging problem of the restoration of recto-verso scans, processed by registration and decorrelation techniques [Bianco *et al.* 2009, Tonazzini *et al.* 2009]. An extension of the approach described in the [Tonazzini *et al.* 2007] to separate the recto and verso patterns is discussed. The goal is twofold: to produce enhanced versions of all the available scans at different wavelengths, and to create a restored visible document that, while cleansed of the unwanted interferences, maintains its useful features as much as possible. An advantage of colour or multispectral imaging over grayscale imaging is that, once two separate recto and verso images are obtained, each of them can be further analyzed, e.g., by the methods proposed in [Tonazzini *et al.* 2004a].

An example of recto-verso capture is shown in Figure III.24. Due to seeping of ink (bleed-through), features like stamp and text are visible in the recto and verso images.

#### III.2.3.1 Double-sided document registration

Recto-verso images are misaligned because of the document repositioning. So, the Fourier-Mellin method is also used to register recto-verso images. All the images from one of the sides must be flipped horizontally before registration. Registering recto-verso pairs is a very challenging task, since the recto and verso sides are different not only with respect to image intensity, but also

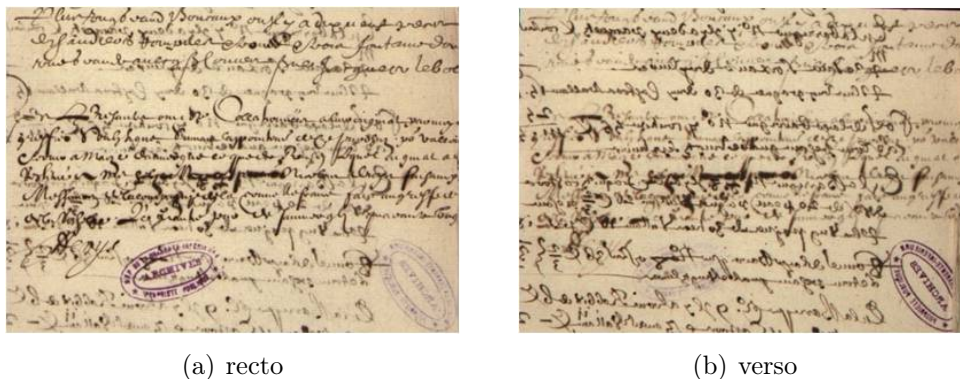


Figure III.24: Recto-verso captures.

topographically. Moreover, the common features on which registration can be based are often very sparse: a correspondence between pixels can only be identified in the restricted regions where a pattern present in one side also appears in the opposite side.

A complete set of registration parameters can thus be very hard to compute. Good results are obtained with the Fourier-Mellin transform and parameter optimization [Bianco *et al.* 2008, Dubois & Pathak 2001]. By applying one of the selected methods, each image to be registered is linked to a set of transformation coefficients (affine transformation). Once all the transformations are computed, the images must be cropped to let them represent the same area and have the same size. As a quality measure for registration of a pair of images, we adopted the normalized mean square error [Fienup 1997] on the correlation coefficient of the transformed pair, defined as follows:

$$E^2 = 1 - \frac{\max_{u,v} \|r_{fg}(u,v)\|^2}{\sum_{x,y} \|f(x,y)\|^2 \sum_{x,y} \|g(x,y)\|^2} \quad (\text{III.22})$$

where,  $f(x,y)$  and  $g(x,y)$  are the registered images, and  $r_{fg}(u,v)$  is their cross-correlation function. For any fixed pair of images, this quality index can be used to evaluate the relative merits of different registration methods. The Fourier-Mellin and the parameter optimization methods have shown perfectly comparable performances in registering recto-verso pairs, which is the most challenging task for our purposes. The results are normally suitable to being processed through the analysis techniques of decorrelation.

As examples, the errors evaluated to register the images shown in Figure III.24 have been 0.178 for Fourier-Mellin and 0.149 for parameter optimization, while for the images in Figure III.25, the values are 0.294 and 0.301, respectively. The registered images of the document shown in Figure III.24, are reported in Figure III.26, overlaid in transparency.



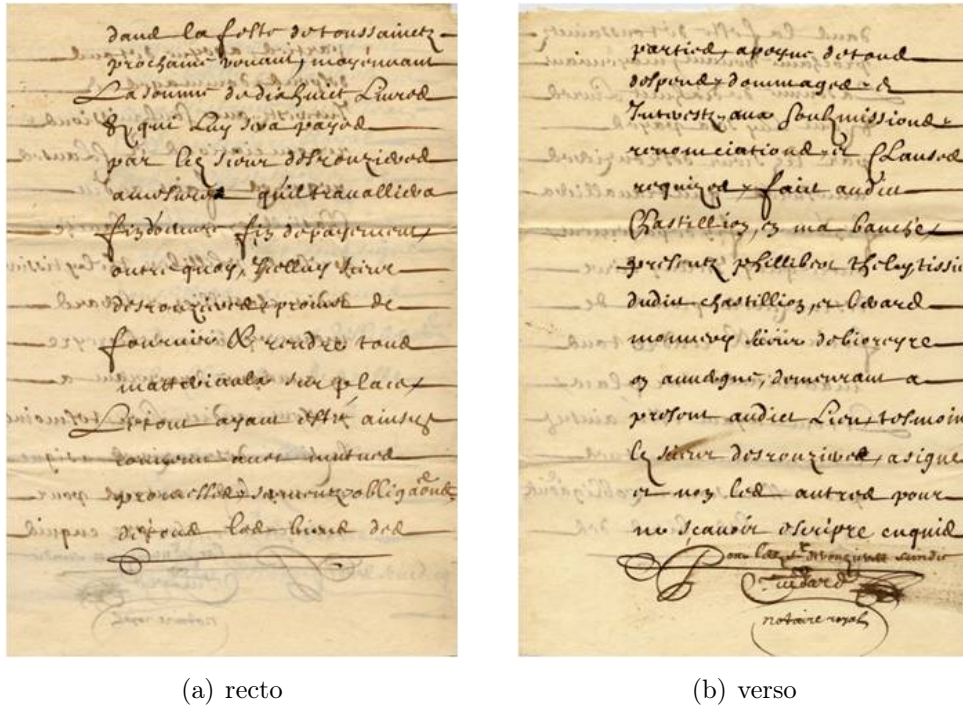
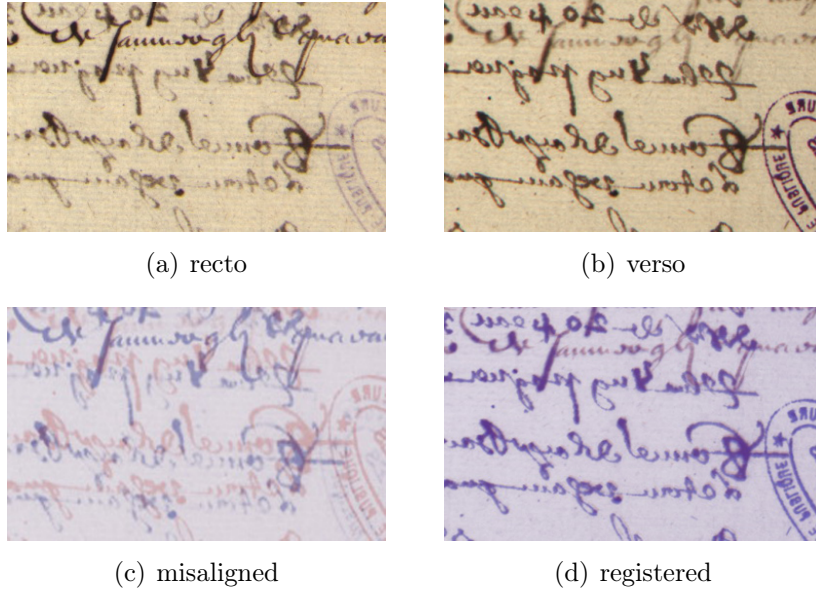


Figure III.25: Recto-verso captures.

### III.2.3.2 Recto-verso decorrelation

The reduction of the bleed-through interference in registered recto-verso grayscale scans is analysed by blind source separation techniques. The data model used in Section III.2.2.4 is now rewritten for the specific recto-verso case. Let us assume that the Fourier-Mellin method has provided us with a set of  $N$  registered views of a recto-verso pair. This assumption includes the grayscale case, with  $N = 1$  (a single pair is available), and any colour or multispectral data set. In [Tonazzini *et al.* 2007] it is demonstrated how in the grayscale case the interfering patterns can be reduced by a second-order blind separation approach. This strategy was motivated by observing that the appearances of the recto and verso sides are normally more correlated than the pure recto and verso patterns. Thus, if the two data images are orthogonalized, the interference of each side in the other should be reduced. Here, a generalization of that approach is introduced, where the data are a multispectral collection of recto-verso pairs, and whose outputs can further be processed and analyzed. In fact, once the two pure recto and verso patterns have been separated, we still have two independent data cubes.

First of all, if these inputs contain the red, green and blue components, then the visible colours of the original document can be reconstructed. Sec-



**Figure III.26:** Recto-verso registration: the misaligned and registered images are shown as overlaid in transparency.

ond, if one of or both the output data cubes can be further thought of as superpositions of uncorrelated patterns, these latter can still be separated. Both these possibilities will be exemplified. To build a model for the data, only two distinct patterns are considered, one in the recto-side and one in the verso-side of the page. Thus, it is admitted that the appearances of the two sides are given, for each channel observation, by  $2 \times 2$  linear mixtures of these patterns:

$$r^k(t) = A_{11}^k s_1^k(t) + A_{12}^k s_2^k(t), \quad t = 1, 2, \dots, T \quad (\text{III.23})$$

$$v^k(t) = A_{21}^k s_1^k(t) + A_{22}^k s_2^k(t), \quad k = 1, 2, \dots, N \quad (\text{III.24})$$

where  $r^k(t)$  and  $v^k(t)$  are the recto and verso appearances, respectively, at the  $k$ -th channel,  $s_1^k(t)$  and  $s_2^k(t)$  are the reflectance maps associated to the  $k$ -th recto and verso pattern components (or sources), respectively,  $t$  is a pixel index, and  $A_{ij}^k$  are unknown mixing coefficients. Physically,  $A_{12}^k/A_{11}^k$  and  $A_{21}^k/A_{22}^k$  represent the verso-to-recto and recto-to-verso attenuations, respectively. These depend on the properties of the transmission medium and other factors, such as ink fading. By grouping all the Eqs. III.23,24 we get a

$2N \times 2N$  block-diagonal linear system with unknown matrix:

$$\begin{bmatrix} r^1(t) \\ v^1(t) \\ \vdots \\ r^N(t) \\ v^N(t) \end{bmatrix} = \begin{bmatrix} A^1 & 0 & \cdots & 0 \\ 0 & A^2 & \cdots & 0 \\ \vdots & \cdots & \ddots & \cdots \\ 0 & \cdots & \cdots & A^N \end{bmatrix} \cdot \begin{bmatrix} s_1^1(t) \\ s_2^1(t) \\ \vdots \\ s_1^N(t) \\ s_2^N(t) \end{bmatrix} \quad (\text{III.25})$$

where the structure of the diagonal blocks is clear from Eqs. III.23,24. If the data vector is denoted by  $x(t)$ , the source vector by  $s(t)$ , and the system matrix (or mixing matrix) by  $A$ , then Eq. III.25 can be written in vector form, as shown the Eq. III.14:

$$\mathbf{x}(t) = A\mathbf{s}(t) \quad (\text{III.26})$$

Reasonable properties of the sources are that, for any possible pair of channels  $(k, h)$  the recto and verso components  $s_1^k(t)$  and  $s_2^h(t)$  are almost uncorrelated, whereas the recto sources  $s_1^k(t)$  and  $s_1^h(t)$  are strongly correlated since they are just different frequency components of the same reflectance function. The same is true for any pair of verso sources,  $s_2^k(t)$  and  $s_2^h(t)$ . In other words, the source covariance matrix has the form

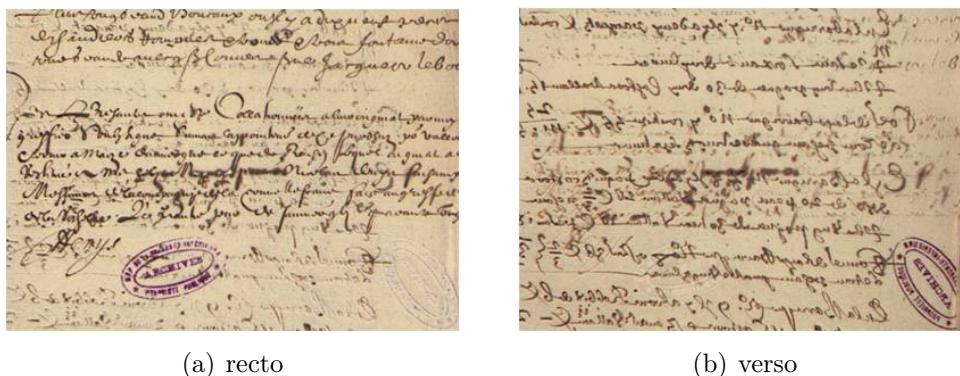
$$R_s = \begin{bmatrix} I & \Sigma_{12} & \Sigma_{13} & \cdots & \Sigma_{1N} \\ \Sigma_{12} & I & \Sigma_{23} & \cdots & \\ \Sigma_{13} & \Sigma_{23} & I & \cdots & \\ & & & \cdots & \\ & & & & I \end{bmatrix} \quad (\text{III.27})$$

Having assumed, without loss of generality, that all the sources have unit variance, matrix  $I$  is the  $2 \times 2$  identity, and

$$\Sigma_{kh} = \begin{bmatrix} \sigma_1^{kh} & 0 \\ 0 & \sigma_2^{kh} \end{bmatrix} \quad (\text{III.28})$$

$\sigma_1^{kh}$  and  $\sigma_2^{kh}$  being the covariances of the  $k - th$  and  $h - th$  channels of the recto and verso sources, respectively. The idea behind the solving system (Eq. III.25) by decorrelating the recto and verso appearances boils down to find a linear transformation  $W_{rv}$  on  $\mathbf{x}$  such that the covariance matrix of the transformed vector has the same structure as the matrix in Eq. III.27. With  $N = 1$ , the problem becomes to diagonalize the  $2 \times 2$  covariance matrix  $R_x$  of the recto and verso grayscale appearances. In a general  $2N \times 2N$  case, however, trying to diagonalize  $R_x$  would be a mistake, since, as shown in Eqs. II.27,28, the source covariance matrix is not diagonal.

