



Universitat de Girona

COMPUTER VISION TECHNIQUES FOR EARLY DETECTION OF SKIN CANCER

Josep QUINTANA PLANA

Dipòsit legal: GI. 1098-2012

<http://hdl.handle.net/10803/82072>

ADVERTIMENT. L'accés als continguts d'aquesta tesi doctoral i la seva utilització ha de respectar els drets de la persona autora. Pot ser utilitzada per a consulta o estudi personal, així com en activitats o materials d'investigació i docència en els termes establerts a l'art. 32 del Text Refós de la Llei de Propietat Intel·lectual (RDL 1/1996). Per altres utilitzacions es requereix l'autorització prèvia i expressa de la persona autora. En qualsevol cas, en la utilització dels seus continguts caldrà indicar de forma clara el nom i cognoms de la persona autora i el títol de la tesi doctoral. No s'autoritza la seva reproducció o altres formes d'explotació efectuades amb finalitats de lucre ni la seva comunicació pública des d'un lloc aliè al servei TDX. Tampoc s'autoritza la presentació del seu contingut en una finestra o marc aliè a TDX (framing). Aquesta reserva de drets afecta tant als continguts de la tesi com als seus resums i índexs.

ADVERTENCIA. El acceso a los contenidos de esta tesis doctoral y su utilización debe respetar los derechos de la persona autora. Puede ser utilizada para consulta o estudio personal, así como en actividades o materiales de investigación y docencia en los términos establecidos en el art. 32 del Texto Refundido de la Ley de Propiedad Intelectual (RDL 1/1996). Para otros usos se requiere la autorización previa y expresa de la persona autora. En cualquier caso, en la utilización de sus contenidos se deberá indicar de forma clara el nombre y apellidos de la persona autora y el título de la tesis doctoral. No se autoriza su reproducción u otras formas de explotación efectuadas con fines lucrativos ni su comunicación pública desde un sitio ajeno al servicio TDR. Tampoco se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR (framing). Esta reserva de derechos afecta tanto al contenido de la tesis como a sus resúmenes e índices.

WARNING. Access to the contents of this doctoral thesis and its use must respect the rights of the author. It can be used for reference or private study, as well as research and learning activities or materials in the terms established by the 32nd article of the Spanish Consolidated Copyright Act (RDL 1/1996). Express and previous authorization of the author is required for any other uses. In any case, when using its content, full name of the author and title of the thesis must be clearly indicated. Reproduction or other forms of for profit use or public communication from outside TDX service is not allowed. Presentation of its content in a window or frame external to TDX (framing) is not authorized either. These rights affect both the content of the thesis and its abstracts and indexes.



Universitat de Girona

PhD Thesis

Computer Vision Techniques for Early Detection of Skin Cancer

Josep Quintana

2012



Universitat de Girona

PhD Thesis

Computer Vision Techniques for Early Detection of Skin Cancer

Josep Quintana

2012

Doctoral Programme in Technology

Supervised by: Rafael Garcia

Work submitted to the University of Girona in fulfilment of the requirements for
the degree of Doctor of Philosophy

Abstract

Although dermoscopy is nowadays a well-established practice for dermatologists in order to diagnose melanocytic tumors, full body exams using dermoscopy involves a long, time-consuming examination for each patient, and images from different explorations have to be compared every few months in order to detect changes over time. The utilization of this technique has demonstrated its utility in reducing the number of biopsies and allowing the early diagnosis of melanoma. Unfortunately, it is a tedious, slow and costly process, since each exploration needs a significant amount of time of an expert. Moreover, human-based visual inspection is prone to errors due to inattention of the physician, and only a limited number of dermatoscopical images are acquired for each patient, leaving several lesions unrecorded on patients with a high number of nevi.

This thesis investigates the problem of developing new computer vision techniques for early detection of skin cancer.

The first part of this work presents a novel methodology to correct color reproduction in dermatological images when different cameras and/or dermoscopes are used. The proposed algorithm includes spectral lighting information, contributing to the state-of-the-art in the dermatological color calibration with a novel formulation.

Next, the problem of automatic full body mapping is addressed by proposing a mosaicing method based on an off-the-shelf digital compact camera and a set of markers. This method increases the possibilities of total body photography by taking the low-resolution images of a whole body exploration and automatically combining them into a high-resolution photomosaic. In this way, a full body exploration is registered into a single image, simplifying the task of comparing different explorations, either analyzing the mosaics of both explorations or comparing the corresponding images between explorations.

The third contribution of this work consists of the development of a full body scanner for acquiring cutaneous images. On one hand,

the scanner reduces the long time-consuming examinations done in dermoscopy explorations, and on the other hand, it increases the resolution of total body photography systems. The proposed optical scanner automatically acquires images of the entire body, and then represents them either in a 2D mosaic or on a dense 3D reconstruction. Throughout this dissertation, the performance of the algorithms is evaluated using real skin images, including dermoscopy, traditional macro photography, and cross-polarized imaging.

Resum

Encara que la dermatòscopia sigui avui en dia una pràctica molt utilitzada pels dermatòlegs a l'hora de diagnosticar tumors melano-cítics, es tracta d'una pràctica tediosa i llarga per a cada pacient, on les imatges de diferents exploracions han de ser comparades en pocs mesos de diferència amb la finalitat de detectar canvis en el temps. La utilització d'aquesta tècnica ha demostrat la seva utilitat en la reducció del nombre de biòpsies, permetent així, la detecció precoç del càncer de pell. Desafortunadament, és un procés lent i costós, ja que per a cada exploració és necessària una quantitat significativa de temps per part d'un expert. Al mateix temps, la inspecció visual és propensa a errors deguts a la falta d'atenció per part del metge, i al limitat nombre d'imatges dermatoscòpiques que es poden adquirir per a cada pacient, restant algunes lesions sense examinar en pacients amb un gran nombre de nevus.

En aquesta tesi s'investiga el desenvolupament de noves tècniques de visió per computador per a la detecció del càncer de pell.

La primera part d'aquest treball presenta una nova metodologia per a la correcció del color en imatges dermatològiques quan s'utilitzen diferents càmeres i/o els dermatoscops. L'algorisme proposat inclou informació espectral sobre la il·luminació emprada, contribuint així amb una nova formulació en l'estat de l'art dins el camp del calibratge de color en dermatologia.

A continuació es proposa una solució al problema del registre automàtic d'imatges de cos complet amb la proposta d'un mètode de *mosaicing* basat en l'ús de càmeres compactes i un conjunt de *markers*. Aquest mètode incrementa les possibilitats de la fotografia de cos complet mitjançant l'acquisició d'imatges a baixa resolució de tot el cos i posteriorment combinant-les automàticament per a l'obtenció d'un fotomosaic d'alta resolució. D'aquesta manera, una exploració de cos complet pot ser registrada en una sola imatge, simplificant així la tasca de comparar diferents exploracions, ja sigui analitzant

els mosaics de les dues exploracions o comparant-ne les imatges corresponents.

La tercera contribució d'aquest treball consisteix en el desenvolupament d'un escàner de cos complet per a l'adquisició d'imatges cutànies. D'una banda l'escàner redueix el llarg temps necessari per a les exploracions dermatoscòpiques, i de l'altre, incrementa la resolució de la fotografia de cos complet. L'escàner òptic proposat adquireix automàticament imatges de cos complet, i després proporciona resultats ja sigui en mosaics 2D o en reconstruccions denses 3D.

Al llarg d'aquesta tesi, el rendiment dels algoritmes s'ha avaluat mitjançant l'ús d'imatges reals de la pell, incloent imatges dermatoscòpiques, imatges macro i imatges amb polarització creuada.

I would like to dedicate this thesis to my parents ...

Acknowledgements

I would like to express my gratitude to the people who have supported me during the elaboration of this PhD thesis. First and foremost to my advisor; Rafa García, for believing in me and encouraging me during this research with his unshakable optimism, inspiration and friendship.

I also thank all the members of the Underwater Vision Lab for their help whenever it was necessary. I want to thank my colleagues Nuno, Tudor, Ricard P., Ricard C., Konstantin, Jordi, Ramon, László, Olivier, Armagan, Quim and Shihav for always being there. I would also like to extend my gratitude to the rest of the members of the Computer Vision and Robotics Group, in particular to Jordi F., Xevi C., Pere, Marc, Lluís, Miki, Joan B., Josep F., Quim, Xavi L., Marta, Robert and Arnau. I should not forget all the PhD members that I meet on the group, Albert T., Sik, Meritxell, Sergio, Luca, Josep A., Mariano, Marc M., Onur, Carles, Eloy, Christian, Albert G., Gerard, Tali, Aggelos, David, Emili, Narcís, Bladimir, Javi, Enric, Andrés, ... and our administrative staff, Joseta, Mariona, Aina, Carla, Mireia and Aric.

Moreover, I want to thank Susana Puig and Josep Malvehy for their continuous support in the project over these years.

To all the co-authors of the papers, for their specific collaboration to carry out our projects.

I cannot finish these acknowledgments without mentioning my family and friends for the hundreds of hours stolen, and for their patience and understanding.

Contents

Contents	ix
List of Figures	xiii
List of Tables	xvii
1 Introduction	1
1.1 Skin Cancer	1
1.1.1 Skin Cancer Causes	4
1.1.2 Skin Cancer Treatment	5
1.2 Diagnostic Techniques for Skin Cancer Detection	5
1.2.1 Non-Specific Skin Imaging Technology	6
1.2.1.1 Visual Detection	6
1.2.1.2 Magnifying Lens	8
1.2.1.3 Wood's Lamp	9
1.2.1.4 Sniffer Dogs	9
1.2.1.5 Baseline Clinical Photography	10
1.2.2 Skin Imaging Technology	11
1.2.2.1 Dermoscopy	11
1.2.2.2 Confocal Scanning Laser Microscopy	22
1.2.2.3 Multispectral Digital Dermoscopy	23
1.2.2.4 Ultrasound	24
1.2.2.5 Optical Coherence Tomography	25
1.2.2.6 Tape Stripping mRNA	26
1.2.2.7 Laser Doppler Perfusion Imaging (LDPI)	28

CONTENTS

1.2.2.8	Electrical Bio-Impedance	28
1.2.2.9	Magnetic Resonance Imaging (MRI)	29
1.2.2.10	Positron Emission Computed Tomography(PET)	30
1.2.2.11	Reflex Transmission Imaging	31
1.3	Scope of Research	31
1.4	Objectives of the thesis	32
1.5	Thesis Outline	33
1.5.1	A new framework for Skin Cancer Detection	33
1.5.2	Document Overview	33
2	Color Calibration	35
2.1	Related Works	36
2.2	Theoretical Model	39
2.2.1	JPEG limitations	44
2.2.2	Custom White Balance (CWB)	46
2.3	Experimental Validation	47
2.4	Discussion	60
3	A novel system for skin mosaicing	65
3.1	Related Works in Image Mosaicing	65
3.1.1	Image Acquisition	66
3.1.2	Image Registration	67
3.1.2.1	Frequency Domain	68
3.1.2.2	Optical Flow	68
3.1.2.3	Feature Based Image Registration	70
3.1.3	Feature detection	72
3.1.4	Feature Description	82
3.1.4.1	Correlation	83
3.1.4.2	SIFT Descriptor	84
3.1.4.3	SURF Descriptor	86
3.1.5	Feature Matching	89
3.1.6	Motion Estimation	90
3.1.6.1	Homography Estimation Methods	92

3.1.6.2	Rejection of Outliers: RANSAC	95
3.1.6.3	Homography Optimization	97
3.1.7	Global Registration and Alignment	98
3.1.7.1	Global Registration	99
3.1.7.2	Global Alignment	99
3.1.7.3	Four Point Transformation	101
3.1.8	Mosaic Blending	102
3.1.8.1	Transition Smoothing Techniques	103
3.1.8.2	Optimal Seam Finding Techniques	103
3.1.8.3	Combined Techniques	104
3.2	A proposal to Skin Image Mosaicing	104
3.2.1	Marker Detection	104
3.2.2	Extraction and description of keypoints.	107
3.2.2.1	GPU-SIFT	108
3.2.3	Initial matching	109
3.2.4	Homography computation	110
3.2.5	Global registration and coordinate frame change	110
3.2.6	Compute 4-point warping	112
3.2.7	Global alignment	112
3.2.8	Mosaic Blending	114
3.3	Results	115
3.3.1	Test Data Set Description and Results	115
3.4	Registration Problems	119
3.4.1	Computational Cost	121
3.5	Conclusions	121
4	Design and development of a new 3D Body Scanner	125
4.1	State-of-the-art in 3D Body Scanning	127
4.1.1	Laser Scanning	128
4.1.2	Coded Structured Light Scanners	129
4.1.3	Photogrammetry	130
4.1.4	Other	132
4.1.5	Scanner Proposal	134

CONTENTS

4.2	Design and development of a cutaneous scanner	135
4.2.1	Horizontal Scanner	135
4.2.1.1	Mechanical Aspects	135
4.2.1.2	Computer Module	137
4.2.1.3	Actuators	138
4.2.1.4	Lighting System	138
4.2.1.5	Image Acquisition Module	138
4.2.2	Vertical Scanner	140
4.3	3D in vivo skin mapping	143
4.3.1	Method	144
4.3.1.1	Image preprocessing	144
4.3.1.2	Estimation of Camera positions	144
4.3.1.3	Dense Reconstruction	145
4.3.1.4	Matching	145
4.3.1.5	Expansion and Filtering	146
4.3.1.6	Surface Reconstruction	146
4.3.1.7	Texture Mapping	147
4.4	Results	147
4.4.1	3D reconstruction results	147
4.4.2	2D - 3D results on the scanner	148
4.5	Conclusions	151
5	Conclusions	155
5.1	Summary of the Thesis	155
5.2	Contributions	157
5.3	Further Work	158
5.4	Related Publications	159
	References	161

List of Figures

1.1	Skin Cancer Types	2
1.2	Skin Cancer Classification	3
1.3	Dermatoscope Schema	12
1.4	Sample set of commercially available dermoscopes.	13
1.5	ABCD rule of dermoscopy	16
1.6	Seven-point checklist	18
1.7	Confocal Scanning Laser Microscopy (CSLM)	23
1.8	Siascan	24
1.9	Optical Coherence Tomography	27
2.1	Elements affecting color reproduction.	36
2.2	Camera Calibration Pipeline	41
2.3	Color Checker	42
2.4	Relative spectral power distribution of DermLite Foto [®]	43
2.5	Pipeline for computing the ΔE	48
2.6	5D Color reproduction accuracy	50
2.7	50D Color reproduction accuracy	51
2.8	G9 Color reproduction accuracy	52
2.9	ΔE error comparison	53
2.10	Color calibration results 1	55
2.11	color calibration results 2	56
2.12	Distribution of the probability	57
2.13	Results: Original Images	59
2.14	Results: JPEG calibrated	60
2.15	Results: RAW Images	61

LIST OF FIGURES

3.1	Mosaicing Pipeline	67
3.2	Feature Based Mosaicing Pipeline	71
3.3	Harris Detector	75
3.4	GSS - Gaussian scale space; DoG - Difference of Gaussians scale space	78
3.5	SIFT local extrema detection	79
3.6	SURF: Approximation of the second-order derivatives	81
3.7	SURF: image scale space	82
3.8	SIFT orientation	85
3.9	SIFT descriptor	87
3.10	SURF orientation assignment	88
3.11	SURF Descriptor	89
3.12	Pair of matched images	91
3.13	Planar transformations	92
3.14	RANSAC - Rejection of outliers	96
3.15	Homography error projection	97
3.16	Graph-based registration	100
3.17	Serratus mountain panorama built using <i>AutoStitchTM</i> software .	102
3.18	Serratus mountain panorama blend using <i>AutoStitchTM</i> software	103
3.19	Finding Markers	106
3.20	Keypoints detection	107
3.21	4 point mosaic warping	113
3.22	Global alignment computation	114
3.23	Mosaic Blending	115
3.24	Rigid Object Mosaic	116
3.25	Back Mosaic	117
3.26	Patient Mosaic	119
3.27	Second Patient Mosaic	120
3.28	Registration Problems	120
3.29	Mosaic Viewer	123

LIST OF FIGURES

4.1	PhotoMAX software for managing patient images and database. Images are linked to an avatar by hand. From Derma Medical Systems (http://www.dermamedicalsystems.com)	126
4.2	3D Body Scanning Classification	127
4.3	Schema of 3D body scanning based on laser scanning.	128
4.4	3D body scanning based on laser scanning	129
4.5	Scheme of 3D body scanning based on Coded Structured Light . .	130
4.6	3D body scanning based on coded structured light	131
4.7	Scheme of 3D body scanning based on photogrammetry	132
4.8	3D body scanning based on Photogrammetry	132
4.9	Other 3D body scanning	133
4.10	Horizontal Scanner Schema	136
4.11	Horizontal Scanner	136
4.12	Bloc diagram of the proposed scanner.	137
4.13	Horizontal Scanner Lighting System	139
4.14	Overlapping images on the Horizontal Scanner	140
4.15	Vertical Scanner Schema	142
4.16	Vertical Scanner	143
4.17	Sample reconstruction pipeline	148
4.18	Final reconstruction	149
4.19	2D Mosaic	150
4.20	3D Point Cloud	151
4.21	3D Mesh of triangles	151
4.22	3D Texturized surface	152
4.23	3D representation back	153

LIST OF FIGURES

List of Tables

1.1	Comparison of mole scanning methods	7
1.2	ABCDE rule for dermatologic diagnosis.	8
1.3	Magnifying Lens	9
1.4	Wood's lamp devices	10
1.5	Computer Aided Systems	11
1.6	List of some available dermoscopes	13
1.7	Various mole scanning systems with Computer Automatic Diagnosis tools	14
1.8	ABCD rule: Calculation of the total dermoscopy score(TDS) . . .	15
1.9	ABCD scoring system for melanocytic neoplasms	17
1.10	CASH algorithm	20
1.11	Blink algorithm	21
1.12	Confocal Scanning Laser Microscopy devices	23
1.13	Multispectral Digital Dermoscopy	24
1.14	Ultrasound frequency	25
1.15	Ultrasound device systems	26
1.16	Optical Coherence Tomography	26
1.17	Tape Stripping mRNA	27
1.18	Laser Doppler Perfusion Imaging (LDPI)	28
1.19	Electrical Bio-Impedance	29
1.20	Magnetic Resonance Imaging (MRI)	30
1.21	Positron Emission Computed Tomography(PET)	31
1.22	Reflex Transmission Imaging	31
2.1	Color reproduction accuracy	49

LIST OF TABLES

2.2	Color matching comparison	53
2.3	Color reproduction accuracy	57
2.4	Color matching comparison	58
2.5	Comparison of different patches on two different lesions	62
2.6	Accuracy measurements	62
3.1	Feature Detectors Classification	72
3.2	GMML 3.0a specifications	111
3.3	4 point warping	113
3.4	Computational Time	122
4.1	3D Reconstruction Results	152

Chapter 1

Introduction

The aim of this research is the development of a reliable set of tools to help in the early diagnosis of skin cancer. This initial chapter starts with a brief introduction about Skin Cancer and detection systems. In addition, a general overview of the thesis is provided including a description of its structure.

1.1 Skin Cancer

Skin cancer is a major public health problem. Each year there are more new cases of skin cancers than the combined incidence of breast, prostate and colon cancers[160]. A study developed in 2010 by the American Cancer Society estimates that in the United States around 74,000 new cases appear each year and 18,000 deaths occur. The new cases correspond to 5% in men and 4% in women of the total number of new cancers in the United States.

In Europe, skin cancer accounted for 35,000 new cases and 9,000 deaths in 2000[23]. Men have a 1 in 37% probability of developing skin cancer during their lifetime with women coming in at 1 in 56. The highest probability for both men and women is in old age, i.e. from 70 years old, on.

However, although skin cancer incidence has increased by 45% between 1992 and 2004, the death rate is one of the lowest, being only 9% in the period 1999 to 2005.

Skin cancer is divided into two main groups; non-melanoma skin cancer

1. INTRODUCTION

(NMSC) and melanoma skin cancer (MM). The second group is the primary disease of whites. Melanoma is classified into superficial spreading melanoma (SSM), nodular melanoma (NM), acral lentiginous melanoma (ALM) and lentigo maligna melanoma (LMM). Non-melanoma skin cancer is the most prevalent cancer among light-skinned people. It is classified into basal cell carcinoma (BCC)(75%), squamous cell carcinoma (SCC)(24%) and other rare types (1%) such as sebaceous carcinoma. Non-melanoma skin cancer is seldom lethal, but in an advanced stage can cause severe disfigurement and morbidity.

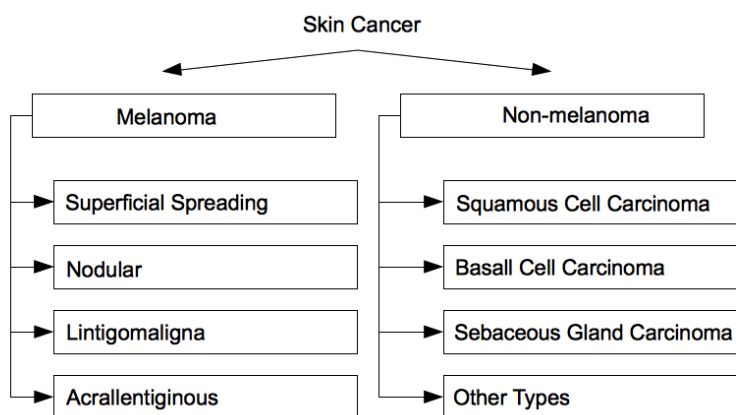


Figure 1.1: Skin Cancer Types

Superficial spreading melanoma (SSM) accounts for 70% of melanoma skin cancers, and arises in a preexisting dysplastic nevus more often than any other type. It has a radial growth phase where the lesion is confined within the epidermis and increases in diameter. The second phase of this kind of lesion is a vertical, downward extension to the dermis, increasing the metastatic potential. SSM evolves in from 1 to 7 years, varying the shade and pigmentation. It occurs equally in men and women, however it commonly appears on the trunk in men and lower extremities in women.

Nodular melanoma (NM), which accounts for 15% to 30% of melanoma skin cancers, is the most aggressive type, predominantly affecting men of around 50. It often arises in new lesions, although it can also occur at the site of preexisting nevus. NM has a radial growth phase in which it can be of a small diameter. The growth phase is predominantly vertical, with a deep invasion associated with a

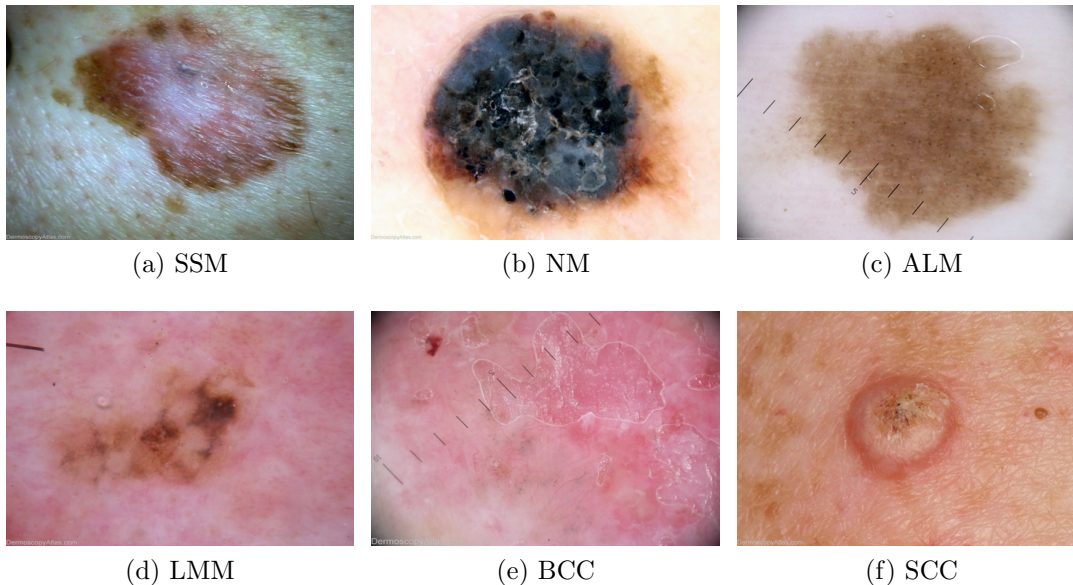


Figure 1.2: Skin Cancer Classification, Melanoma: Superficial Spreading Melanoma (SSM), Nodular Melanoma (NM), Acral Lentiginous Melanoma (ALM) and Lentigo Maligna Melanoma (LMM). Non-melanoma: Basal Cell Carcinoma (BCC) and Squamous Cell Carcinoma (SCC). (Images from of The International Atlas of Dermoscopy, The Skin Cancer College of Australia and New Zealand)

prognosis usually worse than any other type. The lesion characteristically appears as a smooth, shiny, dome-shaped nodule of a uniformly dark black or blue color and has a rapid growth.

Lentigo maligna melanoma (LMM), also known as in situ melanoma, occurs on extremely sun-damaged skin. It is the least aggressive and common type of melanoma skin cancers, accounting for less than 5%. Basically, it affects women in their seventh decade of life and can appear on the head, neck, or dorsum of the hands. It grows slowly for 3 to 15 years in a radial shape, reaching 3 to 6 cm.

Acral Lentiginous Melanoma (ALM) accounts for 5% of all melanoma skin cancers, although it affects 35% to 65% of darkly pigmented individuals (African Americans, Asians and Hispanics). It appears on the palms, soles, nail beds and occasionally on the mucous membranes.

Basal cell carcinoma (BCC) is the most common type of skin cancer. Although

1. INTRODUCTION

it rarely metastasizes or kills, it causes significant destruction and disfigurement. 30% of caucasians may develop a BCC cancer during their lifetime[186]. 80% of all cases are found on the head and neck, recently however, BCC has appeared more on the torso.

Squamous cell carcinoma (SCC) is the second most common cancer of the skin after BCC. It usually occurs in areas exposed to the sun. The risk of metastasis is low, but higher than BCC. 20% to 50% of SCC of the lip and ears have high rates of local recurrence and distant metastasis.

1.1.1 Skin Cancer Causes

Ultraviolet (UV) light exposure, most commonly from sunlight, is overwhelmingly the most frequent cause of skin cancer[189]. Other important factors are:

- Use of tanning booths.
- Immunosuppression-impairment of the immune system, which protects the body from foreign entities, such as germs or substances that cause an allergic reaction. This may occur as a consequence of some diseases or can be due to medications prescribed to combat autoimmune diseases or prevent organ transplant rejection.
- Exposure to unusually high levels of x-rays.
- Contact with certain chemicals-arsenic (miners, sheep shearers, and farmers), hydrocarbons in tar, oils, and soot may cause squamous cell carcinoma.

On the other hand, people who have the greatest risk of developing a skin cancer are[139]:

- People with fair skin, especially types that freckle, sunburn easily, or suffer pain in the sun.
- People with light (blond or red) hair and blue or green eyes.
- Those with certain genetic disorders that deplete skin pigment such as albinism or xeroderma pigmentosum.

-
- People with numerous moles, unusual moles, or large moles that were present at birth.
 - People who have already been treated for skin cancer.
 - People with close family members who have developed skin cancer.
 - People who have at least one severe sunburn early in life.

1.1.2 Skin Cancer Treatment

Treatment depends on the type of skin cancer, its stage and location and the individual's age and overall health. For instance, people with small basal cell carcinomas may require only simple treatment because BCC cancers rarely spread to other parts of the body and are seldom fatal. SCC is the opposite because it tends to spread and therefore requires more treatment. Malignant melanoma may require complicated treatment because of its high risk of spreading.

Biopsy can completely remove some tiny skin cancers, however to eliminate all the cancer cells most skin cancers need additional treatment such as:

- Surgery: Taking out or destroying the cancer.
- Chemotherapy: Giving drugs to kill the cancer cells.
- Radiation therapy: Using powerful energy from x-rays or other sources to destroy the cancer cells.
- Others: Photodynamic therapy, Biological therapy, ...

1.2 Diagnostic Techniques for Skin Cancer Detection

In this section we review different approaches to detect skin cancer, describing their main features and highlighting the differences among them. The key objective is to point out the advantages and disadvantages of these approaches (see Tab. 1.1). We decided to divide the diagnostic techniques into non specific

1. INTRODUCTION

skin imaging and specific skin imaging technology. The first group includes visual detection, magnifying lens, wood's lamp, sniffer dogs and baseline clinical photography, while the second presents dermoscopy, confocal scanning laser microscopy, multi-spectral digital dermoscopy, ultrasound, optical coherence tomography, tape stripping mRNA, laser doppler perfusion imaging (LDPI), electrical bio-impedance, magnetic resonance imaging (MRI), positron emission computed tomography(PET), and reflex transmission imaging. Some algorithms for the most common diagnostic systems are described.

1.2.1 Non-Specific Skin Imaging Technology

This section includes all the diagnostic approaches that do not require any specific dermatological device. Visual detection, magnifying lens, wood's lamp, sniffer dog and baseline clinical photography are presented. These methods, combined with a complete patient history, can provide enough evidence to make a good initial diagnose. A detailed history of a patient should include:

- age and sex
- personal medical history of melanoma or non-melanoma skin cancer
- family medical history of melanoma
- number of nevus
- presence of atypical or dysplastic nevus
- skin type
- tanning habits
- response to sun exposure and evidence of skin damage from the sun

1.2.1.1 Visual Detection

Visual inspection together with the patients history form the basis for early detection of melanoma.

Table 1.1: Comparison of mole scanning methods

Method	Principle	Advantages	Main disadvantages
Visual Inspection and magnifying lens	ABCDE rule is the usual clinical guide for most lesions and EFG is more appropriate for nodular lesions.	Easy to perform.	Limited sensitivity in melanoma diagnosis.
Baseline Clinical Photography	Digital imaging in standardized positions. Nearly whole skin surface visualized.	Identification of 'ugly ducklings'. Identification of new or evolving lesions.	Only gives macroscopic information.
Wood's Lamp	UVL emissions are strongly absorbed by melanin.	Can reveal irregular pigment distributions, defining the letigo maligna melanoma borders.	Limited sensitivity in melanoma diagnosis.
Dog Sniffer	Based on the odor.	Detection in non treated patients.	Few cases reported.
Dermoscopy	Visualization of subsurface anatomic structures of epidermis and upper dermis.	Well-established criteria. Increases diagnostic sensitivity without diminishing specificity. Dermoscopes with polarized and non-polarized light are available.	It requires specialized training.
Multispectral Digital Dermoscopy	Light reflected at different skin depths is collected and analyzed.	Gives visual information of deeper skin layer in comparison with dermoscopy. Automated diagnosis possible.	Needs further evaluation in clinical trials.
Ultrasound and optical coherence tomography	Vertical imaging of the skin.	Monitoring of topical treatment possible.	To date, no diagnostic aids.
Confocal scanning laser microscopy	Horizontal imaging of the skin with laser light that causes no tissue damage. Melanin/melanocytes are a strong source of contrast.	Quasi-histological resolution offers in vivo biopsy, monitoring of treatments, presurgical margin assessment.	It requires specialized training. Limited imaging depth. Involves moving parts, making scanning more difficult. To date, mainly used for research.
Tape stripping mRNA	RNA analysis from suspicious melanocytic lesions.	Pre-screen for suspiciously pigmented lesions.	Under investigation, already creating gene expressions.
Laser doppler perfusion imaging	Doppler effect.	Visualization of cellular and subcellular structures.	To date, mainly used for research.
Electrical bio-impedance	Different electrical impedance of cancer and healthy cells.	Almost 100% sensitivity for in situ and thin melanoma.	Scanning and image analysis takes 7 minutes.
Magnetic resonance imaging	Utilizes radio waves and strong magnets.	Allows differentiation of the different skin layers.	Cost, size, time, and specialized training. Not usable for patients with metal implants.
Positron emission computed tomography	'Warburg Effect'	The entire body can be analyzed.	Expensive and time consuming.
Reflex Transmission Imaging	High resolution ultrasound	Good discrimination.	Hight cost.

1. INTRODUCTION

Two algorithms can be applied to visual detection; the ABCDE rule (**a**symmetry, **b**order **i**rrregularity, **c**olour variegation, large **d**iameter and evolution, see Table 1.2) and the EFG algorithm (**e**levated, **f**irm to touch and **g**rowing progressively over more than a month) which is more appropriate for nodular melanomas that often have a more subtle clinical appearance[83].

Table 1.2: ABCDE rule for dermatologic diagnosis.

A - Asymmetry	Normal moles or freckles are quite symmetrical. In cases of skin cancer, spots do not look the same on both sides.
B - Border	A mole or spot with blurry and/or jagged edges.
C - Color	A mole that has more than one hue is suspicious and needs to be evaluated by a physician. Normal spots are usually one color. This can include lightening or darkening of the mole. Melanoma cells usually continue to produce melanin, which accounts for the cancers appearing in mixed shades of tan, brown and black.
D - Diameter	If the spot is larger than 6mm, it needs to be examined by a doctor. This includes areas that do not have any other abnormalities (color, border, asymmetry).
E - Elevation/Evolving	The mole with an uneven surface raised above the surface. It looks different from neighboring moles or changes in size, shape and/or color.

One single visual inspection fails to detect small melanomas and amelanotic melanomas. Thus for high-risk individuals, a full cutaneous examination supported by total body photography and dermoscopy every six month is recommended as well as patient education for self-examination.

1.2.1.2 Magnifying Lens

The simple use of a magnifying glass can often assist clinicians in differentiating and correctly diagnosing many pigmented lesions[47]. Magnification allows easy visualization of comodo-like openings in seborrheic keratoses and telangiectasias in pigmented basal cell carcinomas, thus helping to exclude melanoma from the differential diagnosis.

Lumio from 3Gen is one of the multiple examples of magnifying lenses available. Lumio has a 75 mm lens offering a magnification of 2× and 40 bright-white

LEDs (light emitting diodes) on a cross-polarization set-up. It is used in dermatology for examining varicose veins, pigmented skin lesions and hair follicles among others. Some of the commercial magnifying lenses used in dermatology are shown in table 1.3

Table 1.3: Magnifying Lens

Company	Country	Website
3Gen LLC	San Juan Capistrano, CA USA	http://www.dermlite.com
Daray [®] Medical	Moira, Derbyshire UK	http://www.daray.co.uk

1.2.1.3 Wood's Lamp

Wood's lamp was invented in 1903 by Robert W. Wood[187]. It was first used in dermatology practice for the detection of fungal infection of hair.

Wood's lamp, emits ultraviolet light at a wavelength of 360 nm that can help to diagnose dermatoses with a characteristic fluorescence (tinea capitis, erythrasma, tinea versicolor, pseudomonas infections, porphyrians, and pigmentary alterations)[181].

The emissions are transmitted into the dermis, where they give a white to blue-white fluorescent color. These emissions are strongly absorbed by melanin, aiding in the diagnosis of lentigo maligna melanomas by revealing irregular pigment distribution, and defining subtle clinical borders. Melanin present in the epidermis (but not dermis) absorbs long-wave UVA. Wood's light accentuates the contrast between pigmented and non-pigmented skin, but, more importantly, it separates hypo-pigmented from totally amelanotic areas (the latter have true white to blue-white fluorescence). We can find a high number of projects in the literature where this technology is used, e.g. [67], [9], [91], [36], [124] and [62]. Some companies still produce Wood's lamps for dermatology (see Tab. 1.4).

1.2.1.4 Sniffer Dogs

Sniffer dogs may be able to detect malignant tumors on the basis of odor. This fact was first described in [183], reporting a case where a dog was able to detect

1. INTRODUCTION

Table 1.4: Wood's lamp devices

Company	Country	Website
Amjo Corp.	West Chester, OH USA	http://www.woodslight.com
Daray [®] Medical	Moira, Derbyshire UK	http://www.daray.co.uk

a malignant melanoma. Since then, similar anecdotal claims are described in the literature of skin cancer[35]. Recently, some studies have presented a more accurate reasoning in the accuracy of sniffer dog detection [184] [10].

1.2.1.5 Baseline Clinical Photography

Comparing images over time is one of the simplest ways of detecting melanomas. Melanomas change in size, shape and/or color whereas normal moles are usually stable[11], this fact allows the detection of suspicious moles. The presence of baseline photography can help in deciding of whether biopsy is needed or not[153]. Baseline photography is very useful for patients with many nevi or with the dysplastic nevus syndrome, due to the time needed to carry out a complete dermoscopic exploration.

The photographic equipment varies from Polaroid cameras or 35 mm cameras in the past to the new digital cameras nowadays. Any camera available on the market can be useful for clinical explorations. Other considerations to take into account are lights, patient position, etc., elements to set up in the acquisition process that can facilitate further comparisons.

In order to help skin self examination, the images can be printed and the patient can compare them by himself, or with the help of another person.

The images can be obtained lesion by lesion or using total body photography. Each system presents advantages and disadvantages. The positions for the patients are defined in the literature in [66], the patient should assume standardized positions under good lighting conditions. Images should be taken of the face, neck, area behind the ears, scalp (in bald individuals), anterior and posterior torso, and the extremities (including palms and soles).

All the photographic images should be appropriately stored for later comparison with new explorations. To this end, there are some commercially avail-

able systems such as Fotofinder, DermoGenius, Dermagraphix and MoleMap (see table 1.5). Using these systems, different explorations can be opened and images compared by superimposing one over the other. For instance, systems such as Molemax allows superimposing images in order to detect changes. In [179] the benefits of topodermatographic sequential images are presented, showing the advantage of detecting subtle clinical changes in the measurement or shape of multi-pigmented skin lesions.

Total Body Photography is well described in [65], where the photographic technique, anatomic poses, camera system and examination procedure are explained in detail. It can reduce the number of biopsies in high-risk patients[86][49]. This approach easily detects new lesions, but it is difficult to observe small changes in nevi over time, and it is possible that some areas of skin may be covered by undergarments or hair and, thus, missed in baseline photographs.

Table 1.5: Computer Aided Systems

Device Names	Company	Website
DermoGenius Ultra	LINOS Photonics, Inc.	http://www.dermogenius.de
MelaFind	Electro-Optical Sciences, Inc.	http://www.melafind.com
SIAscope	Astron Clinica	http://www.astronclinica.com
MoleMax II	Derma Instruments L.P.	http://www.dermamedicalsystems.com
MicroDerm	VisioMED	http://www.zn-ag.com
NevusScan	Romedix	http://www.romedix.com
SolarScan	Polartechnics Limited	http://www.polartechniques.com.au
FotoFinder Derma	Edge Systems Corp.	http://www.edgesystem.net
DBDermo-Mips	University of Siena	http://www.skinlesions.net
VideoCap100	DS Medica	http://www.dsmediagroup.it

1.2.2 Skin Imaging Technology

This section reviews some techniques that use specific devices for detecting skin cancer.

1.2.2.1 Dermoscopy

Dermoscopy (also known as dermatoscopy or epiluminescence microscopy), is a diagnostic technique used worldwide in the identification and diagnosis of nu-

1. INTRODUCTION

merous skin lesions. The literature presents results with increased accuracy in diagnosing melanoma when physicians followed training in Dermoscopy[163].

Dermoscopy is based on examining the pigmented and non-pigmented lesions of the skin by means of a photographic camera equipped with a dermatoscope: a magnifier (typically $10\times$ or $12\times$), a light source (typically a LED ring), and a transparent plate placed on the skin with the objective of making subsurface structures more easily visible when compared to conventional clinical images[7] (Fig 1.3).

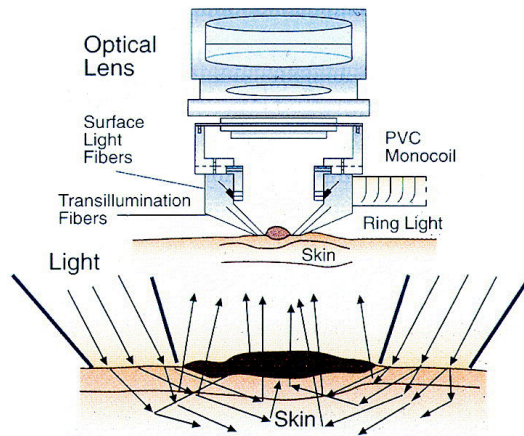


Figure 1.3: Dermoscopy functions by trasillumination of the lesion. This allows for the visualization of sub-surface structures not seen with the naked eye. (From Pizarro A. Techniques and instruments. In: Malveyh J, Puig S, editors. Principles of dermoscopy. Spain: CG Creaciones Graficas S.A.; 2002. p.28)

A large number of dermatoscopes are available on the market[113](see Figure 1.4). These devices can be divided into two main groups: those using an unpolarized light source with an incident angle of 45° (non-polarized dermoscopy [NPD]) and those using use perpendicularly oriented light towards the skin with polarized filters (polarized dermoscopy [PD]).

NPD requires a liquid between the dermatoscopes plate and the skin that is in contact with the tool. The use of the fluid (oil or water) helps to roughen the area and makes the skin surface uneven, reducing the amount of light reflected from the stratum corneum as well as reducing the amount of air between the skin and the device. This effect makes the epidermis essentially translucent and



Figure 1.4: Sample set of commercially available dermoscopes.

allows in vivo visualization of subsurface anatomic structures of the epidermis and papillary dermis that are otherwise not discernible to the unaided eye.

With this effect the skin becomes transparent to light and reflect the subsurface structures.

PD uses two polarizing filters aligned with a difference of 90° between them (known as cross-polarization) allowing 60 to 100 μm penetration of the light under the skin. The backscattered component of the light can reach the surface and be registered with a photographic camera. PD can be divided into Polarized Light Non-contact Dermoscopy (PNCD) and Polarized Light Contact Dermoscopy (PCD).

Table 1.6: List of some available dermoscopes

Dermoscope	Company	Website
DermLite Platinum [®] DermLite 100 [®]	3Gen, LLC USA	http://www.dermlite.com
DermLite Foto [®] DermLite PROtextregistered	3Gen, LLC USA	http://www.dermlite.com
Mini 2000 Dermoscope [®]	Heine Optotechnik, Germany	http://www.heine.com
Delta 10 [®]	Heine Optotechnik, Germany	http://www.heine.com
Delta 20 [®]	Heine Optotechnik, Germany	http://www.heine.com
Alpha+ [®]	Heine Optotechnik, Germany	http://www.heine.com
DermoGenius [®] basic II	LINOS Photonics, Inc. MA	http://www.dermogenius.com
Dino-lite Digital Microscope AM-413-ZT	AnMo Electronics Corporation, Taiwan	http://www.dino-lite.com
AMD-2030	AMD Global Telemedicine, UK	http://www.amdtelemedicine.com

Dermoscopy is usually combined with photographic images of the full body (Total Body Photography) and a computer assisted diagnosis. Table 1.6 shows different commercial devices with their softwares and capabilities. Below we present some of the basic algorithms used in CAD systems for helping dermatologists to diagnose skin cancer (see. Table 1.7).

1. INTRODUCTION

Table 1.7: Various mole scanning systems with Computer Automatic Diagnosis tools

Device	Standardized total body photography	Total body photography comparator	Macroscopic lesion images	Non-polarized dermoscopy	Polarized dermoscopy	Sequential dermoscopic imaging	Computer assisted diagnosis tools	Fully automated diagnosis
Molemax (http://www.equipmed.com)	*	*	*	*	*	*	*	*
Fotofinder (http://www.fotofinder.de)	*	*	*	*	*	*	*	*
DermoGenius Ultra (http://www.biocam.de)			*	*	*	*	*	*
MicroDerm (http://www.visionmedag.com)			*	*	*	*	*	*
Db-Dermo Mips (http://www.skinlesions.net)					*			*
MelaFind (http://www.eosciences.com)								*

Pattern Analysis Pattern recognition has historically been used by clinicians and histopathologists to differentiate benign lesions from malignant neoplasms. In 1987 Pehamberger et al. [134] presented a similar process for dermoscopy, known as pattern analysis. This method uses specific global patterns and combinations of additional local features for the classification of pigmented skin lesions as melanoma, clark, spits/reed or blue nevi[134]. Pattern analysis is the most well-known and reliable method for the dermoscopic diagnosis, but it is difficult to use by non-experienced dermatologists.