In other words, this would force zero correlations between strongly correlated pairs. One way to solve this separation problem in this case is to



**Figure III.27:** Restoration of a pair of RGB recto and verso sides of a real document page.

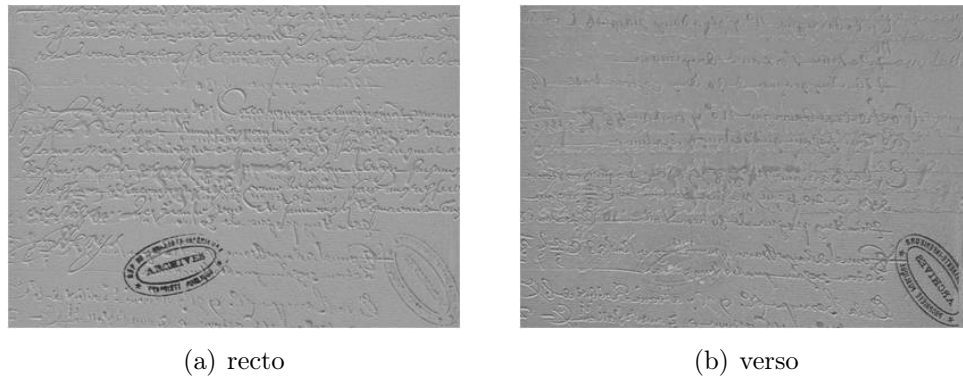
observe, from Eq. III.25, that the mixing model is separable, and thus it can be solved channel by channel as done in the grayscale case by diagonalizing separately the  $N$  channel covariance matrices of size  $2 \times 2$  estimated as

$$R_x^k = \langle (x^k)(x^k)^* \rangle \approx \frac{1}{T} \sum_{t=1}^T [x^k(t)][x^k(t)]^* \quad (\text{III.29})$$

where the superscript  $k$  denotes the channel and the asterisk means transposition. To choose among all the possible diagonalizing matrices, it is assumed that all the diagonal blocks of matrix  $A$  are symmetric [Tonazzini *et al.* 2007]. Indeed, if, as is reasonable to assume that  $A_{11}^k = A_{22}^k$  and  $A_{12}^k/A_{11}^k = A_{21}^k = A_{22}^k$ , then we also have  $A_{12}^k = A_{21}^k$  for each  $k$ . In this case [Cichocki & Amari 2002], applying a symmetric diagonalizing matrix can be proved to be equivalent to perform an independent component analysis separation. Such a matrix can be derived from the following formula:

$$W_s^k = (V_x^k)(\Lambda_x^k)^{-\frac{1}{2}}(V_x^k) \quad (\text{III.30})$$

where  $V_s^k$  is the matrix of the eigenvectors of  $R_x^k$ ,  $\Lambda_x^k$  is the related eigenvalue matrix. By premultiplying all the available recto-verso pairs by matrices  $W_s^k$ , we get a new set of images, whose covariance matrix has the same diagonal blocks as matrix  $R_s$ . The values in the off-diagonal blocks (Eq. III.28) are completely disregarded, which are left completely free, whereas all the cross-covariances of recto and verso images should be zero. At present, studies are trying to take this information into account by also relying on the structure of the mixing matrix. Ideally, the values of the cross-covariances  $\sigma_1^{kh}$  and  $\sigma_2^{kh}$  should also be estimated. As mentioned, when the pages are captured by RGB sensors, the original document colors can then be reconstructed by



**Figure III.28:** Application of symmetric orthogonalization to extract a stamp in the RGB documents shown in Figure III.27.

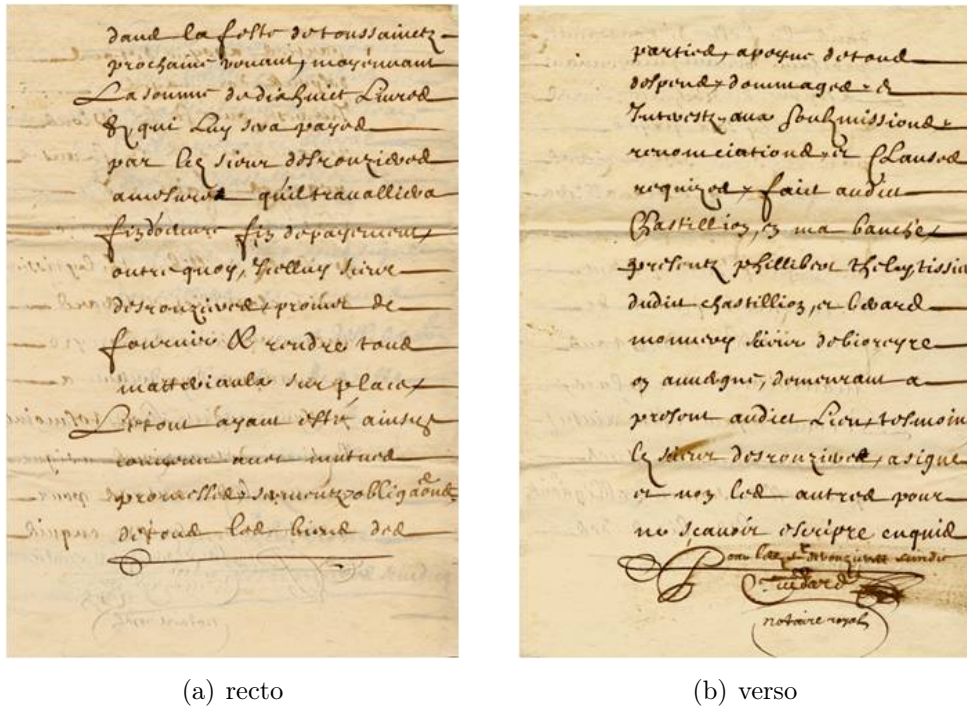
recomposing the color channels. This fact has been exploited to produce restored visible documents that, while cleansed of the unwanted interferences, maintain their original appearances as much as possible.

Experimentation on real degraded documents has been conducted to test the data model described above for the recto-verso scans. Figures III.27,29 show the RGB reconstructions of the recto and verso sides of two manuscripts affected by a strong bleed-through. Since the manuscripts were acquired in RGB, the registration required the alignment of six maps per document. In both cases, a significant attenuation of the bleed-through has been obtained, and the original color has been pretty well recovered. The RGB recto and verso images thus obtained can further be analyzed to extract possible extra uncorrelated patterns. In Figure III.28, the results obtained through symmetric decorrelation applied to the RGB image of Figure III.27 are shown. As is apparent, the stamp pattern and the main manuscript text have been separated very well.

### III.2.4 Image segmentation

The image segmentation is the last step of the enhancement methodology that is proposed in this work. Image segmentation techniques are usually applied to grayscale images, and aim at individuating a threshold gray value, which allows to separate the textured background from the foreground text, i.e. to isolate the written words from the artifacts. The output of segmentation is then a binary image, where text pixels and background pixels are labeled. The segmentation techniques can be of two types: global and local. The global techniques are applied to the whole image without distinction, while the local ones apply different types of thresholding, according to the various regions





**Figure III.29:** Restoration of the pair of RGB recto and verso sides showed in Figure III.28.

of the image. In [Sezgin & Sankur 2004], a performance analysis on several thresholding algorithms is reported. A specific test related to the separation of the text from the background was carried out using 40 documents with different fonts, sizes and typefaces. The document degradations were simulated by blur and speckle noise, as suggested by the degradation models proposed in [Baird 1992]. Among the different local-based methods examined, the one that showed the best performance for document binarization was the method proposed in [Sauvola & Pietaksinen 2000] by Sauvola and Pietaksinen.

In fact, this binarization algorithm was tested with the benchmarking technique and various scenarios against several known binarization techniques in literature [Niblack 1986, Eikvil *et al.* 1991, Bernsen 1986, Parker 1991]. Using the environment factors (such as different degradations) and available document and test image databases, the algorithm results were evaluated and benchmarked first against each other, then against the ground-truth knowledge by visual and benchmark event(s) evaluation processes. The focus was set on documents with textual content and on multi-content documents, i.e. documents having text, graphics, linedrawings and halftones. The test images were selected from a special database of document image categories, comprising over 1000 categorized document images (e.g. article, letter, memo, fax,

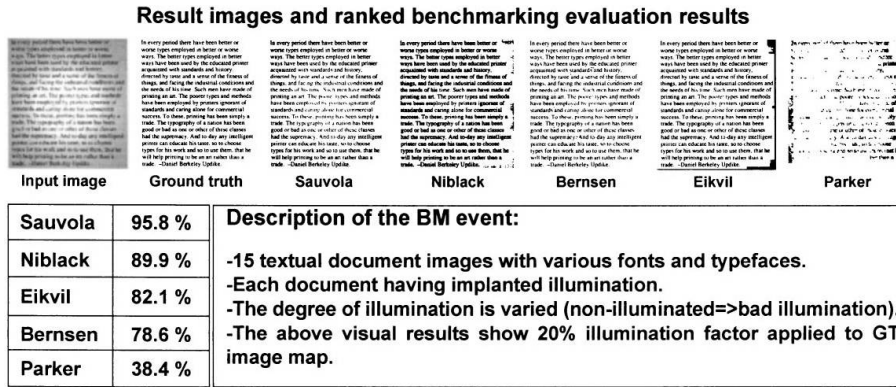


Figure III.30: Visual and numeric results on the comparison algorithms applied to illuminated, textual images [Sauvola & Pietaksinen 2000].

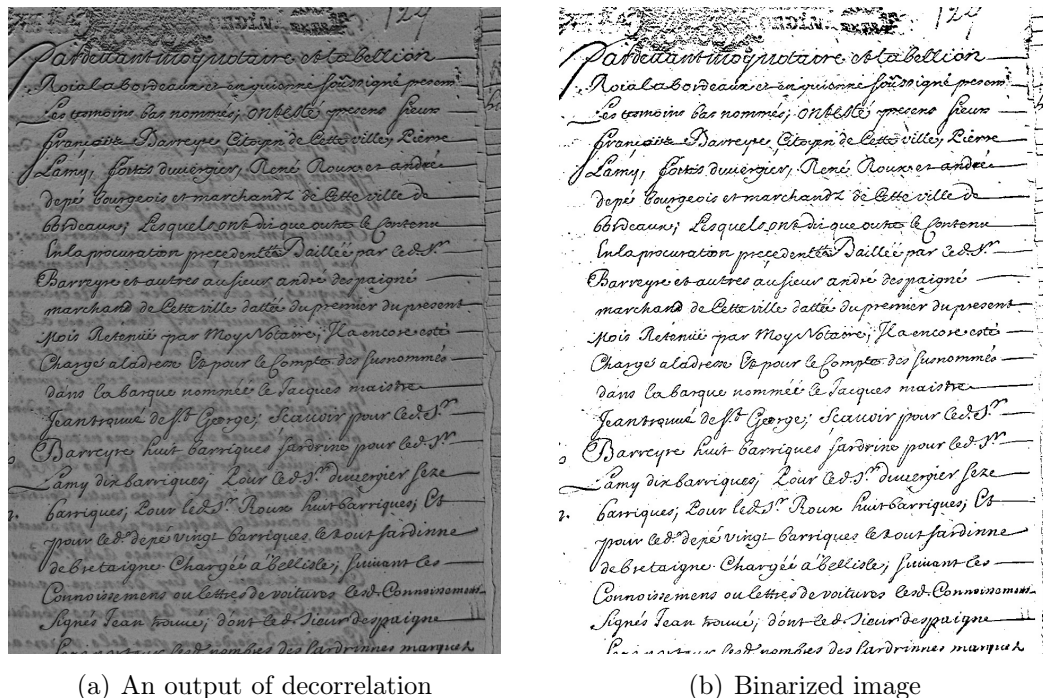
journal, scientific, map, advertisement, etc.) [Sauvola *et al.* 1997]. The numerical test and results presented were obtained using binarization metrics emphasizing the performance in textual image region binarization.

Figure III.30 presents an example benchmarking scene performed on a database of 15 textual document images having illumination. Visual results to a sample input image having 20% of centered illumination defect, an example of a ground-truth image map and the results of the proposed and comparison binarization algorithms. The results show good behaviour of Sauvola’s, Niblack’s and Eikvil’s algorithms, when the limit is set to 80% performance, i.e. the limit where the OCR performance drop is less than 10% using Caere Omnipage OCR package [OCR 1997]. Bernsen suffered of noise that was introduced to binarized result image, while the Eikvil’s threshold ruled some of the darkest areas belong to object pixels. Parker’s algorithm adapted poorly to even small changes in lumination, but had sufficient results with relatively “clean” grey-scale document images.

In the Sauvola algorithm, a threshold  $t(x, y)$  is computed for each pixel, which depends on two quantities,  $m(x, y)$  and  $s(x, y)$ , representing the local mean and standard deviation, respectively, calculated on a square window 10-20 pixels wide:

$$t(x, y) = m(x, y) \left[ 1 + k \left( \frac{s(x, y)}{R} - 1 \right) \right] \tag{III.31}$$

where  $R$  is the maximum value of the standard deviation ( $R = 128$  for a grayscale document), and  $k$  is a parameter that takes positive values in the range  $[0.2, 0.5]$ . Using this formula, the threshold is calculated according to the contrast in the area surrounding each pixel. When the contrast in a small neighborhood of the pixel is high,  $s(x, y) \approx R$  and  $t(x, y) \approx m(x, y)$ . Conversely, when the contrast is low,  $t(x, y)$  becomes smaller than the mean value,



**Figure III.31:** Segmentation with Sauvola algorithm of a decorrelated image.

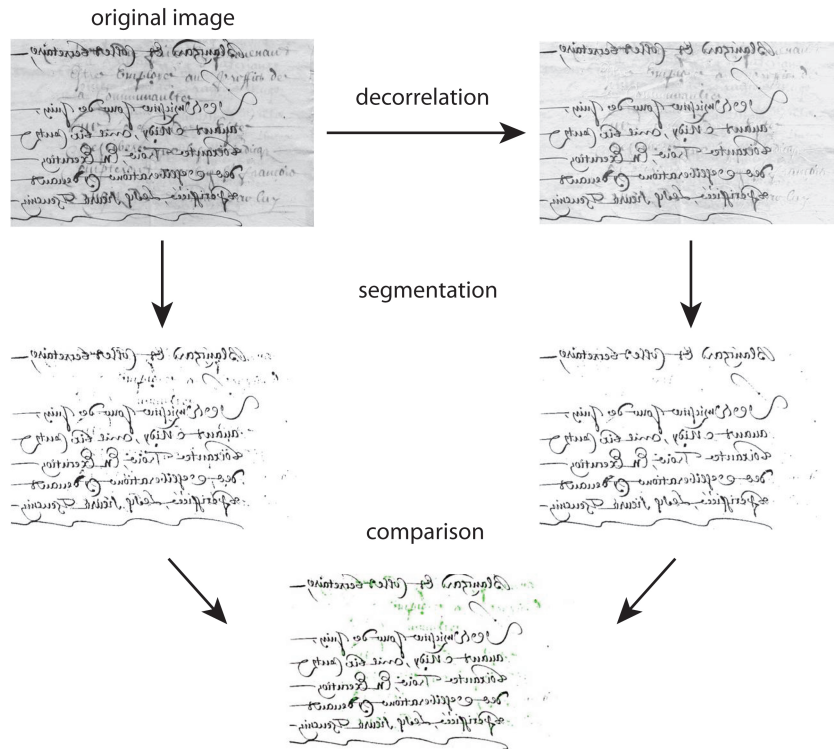
thus allowing the relatively dark regions in the background to be removed successfully. Parameter  $k$  controls the value of the threshold in the local window so that the higher the value of  $k$ , the smaller is the threshold. A value of  $k = 0.5$  is suggested in [Sauvola & Pietaksinen 2000], but some experiments reported in [Badekas & Papamarkos 2005] put in evidence that  $k = 0.34$  gives better results. Anyway, the algorithm is not very sensitive to the value of  $k$  used.

Then choosing a decorrelated image this is processed with Sauvola algorithm. The output of the orthogonalization, for example, is the input to the binarization algorithm in order to segment the main text (Figure III.31). In Figure III.32 an output of the decorrelation is binarized and compared with the output of the Sauvola algorithm as applied directly to the original image. It is apparent that the effectiveness of the decorrelation technique improves the segmentation process.

### III.3 Deblurring

A deblurring technique based on non-blind deconvolution is described: it may be used to improve the image quality of any multispectral channel affected by blur. In fact, any image can often present a blur component due to an





**Figure III.32:** The decorrelation improves the segmentation process.

incorrect focus setting of lens. This problem may occur especially in the infrared acquisition, because the main text may not be visible and the contrast image decreases. In this work, it is proposed to apply the deblurring technique after registration. First, a review of the deblurring techniques suitable for the improvement of the virtual restoration of ancient documents is provided and then a discussion about a possible application of this technique is presented.

### III.3.1 Related works

Deblurring techniques, often called deconvolution methods, aim to improve the image quality in terms of removal of image blur and noise reduction. The amount of a priori information related to the degradation, such as the size or shape of the blurring functions and noise parameters, affects significantly the success of deconvolution. When the blur function is known (non-blind deconvolution), many conventional approaches have been developed to recover the sharp image (see a review in [Banham & Katsaggelos 1997]). A more challenging problem occurs when the blur is unknown (blind image deconvolution). A state-of-the-art survey of blind restoration techniques is given in [Campisi & Egiazarian 2007]. Most of the methods are iterative and

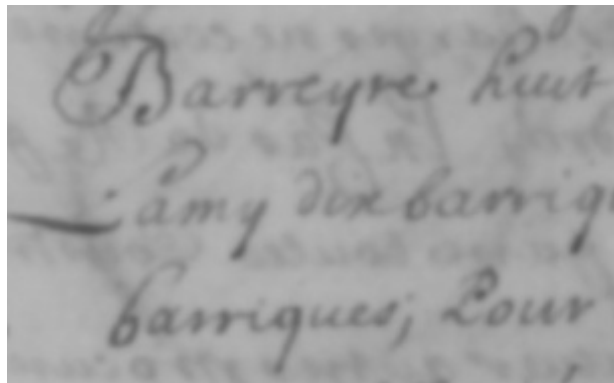
minimize specific functionals. They involve regularization terms based on available prior information (i.e. image characteristics which can be expected) that assures various statistical properties of the image and constrains the estimated image and/or blurs. Regularization is required to improve stability. For images with sharp changes of intensity (e.g. characters printed on paper), an appropriate regularization is based on total variation, as first introduced by Rudin [Rudin *et al.* 1992]. Minimizing such functionals preserves edges and fine details in the image [Chan & Wong 1998, Chan & Wong 2000, You & Kaveh 1999]. Since the blind case is strongly ill-posed, all the methods suffer from convergence and stability problems.

If the images are smooth and homogeneous, an autoregressive model can be used to describe the measuring process. The autoregressive model simplifies the blind problem by reducing the number of unknowns. Several techniques have been proposed to find a solution to the blind case [Lagendijk *et al.* 1990, Reeves & Mersereau 1992, Haindl 2000]. One way to overcome the difficulties of single-image blind deconvolution is to use multiple acquisitions of the same scene and apply the so-called multiframe blind deconvolution (MBD) techniques. A typical example of multiframe acquisition with a single camera is the capture of several pictures of the same object with slightly different focus settings. The lack of information at one frequency on a single image can be substituted by the information at the same frequency from other images.

The blind restoration problem is thus simplified by the availability of different images. Research on intrinsic multiframe methods has been started by Harikumar [Harikumar & Bresler 1999] and Giannakis [Giannakis & Heath 2000]. Such MBD methods overpass the limitations of previous techniques and can recover the blurring functions just from the degraded images. The MBD theory was further developed by Šroubek [Šroubek & Flusser 2005], who proposes a noise-robust method applicable to images misaligned by unknown shifts. Šorel [Šorel *et al.* 2008] recently considered the very challenging problem of shift-variant blind deconvolution. The above methods have been applied to images from the same band, with different sensor settings. This section deal with the issue of multiple images, obtained using several acquisition bands, as in multispectral acquisition.

### III.3.2 Discussion

After the registration, the insufficient sharpness of the multispectral acquired data has to be compensated. Figure III.33 shows a typical example of what can happen if the camera focus is not properly set. The acquired data are then out-of-focus, blurred, and with no sharp details. The wrong focus setting is



**Figure III.33:** Detail of a document image captured in IR band: the image is blurred.

often present due to the manipulation during the data acquisition, especially in the infrared channel.

This type of image degradation can be modeled by convolution with a certain unknown PSF (Point Spread Function). This PSF, which characterizes the image blur degradation, can be different for distinct acquisition modalities. The blurring itself and the variations of the PSF shapes for the acquired multispectral channels make further processing complicated, because the next document analysis steps suppose either no blurring or at least the same PSF for all channels. These reasons led to introduce a deblurring step into the proposed methodology (Section III.2.2). Blind deconvolution is a mathematically ill-posed problem that, however, can be regularized either if more images of the same scene are available (multiframe blind deconvolution) or by using artificial markers in the data acquisition process for PSF estimation. The latter approach is based on the fact that a blurred version of the unit pulse (Dirac delta) represents theoretically the shape of the respective PSF. Although the unit pulse is not exactly achievable, the detected shape can be used as a very good first estimate of the PSF for its further iterative improvement, performed through the deconvolution method on chosen single images. Figure III.34 depicts the acquired markers with three different radii.

In theory, a smaller marker means a better representation of the PSF. However, in severely blurred images, tiny markers are hard to detect due to the limited sensitivities of the cameras, and markers cropped from the image tends to be underestimated. Still it is possible to apply iterative blind deconvolution algorithms, such as in [Campisi & Egiazarian 2007], and use the cropped marker as an initial estimation of the PSF. One option to improve the estimated PSF is to use bigger markers. If the shape of the marker is known and also the corresponding blurred version from the image, it is possible to apply simple algorithms for PSF estimation [Šorel & Šroubek 2009].



**Figure III.34:** Artificial markers with three different radii introduced in an image. The out-of-focus blurring is apparent.

With the estimated PSF, one can apply standard non-blind deconvolution algorithms [Campisi & Egiazarian 2007] with sophisticated edge-preserving regularization [Rudin *et al.* 1992], which is especially good for images of printed text. Another option to improve the PSF estimation, and thus the whole deconvolution process, is to use multiple acquisitions of the same spectral band, but with slightly different focus settings of the camera. Then one can apply MBD methods such as the ones proposed in [Šroubek & Flusser 2005].

Multichannel blind deconvolution estimates the original (sharper) image and the PSF from the blur convolution model:

$$Image_{blurred} = convolution(Image_{original}, PSF) + noise.$$

The method is based on the processing of several input images (multichannel) and it works blindly, e.g. it does not require strong a priori information about the present PSF. The resulting restored image is sharper, with less noise and blur.

To improve the convergence speed of MBD, the cropped markers from the blurred images can be used as initial guesses of the PSFs. The disadvantage of this approach is the necessity to have at least two images in each spectral band, which slows down even further the tedious manual acquisition process. A promising method, which however has still to be fully tested, would use multispectral images as individual channels in a type of semi-MBD approach. Each spectral image would be considered as a single image blind deconvolution problem with cropped markers as the initial PSFs. However, since the spectral bands are heavily correlated (e.g. edges are shared), regularization terms based on total variation have many features in common, and they can be formulated for multispectral images. A similar idea was proposed for blind deconvolution of multimodal medical images in [Šroubek *et al.* 2009]. It is possible to believe that, if properly modified, this technique can be applied to

multispectral document images as well, and could represent a new direction for image deblurring.

## III.4 Conclusions

Ancient document analysis through multispectral imaging has been discussed in this chapter. The analysis of degraded documents is based on the multispectral acquisition performed with a filter-wheel camera. This imaging allows to acquire additional information respect to the conventional RGB capture. A common limitation of such devices is the absence of automatic focusing. Moreover, filter shifting implies a continuous adjustment of the lens during imaging operations, due to the variation of subject focus.

Then a procedure to guide an user in the application of image processing techniques to improve the readability of ancient documents is presented. This procedure includes registration, enhancement and segmentation. These techniques have been applied in both cases of single-sided and double-sided. The experimentations reveal an effective application of the Fourier-Mellin transform for the difficult problem of registering recto-verso scans of ancient documents. Decorrelation techniques based on the blind source approach allow to recover and separate patterns present in a document, like the foreground, to extract symbols and to attenuate the degradations. It has to be note that these technique can improve the result of the binarization process, reducing the presence of interferences.

The basic problem with this kind of approaches to interference reduction is the adequacy of the data models adopted. Indeed, bleed-through and show-through are very complex phenomena and simple linear instantaneous models are not sufficient to account for all the relevant features of the degraded documents. Nevertheless, the simple approach proposed proves to be often useful to mitigate the interferences. Before trying to introduce nonlinearities in the data model, a linear convolutional model is intersting to apply that should be able to account for ink spreading in bleed-through and for the blurring of the show-through pattern due to light diffusion effects within the paper.

Finally, the multichannel blind deconvolution to deblur images acquired in different spectral bands is proposed in order to allow a virtual restoration of ancient documents to be performed. Image sharpness is necessary to further image analysis tasks. The overall procedure could constitute a fast, reliable and effective system to be routinely used in libraries and archives for the enhancement of multispectral scans of degraded documents. The methodology is flexible for use in various contexts of analysis of the document contents, such as the recovery and extraction of hidden or masked patterns.



# Combination of multispectral and 3D imaging

---

*Geometrical alterations of the shape can affect an historical document, altering its original state and reducing its legibility. This chapter presents a framework for virtual restoration of ancient documents, based on a combination of multispectral, 3D imaging and digital image analysis. The proposed framework consists of several steps. First, digital representations of the documents are acquired as multispectral images and 3D surface maps, reconstructed by a structured light technique. A multispectral camera and a digital projector are used in triangular configuration for 2D and 3D data acquisition. Then the multispectral images are registered against possible misalignments, and the 3D surface representation is used to correct geometrical distortions. Document flattening is then performed by 3D surface parameterization and texture mapping. Statistical techniques of decorrelation are applied to extract individual context parts of the document patterns (stamp, text, etc.) and to attenuate interferences. The processed data are then binarized by the proper segmentation technique. Finally, the entire digital object history - all the acquisition and processing steps, with the corresponding parameters - is stored in metadata files. These data can be exploited during a future evaluation of the restoration process, and can be used for either the creation of an expert knowledge database or the extraction of cross-document observations and conclusions.*

## IV.1 Introduction

Often the surface of an ancient document presents irregular deformations due to ageing, atmospheric influences, inappropriate archiving and manipulations (Figure IV.1). The goal of virtual restoration is to move beyond 2D image enhancement techniques, which cannot adequately address distortions that arise in damaged manuscripts. Shape-based distortion, for example, is a common element in damaged manuscripts that cannot be removed via 2D image enhancement algorithms. So it is required to use 3D geometric information together with 2D imaging in order to formulate a restoration framework for removing shape-based (projective) distortion in the image

[Brown & Seales 2001a, Brown & Seales 2001b, Brown & Seales 2001c].

For which concerns 2D data, it has been demonstrated in Chapter III that the multispectral imaging allows to improve the image analysis of degraded documents. The acquired images in different spectral bands have to be registered preliminarily in order to create a data cube where all the information associated to one pixel is also associated to a fixed spatial location. This pixel-to-pixel correspondence is fundamental to apply the next decorrelation techniques, that work locally to each pixel. Also the patterns (foreground, background, interferences) have to be well separated so that the binarization algorithm can eliminate the degradations and select the main text. As for 3D data, the Fringe Projection Technique explained in Section II.4 is an accurate and reliable technique to obtain the 3D shape of an object. The relative setup is very simple, the data acquisition and elaboration processes are quick. Moreover, the 3D model is related to the phase-map calculated by signal processing techniques, so it is possible to avoid the system calibration. The 3D surface representation is then used to correct geometrical distortions. Document flattening is then performed by 3D surface parameterization and texture mapping.

Until now, the issues of digital document restoration by multispectral and 3D imaging have been addressed separately. The objective of this chapter is to study an integrated multispectral/3D system for the restoration of old documents affected by geometrical alterations of the surface.

The deformations that occur to the document's surface can be classified into two categories: rigid and nonrigid. The rigid deformation is any deformation that preserves surface geodesics. The definition concerns the intrinsic properties of the surface, independent of any embedding. For example, bending and folding a manuscript page is a rigid deformation since points on the manuscript remain the same distance apart when measured on the manuscript. Such rigid surface deformations are typical of the damage and buckling caused from decades of handling and mishandling of documents. Years of folding and unfolding of large materials can leave them warped and creased. Other causes include the natural bending of pages, especially near a books binding.

Nonrigid surface deformations do not preserve surface geodesics. Such deformations can cause the overall surface area of the document to shrink and/or expand. This type of deformation can certainly be found in damaged materials, especially those suffering from fire and water exposure. For such damage it may be impossible to determine the extent of the nonrigid deformation without a priori knowledge of the material.

In this chapter will be considered only the rigid surface deformations, the nonrigid effects, although are not addressed, although nonrigid deformations are present, but this component is considered negligible.





Figure IV.1: Geometrical distortions of the pages in an old printed book.

## IV.2 Related works

In literature there are several applications of multispectral imaging for texture mapping of 3D models. Pelagotti *et al.* [Pelagotti *et al.* 2009] have developed a method for automated multispectral texture mapping of different typologies of 3D models for cultural heritage applications. A depth map image is generated from the 3D model. Then, a multispectral view of the object is automatically registered with the depth map image, using a registration method based on the mutual information [Pelagotti *et al.* 2007].

A system for the automatic construction of multispectral 3D models of architecture is proposed in [Brusco *et al.* 2006]. The acquisition instrument integrates an imaging spectrograph and a laser range-finder on the same platform, so that they can simultaneously measure spectral information and the geometry of the scene. A stereoscopic system based on a multispectral camera and a LCD projector uses multispectral information for 3D artwork object reconstruction. A set of images without optical filter is acquired by a filter-wheel camera. The LCD projector emits a luminous vertical line that sweeps the object. Luminous pixels are detected in camera images. Geometrical reconstruction of 3D points is obtained from the different parts of the luminous pattern on the object through triangulation. The document restoration context takes into account the 3D shape information to correct possible geometrical distortions [Brown & Seales 2001a], that appear when the book is wide-open, or because of bad environment conditions and various damages. In previous works a structured-light system (laser scanning) with a high-resolution camera [Brown *et al.* 2007] or a stereo vision system [Yamashita *et al.* 2004] are used to acquire the document shape.

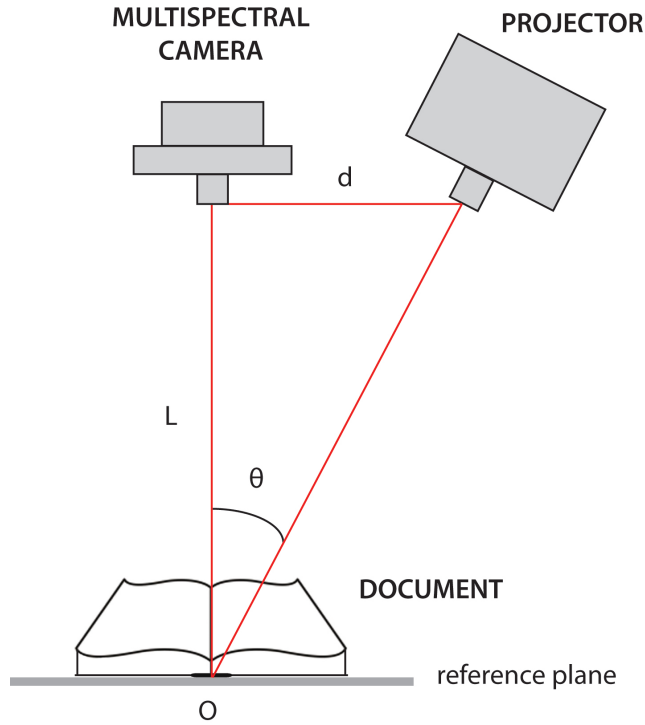


Figure IV.2: Optical configuration of

### IV.3 3D imaging

A digital projector is used in triangular configuration with the multispectral camera (Figure IV.2), in a geometrical setup with crossed optical-axes to perform 3D imaging by Fringe Projection Technique (FPT) [Gorthi & Rastogi 2010]. The optical axes of the devices are placed in the same plane forming an angle  $\theta$  and intersect in the point  $O$  on the reference plane, which is a plane normal to the camera axis and serves as a reference from which object height  $h$  is measured. The distance from camera and projector is called  $d$  and is perpendicular to the distance  $L$  from the camera and the reference plane. If the vertical (or horizontal) fringe pattern are used, this configuration ensures to work with monodimensional problem, in the sense that the variables change along vertical (horizontal) lines.