The complexity of this diagnostic approach is due to the subjective evaluation of multiple criteria and morphological patterns. Another reason is the great variability of the morphological expressions of each parameter. An incomplete or insufficient definition of the criteria results in only modest reproducibility in some

of the dermoscopic variables, and studies have shown that the reproducibility of a criterion is greater in its quantitative evaluation (presence or absence of a given parameter) than in its qualitative evaluation (regularity, distribution, size, shape, etc.).

Pattern analysis improves the rate of correct diagnosis of PSL by 10-30%. Nevertheless, because of problems inherent in the reliability and reproducibility of the diagnostic criteria used in pattern analysis, a number of additional diagnostic methods based on scored algorithms have been introduced in the last few years with the aim of increasing sensitivity in detecting melanoma.

ABCD rule The first approach for simplifying the pattern analysis method is the ABCD rule of dermoscopy (Asymmetry, Borders, Colors and Dermoscopic structures, see. Fig. 1.5), introduced by Stolz et al. [164] in 1994.

This diagnostic method permits a more objective diagnosis of melanoma with the evaluation of only four dermoscopic criteria. Some studies present a specificity of 59% and a sensitivity of 92% in the diagnosis of melanoma, disclosing a high diagnostic accuracy improvement over less experienced observers[103]. For trained dermatologists, the ABCD rule presents less sensitivity and specificity than pattern analysis.

Table 1.8: ABCD rule: Calculation of the total dermoscopy score(TDS)

Criteria	score	* factor	= result
Asymmetry in perpendicular axes: contour, colors and structures	0 - 2	1.3	0 - 2.6
Borders, 8 segments: abrupt ending of pigment pattern	0 - 8	0.1	0 - 0.8
Colours: White, red, light-brown (tan), dark-brown, blue-grey, black	1 - 6	0.5	0.5 - 3.0
Different structural components or Dermoscopic structures (pigment network, structureless areas, dots, aggregated globules, branched streaks)	1 - 5	0.5	0.5 - 2.5

1. INTRODUCTION

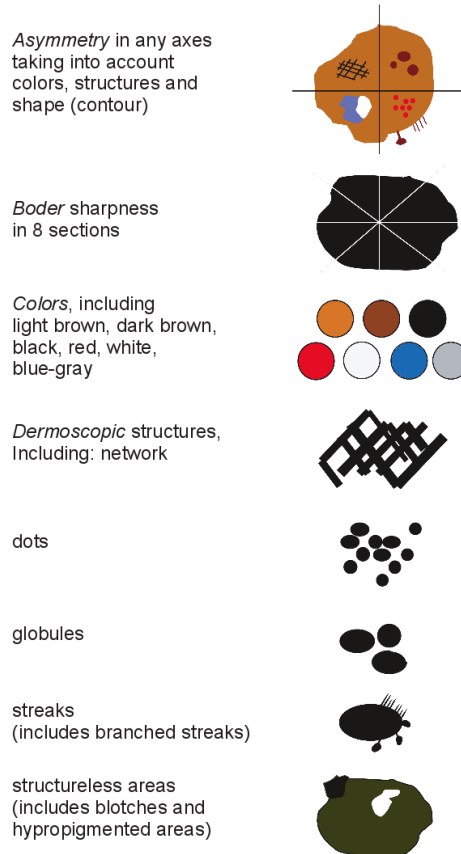


Figure 1.5: Criteria for the calculation of TDS of the ABCD rule of dermoscopy. (From Handbook of Dermoscopy. In: Malvehy J., Puig S., Braun R.P., Marghoob A.A., Kopf A.W.;2006, p.34)

7-Point checklist The 7-Point checklist[6] is an alternative diagnostic approach based on a simplified pattern analysis, using seven standard criteria. The seven criteria are categorized according to their different diagnostic weight into major and minor criteria, the diagnosis of melanoma being suspected when at least two criteria (one major and one minor criterion) are recognized (see Fig. 1.6). In expert hands this method allows correct classification of 95% of melanomas (sensitivity) and 75% of clinically atypical melanocytic nevi (specificity).

- Atypical pigment network (2)

Table 1.9: ABCD scoring system for melanocytic neoplasms

Total score	
Benign melanocytic lesion	< 4.76
Suspicious lesion; close follow-up or excision recommended	4.76-5.45
Lesion highly suspicious for melanoma	> 5.45

- Blue-whitish veil (2)
- Atypical vascular pattern (2)
- Irregular streaks (1)
- Irregular dots/globules (1)
- Irregular blotches (1)
- Regression structures (1)

In order to increase sensitivity, the seven-point rule was revised in 2011 so that each individual item scores 1 (total is 7)[5]. Patients with atypical nevi, any lesion with a score of one, should be carefully examined. Excision of such lesions will pick up many early-stage melanomas.

Menzies Method The Menzies Method[117] was proposed in 1996. It is an algorithm based on the recognition of two negative dermoscopic features (not favouring melanoma diagnosis) and nine positive features (favouring melanoma diagnosis).

The two non-favouring melanoma diagnosis features are symmetry of pattern and presence of a single color. Symmetry of pattern is present if the lesion has symmetry across all axes through the center of the lesion. The considered colors are black, gray, blue, dark brown, tan and red. Here white is not considered as a color. Melanoma is discarded if both criteria are accomplished.

Positive features include blue-white veil, multiple brown dots, pseudopods, radial streaming, scar-like depigmentation, multiple (5-6) colors, multiple blue/grey

1. INTRODUCTION

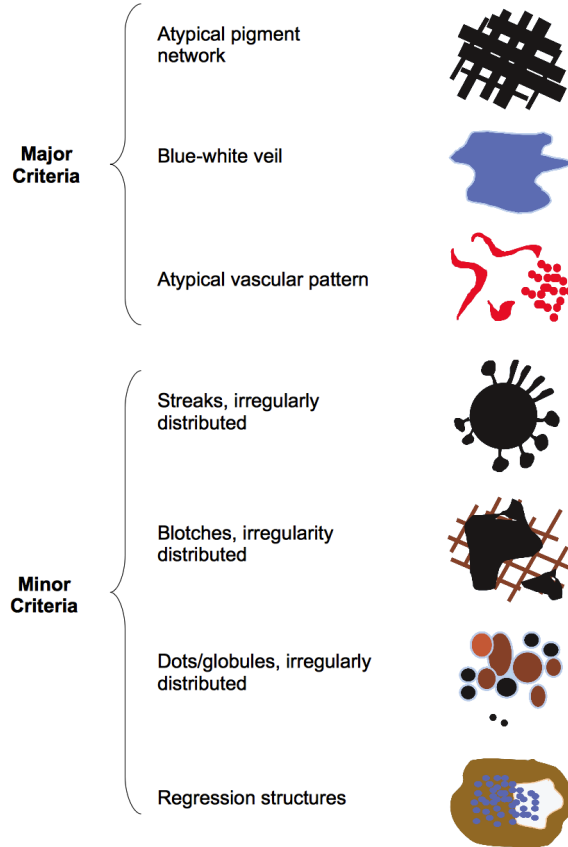


Figure 1.6: Seven-point checklist (From Handbook of Dermoscopy. In: Malvehy J., Puig S., Braun R.P., Marghoob A.A., Kopf A.W.;2006, p.38)

dots and broadened network. Blue-white veil is considered when an irregular, structureless area of confluent blue pigmentation with an overlying white “ground-glass” haze is present. The pigmentation cannot occupy the entire lesion and cannot be associated with red-blue lacunes. Multiple brown dots are considered when focal areas of multiple brown (usually dark brown) dots (not globules) are present. Pseudopods are bulbous and often kinked projections that are found at the edge of a lesion directly connected to either the tumor’s body or a pigmented network and are never seen distributed regularly or symmetrically around the lesion. Radial streaming are finger-like extensions at the edge of a lesion that are not distributed regularly or symmetrically around the lesion.

Scar-like depigmentation covers areas of white, distinct, irregular extensions (true scarring) and should not be confused with hypo or depigmentation due to simple loss of melanin. Peripheral black dots/globules found at or near the edge of the lesion are taken into account. The colors scored are black, gray, blue, dark brown, tan and red. Again white is not counted as a color. The presence of multiple blue/gray dots (not globules) is often described as being “pepper-like” in pattern. A broadened network is a network made up of irregular, thick “cords”, often seen as being focally thicker. The presence of one of these positive features associates the lesion to melanoma.

When used by experts, the Menzies method gave a sensitivity of 92% and a specificity of 71% [117]. In a study based on the evaluation of clinical and dermoscopic pictures taken after a brief dermoscopy training session on the Menzies method, a group of 74 primary care physicians (PCPs) improved sensitivity without a decrease in specificity for the diagnosis of melanoma compared with a control group.

CASH Algorithm The CASH algorithm (Colors, Architecture, Symmetry and Homogeneity) is a simplified version of pattern analysis, designed for use by less experienced dermatologists. The algorithm considers colors, architecture, symmetry and homogeneity[73].

Although there is some overlap between the ABCD Dermatoscopy scoring system and the CASH algorithm, the latter adds architectural disorder as an important component and does not include border sharpness. Architectural order/disorder is a subjective feature that measures the nevus cell proliferation. Benign nevus cells evolve to a finite size in a relatively organized (orderly) and controlled manner. In contrast, melanoma cells caused by a failure in growth control pathways lack the ability to construct an organized lesion.

The number of colors, symmetry, and homogeneity/heterogeneity (the increasing number of dermoscopic structures) are key components in other dermoscopic algorithms[74]. For this reason the authors of the CASH algorithm decided to maintain these features in their algorithm, helping to distinguish benign from malignant melanocytic neoplasms.

Add up the scores for a total CASH score (2 to 17). A score of 7 or less is

1. INTRODUCTION

Table 1.10: CASH algorithm

Suspicion for melanoma	Low	Medium	High
Colors: few vs many Light brown, dark brown, black, red, white, blue Score 1 point for each color	1-2 colors (1-2 points)	3-4 colors (3-4 points)	5-6 colors (5-6 points)
Architecture: order vs disorder Score 0-2 points	None or mild disorder (no points)	Moderate disorder (1 point)	Marked disorder (2 points)
Symmetry vs asymmetry Consider contour, colors and structures Score 0-2 points	Symmetry in 2 axes (no points)	Symmetry in 1 axis (1 point)	No symmetry (2 points)
Homogeneity vs Heterogeneity Consider pigment network, dots/globules, blotches, regression, streaks, blue-white veil, polymorphous vessels Score 1 point for each structure	Only one structure (1 point)	2 structures (2 points)	3 or more structures (3-7 points)

most likely benign, while scores of 8 or more are suspected melanomas. Table 1.10 summarizes the scoring system.

Chaos and Clues Algorithm Chaos (asymmetry of structure and/or colors) and clues is an algorithm designed to detect any type of malignancy in any type of pigmented skin lesion. While pattern analysis techniques focus on classifying a lesion from among the different skin cancer types, the Chaos algorithm only recognizes malignant and non malignant lesions[88].

To this end, they use a three-step procedure. First, they define simple geometric elements and basic patterns created by the lesion. Second, they integrate information gleaned from color. Third, they distinguish specific pigmented skin lesions according to their stereotypical presentation that results from characteristic combinations of elements, patterns and colors. These repeatable combinations of elements, patterns and colors are referred to as “clues”. Each of the three steps is then integrated into an algorithm that directs the practitioner to a diagnosis.

The clues to melanoma are:

- Thick reticular lines.
- Grey or blue structures of any kind.
- Pseudopods or radial lines at the periphery.
- Black dots in the periphery.

-
- Eccentric structureless areas of any color.
 - Polymorphous vascular patterns.
 - White lines.
 - Parallel lines on ridges.

The BLINCK algorithm Even for experienced dermatologists, some kind of skin cancers, such as in situ melanoma, hypo/amelanoctic melanoma, naevoid melanoma and Merkel cell carcinoma are difficult to diagnose with a purely dermoscopic inspection. For this reason, Bourn proposed the BLINK algorithm that uses other clues for diagnosis. Table 1.11 defines the mnemonic, BLINCK, that refers to a logical progression of six clinical questions that should be asked when examining a suspicious skin lesion.

Table 1.11: Blink algorithm

B enign	If not, then consider the following:	
L onely	An ugly duckling	Score 1
I rrregular	Asymmetrical pigmentation pattern or >1 color	Score 1
N ervous	Nervous patient	Score 1
C hange	Changing lesion	
K nown	Known clues to malignancy	Score 1

Known Clues to malignancy are:

- An atypical network
- Segmental streaks
- Irregular black dots, globules, clods
- An eccentric structureless zone
- An irregular blue or grey color
- Polymorphous, arborising, glomerular vessels

1. INTRODUCTION

- A parallel ridge pattern or diffuse irregular brown/black pigmentation in acral lesion

The algorithm is followed only if the lesion is not considered as benign. Each question with a positive answer accounts for 1 in the final score. For “Nervous” and “Change” only a total score of 1 can be either given if one or both of the questions have a positive answer.

If the lesion scores 2 or more, it should be excision biopsied. If there is a score of 0 or 1, the lesion does not require biopsy and may be either reviewed at a later time, digitally monitored, or the patient simply reassured that the lesion is benign.

1.2.2.2 Confocal Scanning Laser Microscopy

Confocal Scanning Laser Microscopy (CSLM) is a non-invasive technique that permits the in vivo examination of the skin at variable depths in horizontal planes[140][112], and is the only in vivo technique that allows the examination of the epidermis and papillary dermis at a resolution approaching histological detail[25]. CSLM works by tightly focusing a low-power laser beam (visible or near infrared) on a specific point on the skin, and detecting only the light reflected from that focal point through a pinhole-sized spatial filter (see Fig. 1.7). This beam is scanned horizontally over a 2-dimensional grid to obtain a horizontal microscopic section. The measured lateral resolution of CSLM images is approximately $0.5\ \mu\text{m}$, and axial resolution is approximately $3\text{-}5\ \mu\text{m}$, while the imaging depth is limited to the level of papillary dermis, $200\text{ to }300\ \mu\text{m}$ [55]. This depth capability is associated with the penetrating depth of the wavelength of the laser light used.

CSLM can be classified into reflectance[29] or fluorescence mode[44]. The former relies on inherent differences in the reflectivity of structures, while the fluorescence mode relies on the differential distribution of endogenous or exogenous fluorescent molecules (fluorophores) to produce contrast in the tissue.

The principal advantage of CLSM is the ability to noninvasively assess the cellular components of intact skin lesions with detail approaching histology. Lesions can be examined temporally to determine the extent of their lateral margins. A

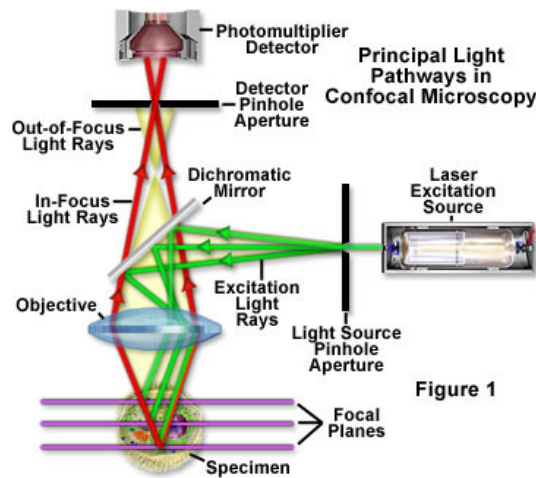


Figure 1.7: Confocal Scanning Laser Microscopy (CSLM). Only the light from the exact focal plane of interest is able to return to the detector via a pinhole, while out of focus light is excluded. (From Science Education Resource Center at Carleton College (SERC))

number of lesions can be examined during the same office visit, using telepathology. CLSM images can be viewed simultaneously by the clinician and the pathologist facilitating the decision of biopsy, clinically follow up, or to excise the lesion. One of the limitations of this technique is the strong contrast attenuation and light scattering caused by hyper-pigmented or hyperkeratotic lesions.

Some examples of CLSM are shown in Tab. 1.12.

Table 1.12: Confocal Scanning Laser Microscopy devices

Device Name	Company
Vivascope 1500, 2500 and 3000 [®]	Lucid Inc, NY, USA
Optiscan F900	Optiscan Pvt Ltd, Australia
LSM 700, 710 and 780	Carl Zeiss MicroImaging, LLC, North America

1.2.2.3 Multispectral Digital Dermoscopy

Multispectral digital dermoscopy consists of a CCD camera, a multispectral source (white lamp), and some simple intermediary optics. The system allows

1. INTRODUCTION

the visualization of the horizontal plane of a given skin lesion down to the papillary dermis, 0.75 mm. from the surface skin[174]. Using computer-based image analysis, the various images can be combined providing information from a range of depths in a lesion (see Fig. 1.8).

This technique allows physicians to analyze features indiscernible to the human eye, probing up to 2 mm below the surface, and limiting physician-to-physician variability. The following table 1.13 enumerates commercial solutions based on this technique.

Table 1.13: Multispectral Digital Dermoscopy

Device Name	Company
SIAscope	Biocompatibles International plc [120]
MelaFind [®]	Electro-Optical Sciences [121]
Stratum	Optiscan Pty Ltd, Melbourne

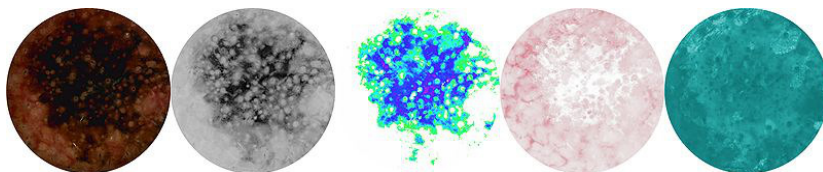


Figure 1.8: Siascan example. These images represent (from left to right) hemoglobin, collagen, melanin, dermal melanin and a dermatoscopic view of a lesion (From: Biocompatibles webpage: <http://siascopy.biocompatibles.com>)

1.2.2.4 Ultrasound

Ultrasound images are created due to the different acoustic properties of tissues. High-frequency sound impulses are transmitted into the skin and then reflected, refracted, or inflected when tissues interface when different acoustic impedance is encountered [82].

There are three modes of ultrasound scans; the A-mode scan that displays a one-dimensional graphic showing the amplitude of the intensity at different levels in the skin tissue, the B-mode scan which creates two-dimensional images

from the brightness level of multiple A-mode scans and is most commonly used in clinical settings, and the C-mode scan whose image is formed in a plane normal to a B-mode image. A gate that selects data from a specific depth from an A-mode line is used, then the transducer is moved in the 2D plane to sample the entire region at this fixed depth. When the transducer traverses the area in a spiral, an area of 100 cm² can be scanned.

Transducers with higher frequency wavelengths are beneficial for diagnosing skin lesions because they allow better resolution of small lesions located near the skin surface[174][113]. However, higher frequency suffers from larger attenuation of the signal, decreasing the depth of penetration. Table 1.14 shows the resolution and penetration for commonly used ultrasonic devices available for dermatological use. The axial resolution is the smallest thickness that can be measured and the lateral resolution refers to the width of the smallest structures that can be seen.

Because of the limited resolution, ultrasound alone is not a reliable diagnostic aid. It is more appropriately used for preoperative management in dermatology, as, for example, in assessing tumor thickness and vascularity. Some of the ultrasound devices used are cited in table 1.15.

Table 1.14: Ultrasound frequency

Frequency (MHz)	Axial Resolution (μm)	Lateral Resolution (μm)	Penetration (mm)
7.5	200	400	>15
10	150	300	>15
20	100	350	7
40	30	94	4
50	39	120	4
100	11	30	2

1.2.2.5 Optical Coherence Tomography

Optical Coherence Tomography (OCT) was originally used to examine eye structures in the 1980s but it is currently widely used in dermatology[180]. OCT is analogous to ultrasound B imaging, except that it uses light rather than sound waves[111]. It uses a fiber-optic Michelson interferometer with a low-coherence length broadband light source, reaching a penetration depth of about 1 mm (de-

1. INTRODUCTION

Table 1.15: Ultrasound device systems

Device Name	Company
DermaScan C	Cortex Technology, Hadsund, Denmark
DUB 20	Taberna pro medicum, Lüneburg, Germany
SSA-340 A	Toshiba Medical Systems, Neuss, Germany
Siemens Sonoline Elegra	Siemens, Erlangen, Germany
AU 4 Idea and AU 5 Idea sonography	Esaote Biomedica, Genoa, Italy
Dermcup 2020 [®]	Atys Medica, St Soucieu en Jarrest, France

pending on the scattering properties of the tissue) (see Fig. 1.9). Lateral resolution is determined by the numeric aperture of the objective[169].

The resolution of OCT does not reach the capabilities of reflectance confocal microscopy of histopathology, however, cellular details can be viewed with the more modern devices (Table 1.16 presents some modern devices). Compared with ultrasound, it has a deeper detection depth and a stronger resolution compared to CSLM[176].

Table 1.16: Optical Coherence Tomography

Device Name	Company
SkinDex-300	ISIS optronics GmbH, Mannheim, Germany
VivoSight [®]	Michelson Diagnostics Ltd., Orpington, Kent, UK
Stratus _{OCT}	Carl Zeiss Ophthalmic Systems, Dublin, CA, USA
Skintell	Agfa Healthcare, Mortsel, Belgium

1.2.2.6 Tape Stripping mRNA

Tape Stripping Messenger Ribonucleic Acid (mRNA) is a non-invasive method that allows for the recovery of cells comprising the upper epidermis. When the tape is removed from the skin, superficial cell layers of the stratum corneum are stripped off and RNA is harvested from these skin samples. These mRNA are analyzed by ribonucleic protection assay (RPA) to differentiate melanoma from benign lesions based on gene expression profiles[122].