Considering a telecentric projection of a fringe pattern (for convenience of discussion), called  $p$  and  $p_r$  the period of fringe projected on the object and the reference plane, respectively, the displacement between two fringes (Figure IV.3) is given by:

$$\Delta h = h_2 - h_1 = \frac{p_r - p}{\tan \theta} - \frac{\Delta p}{\tan \theta} \quad (\text{IV.1})$$

with  $p_r = x_2 - x_4$ ,  $p = x_1 - x_3$ . The fringe period along a line  $\gamma$  orthogonal

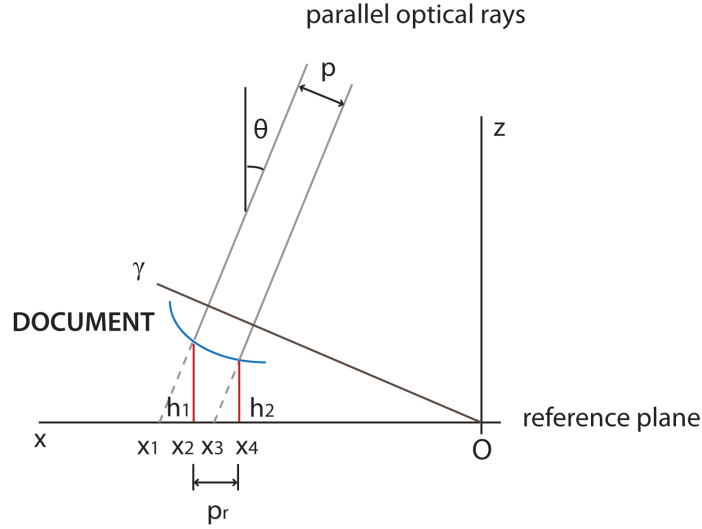


Figure IV.3: Telecentric

to projector axis is constant and related to the phase by:

$$\varphi(\gamma) = \frac{2\pi}{p}\gamma = \frac{2\pi}{p}(x \cos \theta + z \sin \theta) \quad (\text{IV.2})$$

Otherwise the period  $p$  is linearly variable along the  $x$  line and is arbitrary on the object surface. So the pattern image for both object and reference plane can be represented by the Eq. II.9, that are rewritten as follows for object  $g(x, y)$  and reference  $g_0(x, y)$ :

$$g(x, y) = a(x, y) + b(x, y) \cos(2\pi f_0 + \varphi(x)) \quad (\text{IV.3})$$

$$g_0(x, y) = a(x, y) + b(x, y) \cos(2\pi f_0 + \varphi_0(x)) \quad (\text{IV.4})$$

To calculate the height map  $h(x, y)$ , one can write, noting in Figure IV.4 that the triangles HCD and HPE are similar:

$$\frac{h}{\overline{CD}} = \frac{L - h}{d} \quad (\text{IV.5})$$

where  $\overline{CD} = \overline{BD} + \overline{BC}$ . By knowing that  $\varphi = 2\pi f_0 \overline{BD}$  and  $\varphi_0 = 2\pi f_0 \overline{BC}$  we obtain:

$$\overline{CD} = -\frac{\varphi(x) - \varphi_0(x)}{2\pi f_0} = -\frac{\Delta\varphi}{2\pi f_0} \quad (\text{IV.6})$$

where  $\Delta\varphi$  is the difference of phase observed between the patterns  $g(x, y)$  and  $g_0(x, y)$ . Substituting the Eq. IV.6 in the Eq. IV.5, the height distribution is computed by:

$$h(x, y) = \frac{L\Delta\varphi(x, y)}{\Delta\varphi(x, y) - 2\pi f_0 d} \quad (\text{IV.7})$$



of  $\phi(u, v)$  to the  $uv$ -plane is conformal [Courant 1977] if and only if

$$\nabla^2 \phi = \frac{\partial^2 \phi}{\partial u^2} + \frac{\partial^2 \phi}{\partial v^2} \quad (\text{IV.9})$$

where  $\nabla^2$  is the Laplacian operator  $\phi$ . Since the input is a set of discrete 3D points reconstructed on the document's surface, it is necessary looking for an inverse mapping from  $(x, y, z) \rightarrow (u, v)$ . Consider now a 2D function  $f(x, y)$ , parameterized by  $(x, y)$  that returns a 2D point  $(u, v)$ . The 2D Laplacian operator can be satisfied by the Cauchy-Riemann equation [Krantz 1999] such that

$$\frac{\partial f}{\partial x} + i \frac{\partial f}{\partial y} = 0 \quad (\text{IV.10})$$

which implies

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \text{ and } \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x} \quad (\text{IV.11})$$

To convert the problem to fit the form described in Eq. IV.11, the mapping of the 3D surface to a 2D coordinate frame can be reduced to  $\mathbb{R}^2 \rightarrow \mathbb{R}^2$  by considering how to map the triangles of a 3D mesh to their corresponding triangles in the  $uv$ -plane [Leévy *et al.* 2002]. Each 3D triangle is converted to its local 2D coordinate frame aligned by the triangle's normal as shown in Figure IV.5. Using this idea, a mesh  $M = \mathbf{p}_{i \in [1 \ n]}, T_{j \in [1 \ m]}$  is generated, where  $\mathbf{p}_i$  are the  $n/3D$  points acquired on the document's surfaces and serve as the mesh's vertices, and  $T_j$  represents the resulting  $m$  triangles, denoted as a tuple of three mesh vertices. The coordinates of the mesh vertices,  $\mathbf{p}_i$  can be expressed by their local coordinate frame, such that  $\mathbf{p}_i = (\hat{x}_i, \hat{y}_i)$ , which it will write as  $\mathbf{p}_i = (x, y)$ . A triangle-to-triangle mapping is defined by a unique affine transformation between the source and destination triangle. If we consider the affine mapping  $f(\mathbf{p}) \rightarrow \mathbf{q}$ , where  $\mathbf{p} = (x, y)$  and  $\mathbf{q} = (u, v)$ , where  $T_{p_1 p_2 p_3}$  is the source triangle and  $T_{q_1 q_2 q_3}$  is the destination, then the two triangles' vertices are related as

$$f(\mathbf{p}) = \frac{\text{area}(T_{pp_2 p_3}) \mathbf{q}_1 + \text{area}(T_{pp_3 p_1}) \mathbf{q}_2 + \text{area}(T_{pp_1 p_2}) \mathbf{q}_3}{\text{area}(T_{p_1 p_2 p_3})} \quad (\text{IV.12})$$

The partial derivatives (triangle gradient) of this equation is given as follows:

$$\frac{\partial f}{\partial x} = \frac{\mathbf{q}_1(y_2 - y_3) + \mathbf{q}_2(y_3 - y_1) + \mathbf{q}_3(y_1 - y_2)}{2\text{area}(T_{p_1 p_2 p_3})} \quad (\text{IV.13})$$

$$\frac{\partial f}{\partial y} = \frac{\mathbf{q}_1(x_2 - x_3) + \mathbf{q}_2(x_3 - x_1) + \mathbf{q}_3(x_1 - x_2)}{2\text{area}(T_{p_1 p_2 p_3})} \quad (\text{IV.14})$$

This triangle gradient can be used to formulate the Cauchy-Riemann equa-

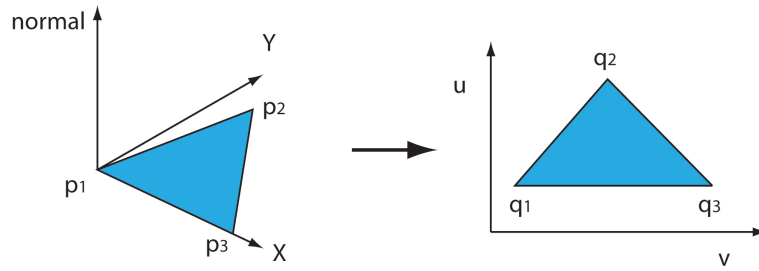


Figure IV.5: 3D triangle to 2D triangle mapping.

tions stated in Eq. IV.11. This can be written compactly in matrix form as follows:

$$\begin{bmatrix} \frac{\partial u}{\partial x} - \frac{\partial v}{\partial y} \\ \frac{\partial v}{\partial x} + \frac{\partial u}{\partial y} \end{bmatrix} = \frac{1}{2A_T} \begin{bmatrix} \Delta x_1 & \Delta x_2 & \Delta x_3 & \Delta y_1 & -\Delta y_2 & -\Delta y_3 \\ \Delta y_1 & \Delta y_2 & \Delta y_3 & \Delta x_1 & \Delta x_2 & \Delta x_3 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ \dots \\ v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (\text{IV.15})$$

where  $\Delta x_1 = (x_3 - x_2)$ ,  $\Delta x_2 = (x_1 - x_3)$ ,  $\Delta x_3 = (x_2 - x_1)$ , and  $\Delta y_i$  is defined similarly from Eq. IV.12;  $A_t$  is the area of the triangle defined by  $T_{p_1 p_2 p_3}$ . Solving this equation will find the appropriate  $(u_i, v_i)$  that are conformal on  $f(x, y) \rightarrow (u, v)$ . Working from this equation, a global system of equations,  $A\mathbf{x} = \mathbf{b}$ , can be written that incorporates all the vertices and triangles in

mesh  $M$ , such that matrix  $A$  and vector  $x$  are defined as

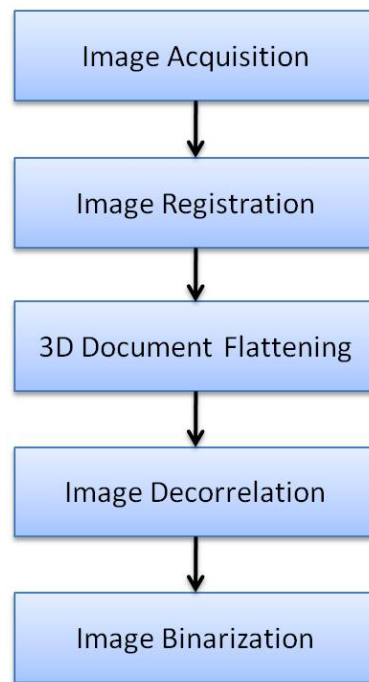
$$\begin{bmatrix}
 a_{j,i} = \begin{cases} \frac{\Delta x}{2A_{T_j}}, & \text{if } \mathbf{p}_i \geq T_j \\ 0, & \text{otherwise} \end{cases} & \vdots & a_{j,2i} = \begin{cases} \frac{-\Delta y}{2A_{T_j}}, & \text{if } \mathbf{p}_i \geq T_j \\ 0, & \text{otherwise} \end{cases} \\
 \dots & \dots & \dots \\
 a_{2j,i} = \begin{cases} \frac{\Delta y}{2A_{T_j}}, & \text{if } \mathbf{p}_i \geq T_j \\ 0, & \text{otherwise} \end{cases} & \vdots & a_{2j,2i} = \begin{cases} \frac{\Delta x}{2A_{T_j}}, & \text{if } \mathbf{p}_i \geq T_j \\ 0, & \text{otherwise} \end{cases}
 \end{bmatrix}
 \begin{bmatrix}
 u_0 \\
 \vdots \\
 u_i \\
 \vdots \\
 u_n \\
 \dots \\
 v_0 \\
 \vdots \\
 v_i \\
 \vdots \\
 v_n
 \end{bmatrix}
 \tag{IV.16}$$

The entries  $a_{j,i}$  of matrix  $A$  can be described as follows: each triangle  $T_j$  occupies two matrix row entries located at row  $j$  and  $2j$ . For each triangle row pair, six entries per row will result from the vertices  $\mathbf{p}_k \geq t_j$  occupying the columns  $k$  and  $2k$  with their corresponding  $\Delta x_k$  and  $\Delta y_k$ , as defined above. This results in  $A$  being a  $2m \times 2n$  matrix, where  $m$  is the number of triangles, and  $n$  is the number of vertices in the 3D mesh. The entries for vector  $\mathbf{x}$  are the desired conformal points  $(u_i, v_i)$ . To obtain a unique solution for the system  $A\mathbf{x} = \mathbf{b}$  up to a similarity, some vertices must be constrained - otherwise, the  $(u_i, v_i)$  could have any arbitrary orientation in the 2D plane. Fixing at least two values will give a unique solution and constraints the orientation in the resulting conformal map.

## IV.5 A framework for document restoration

A combination of multispectral and 3D imaging is proposed to analyze old documents affected by geometrical distortions of the surface. These two imaging techniques provide 2D and 3D data used for document *deskewing* (removing distortions) in order to apply the next enhancement techniques. The framework flow-chart (Figure IV.6) shows the main steps required for document restoration.

Placed a projector Mitsubishi PK20 and the multispectral camera Chroma CX3 at a distance of 50 cm with an angle of  $30^\circ$ , and fixed the imaging system on the document at a distance of about 90 cm, a set of images are acquired in different spectral bands (Figure IV.7). A sinusoidal fringe pattern is projected on the document surface and on the reference plane (removing the document)



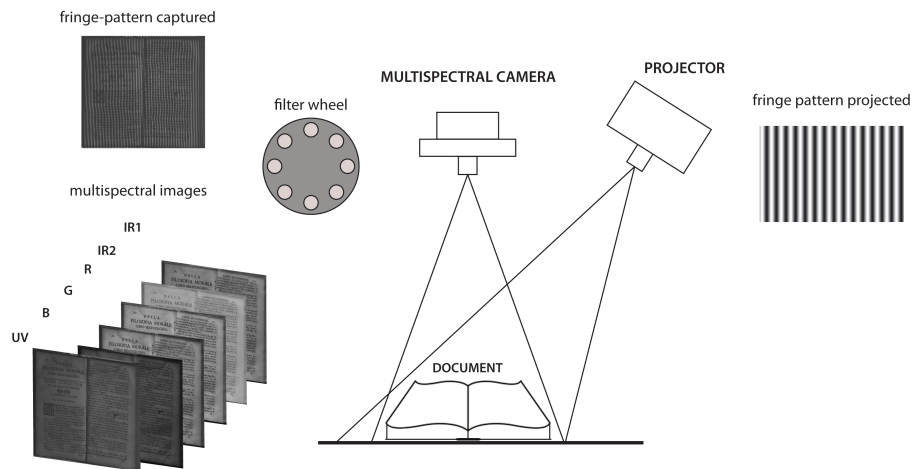
**Figure IV.6:** Framework flowchart.

and the images of reflecting patterns are captured by the camera without filter.

All the acquired images must be preliminarily registered to obtain a correspondence pixel-to-pixel among the multispectral textures and the 3D model calculated as a depth map by Wavelet Transform Profilometry [Dursun *et al.* 2004], as described in Section II.4.1. The Fourier-Mellin method described in Section III.2.2.3 is applied to align the multispectral images and the fringe patterns. In Figure IV.8 are represented the Green channel image and the fringe patterns projected on a document and on a reference plane. First, the multispectral images are aligned with the object fringe pattern, choosing the Green channel as reference, so the geometrical parameters for each transformation (translation, rotation and scaling) are computed. In particular, assuming that the two images of fringe patterns are aligned (this is conceivable since both images are captured without filter) the parameters used to transform the object fringe pattern are also applied to transform the reference fringe pattern. In this manner a data cube is formed by all acquired images.

The Wavelet Transform is applied to both fringe images separately, to extract the wrapped phase map. Subsequently, the phase unwrapping algorithm provide the continuous phase map. The 3D shape of document's surface is obtained by the difference of the two phase map. The textured 3D model is





**Figure IV.7:** Imaging system for 2D/3D data acquisition.

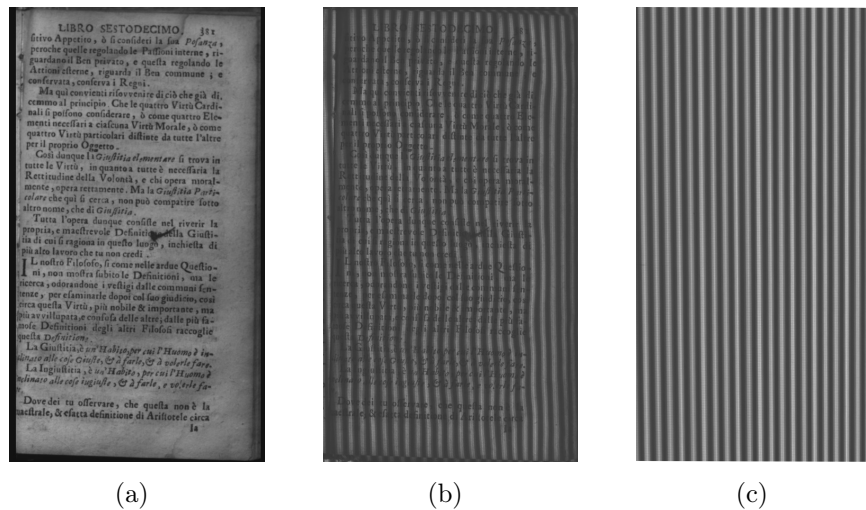
illustrated in Figure IV.9.

3D document flattening step is performed by triangulating the 3D model to obtain a mesh that is mapped in a 2D plane by a conformal parameterization applied to each triangle, as described in previous section. The triangulated mesh relative to 3D model is shown in Figure IV.10. Then the multispectral images can be warped using the coordinates of this plane (Figure.11a). A comparison between an the original image and the restored image is shown in Figure IV.11b.

Then, rectified images can be processed by image enhancement techniques. Statistical techniques of data decorrelation (ICA, PCA, Symmetric Orthogonalization, see Section III.2.2.4) are employed to separate individual context part of document patterns (background, foreground, interfering patterns), to extract symbols, stamps, etc., and to attenuate degradations. Finally, an image binarization algorithm (i.e. Sauvola [Sauvola & Pietaksinen 2000]) can be applied to fully isolate the pattern of interest.

As example, a sign present in a multispectral image is removed processing the image set with the Symmetric Orthogonalization algorithm, to improve the result of the binarization step.

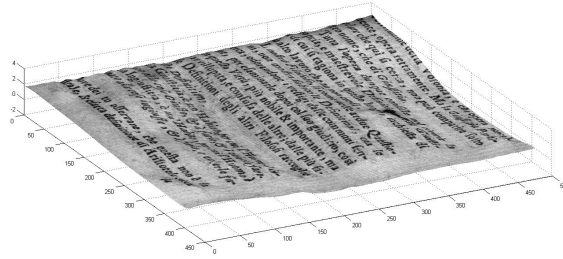
In the last module of our framework the taken steps and corresponding parameters of the document virtual restoration are described and archived. A database is created for further data management, where each document-related image is represented by metadata taking into account the history of the elaboration process and the intrinsic information about the document.



**Figure IV.8:** Greyness channel image (a), object fringe pattern (b) and reference fringe pattern (c).

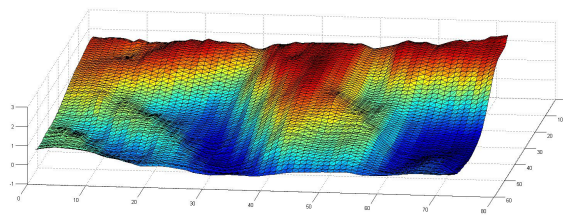
## IV.6 Archiving

The last module of our framework is dedicated to the archiving of the raw and processed data, along with a metadata description of all the imaging and processing tasks applied to each document. Each document-related image is represented by metadata, characterizing the history of the elaboration process and the document intrinsic information. Such information can be exploited in a future evaluation of the digital restoration process and for the creation of an expert knowledge database for the extraction of cross-document observations and conclusions about the document production process, the ageing process, etc., which are important for both historians and restorers. The metadata schema is composed of four main entities: the Creation entry, which provides metadata describing the very art object, the Collection entity, which depicts the physical location where the cultural heritage object is located, the Digital Representation (DR) entity, which represents the digitized versions of the Creation entity (multispectral images, registered images, enhanced images, etc.), and the Operator entity, which describes persons or organizations primarily responsible for creating the content of the record. The proposed metadata schema links a set of attributes to each entity. For example, DR is associated with acquisition attributes, parameters of image registration, relationships between recto and verso images, to name a few. The type and number of attributes and relations is not predefined, so the model is flexible enough to support unforeseen representations. This rich description of acquisitions and of image processing results will support the archiving of acquired



(a)

**Figure IV.9:** Textured 3D model of a document.



(a)

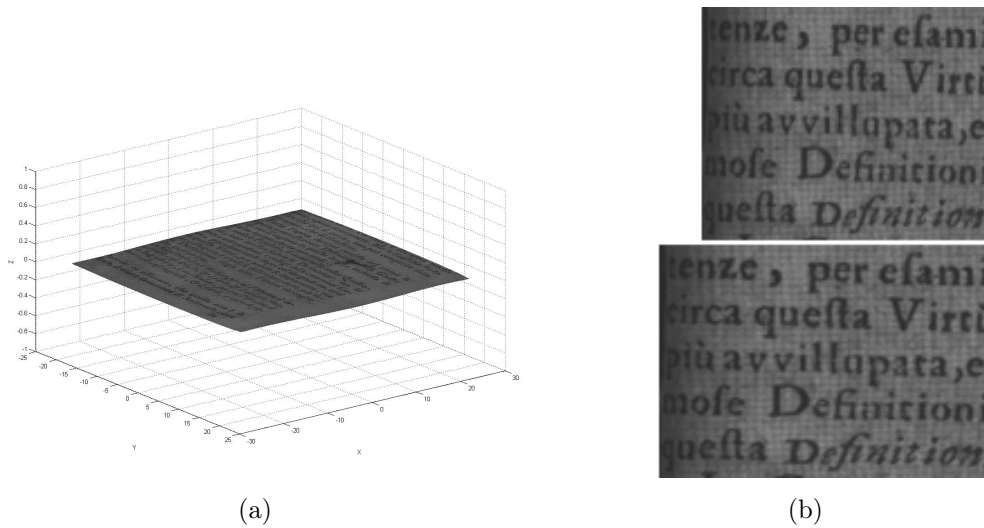
**Figure IV.10:** Triangulated 3D mesh.

images and their retrieval based on the characteristics of the processing. At the same time, the availability of traditional descriptive metadata associated to the Creations, will support content based search, as is usually done in a Digital Library.

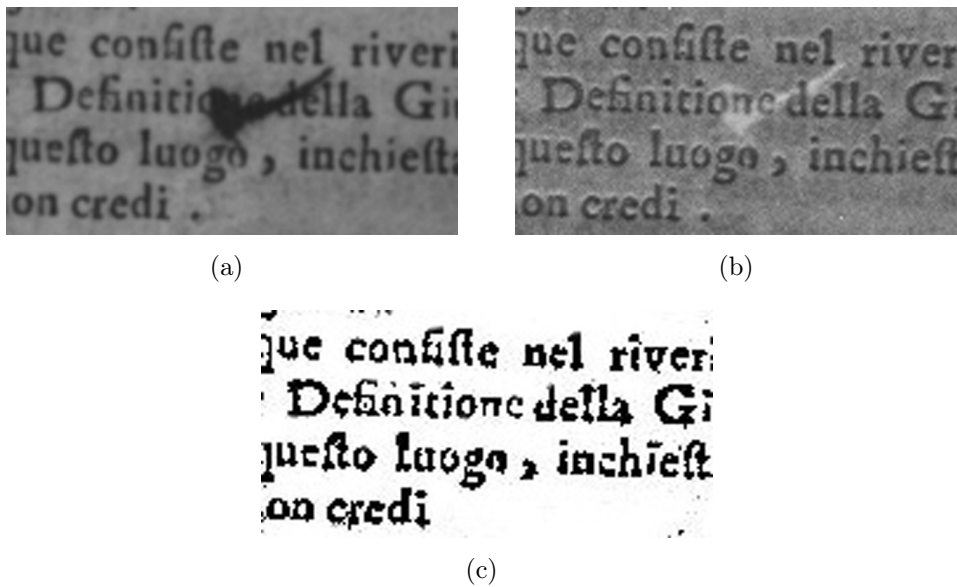
## IV.7 Conclusions

A framework for virtual restoration of ancient documents affected by various alterations is proposed. This framework is based on multispectral and 3D digital representations. The framework consists of several steps, in which image processing techniques are applied. Many of these steps are automatic or require little operator intervention. The geometrical alterations of the document's surface are corrected by surface parameterization while the degradations due to bleed-through, show-through, spots, etc. can be attenuated by decorrelation techniques.

A new fast imaging system is proposed for multispectral and 3D data acquisition that consists of a multispectral camera and a digital projector, arranged in a triangular configuration. Concerning 3D shape reconstruction by fringe projection, the system does not require calibration. However, calibration can be performed if real size information is required.



**Figure IV.11:** The developable surface is mapped on a plane (a). A comparison before (up) and after surface parameterization (down) (b).



**Figure IV.12:** A sign is present in the original image(a). By applying Symmetric Orthogonalization to the multispectral images the sign is removed (b), to improve the binarization step (c).

# Conclusions

2D and 3D digital imaging techniques for Cultural Heritage applications have been investigated in this thesis. Both the techniques are based on the image acquisition that provides data to be elaborated with image and signal processing techniques. The two kind of imaging are applied before separately and also in combination to create a integrated system of 2D and 3D data.

The 3D system proposed for underwater applications using whole-field structured-light techniques, provides interesting results in turbidity conditions. The main drawback of the gray-code technique, that consists in the projecting of several patterns, is overcome by using the Fringe Projection Technique that can project only one pattern. The turbidity of water decreases the image quality and then the performance of the 3D techniques. In this work, solutions have not been proposed for this issue, but certainly could improve the 3D reconstructions, like the use of the polarizers or taking into account a scattering model of the light.

About the 2D imaging application, the virtual restoration of ancient document has been addressed through the multispectral acquisition. Multispectral imaging supports the investigation of ancient manuscripts where the text is hardly visible in conventional RGB images or for the human eye. This information are exploited from the decorrelation techniques to separate patterns and enhance the document aesthetics. The enhancement techniques analysed provide a effective tool to investigate degraded documents, although simple and linear data model have been used.

To extend the analysis in the historical documents topic, a combination of multispectral and 3D imaging has been studied. This is not a new combination of these techniques, but in the document restoration context it has been proposed a new framework to correct 2D and 3D alterations, with automatic and reliable solutions.

Moreover, it has to be note that the Fringe Projection Technique is a good and versatile tool in the discussed applications.



# Bibliography

- [Allais *et al.* 2007] A-G. Allais, V. Brandou, S. Dentrecolas, J-P. Gilliotte and M. Perrier. *IRIS - A vision system to reconstruct natural deep-sea scenes in 3D*. In Seventeenth International Offshore and Polar Engineering Conference, pages 111–118, Lisbon - Portugal, 1-6 July 2007. 14, 30
- [Amann *et al.* 2001] M. C. Amann, T. Bosch, M. Lescure, R. Myllylä and M. Rioux. *Laser ranging: A critical review of usual techniques for distance measurement*. Optical Engineering, vol. 40, no. 1, pages 10–19, 2001. 13
- [Anderson S. 2002] Levoy M. Anderson S. *Unwrapping and Visualizing Cuneiform Tablets*. IEEE Computer Graphics and Applications, vol. 22, pages 82–88, 2002. 7
- [Arias *et al.* 2004] V. Arias, T. Runge-Kuntz, J. Heather M. R. ans van der Elst and R. P. Watson. *Documentation and Virtual Modeling of Archaeological Sites Using a Non-Metric Extreme Wide-Angle Lens Camera*. In International Archives of Photogrammetry and Remote Sensing, vol. XXXV, 2004. 7
- [Avdelidis 2004] A. Avdelidis N.P. Moropoulou. *Review Paper: Applications of infrared thermography for the investigation of historic structures*. J. Cultural Heritage, vol. 5, no. 1, pages 119–127, 2004. 8
- [Badekas & Papamarkos 2005] E. Badekas and N. Papamarkos. *Automatic evaluation of document binarization results*. In Proc. 10th Iberoamerican Congress on Pattern Recognition, page 10051014, Havana, Cuba, 2005. 94
- [Baird 1992] H.S. Baird. *Document image defect models and their uses*. In Proc. ICDAR'92, pages 62–67, 1992. 92
- [Baird 2000] H. S. Baird. *The State of the Art of Document Image Degradation Modeling*. In In Proc. of 4 th IAPR International Workshop on Document Analysis Systems, Rio de Janeiro, pages 1–16, 2000. 65
- [Baldi *et al.* 2002] A. Baldi, F. Bertolino and F. Ginesu. *On the performance of some unwrapping algorithms*. Optics and Lasers in Engineering, vol. 37, pages 313–330, 2002. 48

- [Banham & Katsaggelos 1997] M. Banham and A. Katsaggelos. *Digital image restoration*. IEEE Signal Processing Magazine, vol. 14, no. 2, pages 24–41, 1997. 95
- [Barone & Razionale 2004] S. Barone and A. V. Razionale. *A reverse engineering methodology to capture complex shapes*. In XVI Congreso Internacional de Ingeniería Gráfica, volume 1, pages 1–10, Saragoza - Spain, 2004. 20, 24
- [Barone et al. 2003] S. Barone, A. Curcio and A. V. Razionale. *A structured light stereo system for reverse engineering applications*. In IV Seminario Italo-Espanol, Reverse Engineering Techniques and Applications, volume 1, pages 65–74, Cassino - Italy, 2003. 20, 22
- [Beraldin et al. 1998] J.A. Beraldin, F. Blais, M. Rioux and L. Cournoyer. *Portable digital 3D imaging system for remote sites*. In Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS 98), Monterey - California, 31 May – 3 June 1998. 3
- [Beraldin et al. 1999] J.A. Beraldin, F. Blais, L. Cournoyer, M. Rioux, S.H. El-Hakim, R. Rodella, F. Bernier and N. Harrison. *Digital 3D Imaging System for Rapid Response on Remote Sites*. In Proceedings of the 2nd Intern. Conference on 3D Digital Imaging and Modeling, pages 34–43, Ottawa, 4-8 October 1999. 2, 4
- [Berni M. 1999] Cappellini V. Berni M. Bartolini F. *Image processing tools for the enhancement of the aspect of ancient painting reproductions*. In Proceeding of the 2nd International Congress on Science and Technology for the safeguard of cultural heritage in the Mediterranean basin, Paris - France, 56 July 1999. 4
- [Berssen 1986] J. Berssen. *Dynamic thresholding of grey-level images*. In Proceedings of the Eighth ICPR, pages 1251–1255, 1986. 92
- [Bianco et al. 2008] G. Bianco, A. Tonazzini and E. Salerno. *Assessing automatic registration methods applied to digital analysis of historical documents*. In Abstracts SIMAI 2008, page 24, Rome, Italy, 2008. 86
- [Bianco et al. 2009] G. Bianco, F. Bruno, A. Tonazzini, E. Salerno and E. Console. *Recto-verso Registration, Enhancement and Segmentation of Ancient Documents*. In Virtual Systems and Multimedia, 2009. VSMM '09. 15th International Conference on, pages 131–136, 2009. 85



- [Bianco *et al.* 2010] G. Bianco, F. Bruno, M. Muzzupappa and M. L. Luchi. *Underwater 3D shape reconstruction by fringe projection*. In CAA 2010 Conference, Granada - Spain, 6 - 9 April 2010. 39
- [Bioucas-Dias & Valadao 2005] M. Bioucas-Dias and G. Valadao. *Phase unwrapping: a new max-flow/min-cut based approach*. In IEEE trans, 2005. 48
- [Blais 2004] F. Blais. *Review of 20 years of range sensor development*. Journal of Electronic Imaging, vol. 13, no. 1, pages 231–243, 2004. 15
- [Bouguet 2010] J. Y. Bouguet. *Camera calibration toolbox for Matlab*. <http://www.vision.caltech.edu/bouguetj>, 2010. (Accessed July 2010). 30, 31, 53
- [Bracewell 2000] R.N. Bracewell. *The fourier transform and its applications*. McGraw-Hill, New York, 3rd edition édition, 2000. 73
- [Bräger & Chong 1999] S. Bräger and A. K. Chong. *An application of close range photogrammetry in dolphin studies*. Photogrammetric Record, vol. 16, no. 93, pages 503–517, 1999. 13
- [Brandou *et al.* 2007] V. Brandou, A. G. Allais, M. Perrier, E. Malis, P. Rives, J. Sarrazin and P. M. Sarradin. *3D reconstruction of natural underwater scenes using the stereovision system IRIS*. In OCEANS 2007 - Europe, 2007. 16
- [Brauwers *et al.* 2008] J. Brauwers, N. Schulte and T. Aach. *Multispectral Filter-Wheel Cameras: Geometric Distortion Model and Compensation Algorithms*. IEEE Transactions on Image Processing, vol. 17, no. 12, pages 2368–2380, 2008. 62
- [Brown & Seales 2001a] M.S. Brown and W.B. Seales. *3D Imaging and Processing of Damaged Texts*. In Proc. Assoc. for Computing in the Humanities and Assoc. of Linguistic and Literacy Computing Joint Conf., June 2001. 102, 103
- [Brown & Seales 2001b] M.S. Brown and W.B. Seales. *Digital Atheneum: New Approaches for Preserving, Restoring and Analyzing Damaged Manuscripts*. In Proc. First ACM-IEEE Joint Conf. Digital Library, June 2001. 102
- [Brown & Seales 2001c] M.S. Brown and W.B. Seales. *Document Restoration Using 3D Shape: A General Deskewing Algorithm for Arbitrarily*