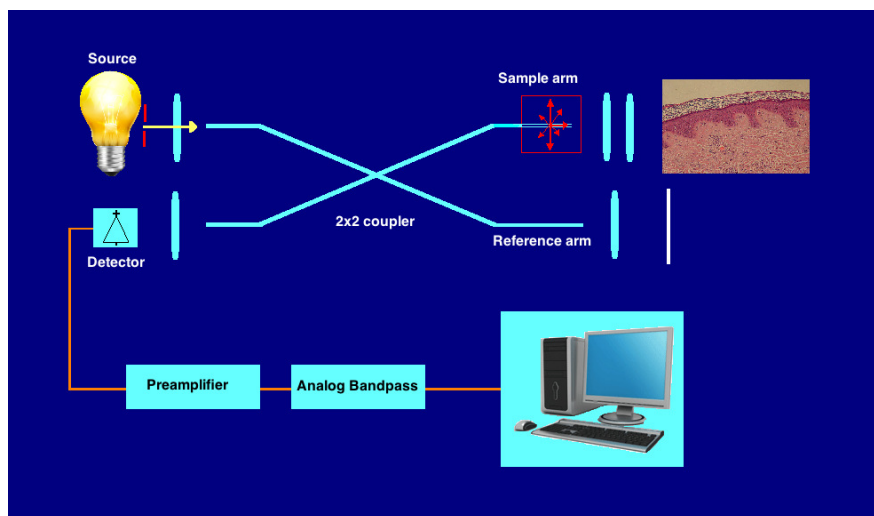


Figure 1.9: Optical Coherence Tomography. The light from a diode is channeled into optical fibers and divided into a reference and a sample beam. The reference beam is reflected by a scanning mirror system, while the sample beam is focused onto the skin. Backscattered photons from the skin recombine with the reference signal and are detected by interference if they match within the short coherence length (from Isis Optronics GmbH, Germany)

In [18] the tape-stripping, the toluidine blue method demonstrated a sensitivity of 68.7% and a specificity of 74.5% over 150 suspicious pigmented lesions.

Clinical trials are currently underway to finalize candidate gene expression profiles for identifying early stage melanomas. Tape stripping will not be a substitute for necessary biopsies, but it is most beneficial as a pre-screen for suspiciously pigmented lesions. Two devices are available (see 1.17).

Table 1.17: Tape Stripping mRNA

Device Name	Company
DermTech	La Jolla, CA. Epidermal Genetic Information Retrieval (EGIR™)
siRNAsense	siRNAsense, Oslo, Norway

1. INTRODUCTION

1.2.2.7 Laser Doppler Perfusion Imaging (LDPI)

Malignant melanomas usually show a higher heterogeneity in their structure and a higher vessel density when compared to benign PSLs because of neovascularization which starts very early during the radial growth phase [13].

The principle of LDPI is the Doppler effect on monochromatic radiation caused by movement of erythrocytes in the micro-vascular network [166]. The output of the LDPI system consists of two different two-dimensional data sets, perfusion and total backscattered light intensity (TLI), with point-to-point correspondence.

LDPI has a sensitivity of almost 100 % and a specificity of 85-90 % [165]. Some commercially available devices are presented in Table 1.18.

Table 1.18: Laser Doppler Perfusion Imaging (LDPI)

Company	Location
Oxford Optronix Ltd.	Oxford, United Kingdom
Moor Instruments	Devon, United Kingdom
Perimed AB	Järfälla, Sweden

1.2.2.8 Electrical Bio-Impedance

Electrical Bio-impedance is a new non-invasive approach, based on the electrical differences between malignant and benign skin lesions. Electrical impedance scanning (EIS) reflects information about cell shape, structure and orientation, integrity of cell membranes, relative proportions of intra and extra cellular fluids and ionic composition.

During an EIS examination, an alternating electric field is created between a voltage source and a measuring probe placed at the site of the suspicious lesion. The detection probe is composed of electrodes that penetrate the stratum corneum. The induced electric current is detected at each sensor element and measured using a trans-impedance measurement technique.

Newer models of electrical impedance have a digital camera along with an automated software analysis tool (see Table 1.19). The electric current is registered

at each sensor element and measured using a trans-impedance technique, while the other electrodes remain at ground potential[54].

A digital picture of the lesion and its surroundings is recorded together with a close-up frame of the lesion. The borders and the axis of the lesion are displayed on the computer, and the software automatically extracts five parameters of each lesion. The asymmetry (A1 and A2) defines the ratio between the non-overlapping area of the lesion when folded over either of the two perpendicular axes, and the total area. The Border (B) defines the ratio between the squared perimeter of the lesion and a factor of the area inside the border. The Color parameter (C) is the standard deviation of the red-gray levels inside the border. The Surface (S) is the area of the lesion in square millimeters. Bioimpedance measurements of suspiciously pigmented lesions are taken at both the center of the lesion and a noninvolved reference skin site. Lesional and reference skin are measured at 5 depth levels, approximately 0.1 mm to 2 mm into the tissue, and data are analyzed by a computer. The entire process takes approximately 7 minutes to complete [69].

Electrical impedance is almost 100% for in situ and thin melanomas [69]. Electrical impedance can also differentiate melanoma from benign nevi with studies demonstrating ranges of 92-100% sensitivity and 67-75% specificity. However, electrical impedance properties of human skin vary significantly with the body location, age, gender and season [1].

Table 1.19: Electrical Bio-Impedance

Device Name	Company
TransScan Ltd.	Migdal Haemek, Israel
SciBase II impedance spectrometer	SciBase AB, Huddinge, Sweden
TS2000M	Mirabel Medical System Ltd., Migdal, Ha'Emek, Israel

1.2.2.9 Magnetic Resonance Imaging (MRI)

MRI scans utilize radio waves and strong magnets instead of X-rays. The working principle is based on the absorption and re-emission of radio waves from tissue

1. INTRODUCTION

protons exposed to a strong magnetic field. Under the influence of radio frequency pulses, a proton returns to a stable low-energy state and emits radio waves that are detected by the coil[30].

Specific imaging devices have been developed that permit high-resolution MRI imaging of the skin (see Table 1.20), allowing differentiation of the stratum corneum, epidermis and dermis in vivo with an image acquisition that requires around 3-4 minutes for a section thickness of 1.2 mm. The epidermis appears as a high-intensity layer, while the dermis appears as hypointense with an irregular interface of sub-dermal fat[60].

MRI scanning has some disadvantages, namely its cost, size, duration of evaluation time, and the need for specialized training. Also, it is not suitable for patients with metal implants.

Table 1.20: Magnetic Resonance Imaging (MRI)

Device Name	Company
Gyrosan Intera	Philips Medical Systems, Amsterdam, Netherlands
Magnetom	Siemens, Erlangen, Germany
1.5T GE Signa HDx scanner	General Electric HealthCare, Waukesha, USA

1.2.2.10 Positron Emission Computed Tomography(PET)

PET is a non-invasive, high-resolution imaging technique used to detect the metastatic spread of melanoma. PET shows great promise in the detection of metastatic cutaneous melanoma and may also prove useful in the secondary prevention of primary melanoma in those individuals at high risk or with a familial disposition[52]. Some PET systems are presented in Table 1.21.

Although rare, primary melanomas have also been found in ocular, gastrointestinal, anorectal, genitourinary, mucosal, leptomeningeal, sinonasal, pulmonary, mediastinal, ovarian, vaginal, and vulvar sites and can represent diagnostic challenges [61][155]. PET may be valuable in detecting these primary melanocytic lesions in non-skin sites as a dermatologist's trained eye and the other diagnostic

techniques described above can only detect those primary melanomas located on the skin.

Table 1.21: Positron Emission Computed Tomography(PET)

Device Name	Company
Discovery VCT	General Electric HealthCare, Waukesha, USA
Biograph mCT	Siemens, Erlangen, Germany

1.2.2.11 Reflex Transmission Imaging

Reflex Transmission Imaging (RTI) is a form of high-resolution ultrasound that can be used in combination with white light digital photography for classification of pigmented lesions [142]. Table 1.22 presents an ultrasound device that can be combined with a digital camera.

Some studies have present good enough results in discriminating among melanomas, seborrheic keratoses, and nevi to potentially reduce the referral of benign tumors by 65% without missing melanomas. Taking into account the limited reports using RTI in the literature and its high cost, its future use as a primary imaging modality for melanoma detection is unclear[141].

Table 1.22: Reflex Transmission Imaging

Device Name	Company
DermaScan Cv3 TM	Cortex Technology, Hadsund, Denmark

1.3 Scope of Research

The Computer Vision and Robotics group (VICOROB) of the University of Girona has been working on the development of skin cancer tools since 2006. Two main directions have been explored: the study and development of algorithms to detect temporal changes of skin diseases and the implementation of a skin scanner technology for acquiring full body imaging.

1. INTRODUCTION

The Spin-off Coronis Computing of the VICOROB research group is involved in the project, helping in the development and research on the acquisition, analysis and processing of medical images. Coronis was started in 2007 by two professors of the University of Girona, Rafael García and Jordi Freixenet who were later joined by Susana Puig and Josep Malvehy, physicians in the Melanoma Unit at the Hospital Clinic of Barcelona.

The cooperation with the Melanoma Unit at the Hospital Clinic in Barcelona provided feedback during the development of this thesis, adding a valuable contribution to all the objectives. The Melanoma Unit was created in 2001 by the Dermatology and Medical Oncology services and has become a reference service around the world.

1.4 Objectives of the thesis

The main goal of this thesis is to obtain a new framework for the diagnosis of cutaneous skin cancer. We focus the interest of this research on three main objectives:

The first goal includes a pre-processing step for **correcting non uniform dermatoscopic images** in order to help physicians using different dermoscopes or cameras, and also to provide a uniform image acquisition useful for teledermatology.

The second goal is to **improve Total Body Photography** providing tools for physicians in order to register different sequences of images in an automated way.

The third goal involves the **implementation of a full body scanner** to acquire cutaneous images. The scanner will increase the resolution of the already existing systems, in order to provide the first medical tool able to acquire images of the whole body in an automatic way with enough resolution to file high specificity and sensitivity of the skin cancer diagnosis.

1.5 Thesis Outline

In the next subsection, we summarize the proposed framework designed according to the objectives mentioned. Subsequently, an overview of this document is described.

1.5.1 A new framework for Skin Cancer Detection

We introduced above the diagnostic techniques for skin cancer detection that are currently in use. We pointed out some of their advantages and disadvantages. Keeping these systems in mind we incorporate as many of their good characteristics as possible and solve their inherent problems.

For this reason, we propose a new framework for skin cancer detection able to resolve the poor resolution of the existing systems by employing high quality image sensors and solving the non-uniform lighting with post-processing techniques. We also want to automatize the acquisition process with images having a closer resolution to the ones described in dermoscopy rather than the ones described in baseline photography.

We aim to provide the first tool for a reliable method to detect skin cancer based on detecting changes.

1.5.2 Document Overview

This thesis is structured according to the mentioned objectives. The introduction showed the existing Diagnostic Systems for dermatology.

Chapters 2, 3 and 4 summarize our proposal describing in greater detail the 3 objectives stated above. Results and conclusions are provided for each of them.

Chapter 2 describes a color correction pipeline applied to dermatoscopical images in order to unify the images obtained with different cameras and dermatoscopes.

Chapter 3 covers the skin mosaicing system proposed to facilitate the task of registering different total body images in order to enable a temporal study of the patient.

1. INTRODUCTION

Thus, in **Chapter 4**, a new proposal for full body image acquisition is proposed. The chapter describes the hardware and software used in our proposal and shows preliminary results in 2D and 3D reproductions of the body.

Finally, the thesis ends with **Chapter 5**, where the conclusions are summarized and further work is pointed out. Moreover, in this chapter, a list of publications related to this thesis is included.

Chapter 2

Color Calibration

Different approaches for color calibration have been proposed in order to improve accuracy in the diagnosis of melanoma. In this chapter, we first describe some of them and secondly, we present our developed method for improving their results.

The problem to be solved is that images recorded using dermatoscopes and commercial digital cameras present differences in the colors recorded. So, they present important differences in the image depending on the camera, the different dermatoscopic systems, the computer as well as the screen used by the physician (see Fig. 2.1). The camera is affected by various factors such as the non-linearities of the sensor, the management of the white reference point used (white balancing) and the storage format. Moreover, the dermatoscopes on the market at the moment present different characteristics in the spectrum and uniformity of the lighting, because the illumination of the images acquired with each device is different, and also if the dermatoscope is in direct contact with the skin or has any liquid between it and the skin, it will cause subtle color differences in the images acquired. On the computer, every graphical card has different properties such as the gain incorporated in the R, G and B output channels. Also graphical cards can have different output signals (analog or digital) and a different number of bits. These differences will generate a different signal for the screen. Some visualization software may also make adjustments to the images to increase image contrast, modifying the colors of the images originally acquired by the camera sensor. Unfortunately, the corrected images may have a different appearance on different LCD monitors. This issue should be addressed by monitor calibration,

2. COLOR CALIBRATION

ensuring correct colors according to the reproducible gamut of the screen.

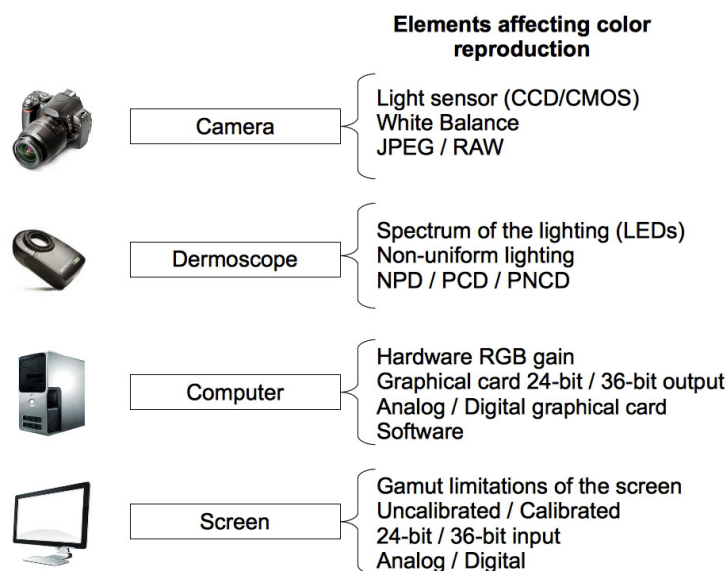


Figure 2.1: Elements affecting color reproduction.

Also, due to their nature, dermatoscopes also reduce the number of observable colors that can be noticed, particularly for the whitish-blue hue[135]. The differences between acquisition systems makes the task of detecting melanoma difficult if dermatologists have not been trained with the same camera and dermatoscope they normally use and makes the use of color information with automatic image processing techniques impossible.

2.1 Related Works

Color calibration is a well-known problem in the computer vision community [147]. Several methods have been proposed to solve it. Basically, they can be divided into three main groups:

1. Methods that directly compute the spectral sensitivity functions for the R, G and B channels.

-
2. Methods that compute the response of the device by taking pictures of a known set of color samples.
 3. Methods that use a variety of other color models.

The first group computes the spectral sensitivity for each channel (R, G, B) [12], using a spectrophotometer and a light source. The light source produces a series of spectral distributions with a narrow bandwidth. These lights are acquired at the same time by the spectrophotometer and the camera, and both responses are used to estimate the camera sensitivity curve. Accurate response curves can also be measured with high spectral resolution using a monochromator or tunable laser illumination over the full visible spectrum.

The second group relates the RGB values of the device with the XYZ values already known from a color chart (e.g. Gretag Macbeth [116]) and with a known lighting system (typically D65). CIE 1931 XYZ color space, created by the International Commission on Illumination (CIE) in 1931 is a commonly used standard, and serves as the basis from which many other color spaces are defined. The relation between the pixel RGB values and XYZ will be computed using a polynomial regression method.

In the dermatological context, Haeghen and Naeyaert[64] presented a method for color correction in clinical images. Their method requires a visible color chart in the clinical image and the color correction is computed for each image. In a previous work by the same research group, Haeghen et al.[63] tested different non-linear operators for estimating the relationship between the camera and the calibration chart. As opposed to the work presented by Haeghen and Naeyaert[64], the approach proposed in [63] was designed to be used on their custom imaging system, which is similar to a dermatoscope. The images were extracted from the video frames provided by the camera and they needed to tune some offsets on the camera to adjust some settings, such as the gamma factor.

A color calibration system for clinical images of human wounds was described by Van Poucke et al.[138]. Their system uses the JPEG values of two commercial digital cameras, one Canon D10 and a Nikon D200. The procedure described in the paper is similar to that proposed by Haeghen and Naeyaert[64]. The system first performs an inverse gamma correction for obtaining lineal RGB values, then

2. COLOR CALIBRATION

a linearizing LUT is applied for gray-balancing, and finally a polynomial transformation (linear, quadratic or cubic) is performed for each color channel, relating the acquired Gretag Macbeth[®] color patches to the theoretical spectrophotometric values. The output image is gamma corrected for further processing.

Grana et al.[59][58] performed color calibration on a FotoFinder[®] video microscope. Their method first corrects the effects of non-uniform illumination of the instrument, then a gamma value is estimated to compensate for the visualization effects introduced by the apparatus and finally, different non-linear operators are applied for computing the calibration. Once XYZ color values have been obtained, Grana et al. measure Euclidean distances between the average lesion and skin RGB values and histogram intersection percentages.

Maglogiannis et al.[110] presented their own system for clinical dermatological imaging, composed of two CCD cameras and a polarized lighting system. The system first acquires an image of a Gretag Macbeth[®] color chart. Then, the black/white and grayscale patches are corrected and an image averaging is performed. Finally the values obtained and the reference RGB values of the patches are interpolated, obtaining a Look Up Table (LUT). This table is then used on real images for mapping the skin colors. This method also performs radial lens distortion correction in both cameras finally obtaining the 3D reconstruction of the skin.

The benefits of RAW with respect to JPEG are presented in the Wighton et al.[182] work in which the authors compared the accuracy and precision values for a Canon camera acquiring RAW images with those of a Sony system that only provided JPEG images. In their work, Wighton et al. [182] performed color and lighting correction and also presented a method able to correct chromatic aberrations in JPEG and RAW images.

The third group comprehends color correction methods using other color models. Iyatomi et al.[80] [81] developed a fully automated method for dermoscopy, making use of the HSV color system. This method performs normalization on hue, saturation and intensity channels independently. The normalization is computed with filters generated with statistical studies of already known datasets. The performance of this method depends on the quality of the dataset used, and how similar the new images are, compared with those in the dataset. This

method does not require a color chart, and can ensure the constancy of images. However, the restored colors may be very different from the real ones, since there is no *ground-truth* calibration. Moreover, when two images are different (due to any change over time) the system would tend to minimize this difference.

With the goal of detecting lesion borders in dermatoscopic images, Schaefer et al.[151] [152] performed a color normalization, named Automatic Color Equalization (ACE), as a pre-processing step. This algorithm consists of two main stages: (1) chromatic/spatial adjustment and (2) dynamic tone reproduction scaling. The first step performs color normalization and image contrast enhancement, while the second performs accurate tone mapping and lighting constancy. While contrast is increased between the lesion and the background skin, the contrast within the lesion is decreased. However, although this characteristic helps their algorithm to distinguish between the lesion and the skin, the fact that the lesion's contrast is decreased does not allow further analysis in the final image. They propose returning to the original image once the segmentation has been computed. At this point, if any further study is needed, colors should be corrected using a color calibration process (such as the one proposed in this Chapter).

2.2 Theoretical Model

Typically, physicians acquire images with digital cameras set (by default) to auto mode and auto white balance. This camera setup tries to fix the white point for each image with the data obtained from the scene. This may cause each image to look different even though the camera and the dermatoscope are the same because the white reference point may be different for each image. Moreover, the camera also tries to change the exposure time and aperture value automatically to obtain the best-fitting histogram. Finally, when storing JPEG images, the camera carries out an additional internal manipulation of the bitmap, an unknown black-box that varies from one manufacturer to another. For all these reasons, we propose a new method to acquire and correct images and to obtain more realistic colors and make them identical on different acquisition systems.

We propose a new calibration pipeline (see Figure 2.2) that starts by acquiring RAW instead of JPEG images. RAW images have been selected since they provide

2. COLOR CALIBRATION

linear values directly, and non-linear correction is not needed. After the images are acquired, they are transformed from the specific camera format (in our case CR2) to a standard raster graphics format such as TIFF. This transformation can be done using *dcraw* (free software for converting RAW images to TIFF). With this program, it is possible to obtain the original linear values of the camera without any extra transformation.

Calibration images are acquired using the camera Av mode. This option is selected to facilitate the acquisition of dark and light images without saturating the camera's dynamic range.

After transforming the images and before starting the calibration procedure, the different exposure times for each image must be compensated. This normalization is basically carried out in two steps. First the image values are divided by their exposure time (Eq. 2.1) and then divided by the maximum exposition of the full image sequence.

$$R = \frac{R}{E}, \quad G = \frac{G}{E}, \quad B = \frac{B}{E} \quad (2.1)$$

where E is equal to the exposure time.

Once the images are normalized, the calibration procedure takes place. This process relates the camera-dependent color space with an independent color space (XYZ). Our system relates the 24 colors from a Gretag Macbeth color checker chart or from a Digital Color Checker SG (see Figure 2.3) obtained by the acquisition system with the expected XYZ values measured by a spectrophotometer. The XYZ values are often referred with ideal daylight, with the standard CIE D65 light source approximated during the computation in the spectrophotometer. The correct way is to apply the spectral reflectivity of the sample and multiply it by the spectral distribution of the light source, being in our case, the spectral data of the LED used in the dermatoscope (see Figure 2.4).