- Warped Documents*. In Proc. Int'l Conf. Computer Vision (ICCV 01), July 2001. 102
- [Brown *et al.* 2007] M. S. Brown, M. Sun, R. Yang, L. Yun and W. B. Seales. *Restoring 2d content from distorted documents*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 29, page 19041916, 2007. 103, 106
- [Brusco *et al.* 2006] N. Brusco, S. Capeleto, M. Fedel, A. Paviotti, E. Zanella, L. Poletto, G.M. Cortelazzo and G. Tondello. *A Range Camera Collecting Multi-Spectral Texture for Architecture Applications*. In 3D Data Processing, Visualization, and Transmission, Third International Symposium on, pages 978–985, june 2006. 103
- [Bythell *et al.* 2001] J. Bythell, P. Pan and J. Lee. *Three-dimensional morphometric measurements of reef corals using underwater photogrammetry techniques*. Coral Reefs, vol. 20, no. 3, pages 193–199, 2001. 13
- [Campisi & Egiazarian 2007] P. Campisi and K. Egiazarian. *Blind image deconvolution, theory and application*. CRC Press, Boca Raton, 2007. 95, 97, 98
- [Canciani *et al.* 2003] M. Canciani, P. Gambogi, F. G. Romano, G. Cannata and P. Drap. *Low cost digital photogrammetry for underwater archaeological site survey and artifact insertion. The case study of the dolia wreck in secche della Meloria, Livorno-Italia*. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. XXXIV, 2003. 13
- [Chan & Wong 1998] T. Chan and C. Wong. *Total variation blind deconvolution*. IEEE Trans. Image Processing, vol. 7, no. 3, pages 370–375, 1998. 96
- [Chan & Wong 2000] T. Chan and C. Wong. *Convergence of the alternating minimization algorithm for blind deconvolution*. Linear Algebra Appl., vol. 316, no. 1-3, pages 259–285, 2000. 96
- [Chapman *et al.* 2006] P. Chapman, G. Conte, P. Drap, P. Gambogi, F. Gauch, K. Hanke, L. Longand, V. Loureiro, O. Papini, A. Pascoal and D. Richards J. & Roussel. *VENUS: Virtual Exploration of Underwater Sites*. In 7th International Symposium on Virtual Reality : Archaeology and Cultural Heritage', VAST, 2006. 3

- [Chen *et al.* 1996] W. Chen, Y. Tan and H. Zhao. *Automatic analysis technique of spatial carrier-fringe patterns*. Optics and Lasers in Engineering, vol. 25, no. 2–3, pages 111 – 120, 1996. Optical NDE Methods in the People’s Republic of China (PRC). 46
- [Chen *et al.* 2000] F. Chen, G. M. Brown and M. Song. *Overview of three-dimensional shape measurement using optical methods*. Optical Engineering, vol. 39, no. 1, pages 10–22, 2000. 15
- [Cheriet & Moghaddam 2008] M. Cheriet and R. F. Moghaddam. *Degradation modeling and enhancement of low quality documents*. In in Proc. WOSPA 2008, 2008. 77
- [Cichocki & Amari 2002] A. Cichocki and S.-I. Amari. Adaptive blind signal and image processing. Wiley, New York, 2002. 82, 84, 90
- [Clark *et al.* 2003] M. R. Clark, D. M. McCann and M. C. Forde. *Application of infrared thermography to the non-destructive testing of concrete and masonry bridges*. NDT & E International, vol. 36, no. 4, pages 265 – 275, 2003. 8
- [Cocito *et al.* 2003] S. Cocito, S. Sgorbini, A. Peirano and M. Valle. *3D reconstruction of biological objects using underwater video technique and image processing*. Journal of Experimental Marine Biology and Ecology, vol. 297, no. 1, pages 57–70, 2003. 13, 17
- [Cod 2010] <http://codexsinaiticus.org/en/>, 2010. (Accessed July 2010). 60
- [Cortelazzo G.M. 1999] Marton F. Cortelazzo G.M. *3D modeling for cultural heritage applications*. In Proceedings of the 2nd Intern. Congress on Science and Technology for the safeguard of cultural heritage in the Mediterranean basin, Paris - France, 5-6 July 1999. 2
- [Cotte & Dupraz 2006] P. Cotte and D. Dupraz. *Spectral imaging of Leonardo Da Vincis Mona Lisa: A true color smile without the influence of aged varnish*. In CGIV Final Program and Proceedings, 2006. 3
- [Courant 1977] R. Courant. Plateau’s problem, dirichelt’s principle, conformal mapping and minimal surfaces, chapitre 3, pages 95–134. Springer, 1977. 107
- [Creagh & Bradley 2007] D. Creagh and D. Bradley. Physical techniques in the study of art, archaeology and cultural heritage. Elsevier, 2007. ix, 1, 2

- [Creath & Wyant 1992] K. Creath and J. C. Wyant. Moiré and fringe projection techniques, chapitre 16, pages 653–686. John Wiley and Sons, 1992. 20
- [Cupitt *et al.* 1996] J. Cupitt, K. Martinez and D. Saunders. *Methodology for Art Reproduction in Colour: the MARC project*. Computers and the History of Art, vol. 6, no. 2, pages 1–20, 1996. 3
- [Dagleish *et al.* 2004] F. R. Dagleish, S. Tetlow and R. L. Allwood. *Laser-assisted Vision Sensor for AUV Navigation*. In Proc. 14th International Offshore and Polar Engineering Conference, volume 2, pages 401–408, Toulon - France, 23-28 May 2004. 18
- [Dagleish *et al.* 2006] F. R. Dagleish, F. M. Caimi, C. H. Mazel and J. M. Glynn. *Extended range underwater optical imaging architecture*. In IEEE Oceans 2006, Boston - MA, September 2006. 19
- [Dano 2003] P. Dano. *Joint Restoration and Compression of Document Images with Bleed-through Distortion*, 2003. Master Thesis, Ottawa-Carleton Institute for Electrical and Computer Engineering, School of Information Technology and Engineering, University of Ottawa. 77
- [Dazheng *et al.* 2006] F. Dazheng, W. Nan and L. Baoquan. *A Region-growing Phase Unwrapping Approach Based on Local Frequency Estimation for Interferometric SAR*. In IEEE, ICSP2006 Proceedings, 2006. 48
- [Desmangles 2003] A. I. Desmangles. *Extension of the fringe projection method to large objects for shape and deformation measurement*. PhD thesis, Ecole Polytechnique Federale de Lausanne, 2003. 16
- [Drap *et al.* 2007] P. Drap, D. Scaradozzi, P. Gambogi, L. Long and F. Gauche. *Photogrammetry for virtual exploration of underwater archeological sites*. In XXI International Scientific Committee for Documentation of Cultural Heritage (CIPA) Symposium, Athens, Greece, October 2007. 10, 17
- [Drira *et al.* 2006] F. Drira, F. Le Bourgeois and H. Emptoz. *Restoring ink bleed-through degraded document images using a recursive unsupervised classification technique*. In In Proc. 7th Workshop on Document Analysis Systems, pages 38–49, 2006. 77
- [Drira 2006] F. Drira. *Towards restoring historic documents degraded over time*. In Document Image Analysis for Libraries, 2006. DIAL '06. Second International Conference on, april 2006. 66

- [Dubois & Pathak 2001] E. Dubois and A. Pathak. *Reduction of bleed-through in scanned manuscript documents*. In in Proc. IS&T Image Processing, Image Quality, Image Capture Systems Conference, page 177180, 2001. 77, 86
- [Dursun *et al.* 2004] A. Dursun, S. Özder and F. Necati Ecevit. *Continuous wavelet transform analysis of projected fringe patterns*. Measurement Science and Technology, vol. 15, no. 9, page 1768, 2004. 45, 110
- [Easton *et al.* 2003] R.L. Easton, K.T. Jr. Knox and W.A. Christens-Barry. *Multispectral imaging of the Archimedes palimpsest*. Appl. Imagery Pattern Recognition Workshop Proceedings 2003, pages 15–17, 2003. 7, 60, 67, 68
- [Edge 2006] M. Edge. Dome port theory, chapitre 5, page 314. Elsevier Ltd, 2006. 26
- [Eikvil *et al.* 1991] L. Eikvil, T. Taxt and K. Moen. *A fast adaptive method for binarization of document images*. In International Conference on Document Analysis and Recognition, ICDAR '91, pages 435–443, France, 1991. 92
- [Eur 2010] [http://www.mega-vision.com/cultural\\_heritage.html](http://www.mega-vision.com/cultural_heritage.html), 2010. (Accessed July 2010). 60
- [Europeana 2010] Europeana. <http://www.europeana.eu>, 2010. (Accessed September 2010). 3
- [F. *et al.* 2007] Br'mand F., P. Doumalin, J.C. Dupré, F. Hesser and Valle V. *Optical Techniques for Relief Study of Mona Lisas Wooden Support*. In E. E. Gdoutos, editeur, Experimental Analysis of Nano and Engineering Materials and Structures, pages 67–68. Springer Netherlands, 2007. 9
- [Feng & Manmatha 2005] S. L. Feng and R. Manmatha. *Classification Models for Historical Manuscript Recognition*. Document Analysis and Recognition, International Conference on, pages 528–532, 2005. 69
- [Feng *et al.* 2007] X. Feng, W. Jicang, Z. Lei and L. Xiaoling. *A new method about placement of the branch cut in two-dimensional phase unwrapping*. In Synthetic Aperture Radar, 2007. APSAR 2007. 1st Asian and Pacific Conference on, pages 755 –759, nov. 2007. 48

- [Fermo *et al.* 2002] P. Fermo, F. Cariati, M. Cipriani C. Canetti, G. Padeletti, B. Brunetti and A. Sgamellotti. *The use of small-angle X-ray scattering (SAXS) for the characterisation of lustre surfaces in Renaissance majolica*. Applied Surface Science 185, 2002. 8
- [Fienup 1997] J. R. Fienup. *Invariant error metrics for image reconstruction*. Appl. Opt., vol. 36, no. 32, pages 8352–8357, Nov 1997. 86
- [Firoozfam & Negahdaripour 2003] P. Firoozfam and S. Negahdaripour. *A Multi-Camera Conical Imaging System for Robust 3D Motion Estimation, Positioning and Mapping from UAVs*. In AVSS '03: Proceedings of the IEEE Conference on Advanced Video and Signal Based Surveillance, page 99, Washington, DC, USA, 2003. IEEE Computer Society. 17
- [Flynn 1997] T.J. Flynn. *Two-dimensional phase unwrapping with minimum weighted discontinuity*. J. Opt. Soc. Am., vol. 14, no. 10, page 26922701, 1997. 48
- [Fontana *et al.* 2004] R. Fontana, M.C. Gambino, G. Gianfrate, M. Greco, E. Pampaloni and L. Pezzati. *A 3D scanning device for architectural survey based on Time-Of-Flight technology*. PROCEEDINGS OF THE SOCIETY OF PHOTO-OPTICAL INSTRUMENTATION ENGINEERS (SPIE), vol. 5457, pages 393–400, 2004. 9
- [Foresti 2001] G. L. Foresti. *Visual inspection of sea bottom structures by an autonomous underwater vehicle*. IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, vol. 31, no. 5, pages 691–705, 2001. 13
- [Fotoscientifica 2010] Fotoscientifica. <http://www.fotoscientificarecord.com/>, 2010. (Accessed July 2010). 68
- [Ghiglia & Pritt 1998] D. C. Ghiglia and M. D. Pritt. *Two-dimensional phase unwrapping*. John Wiley and Sons, 1998. 47
- [Giannakis & Heath 2000] G. Giannakis and R. Heath. *Blind identification of multichannel FIR blurs and perfect image restoration*. IEEE Trans. Image Processing, vol. 9, no. 11, pages 1877–1896, 2000. 96
- [Godin *et al.* 1999] G. Godin, J.A. Beraldin, L. Cournoyer, M. Rioux and F. Blais. *NRC Participation in the Digital Michelangelo Project*. In Proceeding of Canada - Italy Workshop: Heritage Applications of 3D Digital Imaging, Ottawa - Canada, October 1999. 3

- [Godin *et al.* 2002] G. Godin, J-A. Beraldin, J. Taylor, L. Cournoyer, M. Rioux, S. El-Hakim, R. Baribeau, F. Blais, P. Boulanger, J. Domey and M. Picard. *Active optical 3D imaging for heritage applications*. IEEE Computer Graphics and Applications, vol. 22, no. 5, pages 24–36, 2002. 19
- [Google 2010] Google. <http://books.google.com/>, 2010. (Accessed September 2010). 3
- [Gorthi & Rastogi 2010] S.S. Gorthi and Pramod Rastogi. *Fringe projection techniques: Whither we are?* Optics and Lasers in Engineering, vol. 48, no. 2, pages 133 – 140, 2010. Fringe Projection Techniques. 104
- [Gracias & Santos-Victor 2000] N. Gracias and J. Santos-Victor. *Underwater video mosaics as visual navigation maps*. Computer Vision and Image Understanding, vol. 79, no. 1, pages 66–91, 2000. 13
- [Griesser *et al.* 2000] M. Griesser, A. Denker, H. Musner and K. H. Maier. *Non-destructive investigation of paint layer sequences*. In Tradition and innovation: advances in conservation, Proceedings of the IIC Melbourne Congress, Melbourne, London IIC, 10-14. Oct. 2000. 8
- [Grun *et al.* 2004] A. Grun, F. Remondino and L. Zhang. *Photogrammetric Reconstruction of the Great Buddha of Bamiyan, Afghanistan*. The Photogrammetric Record, vol. 19, 2004. 10
- [Gupta *et al.* 2008] M. Gupta, S. G. Narasimhan and Y. Y. Schechner. *On controlling light transport in poor visibility environments*. In IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2008., pages 1–8, Anchorage, AK, June 2008. 14
- [Haindl 2000] M. Haindl. *Recursive model-based image restoration*. In Proceedings of the 15th International Conference on Pattern Recognition, vol. III, pages 346–349, 2000. 96
- [Hardeberg *et al.* 1999] J. Y. Hardeberg, F. Schmitt, H. Brettel, J.P. Crettez and H. Maitre. *Multispectral Image Acquisition and Simulation of Illuminant Changes*. In in Colour Imaging: Vision and, pages 145–164. John Wiley Sons, Ltd, 1999. 58
- [Harikumar & Bresler 1999] G. Harikumar and Y. Bresler. *Perfect blind restoration of images blurred by multiple filters: Theory and efficient algorithms*. IEEE Trans. Image Processing, vol. 8, no. 2, pages 202–219, 1999. 96



- [Harsdorf *et al.* 1999] S. Harsdorf, R. Reuter and S. Tonebon. *A reverse engineering methodology to capture complex shapes*. In Proceedings of the SPIE, volume 3821, pages 378–383, 1999. 18
- [Hartley & Zisserman 2000] R. I. Hartley and A. Zisserman. *Multiple view geometry in computer vision*. Cambridge University Press, ISBN: 0521623049, 2000. 17
- [Harvey & Shortis 1995] E. Harvey and M. Shortis. *A system for stereo-video measurement of sub-tidal organisms*. Marine Technology Society Journal, vol. 29, no. 4, pages 10–22, 1995. 16
- [Havermans *et al.* 2003] J. Havermans, H.A. Aziz and H. Scholten. *Non destructive detection of iron gall inks by means of multispectral imaging. Part 1: Development of the detection system*. Restaurator, vol. 24, pages 55–60, 2003. 7
- [He & Seet 2004] D. M. He and G. G. L. Seet. *Divergent-beam Lidar imaging in turbid water*. Optics and Lasers in Engineering, vol. 41, no. 1, pages 217–231, 2004. 7, 14, 19
- [Heredia-Ortiz & Patterson 2003] M. Heredia-Ortiz and E. A. Patterson. *On the industrial applications of Moiré and fringe projection techniques*. Strain, vol. 39, no. 3, pages 95–100, 2003. 19
- [Hill 1998] B. Hill. *Multispectral color technology: A way towards high definition color image scanning and encoding*. Electronic Imaging, pages 2–13, 1998. 7
- [Hiller & Wess 2006] J.C. Hiller and T.J. Wess. *The use of small-angle X-ray scattering to study archaeological and experimentally altered bone*. Journal of Archaeological Science, vol. 33, pages 560–572, 2006. 8
- [Hobson & Watson 1999] P. R. Hobson and J. Watson. *Accurate three-dimensional metrology of underwater objects using replayed real images from in-line and off-axis holograms*. Measurement Science and Technology, vol. 10, no. 12, pages 1153–1161, 1999. 18
- [Hobson & Watson 2002] P. R. Hobson and J. Watson. *The principles and practice of holographic recording of plankton*. Journal of Optics A: Pure and Applied Optics, vol. 4, no. 4, pages S34–S49, 2002. 18