XYZ values can be computed by Eq. (2.2) if the light source is taken into consideration. This equation uses the CIE 1931 color-matching function tables ($\hat{x}(\lambda_i)$, $\hat{y}(\lambda_i)$ and $\hat{z}(\lambda_i)$) that correspond to the spectral sensitivity curves of the three linear light detectors that yield the CIE XYZ tristimulus values X , Y , and Z , the reflectance of each patch ($s(\lambda)$), and the spectral distribution of the

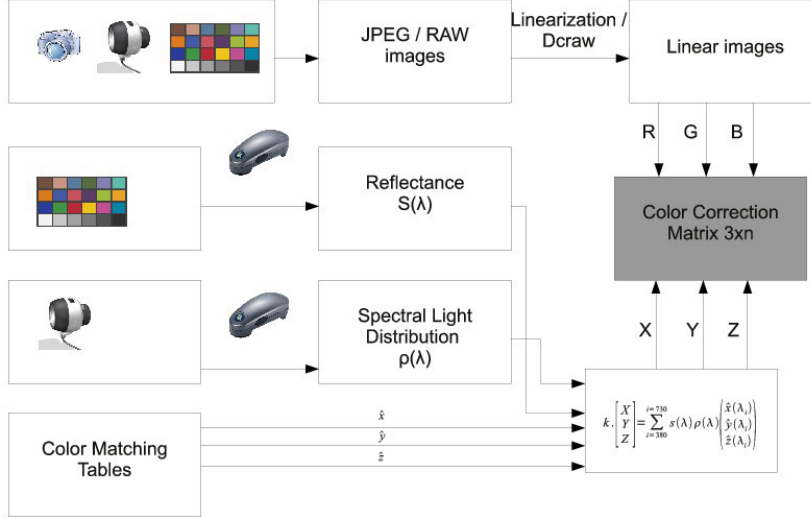


Figure 2.2: Camera Calibration Pipeline. In the first step, the 24 color chart images are acquired. Next the XYZ values are computed using the color reflectance values, the dermatoscope spectral light distribution, and the color matching tables. Finally, the color correction matrix is computed to relate both colorspace.

dermatoscope's LED light ($\rho(\lambda)$). K is used to make Y equal to 100 for the white reference.

$$k \cdot \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}_{LED} = \sum_{i=380}^{i=730} s(\lambda_i) \rho(\lambda_i) \begin{pmatrix} \hat{x}(\lambda_i) \\ \hat{y}(\lambda_i) \\ \hat{z}(\lambda_i) \end{pmatrix} \quad (2.2)$$

Given the RGB values measured by the camera (v) and the XYZ values already computed (v_{ref}), it is possible to compute the mathematical relationship (M) between them (Eq. 2.3).

$$v_{ref} = Mv \quad (2.3)$$

Eq. 2.3 can be extended using Eq. 2.4, where v_{ref} is formed by the n^{th} X, Y, Z triplets, one for each Gretag Macbeth patch, v is formed by the n^{th} R,G,B triplets

2. COLOR CALIBRATION



Figure 2.3: Color Checker. The Gretag Macbeth[®] color checker chart and the Digital Color Checker. SG[®] chart

and α_{ij} are the elements of the correction matrix.

$$\begin{bmatrix} v_{ref}^1 & \cdots & v_{ref}^n \end{bmatrix} = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{bmatrix} \begin{bmatrix} v^1 \cdots v^n \end{bmatrix} \quad (2.4)$$

Eq. 2.4 can be rearranged into a linear algebraic formulation:

$$\begin{bmatrix} v^1 \\ \vdots \\ v^n \end{bmatrix} = \begin{bmatrix} \alpha_{11} \\ \alpha_{12} \\ \alpha_{13} \\ \alpha_{21} \\ \alpha_{22} \\ \alpha_{23} \\ \alpha_{31} \\ \alpha_{32} \\ \alpha_{33} \end{bmatrix} \begin{bmatrix} v_{ref}^1 \\ \vdots \\ v_{ref}^n \end{bmatrix} \quad (2.5)$$

or:

$$A_{3n \times 9} X_{9 \times 1} = b_{3n \times 1} \quad (2.6)$$

where A and b are known and x can be computed using SVD or least squares.

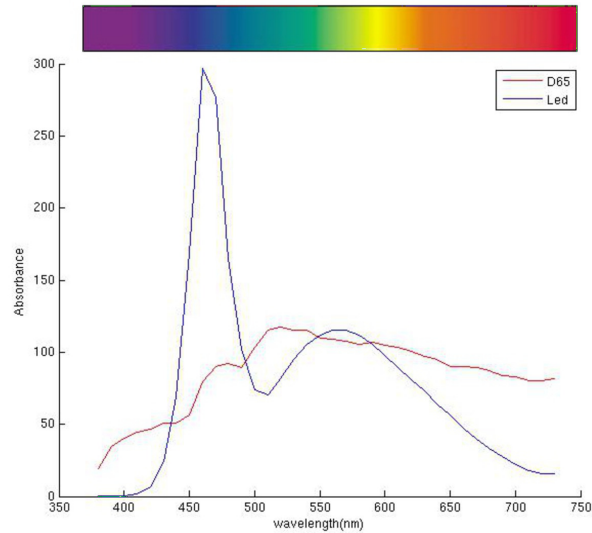


Figure 2.4: Relative spectral power distribution of Dermlite Foto[®] acquired with X-Rite i1[®] Pro spectrophotometer(blue) and illuminant D65 (red). The y axis shows the relative spectrum power while the x axis shows the wavelength (nm).

Is it also possible to use a weighted version of Least Squares, adding weights to colors similar to the skin and the pigmented lesions, but then the final results will be more accurate on the skin color reproduction test but highly inaccurate on the colors not presents or less weighted in the calibration.

The previous equations take into account only the R , G and B values for each camera in order to compute the relation between colorspace. Other works such as Hong et al. [75], suggests the use of more parameters to compute the M matrix. These parameters represent the covariance relationship between color spaces.

We tested our approach using the following polynomials:

2. COLOR CALIBRATION

$$\begin{aligned}
M_{[3 \times 3]} &: [R \ G \ B] \\
M_{[3 \times 10]} &: [R \ G \ B \ RG \ RB \ GB \ R^2 \ G^2 \ B^2] \\
M_{[3 \times 10]} &: [R \ G \ B \ RG \ RB \ GB \ R^2 \ G^2 \ B^2 \ 1] \\
M_{[3 \times 11]} &: [R \ G \ B \ RG \ RB \ GB \ R^2 \ G^2 \ B^2 \ RGB \ 1] \\
M_{[3 \times 20]} &: [R \ G \ B \ RG \ RB \ GB \ R^2 \ G^2 \ B^2 \ RGB \ R^2G \ G^2B \\
&\quad B^2R \ R^2B \ G^2R \ B^2G \ R^3 \ G^3 \ B^3 \ 1] \\
M_{[3 \times 35]} &: [R \ G \ B \ RG \ RB \ GB \ R^2 \ G^2 \ B^2 \ RGB \ R^2G \ G^2B \\
&\quad B^2R \ R^2B \ G^2R \ B^2G \ R^3 \ G^3 \ B^3 \ R^3G \ R^3B \ G^3R \ G^3B \\
&\quad B^3R \ B^3G \ R^2GB \ RG^2B \ RGB^2 \ RGB^2 \ R^2G^2 \ R^2B^2 \\
&\quad G^2B^2 \ R^4 \ G^4 \ B^4 \ 1]
\end{aligned} \tag{2.7}$$

These polynomials will increase the size of the M matrix in a range from 9 to 105 parameters.

Although RAW values would provide the best image quality since they provide high bit depth (e.g. 12 bits per channel) and accurate linear RGB values, many cameras produce only JPEG images. Therefore, when RAW values are not available, we propose two different solutions. First, color calibration should be carried out using the JPEG values from the camera. Second, a white patch of the calibration pattern should be used to set the custom white balance (CWB) setting of the camera.

In any case, when images are corrected, we propose an automatic exposure correction of the images based on the Minimal Information Loss algorithm. The dynamic range of each image is stretched by the same prescribed value that is applied for pixel intensities in the logarithmic domain. In this way we obtain equally well-exposed images while the uniformity of their color appearance is retained (up to what the color calibration is able to provide).

2.2.1 JPEG limitations

The camera can be calibrated using the JPEG values and the same procedure as that for the RAW values. The only difference is the way the input data is preprocessed. A JPEG image has two main sources of nonlinearity in its values.

First of all, the acquired images are affected by the sensor's non-linear response. These nonlinearities are particular for each camera and affect the final results in different ways. Second, manufacturers build visual effects into the cameras such as contrast or gamma-curve enhancements or increased color saturation to enhance the image's appearance. For amateur usage these effects typically produce a pleasant appearance, but in our case they are unknown factors that affect the final results.

Fortunately, the two main sources of nonlinearity described below are significantly less important than the third effect, which is the gamma curve that is finally applied to the image. The non-linear gamma value is used in all JPEG images to compensate for the displays. This power function also compensates for the nonuniformity of the luminance sensitivity of the human eye. The eye perceives approximate light intensity uniformly on a logarithmic scale, and displays are produced following this nonlinear function. In this way, uniform luminance steps can be ensured over a long scale (1,2, ...255), and with fine tone transitions. Therefore, for a JPEG image, we need to obtain linear RGB values through the inversion of this built-in gamma nonlinearity. The rules of nonlinearity are defined in the sRGB¹ conversion formulas.

To remove these nonlinearities, three alternative procedures can be applied:

1. Exploit the non-linear sRGB to the linear RGB algorithm (Alg. 1) taking into account a standard gamma value (2.4).
2. Estimate the CCD response curve using the Debevec algorithm [43].
3. Compute the camera response curve by fitting a function between the images acquired from a known gray scale chart and the expected values of this chart.

¹sRGB is a color standard defined by the International Electrotechnical Commission (IEC) as IEC 61966-2-1 (IEC, 1999), and involves recommended parameters for displaying photographic images on a monitor.

2. COLOR CALIBRATION

Algorithm 1 Non-linear sRGB to linear RGB

```

if R,G,B ≤ 0.04045 then
  sR = R ÷ 12.92
  sG = G ÷ 12.92
  sB = B ÷ 12.92
else
  sR = - ( (-R+0.055) / 1.055 )2.4
  sG = - ( (-G+0.055) / 1.055 )2.4
  sB = - ( (-B+0.055) / 1.055 )2.4
end if

```

2.2.2 Custom White Balance (CWB)

The simplest way to perform color correction is to fix a CWB on the camera. This is a typical option on all digital cameras. It fixes the white point of the camera to match the light source. We propose this method because it is fast and can be executed in real time on the camera. Physicians who use a CWB will be able to see the images from anterior explorations and the new images simultaneously as they acquire each new image. Different images from different cameras will have the same white point value, although other colors may already be shifted away from their real values.

CWB scales color values of all the pixels in an image, so that a *white reference* specified by the user becomes white on the image. In RGB space, a CWB performs the operation defined by (Eq. 2.8).

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 255/R'_w & 0 & 0 \\ 0 & 255/G'_w & 0 \\ 0 & 0 & 255/B'_w \end{bmatrix} \begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} \quad (2.8)$$

where R'_w , G'_w and B'_w are the red, green, and blue components of a pixel believed to be a white surface in the image before the color balancing. The value 255 assumes that images are only 8-bit.

2.3 Experimental Validation

To validate our method, we performed some experiments using various dermatoscopes and cameras. These experiments compare our method with those described in the literature.

Our first validation consisted of obtaining objective results describing the precision of our method. In this validation, we first acquired 24 images from the Gretag Macbeth color chart using three different cameras (a *Canon 5D Mark II*, a *Canon 50D* and a *Canon G9*) with the same dermatoscope (*Dermlite Pro*). These images were used to obtain a set of 12 calibration matrices for each camera. These matrices correspond to the polynomials defined in Eq. 2.7, corresponding to 3, 9, 10, 11, 20 and 35 parameters, respectively. Every matrix is then solved under two different assumptions: first without taking into account the light spectral characteristics, and next by applying our method, which incorporates the specific spectral properties of the light.

Once the calibration step was done, we computed the differences using the 14 skin colors of the *Digital Color Checker SG* not used in the calibration step. These patches were corrected using the calibration matrices and then compared with the values obtained from the spectrophotometer. We also compared these results with the ones obtained by performing the custom white balance.

To compute the difference between two colors, we converted both colors into the CIELab color space (CIE, 1986), and then we computed the Euclidean distance between the colors. CIE $L^*a^*b^*$ color triplets (L^*,a^*,b^*) are a nonlinear transformation of CIE XYZ tristimulus values (X,Y,Z), which are an attempt to obtain a perceptually uniform color space. The transformation from XYZ to Lab can be computed from Eq. 2.9.

$$\begin{aligned}
 L^* &= 116f(Y/Y_w) - 16 \\
 a^* &= 500(f(X/X_w) - f(Y/Y_w)) \\
 b^* &= 200(f(Y/Y_w) - f(Z/Z_w)) \\
 f(t) &= \begin{cases} 1.787t + \frac{16}{116}, & \text{if } t < 0.008856 \\ t^{1/3}, & \text{if } 0.008856 \leq t \leq 1 \end{cases}
 \end{aligned} \tag{2.9}$$

2. COLOR CALIBRATION

where L^* is the luminance and a^* b^* are the perceptually uniform red/green and yellow/blue opponent channels, respectively. X_w, Y_w and Z_w are the tristimulus of the white point. Color differences, ΔE_{ab}^* (CIE, 1986), are defined as the Euclidean distance between two colors (L_1^*, a_1^*, b_1^*) and (L_2^*, a_2^*, b_2^*) .

$$\Delta E_{ab}^* = \sqrt{(\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2} \quad (2.10)$$

According to the literature [24], ΔE differences lower than 4 are not distinguishable to the human eye. In the printing industry, the accepted ΔE_{ab}^* average is 6, while in video camera work and television it is normally set to 5 [37].

Figure 2.5 illustrates the pipeline used for these testing experiments, and Figure 2.9 shows the ΔE values obtained for each matrix and method. The results show that our proposed method performs better than assuming daylight as other projects in the literature do. On the other hand, it also shows how more complex polynomials do not always perform better than simpler ones.

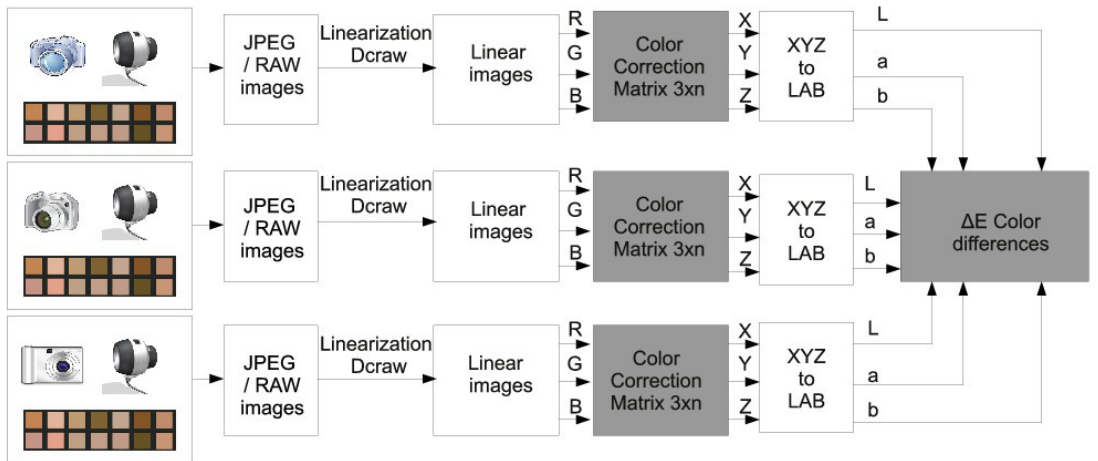


Figure 2.5: Pipeline for computing the ΔE error for the skin color chart using various cameras and dermatoscopes.

Tables 2.1 and 2.2 report the accuracy in color reproduction for a set of 14 skin-color patches that have not been used for color calibration, since the calibration has been carried out by using the 24 standard color patches of the

		Model					
		3	9	10	11	20	35
5D	D65	2.5	4.7	3.7	3.7	11.8	63.8
	SC	2.6	4.5	3.2	2.8	12.2	62.2
G9	D65	4.0	5.3	4.4	3.9	7.5	60.9
	SC	4.0	5.7	4.5	3.9	10.4	58.6
50D	D65	4.1	4.6	4.2	3.9	15.9	70.5
	SC	3.4	4.4	3.5	3.6	13.0	69.0

Table 2.1: Color reproduction accuracy (measured in ΔE) for the 14-skin color patches, after calibration using the 24 Gretag Macbeth color chart. Note that none of the 14 tested colors are included in the color calibration chart.

Gretag Macbeth ColorChecker. It is important to note that this approach of using some colors to calibrate, and using a different set of colors to quantify the accuracy of calibration, is more realistic than just using the same colors for calibration and verification, as done by some methods in the literature [59].

In table 2.1, the ΔE error values are shown for the state-of-the-art approach (D65) versus our proposal that models the light spectrum (SC). Our approach performs better in most cases. On the other hand, it can be observed that higher order polynomials using 20 and 35 parameters never perform better than lower order ones. The 3-parameter polynomial provides the best results for the 2 reflex cameras (5D and 50D), while the 11-parameter polynomial obtains the best results for the G9 camera. It is especially remarkable that our method (which uses RAW data) fails more often in the G9 camera, while it performs much better for the two reflex apparatus. This probably means that the RAW data provided by a compact camera is not as accurate as the RAW images acquired by higher quality cameras.

Figures 2.6, 2.7 and 2.8, presents the ΔE error values for each camera. Each one of the 24 Gretag Macbeth patches and the 14 skin patches are represented for the state-of-the-art approach (D65) and for our proposal that models the light spectrum (SC). The 3-parameter polynomial is the one used for calibrating the cameras. These figures shows that the error distribution is similar between both approaches, but in mean is smaller on the SC approach.

Table 2.2 compares how similar two images are when acquired with different

2. COLOR CALIBRATION

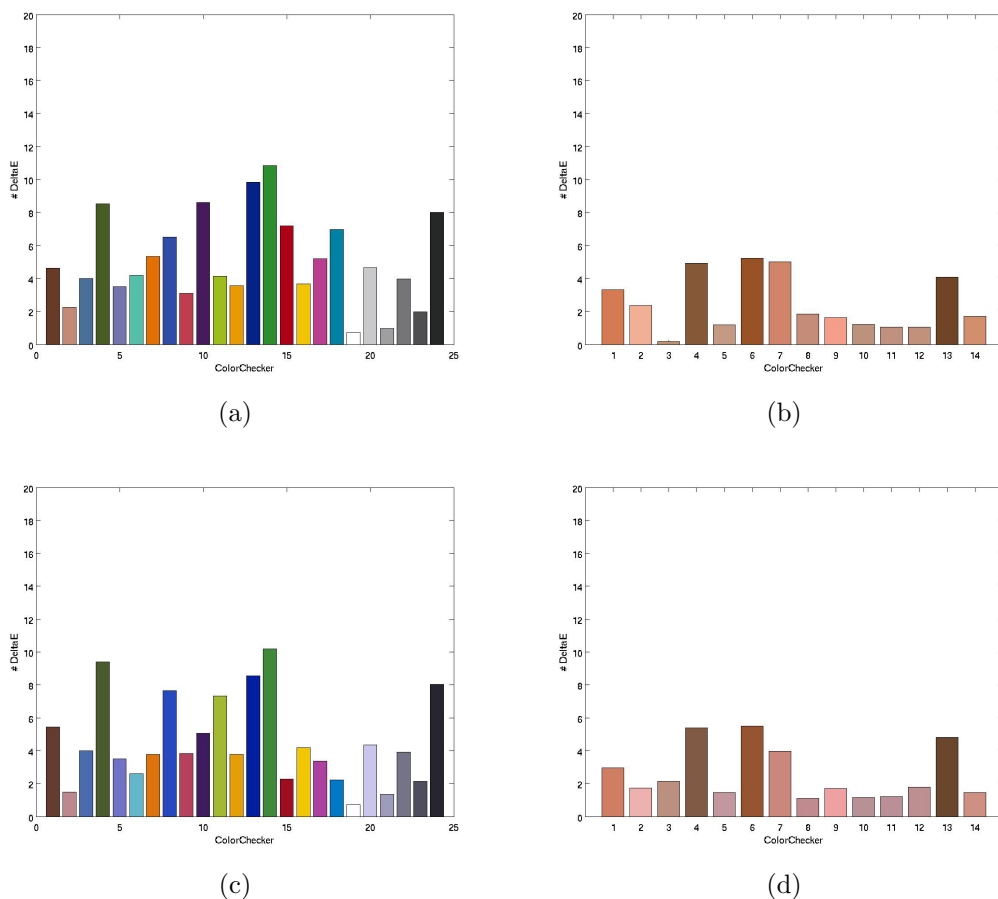


Figure 2.6: ΔE error comparison for the Canon 5D Mark-II camera. Calibration is performed from the 24 Gretag Macbeth patches and results are applied to the same 24 color patches and the 14 skin patches. (a) Error between the 24 calibrated colors and the expected ones using D65. (b) Error between the 14 calibrated skin colors and the expected ones using D65. (c) Error between the 24 calibrated colors and the expected ones using SC. (d) Error between the 14 calibrated skin colors and the expected ones using SC.

cameras. The accuracy is measured in ΔE differences for the 14-skin color patches after calibration using the 24 Gretag Macbeth color chart. Again, none of the 14 tested colors are included in the color calibration chart. As already stated above, higher order polynomials should be avoided since they provide large ΔE values, i.e., the same lesion acquired with two different cameras would have a

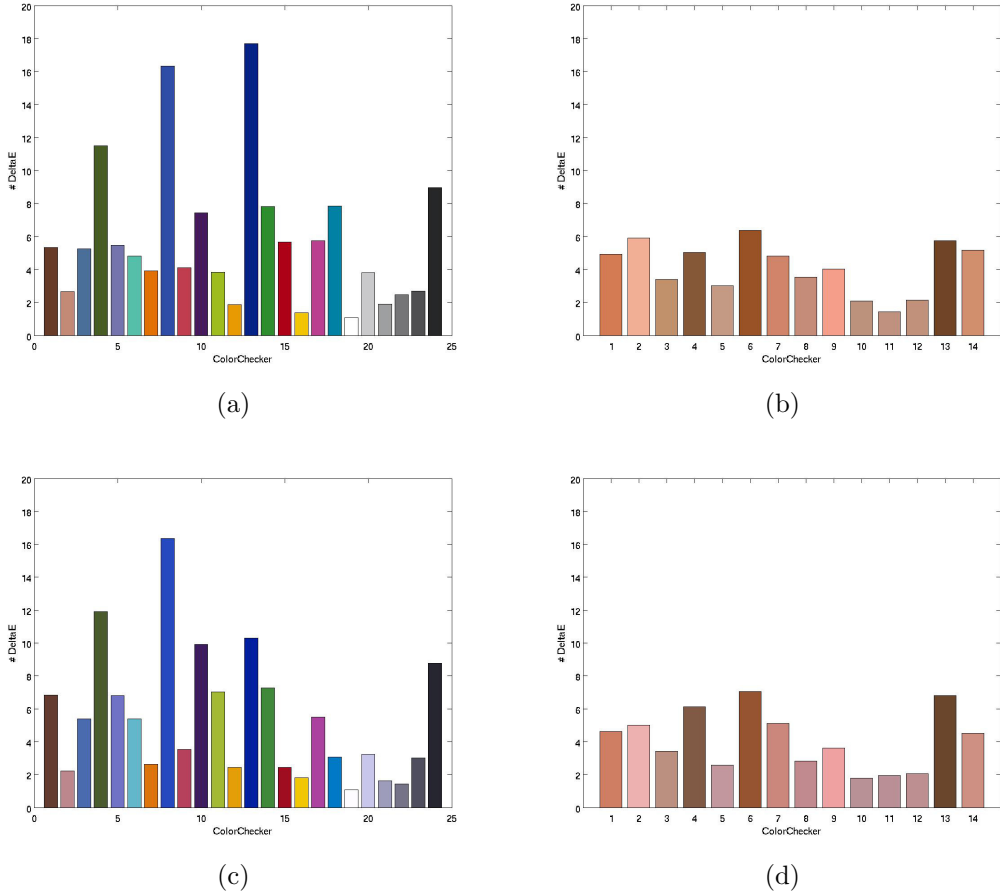


Figure 2.7: ΔE error comparison for the Canon 50D. Calibration is performed from the 24 Gretag Macbeth patches and results are applied to the same 24 color patches and the 14 skin patches. (a) Error between the 24 calibrated colors and the expected ones using D65. (b) Error between the 14 calibrated skin colors and the expected ones using D65. (c) Error between the 24 calibrated colors and the expected ones using SC. (d) Error between the 14 calibrated skin colors and the expected ones using SC.

very different color appearance. In this comparison, our method (SC) always performs better than the state-of-the-art (D65). Moreover, reflex cameras provide better results, independently of the method used (CWB, D65 or SC). Considering the best possible result (marked in boldface), table 2.2 shows that our method improves between 0.1 and 0.9 ΔE with respect to D65.