- [Howell & Tasnim 2005] G. M. E. Howell and M. Tasnim. *Diagnostic Raman spectroscopy for the forensic detection of biomaterials and the preservation of cultural heritage*. Analytical and Bioanalytical Chemistry, vol. 382, pages 1398–1406, 2005. 9
- [Hyvärinen *et al.* 2001] A. Hyvärinen, J. Karhunen and E. Oja. Independent component analysis. John Wiley, New York, 2001. 79
- [Jaffe & Dunn 1988] J. S. Jaffe and C. Dunn. *A model-based comparison of underwater imaging systems*. Proceedings of the Ocean Optics IX, vol. 925, pages 344–350, 1988. 14, 18
- [Jaffe *et al.* 2001] J. S. Jaffe, K. D. Moore, J. McLean and M. P. Strand. *Underwater optical imaging: Status and prospects*. Oceanography, vol. 14, no. 3, pages 64–75, 2001. 15
- [Jaffe 1990] J. S. Jaffe. *Computer modeling and the design of optimal underwater imaging systems*. IEEE Journal of Oceanic Engineering, vol. 15, no. 2, pages 101–111, 1990. 17, 18
- [Jalkio *et al.* 1985] J. A. Jalkio, R. C. Kim and S. K. Case. *Three dimensional inspection using multistriple structured light*. Optical Engineering, vol. 24, pages 966–974, December 1985. 18
- [Jenkin *et al.* 2007] M. Jenkin, A. Hogue, A. Germam, S. Gill, A. Topol and S. Wilson. *Underwater surface recovery and segmentation*. In Proceedings of the 6th IEEE International Conference on Cognitive Informatics, ICCI 2007, pages 373–380, 2007. 16
- [Karpel & Schechner 2004] N. Karpel and Y. Y. Schechner. *Portable polarimetric underwater imaging system with a linear response*. In Proceedings of SPIE - The International Society for Optical Engineering, volume 5432, pages 106–115, 2004. ix, 28
- [Kemao 2004] Q. Kemao. *Windowed Fourier Transform for Fringe Pattern Analysis*. Appl. Opt., vol. 43, no. 13, 2004. 44
- [Kennedy *et al.* 2004] C.J. Kennedy, J.C. Hiller, D. Lammie, M. Drakopoulos, M. Vest, W.P. Cooper M. Adderley and T.J. Wess. *Microfocus X-ray diffraction of historical parchment reveals variations in structural features through parchment cross-sections*. NanoLetters, vol. 4, no. 8, 2004. 8

- [Klein *et al.* 2008] M. E. Klein, B. J. Aalderink, R. Padoan, G. de Bruin and Ted A.G. Steemers. *Quantitative Hyperspectral Reflectance Imaging*. Sensors, vol. 8, no. 9, pages 5576–5618, 2008. x, 60
- [Klepsvik *et al.* 1990] J. O. Klepsvik, H. O. Torsen and K. Thoresen. *Laser imaging technology for subsea inspection: principles and applications*. In IRMI/ROV90 Conference, 1990. 15
- [Knox 1998] K. Knox. *Show-Through Correction for Two-Sided Documents*, November 1998. United States Patent 5,832,137. 77
- [Kocak & Caimi 2005] D. M. Kocak and F. M. Caimi. *The current art of underwater imaging - with a glimpse of the past and vision of the future*. Marine Technology Society Journal, vol. 39, no. 3, pages 5–26, 2005. 13, 15
- [Kocak *et al.* 2008] D. M. Kocak, F. R. Dalgleish, F. M. Caimi and Y. Y. Schechner. *A focus on recent developments and trends in underwater imaging*. Marine Technology Society Journal, vol. 42, no. 1, pages 52–67, 2008. 13, 15
- [Krantz 1999] S.C. Krantz. *The cauchy-riemann equations, handbook of complex analysis*. Birkhauser, 1999. 107
- [Krattenthaler *et al.* 1993] W. Krattenthaler, K. J. Mayer and H. P. Duwe. *3D-surface measurement with coded light approach*. In Proceedings of the 17th meeting of the Austrian Association for pattern recognition on Image analysis and synthesis, pages 103–114, Munich, Germany, Germany, 1993. R. Oldenbourg Verlag GmbH. 20
- [Kwiatkowska *et al.* 2008] E. Kwiatkowska, M. Sylwestrzak, B. J. Rouba, L. Tyminska-Widmer, M. Iwanicka and P. Targowski. *Optical Coherence Tomography for non-destructive investigations of structure of easel paintings*. In 1st Canterbury Workshop on Optical Coherence Tomography and Adaptive Optics, volume 7139. SPIE, 2008. 9
- [Legendijk *et al.* 1990] R. Legendijk, J. Biemond and D. Boekee. *Identification and restoration of noisy blurred images using the expectation-maximization algorithm*. IEEE Trans. Acoust. Speech Signal Process., vol. 38, no. 7, pages 1180– 1191, 1990. 96
- [Leévy *et al.* 2002] B. Leévy, S. Petijean, R. Nicholas and J. Maillot. *Least Squares Conformal Maps for Automatic Texture Atlas Generation*. In Proc. Int’l Conf. Computer Graphics and Interactive Techniques, pages 362–371, 2002. 107

- [Leone *et al.* 2006] A. Leone, G. Diraco and C. Distanto. *Stereoscopic system for 3-D seabed mosaic reconstruction*. In Proceedings - International Conference on Image Processing, ICIP, volume 2, pages II541–II544, 2006. 16
- [Lettner *et al.* 2007] M. Lettner, R. Diem M. Sablatnig, P. Kammerer and H. Miklas. *Registration of Multispectral Manuscript Images as Prerequisite for Computer Aided Script Description*. In 12th Computer Vision Winter Workshop, St. Lambrecht, Austria, February 2007. 68
- [Lettner *et al.* 2008] M. Lettner, M. Diem, R. Sablatnig and H. Miklas. *Registration and enhancing of multispectral manuscript images*. In 16th European Signal Processing Conference (EUSIPCO 2008), Lausanne - Switzerland, 2008. 60
- [Levoy 1999] M. Levoy. *The Digital Michelangelo Project*. In Proceedings of the Second International Conference on 3D Digital Imaging and Modeling, Ottawa, Canada, 5-8 October 1999. 2
- [Lewis *et al.* 1999] G. D. Lewis, D. L. Jordan and P. J. Roberts. *Backscattering target detection in a turbid medium by polarization discrimination*. Applied Optics, vol. 38, no. 18, pages 3937–3944, 1999. 19
- [Leydier *et al.* 2004] Y. Leydier, F. Le Bourgeois and H. Emptoz. *Serialized unsupervised classifier for adaptive color image segmentation: application to digitized ancient manuscripts*. In in Proc. Int. Conf. on Pattern Recognition, page 494497, 2004. 77
- [Leydier *et al.* 2005] Y. Leydier, F. Le Bourgeois and H. Emptoz. *Omnilingual Segmentation-free Word Spotting for Ancient Manuscripts Indexation*. Document Analysis and Recognition, International Conference on, pages 533–537, 2005. 69
- [Lin & Lin 1999] C.F. Lin and C.Y. Lin. *A new approach to high precision 3-D measuring system*. Image Vision Comput., vol. 17, no. 11, pages 805–814, 1999. 46
- [Liu 2001] H. T. Liu. *A video-based stereoscopic imaging and measurement system (SIMS) for undersea applications*. In OCEANS, 2001. MTS/IEEE Conference and Exhibition, volume 1, pages 275–286, Honolulu, HI, 2001. 30
- [Lognoli *et al.* 2003] D. Lognoli, G. Cecchi and I. Mochi. *Fluorescence lidar imaging of the cathedral and baptistery of Parma*. Applied Physics B:Lasers and Optics, vol. 76, no. 4, page 457465, 2003. 8

- [Lowe 2004] D. G. Lowe. *Distinctive Image Features from Scale-Invariant Keypoints*. International Journal of Computer Vision, vol. 60, no. 2, pages 91–110, 2004. 16
- [Ma *et al.* 2003] Y. Ma, S. Soatto, J. Kosecka and S. Sastry. An invitation to 3d vision: From images to geometric models. Springer Verlag, 2003. 16
- [Manmatha & Rothfeder 2005] R. Manmatha and Jamie L. Rothfeder. *A Scale Space Approach for Automatically Segmenting Words from Historical Handwritten Documents*. IEEE Trans. Pattern Anal. Mach. Intell., vol. 27, no. 14, pages 1212–1225, August 2005. 69
- [Mansfielda *et al.* 2002] J. R. Mansfielda, M. Attasa, C. Majzelsb, E. Cloutisc, C. Collinsd and H. H. Mantscha. *Near infrared spectroscopic reflectance imaging: a new tool in art conservation*. Vibrational Spectroscopy, vol. 28, page 5966, 2002. 7
- [Mansouri & Marzani 2005] A. Mansouri and F. S. Marzani. *Optical Calibration of a Multispectral Imaging System Based on Interference Filters*. Opt. Eng., vol. 44, no. 2, 2005. 62
- [Mansouri *et al.* 2007] A. Mansouri, A. Lathuili're, F. S. Marzani, Y. Voisin and Gouton P. *Toward a 3D multispectral scanner: an application to multimedia*. IEEE MultiMedia, vol. 14, no. 1, pages 40–47, January 2007. 10
- [Martinez *et al.* 2002] K. Martinez, J. Cupitt, D. Saunders and R. Pillay. *Ten years of art imaging research*. Proceedings of the IEEE, vol. 90, no. 1, pages 28–41, 2002. 68
- [Martinez 1991] K. Martinez. *High Resolution Digital Imaging of Paintings: The Vasari Project*. Microcomputers for Information Management, vol. 8, no. 4, pages 277–283, December 1991. 3
- [Melessanaki *et al.* 2001] K. Melessanaki, V. Papadakis, C. Balas and Anglos D. *Laser induced breakdown spectroscopy and hyper-spectral imaging analysis of pigments on an illuminated manuscript*. Spectrochim. Acta B, vol. 56, pages 2337–2346, 2001. 7
- [Menci-Software 2010] Menci-Software. <http://www.menci.com>, 2010. (Accessed July 2010). 17
- [Mikhail *et al.* 2001] E. M. Mikhail, J. S. Bethel and J. C. McGlone. Introduction to modern photogrammetry. John Wiley and Sons, 2001. 16

- [Muller *et al.* 2006] M. Muller, B. Murphy, M. Burghammer, C. Riekell, M. Roberts, M. Papiz, D. Clarke, J. Gunneweg and E. Pantos. *Identification of ancient textile fibres from Khirbet Qumran caves using synchrotron radiation microbeam diffraction*. Spectrochimica Acta B: Atomic Spectroscopy, 2006. 8
- [Nan & Dazheng 2004] W. Nan and F. Dazheng. *inSAR phase unwrapping algorithm using branch-cut and finite element*. In ICSP 04 proceeding, IEEE, 2004. 48
- [Narasimhan *et al.* 2008] S. G. Narasimhan, S. K. Nayar, B. Sun and S. J. Koppal. *Structured light in scattering media*. In SIGGRAPH Asia '08: ACM SIGGRAPH ASIA 2008 courses, pages 1–8, New York, NY, USA, 2008. ACM. 19
- [Navarro *et al.* 2009] S. Navarro, A.E. Segui, C. Portales, J.L. Lerma, T. Akasheh and N. Haddad. *Integration of TLS Data and Non-metric Imagery to Improve Photo Models and Recording - A Case Study on Djin Block No. 9, Petra (Jordan)*. Virtual Systems and MultiMedia, International Conference on, pages 58–63, 2009. 10
- [Negahdaripour *et al.* 2001] S. Negahdaripour, H. Zhang, P. Firoozfam and J. Oles. *Utilizing panoramic views for visually guided tasks in underwater robotics applications*. In Oceans Conference Record (IEEE), volume 4, pages 2593–2600, 2001. 17
- [Negahdaripour *et al.* 2002] S. Negahdaripour, H. Zhang and X. Han. *Investigation of photometric stereo method for 3D shape recovery from underwater imagery*. In Oceans Conference Record (IEEE), volume 2, pages 1010–1017, 2002. 17
- [Niblack 1986] W. Niblack. *An introduction to image processing*, pages 115–116. Prentice-Hall, Englewood Cliffs, NJ, 1986. 92
- [Nishida & Suzuki 2003] H. Nishida and T. Suzuki. *A Multiscale Approach to Restoring Scanned Color Document Images with Show-Through Effects*. In in Proc. ICDAR 2003, 2003. 76
- [Novati *et al.* 2005] G. Novati, P. Pellegrini and R. Schettini. *An affordable multispectral imaging system for the digital museum*. International Journal on Digital Libraries, vol. 5, no. 3, pages 167–178, 2005. 8
- [OCR 1997] Caere Ominpage OCR. *Users Manual*. Caere Corp., 1997. 93

- [Ohta *et al.* 1980] Y. Ohta, T. Kanade and T. Sakai. Computer graphics, vision, and image processing, chapitre 13, page 222241. 1980. 82
- [Oxf 2003] Oxford dictionary of english. Oxford University Press, 2th édition, 2003. 5
- [P. Hellier 2000] C. Barillot P. Hellier. *Coupling dense and landmark-based approaches for non rigid registration*. Rapport technique, IRISA research report, France, 2000. 73
- [Padoan *et al.* 2008] R. Padoan, Th. A.G. Steemers, M.E. Klein, B.J. Aalderink and G. de Bruin. *Quantitative Hyperspectral Imaging of Historical Documents: Technique and Application*. ART Proceedings, 2008. 7
- [Parker 1991] J. Parker. *Gray level thresholding on badly illuminated images*. In IEEE Trans. Pattern Anal. Mach. Intell., volume 13, pages 813–819, 1991. 92
- [Pelagotti *et al.* 2006] A. Pelagotti, L. Pezzati, A. Piva and A. Del Mastio. *Multispectral UV fluorescence analysis of painted surfaces*. In 14th European Signal Processing Conference (EUSIPCO 2006), Florence - Italy, September 4-8 2006. 7
- [Pelagotti *et al.* 2007] A. Pelagotti, A. Del Mastio and A.V. Rationale. *Active and passive sensors for art works analysis and investigations*. In Videometrics IX, volume 6491, page 64910R. SPIE, 2007. 103
- [Pelagotti *et al.* 2009] A. Pelagotti, A. Del Mastio, F. Uccheddu and F. Remondino. *Automated multispectral texture mapping of 3D models*. In 17th European Signal Processing Conference (EUSIPCO), Glasgow - Scotland, 24–28 August 2009. 103
- [Perardi *et al.* 2003] A. Perardi, L. Appolonia and P. Mirti. *Non-destructive in situ determination of pigments in 15th century wall paintings by Raman microscopy*. Analytica Chimica Acta, vol. 480, no. 2, pages 317 – 325, 2003. 9
- [Pizarro *et al.* 2004] O. Pizarro, R. Eustice and H. Singh. *Large area 3D reconstructions from underwater surveys*. In Ocean '04 - MTS/IEEE Techno-Ocean '04: Bridges across the Oceans - Conference Proceedings, volume 2, pages 678–687, 2004. 17