2. COLOR CALIBRATION

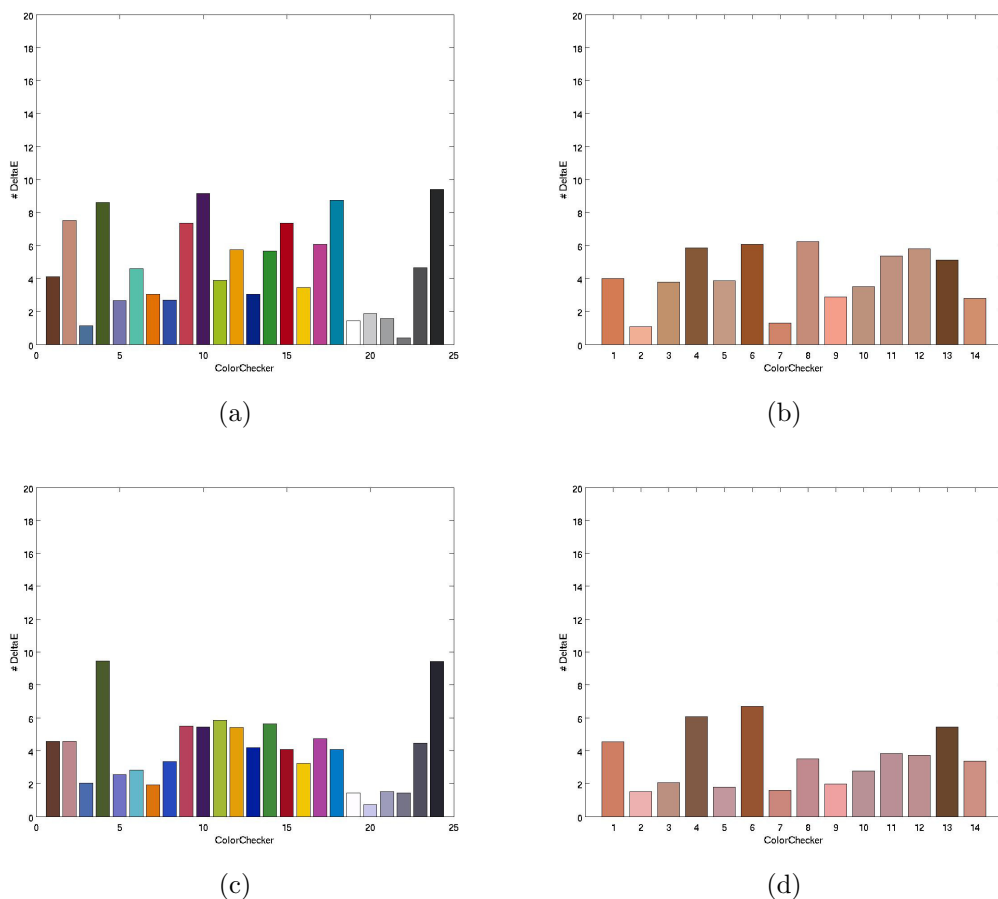


Figure 2.8: ΔE error comparison for the Canon G9. Calibration is performed from the 24 Gretag Macbeth patches and results are applied to the same 24 color patches and the 14 skin patches. (a) Error between the 24 calibrated colors and the expected ones using D65. (b) Error between the 14 calibrated skin colors and the expected ones using D65. (c) Error between the 24 calibrated colors and the expected ones using SC. (d) Error between the 14 calibrated skin colors and the expected ones using SC.

The second test consists of acquiring skin images using two compact cameras (Canon G7 and Canon A640) equipped with polarized and non-polarized dermatoscopes. First, these images were corrected with the parameters obtained through the calibration and, next, non-uniform lighting was corrected, applying the procedure described below. For calibration, an image of a white surface is

		Model					
		3	9	10	11	20	35
		CWB					
		6.32					
50D-5D	D65	2.9	3.5	3.5	3.8	8.6	64.0
	SC	2.8	3.1	3.3	3.7	13.6	64.1
		CWB					
		17.2					
50D-G9	D65	5.1	4.9	4.9	5.1	14.3	49.7
	SC	5.0	4.3	4.4	4.4	17.6	49.7
		CWB					
		13.9					
5D-G9	D65	4.2	4.5	4.4	5.2	12.5	42.6
	SC	4.1	3.7	3.5	4.3	12.9	42.6

Table 2.2: Color matching comparison using two different cameras. Given 3 cameras (5D, 50D and G9), the 3 possible comparisons are illustrated. In every comparison, we include the use of custom white balance (CWB), state-of-the-art (D65) and our method (SC). Differences are measured in ΔE for the 14-skin color patches after calibration using the 24 Gretag Macbeth color chart patches. Note that none of the 14 tested colors are included in the color calibration chart.

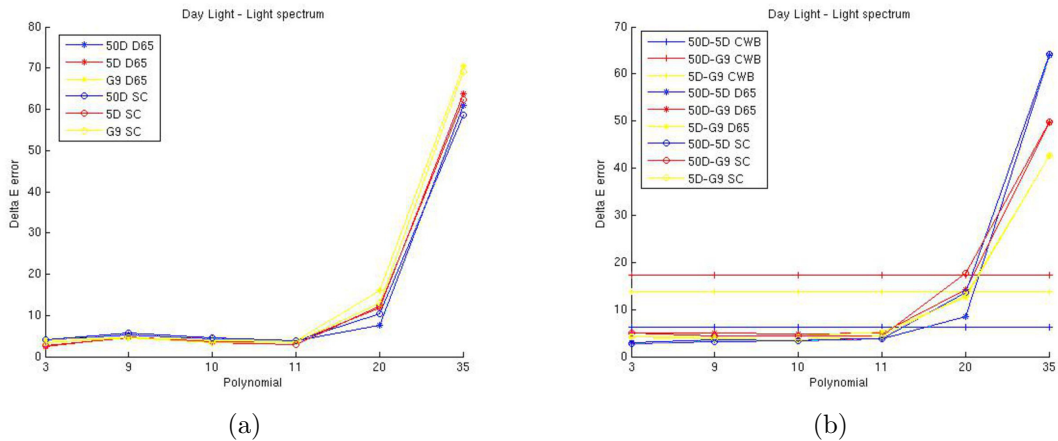


Figure 2.9: ΔE error comparison. Calibration is done with the 24 Gretag Macbeth patches and results are applied to the 14 skin patches. (a) The error between the calibrated colors and the expected ones. (b) The error between the two cameras, with calibrated images and using the custom white balance.

acquired first, then the mean of a small central patch is computed, assuming uniform lighting for that central patch. This value will be used as the expected

2. COLOR CALIBRATION

one for the entire image. For this reason, we create a mask, Eq. 2.11, and apply this mask to each image according to Eq. 2.12.

$$Mask(x, y) = \frac{R(x, y) - Mean}{Mean} \quad (2.11)$$

$$OutputImage(x, y) = \frac{InputImage(x, y)}{1 + Mask(x, y)} \quad (2.12)$$

Finally, the images are transformed from the XYZ color space to RGB using Eq. 2.13, and the linear values obtained are corrected with a “standard ”gamma ($\gamma = 2.4$, see Eq. 2.14).

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 3.2410 & -1.5374 & -0.4986 \\ -0.9692 & 1.8760 & 0.0416 \\ 0.0556 & -0.2040 & 1.0570 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad (2.13)$$

$$C_{sRGB} = \begin{cases} 12.92C_{linear}, & C_{linear} \leq 0.0031308 \\ (1 + a)C_{linear}^{1/2.4} - a, & C_{linear} > 0.0031308 \end{cases} \quad (2.14)$$

where C corresponds to each color channel.

Figures 2.10 and 2.11 show the difference between calibrated and non-calibrated real images. The first figure shows the differences between the three systems used using auto white balance, custom white balance, and calibrated cameras. The second figure shows the difference between custom white balance and calibrated images using JPEG images and RAW images. The images corrected from RAW values show the best results and, at the same time, have colors closer to the real ones and the closest colors between the different cameras. Without using RAW values the images improve in quality but are significantly less than when RAW images are used.

The third experiment quantifies the accuracy for the 24 corrected images that correspond to each Gretag Macbeth[®] color patch. This test quantifies the accuracy of the color calibration process by measuring how different corrected colors are from the ones that can be perceived with the naked eye. This test was performed using three commercial digital cameras (Canon 5D Mark II, Canon 50D and Canon G10). The images were obtained using Custom White Balance, setting the camera to the Av mode (fixed camera aperture and variable time exposure)

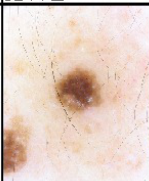
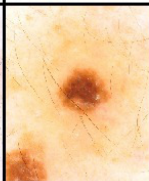

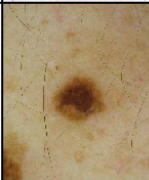
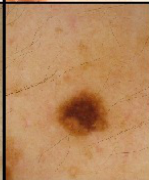
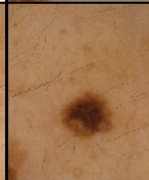



	AWB	CWB	Calibrated
Canon A640 + Polarized dermatoscope			
Canon A640 + Non Polarized dermatoscope			
Canon G7 + Polarized dermatoscope			

Figure 2.10: Skin images using the a640 with polarized and non-polarized light and the G7 with polarized light. The first column shows the original images with auto white balance, the second column shows the improvements using custom white balance, and the last column shows results of a color calibration using JPEG values.

and fixing the ISO value to the minimum available for each camera (100 for 5D and 50D, and 80 for the G10). The color of each image is adjusted using custom white balance. For the Canon cameras taking an image of a white reference enables this setting for the JPEG images, RAW images are independent on the white point selected, and depend exclusively on the camera and the available lighting.

The dermatoscopes used in this study are the Dermlite DL3, the Dermlite II Fluid and the Dermlite Foto (3Gen LLC, Dana Point, California, USA). The first one consists of a photographic lens attachment, and uses 21 light-emitting diodes (LEDs) in cross-polarized mode and 7 non-polarized LEDs for immersion fluid dermoscopy (not used in our experiment). The Dermlite II Fluid uses 16 or 32 LEDs with non-polarized lighting, allowing physicians to perform immersion fluid dermoscopy. The DermLite Foto uses a cross-polarized system with 24 cold white

2. COLOR CALIBRATION

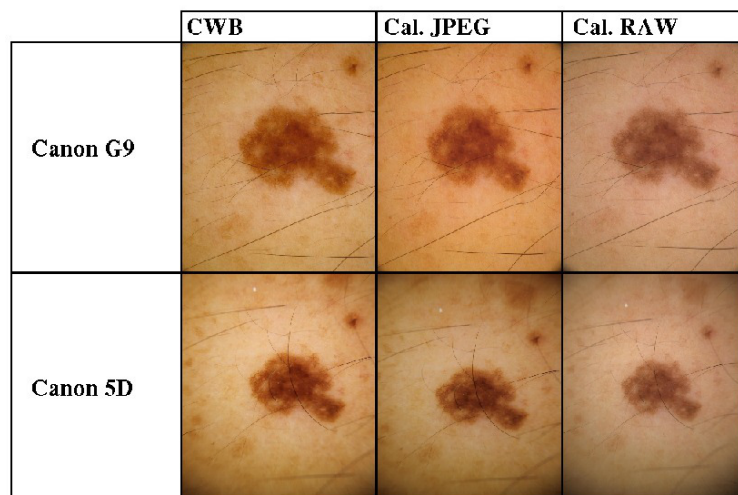


Figure 2.11: Skin images using the G9 and 5D with a polarized dermatoscope. The first column shows the custom white balance images, the second, the calibrated images using JPEG, and the third, the images obtained after calibrating with the RAW values.

light emitting LEDs; the system is combined with a digital camera. During the experiment, batteries were always fully charged to avoid illumination changes.

Table 2.3, shows the ΔE error values between each of the 3 cameras and the 3 dermatoscopes compared with the values obtained using a spectrophotometer. This table clearly shows the large initial input error of each image. The differences in the ΔE error are statistically significant between the original images before calibration and the JPEG images after color correction ($P < 0.001$), as well as between the original images and their color-corrected counterpart using RAW images ($P < 0.001$). The improvement between RAW and JPEG images is also statistically significant ($P < 0.001$).

The corrected images are still 1.5 ΔE worse (on average) than the ones calibrated using RAW values.

Figure 2.12 presents the highest probability of having images corrected with less error than the original ones. Almost 84% of the JPEG corrected images present less than 4 ΔE errors, while if we consider RAW images, 99% of them

Table 2.3: Color reproduction accuracy using three different cameras. Given 3 cameras (5D, 50D and G10), and 3 dermatoscopes (DermLite DL3, DermLite II Fluid and DermLite Foto) the 9 possible comparisons are illustrated. In every comparison, we include error before calibration (Before), after calibration with JPEG images (after JPEG) and after RAW calibration (After RAW). Differences are measured in ΔE for the 24 Gretag MacBeth color patches.

	Color Correction	DermLite DL3	DermLite II Fluid	DermLite Foto
5D	Before	27.07	72.88	23.43
	After JPEG	2.90	1.18	3.93
	After RAW	1.67	0.89	0.90
50D	Before	31.81	30.38	24.41
	After JPEG	1.83	1.87	1.09
	After RAW	0.75	0.82	1.24
G10	Before	32.84	28.02	27.51
	After JPEG	2.98	1.10	2.19
	After RAW	0.97	0.64	0.61

are within the same threshold.

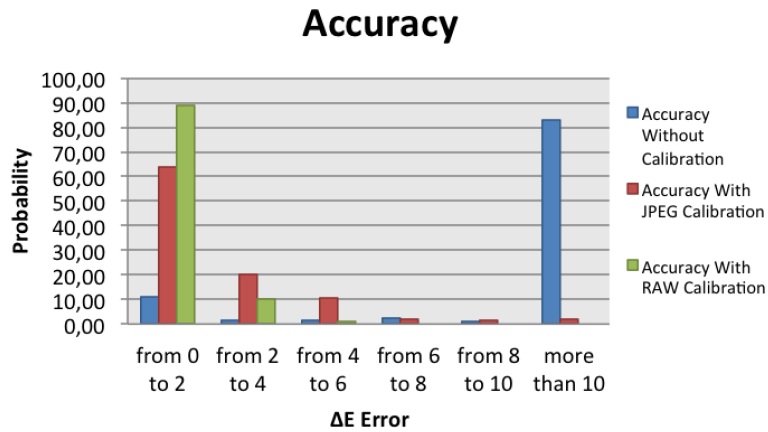


Figure 2.12: Distribution of the probability of having n of the 24 Gretag Macbeth[®] colors with a given difference between them and the expected spectrophotometric value. Images are compared with and without calibration.

Most of the corrected colors present small ΔE values on the JPEG and RAW

2. COLOR CALIBRATION

corrected images when comparing their values with the spectrophotometric ones. However, note that a few of them reach the value of 12 ΔE on JPEG and 6 ΔE on RAW images.

Table 2.4 shows a comparative analysis of the accuracy of our method comparing different cameras with the same dermatoscopes. We can see that the error is always smaller when RAW values are used. The differences in the ΔE error are statistically significant between the original and JPEG images ($P < 0.001$), between the original and RAW images ($P < 0.001$) and between RAW and JPEG images ($P < 0.01$).

Table 2.4: Color matching comparison using three different cameras. Given 3 cameras (5D, 50D and G10), and 3 dermatoscopes (DermLite DL3, DermLite II Fluid and DermLite Foto) the 9 possible comparisons are illustrated. In every comparison we include error before calibration (Before), after calibration with JPEG images (after JPEG) and after RAW calibration (After RAW). Differences are measured in ΔE for the 24 Gretag MacBeth color patches.

	Color Correction	DermLite DL3	DermLite II Fluid	DermLite Foto
5D	Before	8.83	8.75	10.83
	After JPEG	1.53	0.97	4.50
	After RAW	1.02	0.47	1.88
50D	Before	11.07	11.33	15.87
	After JPEG	1.15	0.90	2.43
	After RAW	0.93	1.30	1.08
G10	Before	13.20	10.60	11.77
	After JPEG	1.95	1.37	2.62
	After RAW	0.45	1.15	0.92

The forth experiment presents some results on in vivo images, although, note that spectrophotometric measurements were not made on them due to the difficulty of measuring exactly the same area with the spectrophotometer and the camera. For in vivo images, we only tested how constant the color of each image is with respect to the colors of images acquired with different dermatoscopes or cameras (Figures 2.13, 2.14 and 2.15).

Table 2.5 presents numerical results for the two in vivo images. We decided first to check the accuracy of our method by looking at a part of the skin that does not contain the lesion and then comparing the lesion part of the images.

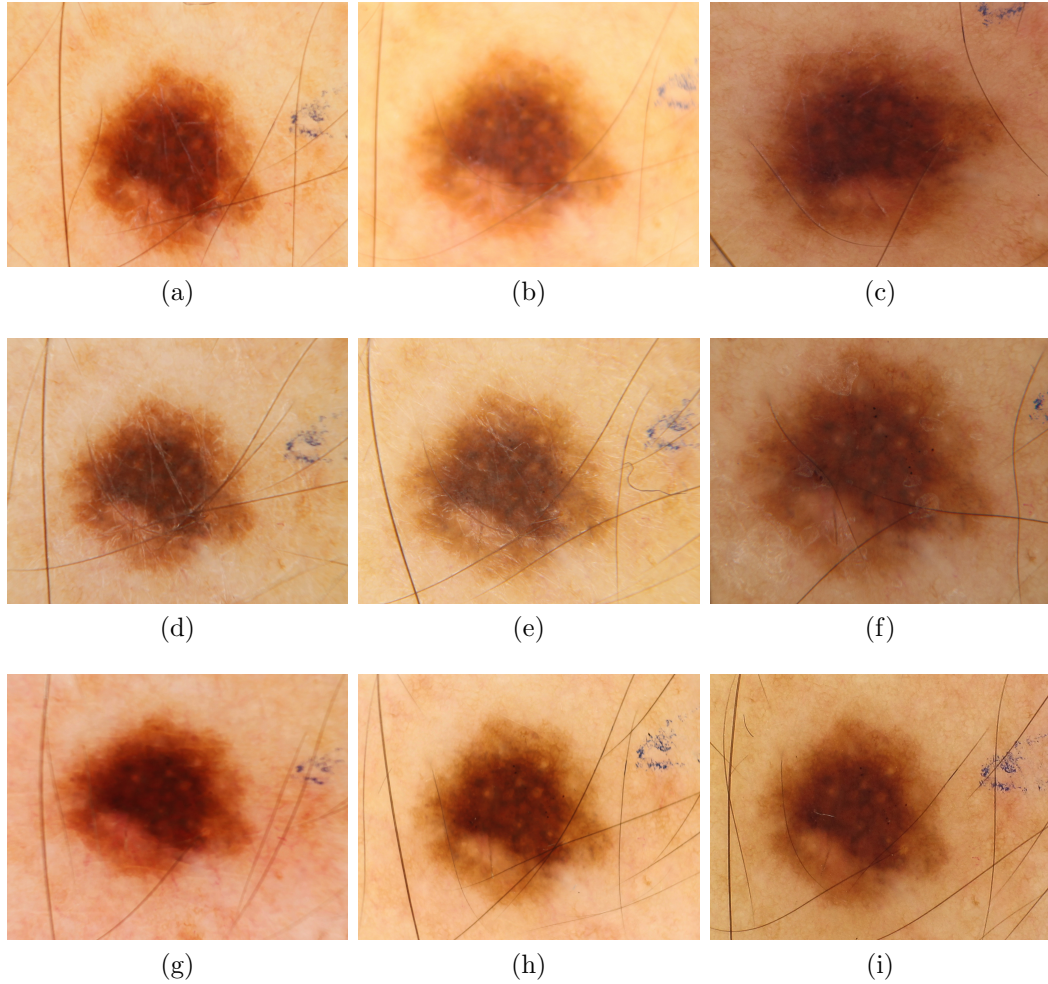


Figure 2.13: Original images. Each column corresponds to a Canon 5D, 50D and G10, respectively. The rows correspond to DermLite DL3, DermLite II Fluid and DermLite Foto

Results show an improvement of 5 ΔE when RAW images are used comparing them with the original images. Also we observe a small improvement with JPEG corrected images.