- [Posdamer & Altschuler 1982] J. L. Posdamer and M. D. Altschuler. *Surface measurement by space-encoded projected beam systems*. Computer Graphics and Image Processing, vol. 18, no. 1, pages 1 – 17, 1982. 20
- [Quan *et al.* 2001] C. Quan, X. Y. He, C. F. Wang, C. J. Tay and H. M. Shang. *Shape measurement of small objects using LCD fringe projection with phase shifting*. Optics Communications, vol. 189, no. 1-3, pages 21 – 29, 2001. 47
- [Rapantzikos & Balas 2005a] K. Rapantzikos and C. Balas. *Hyperspectral imaging: potential in non-destructive analysis of palimpsests*. In ICIP (2), pages 618–621, 2005. 60
- [Rapantzikos & Balas 2005b] K. Rapantzikos and C. Balas. *Hyperspectral imaging: potential in non-destructive analysis of palimpsests*. In Image Processing, 2005. ICIP 2005. IEEE International Conference on, volume 2, pages II – 618–21, 2005. 69
- [Rath & Manmatha 2003] T. M. Rath and R. Manmatha. *Features for Word Spotting in Historical Manuscripts*. Document Analysis and Recognition, International Conference on, vol. 1, page 218, 2003. 69
- [Reddy & Chatterji 1996] B. S. Reddy and B. N. Chatterji. *An FFT-based technique for translation, rotation and scale-invariant image registration*. IEEE Trans. on Image Processing, vol. 5, no. 8, pages 1266–1271, 1996. 73
- [Reeves & Mersereau 1992] S. Reeves and R. Mersereau. *Blur identification by the method of generalized cross-validation*. IEEE Trans. Image Processing, vol. 1, no. 3, pages 301–311, 1992. 96
- [Remondino & El-Hakim 2006a] F. Remondino and S. El-Hakim. *Image-based 3D modelling: a review*. The Photogrammetric Record, vol. 21, no. 115, pages 269–291, September 2006. 9, 10
- [Remondino & El-hakim 2006b] F. Remondino and S. El-hakim. *Image-based 3D modelling: A review*. Photogrammetric Record, vol. 21, no. 115, pages 269–291, 2006. 15, 17
- [Rezaie & Srinath 1984] B. Rezaie and M.D. Srinath. *Algorithms for fast image registration*. IEEE Transactions on Aerospace and Electronic Systems, vol. 20, page 716728, 1984. 73



- [Ribés *et al.* 2005] A. Ribés, 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. 8, pages 563–573, 2005. 60
- [Rives & Borrelly 1997] P. Rives and J.-J. Borrelly. *Underwater pipe inspection task using visual servoing techniques*. In *IEEE International Conference on Intelligent Robots and Systems*, volume 1, pages 63–68, Grenoble - France, September 1997. 18, 39
- [Rizzi *et al.* 2006] A. Rizzi, F. Voltolini, F. Remondino, S. Girardi and L. Gonzo. *16 Cultural Heritage OPTICAL MEASUREMENT TECHNIQUES FOR THE DIGITAL PRESERVATION, DOCUMENTATION AND ANALYSIS OF CULTURAL HERITAGE*, 2006. 10
- [Rocchini *et al.* 2001a] C. Rocchini, P. Cignoni, C. Montani, P. Pingi and R. Scopigno. *A low cost 3D scanner based on structured light*. *Computer Graphics Forum*, vol. 20, no. 3, pages C/299–C/308, 2001. 19
- [Rocchini *et al.* 2001b] C. Rocchini, P. Cignoni, C. Montani, P. Pingi, R. Scopigno, R. Fontana, M. Greco, E. Pampaloni, L. Pezzati, M. Cygielman, R. Giachetti, G. Gori, M. Miccio and R. Pecchioli. *3D Scanning the Minerva of Arezzo*. In *IN ICHIM2001 CONF. PROC., VOL.2*, pages 265–272, 2001. 9
- [Roche *et al.* 2006] H. Roche, R. Tadeu, C. Camerini, C. Maia and Alves Marinho C. *Gamma ray system operated by robots for underwater pipeline inspection*. In *ECNDT 2006*, Berlin - Germany, 2006. 39
- [Roman *et al.* 2010] C. Roman, G. Inglis and J. Rutter. *Application of structured light imaging for high resolution mapping of underwater archaeological sites*. In *IEEE Oceans*, Sydney, 2010. 18
- [Rudin *et al.* 1992] L. Rudin, S. Osher and E. Fatemi. *Nonlinear total variation based noise removal algorithms*. *Physica D*, vol. 60, pages 259–268, 1992. 96, 98
- [Saito *et al.* 1995] H. Saito, H. Kawamura and M. Nakajima. *3D shape measurement of underwater objects using motion stereo*. In *IECON Proceedings (Industrial Electronics Conference)*, volume 2, pages 1231–1235, 1995. 17
- [Salvi *et al.* 2004] J. Salvi, J. Pages and J. Batlle. *Pattern codification strategies in structured light systems*. *Pattern Recognition*, vol. 37, no. 4, pages 827–849, 2004. 19



- [Sansoni *et al.* 1997] G. Sansoni, S. Corini, S. Lazzari, R. Rodella and F. Docchio. *Three-dimensional imaging based on Gray-code light projection: Characterization of the measuring algorithm and development of a measuring system for industrial applications*. Applied Optics, vol. 36, no. 19, pages 4463–4472, 1997. 20
- [Sauvola & Pietaksinen 2000] J. Sauvola and M. Pietaksinen. *Adaptive document image binarization*. Pattern Recogn, vol. 33, pages 225–236, 2000. xi, 92, 93, 94, 111
- [Sauvola *et al.* 1997] J. Sauvola, S. Haapakoski, H. Kauniskangas, T. SeppäKinen, M. PietikaKinen and D. Doermann. *A distributed management system for testing document image analysis algorithms*. In 4th ICDAR, pages 989–995, Germany, 1997. 93
- [Scharstein & Szeliski 2001] D. Scharstein and R. Szeliski. *A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms*. International Journal of Computer Vision, vol. 47, pages 7–42, 2001. 16
- [Schechner & Karpel 2004] Y. Y. Schechner and N. Karpel. *Clear underwater vision*. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, volume 1, pages I536–I543, 2004. 6, 14, 19
- [Schewe *et al.* 1996] H. Schewe, E. Moncreiff and L. Gruendig. *Improvement of fish farm pen design using computational structural modelling and large-scale underwater photogrammetry*. International Archives Photogrammetry and Remote Sensing, vol. 31, no. B5, pages 524–529, 1996. 30
- [Sezgin & Sankur 2004] M. Sezgin and B. Sankur. *Survey over image thresholding techniques and quantitative performance evaluation*. Journal of Electronic Imaging, vol. 13, no. 1, pages 146–165, January 2004. 92
- [Sharma 2001] G. Sharma. *Show-through cancellation in scans of duplex printed documents*. IEEE Trans. on Image Processing, vol. 10, page 736754, 2001. 77, 79
- [Shi *et al.* 2005] Z. Shi, S. Setlur and V. Govindaraju. *Text extraction from gray scale historical document images using adaptive local connectivity map*. In Document Analysis and Recognition, 2005. Proceedings. Eighth International Conference on, volume 2, pages 794 – 798, 2005. 69

- [Shortis *et al.* 2007] M. R. Shortis, E. S. Harvey and J. W. Seager. *A review of the status and trends in underwater videometric measurement*. In SPIE Conference 6491, Videometrics IX, page 26, 2007. 15
- [Surrel 2000] Y. Surrel. *Fringe Analysis*. In Pramod Rastogi, editeur, *Photomechanics*, volume 77 of *Topics in Applied Physics*, pages 55–102. Springer Berlin / Heidelberg, 2000. 41
- [Takeda & Mutoh 1983] M. Takeda and K. Mutoh. *Fourier transform profilometry for the automatic measurement of 3-D object shapes*. *Appl. Opt.*, vol. 22, no. 24, pages 3977–3982, Dec 1983. 42
- [Tan *et al.* 2005] C. Tan, G. Seet, A. Sluzek and D. He. *A novel application of range-gated underwater laser imaging system (ULIS) in near-target turbid medium*. *Optics and Lasers in Engineering*, vol. 43, no. 9, pages 995–1009, 2005. 19
- [Tang *et al.* 2001] C.C. Tang, E.J. MacLean, M.A. Roberts, D.T. Clarke, E. Pantos and A.J.N.W. Prag. *The study of Attic black gloss sherds using synchrotron X-ray diffraction*. *Journal of Archaeological Science*, vol. 28, no. 10, 2001. 8
- [Taylor *et al.* 1975] K. K. Taylor, M. J. Cotter and E. V. Sayre. *Neutron Activation Autoradiography as a Technique for Conservation Examination of Paintings*, 1975. *Bulletin of the American Institute for Conservation of Historic and Artistic Works*. 8
- [Tetlow & Allwood 1994] S. Tetlow and R. L. Allwood. *The use of a laser stripe illuminator for enhanced underwater viewing*. In *Proc. SPIE*, volume 2258, pages 547–55, 1994. 18
- [Tetlow & Allwood 1995] S. Tetlow and R. L. Allwood. *Development and applications of a novel underwater laser illumination system*. *Underwater Technology*, vol. 21, pages 13–20, 1995. 18
- [Tetlow & Spours 1999] S. Tetlow and J. Spours. *Three-dimensional measurement of underwater work sites using structured laser light*. *Measurement Science and Technology*, vol. 10, no. 12, pages 1162–1167, 1999. 14, 39
- [Tomasi & Kanade 1992] C. Tomasi and T. Kanade. *Shape and motion from image streams under orthography: a factorization method*. *Int. J. Comput. Vision*, vol. 9, no. 2, pages 137–154, 1992. 17

- [Tonazzini *et al.* 2004a] A. Tonazzini, L. Bedini and E. Salerno. *Independent component analysis for document restoration*. Int. J. Doc. Anal. Recognit., vol. 7, pages 17–27, March 2004. 3, 68, 69, 79, 85
- [Tonazzini *et al.* 2004b] A. Tonazzini, E. Salerno, M. Mochi and Bedini L. *Blind Source Separation Techniques for Detecting Hidden Texts and Textures in Document Images*. In ICIAR (2), pages 241–248, 2004. 84
- [Tonazzini *et al.* 2004c] A. Tonazzini, S. Vezzosi and L. Bedini. *Analysis and recognition of printed characters in ancient documents*. Int. Journal on Document Analysis and Recognition, vol. 6, pages 236–247, 2004. 7, 69
- [Tonazzini *et al.* 2007] A. Tonazzini, E. Salerno and L. Bedini. *Fast correction of bleed-through distortion in grayscale documents by a Blind Source Separation technique*. IJDAR, vol. 10, pages 17–25, June 2007. 85, 87, 90
- [Tonazzini *et al.* 2009] A. Tonazzini, G. Bianco and E. Salerno. *Registration and Enhancement of Double-Sided Degraded Manuscripts Acquired in Multispectral Modality*. In Document Analysis and Recognition, 2009. ICDAR '09. 10th International Conference on, pages 546–550, July 2009. 85
- [Triggs *et al.* 2000] B. Triggs, P. F. McLauchlan, R. I. Hartley and A. W. Fitzgibbon. *Bundle Adjustment - A Modern Synthesis*. In ICCV '99: Proceedings of the International Workshop on Vision Algorithms, pages 298–372, London, UK, 2000. Springer-Verlag. 17
- [Turner 1993] J. Turner. *Development of an operational digital photogrammetric system for the North Sea oil and gas industry*. In Proceedings of SPIE - The International Society for Optical Engineering, volume 1820, pages 136–144, 1993. 30
- [Unesco 1972] Unesco. *Art.1 - Convention concerning the protection of the world cultural and natural heritage*, 16 November 1972. 1
- [Unesco 2001] Unesco. *Convention on the protection of the underwater cultural heritage*, 2 November 2001. 2
- [Virtuale 2010] Rinascimento Virtuale. [www.bml.firenze.sbn.it/rinascimentovirtuale](http://www.bml.firenze.sbn.it/rinascimentovirtuale), 2010. (Accessed July 2010). 68

- [Šorel & Šroubek 2009] M. Šorel and F. Šroubek. *Space-variant deblurring using one blurred and one underexposed image*. In In Proc. Proceedings of the IEEE 16th International Conference on Image Processing ICIP 2009, Cairo, Egypt, 2009. 97
- [Šorel *et al.* 2008] M. Šorel, and J. Flusser. *Space-variant restoration of images degraded by camera motion blur*. IEEE Transactions on Image Processing, vol. 17, no. 2, pages 105–116, 2008. 96
- [Šroubek & Flusser 2005] F. Šroubek and J. Flusser. *Multichannel blind deconvolution of spatially misaligned images*. IEEE Trans. Image Processing, vol. 14, no. 7, pages 874–883, 2005. 96, 98
- [Šroubek *et al.* 2009] F. Šroubek, M. Šorel, J. Boldyš and J. Šroubek. *PET Image Reconstruction Using Prior Information from CT or MRI*. In In Proc. Proceedings of the IEEE 16th International Conference on Image Processing ICIP 2009, Cairo, Egypt, 2009. 98
- [Wang & Tan 2001] Q. Wang and C. L. Tan. *Matching of double-sided document images to remove interference*. In in Proc. IEEE CVPR 2001, 2001. 77
- [Wang *et al.* 2003] Q. Wang, L. Xia T. Li and C. L. Tan. *Document Image Enhancement Using Directional Wavelet*. In in Proc. IEEE Conf. Computer Vision Pattern Recognition, volume 2, page 534539, 2003. 76
- [Ware *et al.* 2000] G.A. Ware, D.M. Chabries, R.W. Christiansen and C.E. Martin. *Multispectral document enhancement: ancient carbonized scrolls*. In Geoscience and Remote Sensing Symposium, Proceedings. IGARSS 2000, volume 6, pages 2486 –2488, 2000. 68
- [Weibring *et al.* 2001] P. Weibring, T. Johansson and H. Edner. *Fluorescence lidar imaging of historical monuments*. Applied Optics, vol. 40, no. 33, page 61116120, 2001. 8
- [Weiner *et al.* 1980] S. Weiner, Z. Kustanovich, E. Gil-Av and W. Traub. *Dead Sea Scroll parchments: unfolding of the collagen molecules and racemization of aspartic acid*. Nature, vol. 287, pages 820–823, 1980. 8
- [Wolf 2006] C. Wolf. *Document ink bleed-through removal with two hidden Markov random fields and a single observation field*. Report technique, Laboratoire d’Informatique en Images et Syst’ mes d’Information, INSA de Lyon France, 2006. 77

- [Woll *et al.* 2005] A. R. Woll, D. H. Bilderback, S. Gruner, N. Gao, R. Huang, C. Bisulca and J. Mass. *Confocal X-ray Fluorescence(XRF) Microscopy: A New Technique for the Nondestructive Compositional Depth Profiling of Paintings*. In Materials Issues in Art and Archaeology VII: Materials Research Society Symposium, volume 3, Boston - USA, November 30-December 2005. 8
- [Yamashita *et al.* 2004] A. Yamashita, A. Kawarago, T. Kaneko and K. T. Miura. *Shape reconstruction and image restoration for non-flat surfaces of documents with a stereo vision system*. In Pattern Recognition, International Conference on, page 482485, 2004. 103
- [Yamashita *et al.* 2007] A. Yamashita, M. Fujii and T. Kaneko. *Color Registration of Underwater Images for Underwater Sensing with Consideration of Light Attenuation*. In ICRA, pages 4570–4575, 2007. 35
- [Yosef *et al.* 2004] I.B. Yosef, K. Kedem, I. Dinstein, M. Beit-Arie and E. Engel. *Classification of Hebrew calligraphic handwriting styles: preliminary results*. In Document Image Analysis for Libraries, 2004. Proceedings. First International Workshop on, pages 299 – 305, 2004. 69
- [You & Kaveh 1999] Y.-L. You and M. Kaveh. *Blind image restoration by anisotropic regularization*. IEEE Trans. Image Processing, vol. 8, no. 3, pages 396–407, 1999. 96
- [Yuh 2000] J. Yuh. *Design and control of autonomous underwater robots: A survey*. Autonomous Robots, vol. 8, no. 1, pages 7–24, 2000. 13
- [Zhang & Negahdaripour 2003] H. Zhang and S. Negahdaripour. *On reconstruction of 3D volumetric models of reefs and benthic structures from image sequences of a stereo rig*. In Oceans Conference Record (IEEE), volume 5, pages 2553–2559, 2003. 17
- [Zhong & Weng 2004] J. Zhong and J. Weng. *Spatial Carrier-Fringe Pattern Analysis by Means of Wavelet Transform: Wavelet Transform Profilometry*. Applied Optics IP, vol. 43, pages 4993–4998, 2004. 46
- [Zitová & Flusser 2003] B. Zitová and J. Flusser. *Image registration methods: a survey*. Image and Vision Computing, vol. 21, no. 11, pages 977–1000, October 2003. 70