2. COLOR CALIBRATION

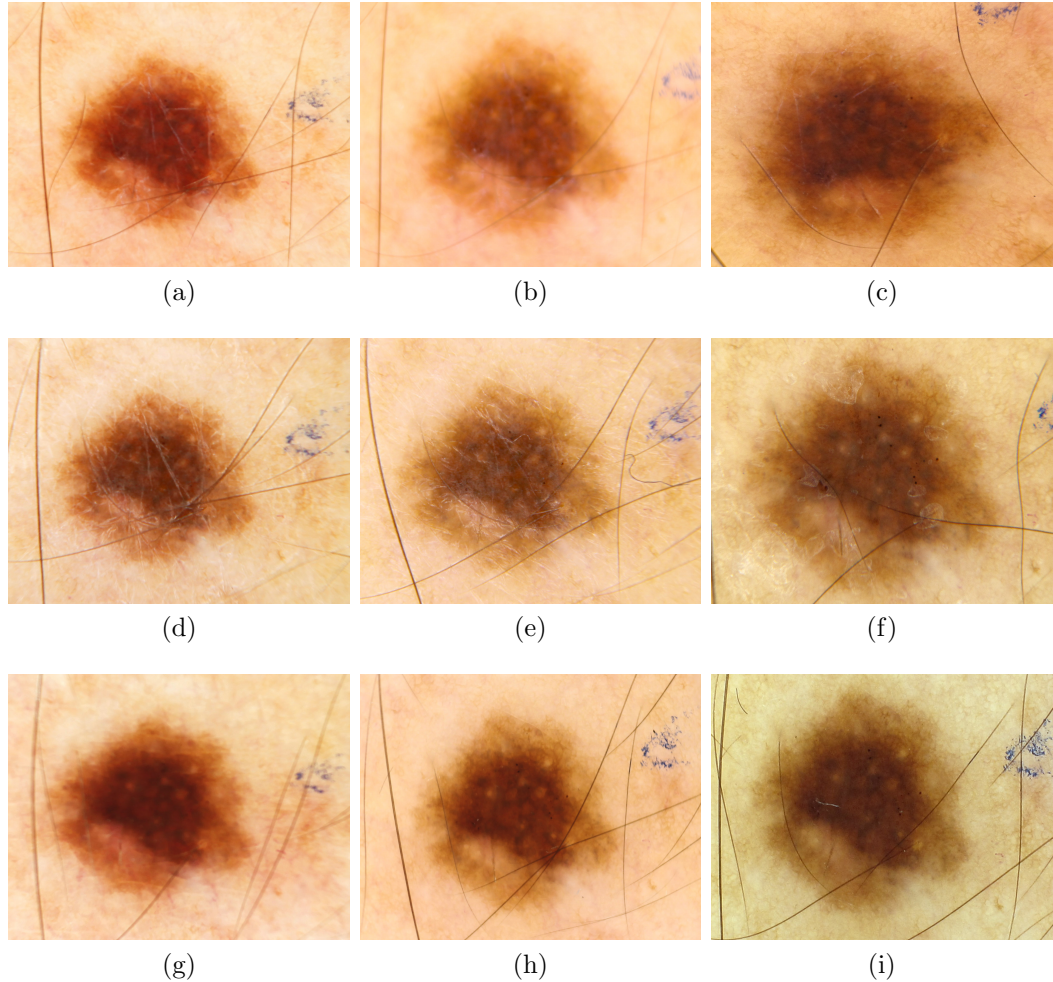


Figure 2.14: Calibrated images using JPEG values. Each column corresponds to a Canon 5D, 50D and G10, respectively. The rows correspond to DermLite DL3, DermLite II Fluid and DermLite Foto

2.4 Discussion

The relevance of color calibration in dermatological images has previously been discussed in the literature. In this work, we show that our color calibration methodology provides accurate color reproduction, improving the results presented in the literature up to date. Some studies have reported the importance of color calibration, both for clinical and dermoscopic images. Our procedure per-

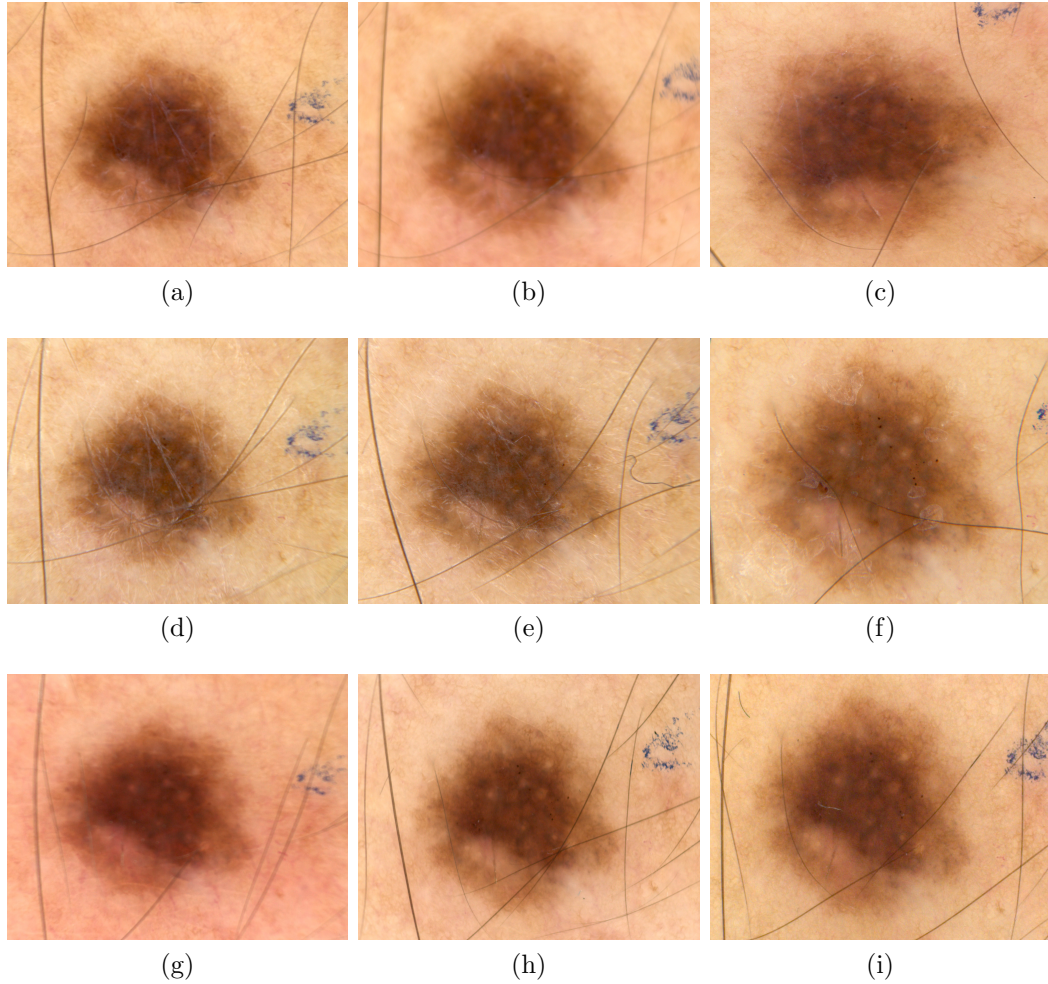


Figure 2.15: Calibrated images using RAW values. Each column corresponds to a Canon 5D, 50D and G10, respectively. The rows correspond to DermLite DL3, DermLite II Fluid and DermLite Foto

forms color calibration in dermoscopic images, independently of the dermatoscope and camera used. Moreover, our approach is able to work with RAW or JPEG images, although our results prove that color correction is much more effective when RAW is used as input. This is due to the fact that the camera introduces a gamma correction in the JPEG as well as other unknown adjustments to enhance the image, adapting it to the spectral response of the human eye. One of the main differences in our work with respect to all other reported work is the use of

2. COLOR CALIBRATION

Table 2.5: Comparison of different patches on two different lesions. Each value corresponds to the average ΔE error after combining all the dermatoscopes and cameras

	Lesion 1		Lesion 2		Total
	Skin	Lesion	Skin	Lesion	Skin + Lesion
Before calibr.	12.92	16.50	14.98	19.26	15.92
After JPEG calibr.	9.52	14.53	12.76	16.80	13.90
After RAW calibr.	9.04	8.42	9.21	11.44	9.53

spectral information of the lighting system in the dermatoscope to improve color calibration accuracy.

Although this method is devised for RAW images, a simple but powerful solution has also been presented for low-cost digital cameras producing JPEG images, which are “black boxes”, colorimetrically speaking. Thus, in the case of JPEG images, we have shown that “custom white balance” produces an initial improvement over the use of “auto white balance”.

The final results show the importance of using the new features proposed in this paper, and the advantages/disadvantages of using different polynomial orders. We have illustrated how important it is to use the classic 3×3 calibration matrix to obtain good results in the calibration of unknown colors and also to preserve the relationship between color channels. Our proposal shows the importance of computing the lighting spectral sensitivity when images captured from different cameras are compared, obtaining improvements between 0.1 and 0.9 ΔE over day light assumption (state-of-the-art).

Table 2.6: Accuracy measurements for our method with respect to previous works.

	Accuracy	Maximum
Haeghen et al.[63]	6.20	13.30
Wighton et al.[182]	10.00	23.60
Wighton et al.[182]	6.40	14.80
Proposed JPEG Cal.	2.16	12.35
Proposed RAW Cal.	0.95	5.22

It is evident from our results that the error is considerably reduced for both

the Gretag MacBeth[®] color checker chart and the in vivo images.

The error in the original JPEG images is due to a number of facts, namely: the different lighting used by each dermatoscope, the different way of processing the values of each camera, and the non-linearities in the JPEG image forming process.

Although calibration using JPEG images partially compensates for some of these effects, it is not able to eliminate all of them. Using RAW values (whenever this option is available in the camera) considerably reduces the final error, even when certain colors may be reproduced with a noticeable inaccuracy.

Moreover, by comparing the accuracy of our method with respect to the state of the art (see Table 2.6), it has been proved that we provide better calibration results using both JPEG and RAW images in overall accuracy as well as regarding the worst corrected value. Haeghen and Naeyaert's[63] previously published study presents a camera color calibration using JPEG data with a mean error of 6.20 ΔE , Whigton et al.[182] recently presented a study with a mean error of 10.00 and 6.40, the first error using JPEG values and the second with RAW images. Our error of 2.16 and 0.95 shows the benefit of using spectral information of the dermatoscope lighting.

Our results illustrate the importance of using RAW values as opposed to JPEG, since JPEG is a *black box*[37] where pixel values may experience changes due to unknown non-linear image manipulations carried out by the camera, thus being less reliable for diagnosis. This is supported by the results presented in this paper, where JPEG images produce larger errors than RAW images. It is important to note, moreover, that even a perfect calibration would not ensure accurate color reproduction in real world JPEG images due to the non-linear procedures applied, as opposed to RAW images.

The procedure described in this paper shows how to improve the color constancy in different cameras and dermatoscopes. However, as some colors could not be perfectly corrected, we finish with a tool that can be used by physicians for improving their diagnosis, both with their own images and when using tele-dermatology. Finally, it should be noted that our software could be also used as pre-processing for automatic mole classification when color information is required.

2. COLOR CALIBRATION

Chapter 3

A novel system for skin mosaicing

This chapter introduces a novel methodology for skin exploration. Our state-of-the-art discovered various techniques based on total body photography lacking enough resolution for adequate clinical practice. Our proposal takes the low resolution images of a total body exploration and merges them into a single, higher-resolution image. This technique is known as image mosaicing. In this way, a full body exploration is registered into a single image, simplifying the task of comparing different explorations, either analyzing the mosaics of both explorations or comparing the corresponding images between explorations.

3.1 Related Works in Image Mosaicing

Image mosaicing[27] (also known in the literature as image stitching[99] or image montage[40]) is used to combine tens or hundreds of images to make a single wide-angle or panoramic view, usually bigger than the initial images. This combination of images is applied to several different fields such as document analysis[125], augmented reality[167], scene stitching[168][26], etc, performing panoramic photography, super-resolution imaging, object insertion, texture synthesis, virtual environments, vision based navigation systems, etc.

Mosaics could be used to build a map with the location of detected moles, showing their positions in the map and tracking them on time for studying the evolution in size, color and shape.

3. A NOVEL SYSTEM FOR SKIN MOSAICING

In our case, constructing a map of the human skin is based on building a global mosaic from several images of the body. The mosaic is usually created without any information about the camera position at each shot. Therefore, all the known information provided for generating the mosaic is retrieved from the images, and only the order of the images is known. To build the mosaic without human intervention, the estimation procedure should be performed in a robust and fully automatic way. The quality of the texture image matching depends on the images and the transformation presented between them.

Fig. 3.1 presents the mosaicing pipeline, composed of image acquisition, registration, prediction of non-consecutive overlaps, global registration and rendering. Each step is described below.

Skin images normally suffer from acquisition noise, non-uniform illumination, lack of texture and non-rigid deformations of the body. These problems increase the difficulty of solving the registration problem. The camera can be shifted and rotated relative to all the three axes between images. Moreover, changes in scale, illumination and perspective can be present.

The registration process tries to recover the position of the camera in each acquisition procedure. Once these camera positions are computed, is it possible to stitch all the images into a common frame. As the positions are incrementally computed, some errors could be generated by small differences accumulating at each camera position. For this reason, non-consecutive image pairs are located and matched. The mosaic with the new correspondences is then re-aligned. This process is repeated iteratively until no more new overlapping image pairs are found. Finally, the images are merged onto a single frame. Blending algorithms solve the intensity value for the overlapping areas.

3.1.1 Image Acquisition

The mosaicing pipeline in Fig. 3.1 starts with the acquisition of the images. These images should have enough quality to be clinically useful, enough at least to detect changes. This quality is of the utmost importance in order to obtain good results in the next steps. Having enough resolution at the scale of the nevus of the patient's body is required because we aim to monitor every single nevus as

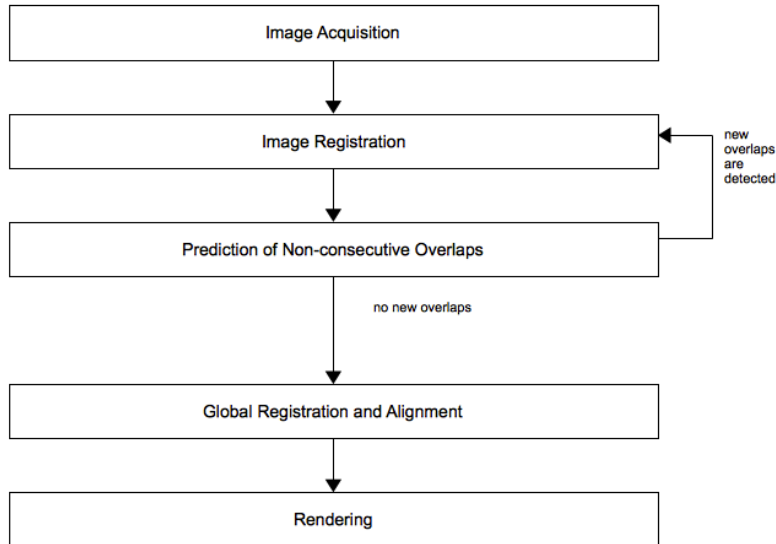


Figure 3.1: Mosaicing Pipeline: Steps needed to solve the image stitching problem.

well as the areas of the body where no previous nevi are present, i.e. the full body. In the first case, the intention is to monitor the evolution of this single nevus, while, in the second case, the goal is to see if any new nevi appear.

3.1.2 Image Registration

Determining the transformation that takes place between images as the viewpoint of the camera changes is an essential problem in image mosaicing. This is widely known as the image registration problem[26].

Image registration has been greatly discussed in the literature where many authors have proposed methods to tackle this problem [167][41]. Broadly speaking, these methods can be classified into: frequency domain based methods (Fourier transform), dense methods (optical flow) and sparse methods (feature based). We describe these methods in the following pages.

3. A NOVEL SYSTEM FOR SKIN MOSAICING

3.1.2.1 Frequency Domain

Frequency-based methods were used in phase-correlation in order to estimate the shifts between two images. These methods were extended in order to account for rotation and scale transformations[143] as well as affine transformations[185] using log-polar coordinates.

3.1.2.2 Optical Flow

Optical Flow methods estimate the disparity (apparent motion) of pixels between pairs of images. Usually, optical flow uses the Brightness Constancy Model (BCM), in which it is assumed that the photometric properties (intensity and color) remain constant.

There are two main approaches in estimating the optical flow: global methods such as Horn-Schunck[76], which yield dense flow fields, and local methods such as Lucas-Kanade[107][108] that produce non-dense regularized grid flow fields that are more robust to noise.

Lucas-Kanade is one of the most widely used methods based on the local Taylor series approximation using partial spatial and temporal derivatives, and is based on the following equations:

$$I(x + \delta x, y + \delta y, t + \delta t) = I(x, y, t) + \frac{\partial I}{\partial x} \delta x + \frac{\partial I}{\partial y} \delta y + \frac{\partial I}{\partial t} \delta t + \xi \quad (3.1)$$

where $I(x, y, t)$ is the pixel intensity at coordinates (x, y) at time t , and ξ is a reminder (small enough to be ignored). Making use of the BCM assumption through images, we have

$$\frac{\partial I}{\partial x} \delta x + \frac{\partial I}{\partial y} \delta y + \frac{\partial I}{\partial t} \delta t = 0 \quad (3.2)$$

or

$$\frac{\partial I}{\partial x} \frac{\partial x}{\partial t} + \frac{\partial I}{\partial y} \frac{\partial y}{\partial t} + \frac{\partial I}{\partial t} = 0 \quad (3.3)$$

therefore,

$$\frac{\partial I}{\partial x}V_x + \frac{\partial I}{\partial y}V_y = -\frac{\partial I}{\partial t} \quad (3.4)$$

Using I_x , I_y and I_t as the spatial and temporal derivatives, we obtain $-I_t = I_xV_x + I_yV_y$ or simply $-I_t = \nabla I \cdot \vec{V}$, which is an equation that imposes a single constraint with two unknowns, thus not solvable as is. However, assuming constant flow within small windows, for instance over 3×3 pixels, we can obtain a set of 9 equations:

$$I_{x11} \cdot V_x + I_{y11} \cdot V_y = -I_{t11} \quad (3.5)$$

$$I_{x12} \cdot V_x + I_{y12} \cdot V_y = -I_{t12} \quad (3.6)$$

$$I_{x13} \cdot V_x + I_{y13} \cdot V_y = -I_{t13} \quad (3.7)$$

$$\vdots \quad (3.8)$$

$$I_{x33} \cdot V_x + I_{y33} \cdot V_y = -I_{t33} \quad (3.9)$$

Therefore, we can construct an over-determined system of $3 \times 3 = 9$ equations:

$$\begin{pmatrix} I_{x11} & I_{y11} \\ I_{x12} & I_{y12} \\ \vdots & \vdots \\ I_{x33} & I_{y33} \end{pmatrix} \cdot \begin{pmatrix} V_x \\ V_y \end{pmatrix} = \begin{pmatrix} -I_{t11} \\ -I_{t12} \\ \vdots \\ -I_{t33} \end{pmatrix} \quad (3.10)$$

This over-determined system is expressed by $B^T \cdot \vec{v} = B^T(-b)$ and, therefore, $\vec{v} = (B^T B)^{-1} B^T(-b)$. Hence,

$$\begin{pmatrix} V_x \\ V_y \end{pmatrix} = \begin{pmatrix} \sum I_{xij}^2 & \sum I_{xij} \cdot I_{xij} \\ \sum I_{xij} \cdot I_{xij} & \sum I_{yij}^2 \end{pmatrix}^{-1} \cdot \begin{pmatrix} -\sum I_{xij} \cdot I_{tij} \\ -\sum I_{yij} \cdot I_{tij} \end{pmatrix} \quad (3.11)$$

Local optical flow methods yield a vector direction for each considered patch in the image.

Some approaches present better alternatives to BCM. These alternatives assume linear changes in illumination, like Generalized Dynamic Image Model

3. A NOVEL SYSTEM FOR SKIN MOSAICING

(GDIM)[126][129], and color[109][128].

Due to the problem's formulation, optical flow methods can not deal with disparities that exceed 1 pixel. For solving this problem, multi-resolution approaches should be used[127]. In this way, images are gradually reduced and the optical flow is computed from coarse levels towards fine levels. The drawbacks here include the computational time (optical flow needs to be computed at each level) and the maximum pixel disparity has to be a known a priori in order to set the number of decimation levels. Multi-resolution approaches are very sensitive to noise, since errors in the estimation of optical flow at coarse levels will propagate to fine levels.

3.1.2.3 Feature Based Image Registration

Feature based techniques rely on locating features in both images and using these features to obtain the transformation parameters for registering the two images. These features must be: (i) repetitive - features can be correctly tracked through images even if their focal point is different or the lighting conditions have changed and (ii) discriminative - they can be singly matched in the images.

The performance of these techniques depends on several factors, such as the area of overlap between images and to what extent it is possible to model the image distortions with simple geometric transformations. Further, image quality, affected by degradations such as noise contamination and blurring as well as image characteristics such as smooth/textured areas or similarity of different areas also play a role in the techniques' performance.

Feature-based image registration involves four steps (see Fig. 3.2):

Feature detection. This is the process of extracting salient points that correspond to corners, boundaries, highly textured patches or regions differing in intensity need to be detected independently in every image. There are various extraction techniques. The primary issue of these algorithms is to detect points invariant to the 2D image deformations resulting from 3D camera motion. These deformations include translation, rotation, scaling, shear and projective distortion. In this work, a number of point detectors will be tested with skin images: Harris, SIFT and SURF, respectively.

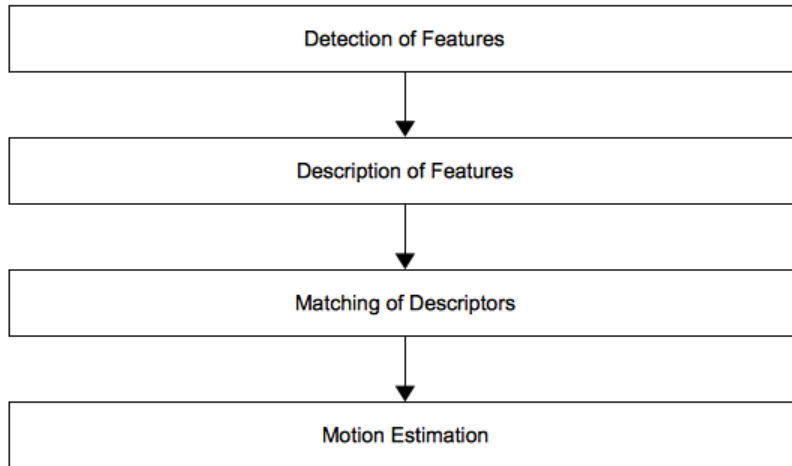


Figure 3.2: Feature Based Mosaicing Pipeline: Steps needed to solve image stitching problem in feature based methods.

Feature description. Once a set of keypoints are detected in each image, an invariant descriptor based on the intensity information of the point neighborhood is computed. The descriptor should be distinctive for each detected point, and at the same time it should be robust to noise, detection errors, and geometric and photometric deformations (lighting, color and contrast changes between images).

Feature matching. The first step toward the estimation of the image registration parameters consists of finding feature correspondences. This is referred to as the matching problem and has different solutions in the literature, like methods using spatial relations, invariant descriptors, etc [190].

Motion Estimation. Once the feature correspondence has been established, the next step is to solve for the parameters of some global transformation. Usually, this involves finding the translation, rotation and scale parameters to transform one image to another.

The next sections will describe each step in detail.

3. A NOVEL SYSTEM FOR SKIN MOSAICING

Feature detector	Edge	Corner	Blob
Canny[31]	x		
Sobel[137]	x		
Harris & Stephens[70]	x	x	
SUSAN [159]	x	x	
Shi & Tomasi[154]		x	
Level curve curvature[87][90]		x	
FAST[146]		x	
Laplacian of Gaussian[101]		x	x
Difference of Gaussians[106]		x	x
Determinant of Hessian[101]		x	x
MSER[115]			x
Grey-level blobs[100]			x

Table 3.1: Classification of Feature Detectors : Some algorithms for detecting features are classified depending on the detected element (Edge, Corner or Blob).

3.1.3 Feature detection

The points detected at the detection step should have the property of repeatability. This means that the same feature should be detected in two or more different images of the same scene. If the same point is detected in two different images, their matching can be performed. The points are distinguished by an interest measure, called *cornerness*. This term is used for all the detectors for simplicity, but not all the detectors search corners, blobs and ridge structures are also detected. In these cases, the interest points are called *keypoints* or *features*. Taking into account this difference, three main groups are formed, see Table 3.1.

Some of these detectors will be explained in detail.

Harris Corner Detector This detector was introduced by Chris Harris and Mike Stephens in 1988 [70] and defines a corner as a point with low self-similarity. The algorithm checks if a corner is present in any pixel of the image by considering how similar a patch centered on the pixel is to the patches centered on his neighbors, see Eq.3.12. The similarity is measured using the sum of squared

differences, *SSD*, between the two patches. Lower values obtained with *SSD* describes more similarity between patches. The Harris corner detector presents invariance to rotation, illumination variation and image noise.

$$c(x, y) = \sum_w [I(x_i, y_i) - I(x_i + \Delta x, y_i + \Delta y)]^2 \quad (3.12)$$

Where $I(\cdot, \cdot)$ denotes the image function and (x_i, y_i) are the points in the window W (Gaussian) centered on (x, y) . The shifted image is approximated by a Taylor expansion truncated to the first order term, where I_x and I_y denote the partial derivatives in x and y , respectively.

$$I(x_i + \Delta x, y_i + \Delta y) \approx I(x_i, y_i) + [I_x(x_i, y_i)I_y(x_i, y_i)] \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \quad (3.13)$$

After substituting Eq. 3.13 into Eq. 3.12 yields,

$$c(x, y) = \sum_w [I(x_i, y_i) - I(x_i + \Delta x, y_i + \Delta y)]^2 \quad (3.14)$$

$$= \sum_w \left(I(x_i, y_i) - I(x_i, y_i) - [I_x(x_i, y_i)I_y(x_i, y_i)] \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \right)^2 \quad (3.15)$$

$$= \sum_w \left(-[I_x(x_i, y_i)I_y(x_i, y_i)] \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \right)^2 \quad (3.16)$$

$$= \sum_w \left([I_x(x_i, y_i)I_y(x_i, y_i)] \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \right)^2 \quad (3.17)$$

$$= [\Delta x \Delta y] \begin{bmatrix} \sum_w (I_x(x_i, y_i))^2 & \sum_w I_x(x_i, y_i)I_y(x_i, y_i) \\ \sum_w I_x(x_i, y_i)I_y(x_i, y_i) & \sum_w (I_y(x_i, y_i))^2 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \quad (3.18)$$

$$= [\Delta x \Delta y] C(x, y) \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \quad (3.19)$$

where matrix $C(x, y)$ captures the intensity structure of the local neighborhood. The geometric interpretation of the gray levels is encoded in the eigenvectors and eigenvalues of the matrix. $C(x, y)$ is symmetric and has two nonnegative eigenval-

3. A NOVEL SYSTEM FOR SKIN MOSAICING

ues λ_1 and λ_2 , so it can be expressed by rotation of the coordinate axes shown in Eq. 3.20, where $C_T(x, y)$ is the matrix $C(x, y)$ after the applied transformations.

$$C_T(x, y) = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \quad (3.20)$$

The eigenvalues of $C(x, y)$ form a rotationally invariant descriptor of the pixel neighborhood, depending on the magnitude of the eigenvalues. Three cases are described:

1. $\lambda_1 \approx 0$ and $\lambda_2 \approx 0$

The local auto-correlation function is approximately flat. This means that the changes in $c(x, y)$ have few changes in any direction, i.e. the windowed image region has constant intensity, with no features of interest.

2. ($\lambda_1 \approx 0$ and $\lambda_2 \gg 0$) or ($\lambda_1 \gg 0$ and $\lambda_2 \approx 0$)

If one eigenvalue is high and the other low, the local auto-correlation function is ridge-shaped. In one direction, only small changes are produced and in the other (orthogonal to the ridge) significant changes are detected. This indicates the presence of an edge.

3. $\lambda_1 \gg 0$ and $\lambda_2 \gg 0$

If both eigenvalues are high, the local auto-correlation function is sharply peaked, shifts of the window in any direction will produce significant differences in the function. This indicates a corner.

Fig. 3.3 displays the three cases described, namely when the windowed image region is flat (a), when there is an edge (b) and when there is a corner (c).

Hessian blob detector. This was one of the first image feature detectors. Proposed by Beaudet in 1978[16], it represents the basis for many recent corner detectors.

The Baudet operator is a rotationally invariant measurement of cornerness given by the determinant of the Hessian matrix H , which represents a second-order partial derivative of an image I :

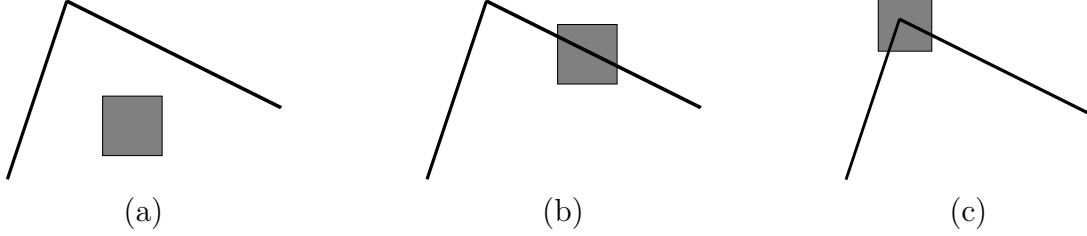


Figure 3.3: Harris Detector: (a) flat region, (b) edge and (c) corner.

$$H(x, y) = \begin{bmatrix} I_{xx}(x, y) & I_{xy}(x, y) \\ I_{xy}(x, y) & I_{yy}(x, y) \end{bmatrix} \quad (3.21)$$

where I_{xx} , I_{xy} and I_{yy} are the second partial derivatives of the image intensity function.

The second derivatives used in the Hessian matrix correspond to blobs and ridges, being represented by the local maxima of $K_{Hessian}$:

$$K_{Hessian}(x, y) = I_{xx}(x, y)I_{yy}(x, y) - I_{xy}^2(x, y) \quad (3.22)$$

Harris affine and Hessian affine detectors. These are robust to image noise and invariant to rotation and lighting changes. However, none of them is invariant to scale and affine transformations[118]. This makes them ineffective in wide base-line image registration where changes in the camera's viewpoint can induce significant geometric transformations.

Mikolajczyk et al.[119] proposed modifications for Harris and Hessian feature extractors, making them invariant to scale changes and affine transformations. For dealing with the scale changes, they propose the use of a scale selection method based on the Laplacian, with the idea of selecting a scale that is characteristic to the local structure. For this purpose, the Harris autocorrelation matrix is modified to include scale information:

$$C_{Affine}(x, \sigma_I, \sigma_D) = \sigma_D^2 g(\sigma_I) * \begin{bmatrix} I_x^2(x, \sigma_D) & I_x I_y(x, \sigma_D) \\ I_x I_y(x, \sigma_D) & I_y^2(x, \sigma_D) \end{bmatrix} \quad (3.23)$$

where $g(\sigma_I)$ is the Gaussian kernel of scale σ_I and $x = (x, y)$. $I(x)$ is the Gaussian-

3. A NOVEL SYSTEM FOR SKIN MOSAICING

smoothed image, and the $*$ operator denotes convolution. $I_x(x, \sigma_D)$ and $I_y(x, \sigma_D)$ are the derivatives in their respective direction applied to the smoothed image and calculated using a Gaussian kernel with scale σ_D . The σ_I parameter determines the current scale at which the Harris corner points are detected.

The local image derivatives are computed using Gaussian kernels of scale σ_D and averaged by smoothing with a Gaussian window of scale σ_I .

In the case of the Hessian feature extractor, the second order matrix becomes:

$$H_{Affine}(x, \sigma_D) = \begin{bmatrix} I_{xx}(x, \sigma_D) & I_{xy}(x, \sigma_D) \\ I_{xy}(x, \sigma_D) & I_{yy}(x, \sigma_D) \end{bmatrix} \quad (3.24)$$

where $I_{xx}(x, \sigma_D)$ and $I_{xy}(x, \sigma_D)$, $I_{yy}(x, \sigma_D)$ are the second-order partial derivatives of the image, σ_D is the Gaussian scale at which the second partial derivatives of the image are computed.

The affine shape of the neighborhood around the feature points in both the Harris and Hessian cases is estimated using an iterative method using the eigenvalues of the second moment matrix.

SIFT Detector: Difference of Gaussians. This detector, descriptor and matching algorithm was introduced by Lowe in 2004[106]. The SIFT detector is invariant to scale, hence it incorporates a multi-scale representation of the image. To increase the speed with respect to Hessian algorithms, Lowe introduced an approximation of the Laplacian of Gaussian by a Difference of Gaussians, called *DoG*. DoG is used to compute the spatial coordinates of the feature and the scale selection.

To detect the features, two main steps are performed: the scale-space local extrema detection and keypoint localization.

Scale-space extrema detection. The first stage of this algorithm is to identify the locations and scales that can be repeatedly assigned under differing views of the same object. Detecting locations invariant to scale change of the image can be accomplished by searching for stable features across all possible scales, using a continuous function of scale, known as scale space.

The Difference of Gaussians is calculated from the Gaussian scale space $L(x, y, \sigma)$

that is produced from the convolution of a variable-scale Gaussian, $G(x, y, \sigma)$, with an input image, $I(x, y)$:

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (3.25)$$

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \quad (3.26)$$

The Difference of Gaussians function $D(x, y)$, convolved with the image $I(x, y)$ is computed from the difference of two nearby scales of $L(x, y, \sigma)$ separated by a constant multiplicative factor k as shown in Eq. 3.27.

$$D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) = L(x, y, k\sigma) - L(x, y, \sigma) \quad (3.27)$$

The smoothed images $L(x, y, \sigma)$ need to be computed in any case for the scale space feature description, and D can therefore be computed by simple image subtraction, which is much faster than computing the LoG space, which involves second-order derivation. The difference-of-Gaussian function provides a close approximation to the scale-normalized Laplacian of Gaussian, $\sigma^2 \nabla^2 G$, demonstrated in Fig.3.4. Lowe derives the relationship between D and $\sigma^2 \nabla^2 G$ from the heat diffusion (see Eq.3.28), parametrized in terms of σ .

$$\frac{\partial G}{\partial \sigma} = \sigma^2 \nabla^2 G \quad (3.28)$$

The term $\nabla^2 G$ can be computed from the finite difference approximation to $\partial G / \partial \sigma$, using the difference of nearby scales at $k\sigma$ and σ :

$$\sigma^2 \nabla^2 G = \frac{\partial G}{\partial \sigma} \approx \frac{G(x, y, k\sigma) - G(x, y, \sigma)}{k\sigma - \sigma} \quad (3.29)$$

and therefore,

$$G(x, y, k\sigma) - G(x, y, \sigma) \approx (k - 1)\sigma^2 \nabla^2 G \quad (3.30)$$

The σ^2 scale normalization required for the scale-invariant Laplacian is already included in Eq. 3.30. The factor $(k - 1)$ is a constant over all scale and does not

3. A NOVEL SYSTEM FOR SKIN MOSAICING

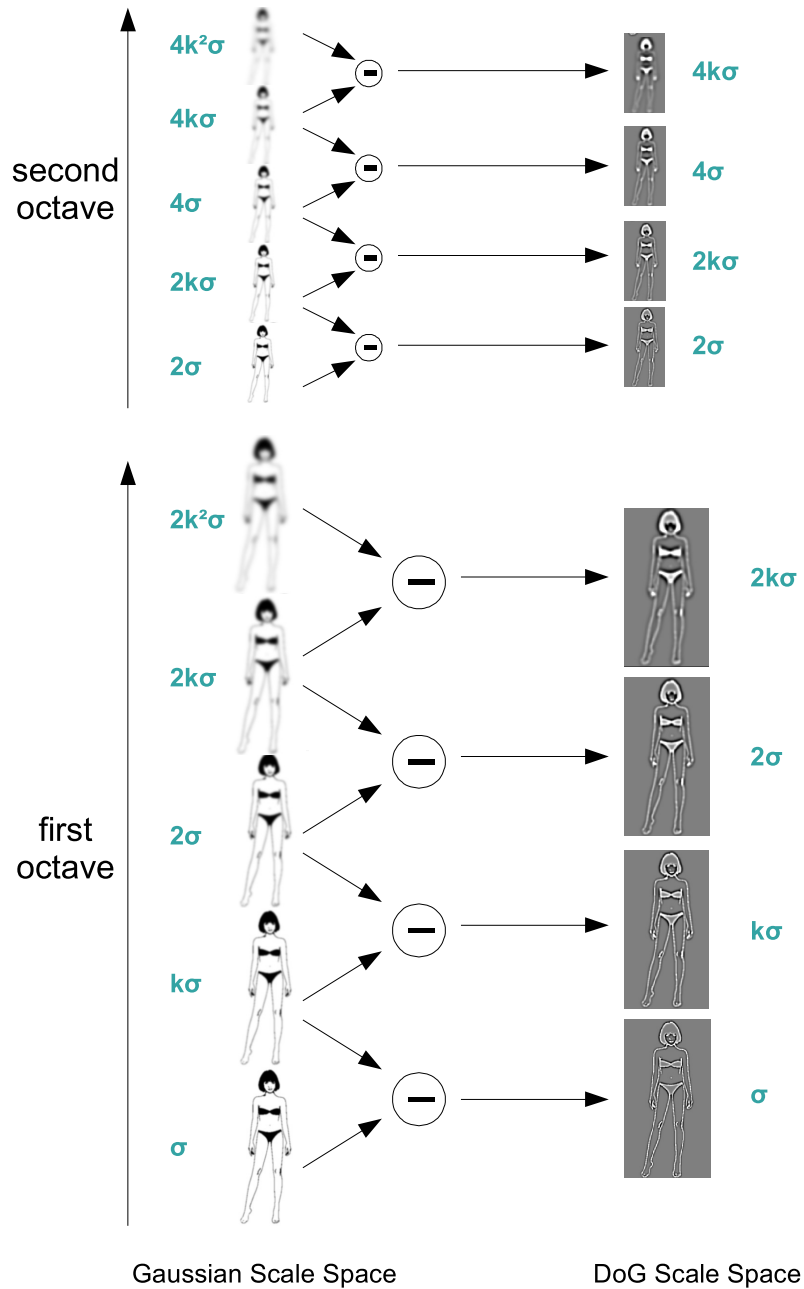


Figure 3.4: GSS - Gaussian scale space; DoG - Difference of Gaussians scale scape

influence the local extrema location. The approximation error will go to zero as k tends to 1, but in practice, this effect has no effect on the stability of extrema detection.

An example of constructing $D(x, y, \sigma)$ is shown in Fig.3.4, where the initial image is incrementally convolved with Gaussians to produce images separated by a constant k factor in scale space, shown stacked in the left column. $s + 3$ blurred images are produced for each octave and final extrema detection covers a complete octave. The adjacent image scales are combined producing the difference-of-Gaussian images shown on the right. Once a complete octave has been computed, the Gaussian images are re-sampled having twice the initial value of σ (2 images from the top of the stack) by taking each second pixel in each row and column. The accuracy sampling with respect to σ is the same as starting from the previous octave and the computational cost is reduced.

Keypoint localization In order to detect the local maxima and minima of $D(x, y, \sigma)$ at each sample point, this point is compared with its eight neighbors in the current image and nine neighbors in the scales above and below. The selected points are the ones that are larger or smaller than all of their neighbors.

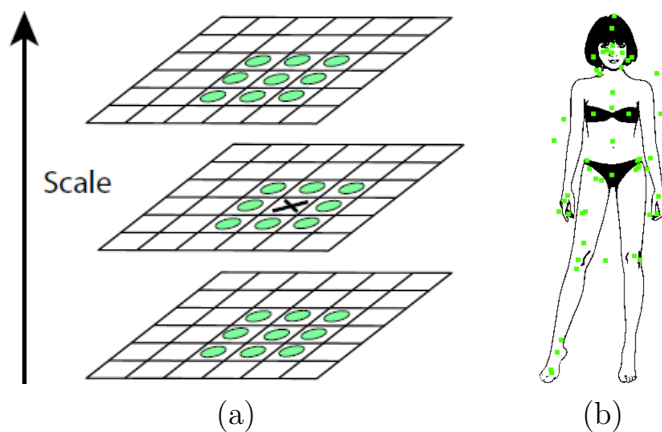


Figure 3.5: SIFT local extrema detection: (a) Maxima and minima of the difference-of-Gaussian images are detected by comparing a pixel (marked with X) to its 26 neighbors in 3×3 regions at the current and adjacent scales (marked with circles). (b) Image with detected SIFT keypoints.

Once a keypoint candidate is detected, a detailed data for location, scale, and ratio is performed. To determine the location of the point, Brown and Lowe developed a 3D quadratic function for fitting a function to the local sample points

3. A NOVEL SYSTEM FOR SKIN MOSAICING

to compute the interpolated location of the maximum, providing a good stability and matching for the keypoint. Their approach uses the Taylor expansion function, Eq. 3.31 (up to the quadratic terms) of $D(x, y, \sigma)$, shifted so that the origin is at the sample point.

$$D(x) = D + \frac{\partial D^T}{\partial x}x + \frac{1}{2}x^T \frac{\partial^2 D}{\partial x^2}x \quad (3.31)$$

The location of the local extremum (\hat{x}) is determined by taking the derivative of this function with respect to x and setting it to zero Eq. 3.32.

$$\hat{x} = -\frac{\partial^2 D^{-1}}{\partial x^2} \frac{\partial D}{\partial x} \quad (3.32)$$

Substitution of \hat{x} from Eq. 3.32 into Eq. 3.13 gives a function for evaluating the SIFT *saliency* $D(\hat{x})$, shown in Eq. 3.33.

$$K_{SIFT} = D(\hat{x}) = D + \frac{1}{2} \frac{\partial D^T}{\partial x} \hat{x} \quad (3.33)$$

SURF Detector: Fast Hessian. The SURF detector was proposed by Bay *et al.* in 2006 [15] and later extended in [14]. Like SIFT, the SURF algorithm also has three main parts, namely detection, description and matching.

The SURF detector is based on the determinant of the Hessian matrix for selecting the location of the keypoint and its scale. Given a point $x = (x, y)$ in an image I , the Hessian matrix $H(x, \sigma)$ in x at scale σ is defined by Eq. 3.34, where L_{xx} , L_{xy} and L_{yy} are the convolutions of the Gaussian second-order derivatives with the image I at the point (x, y) .

$$H(x, \sigma) = \begin{bmatrix} L_{xx}(x, y, \sigma) & L_{xy}(x, y, \sigma) \\ L_{xy}(x, y, \sigma) & L_{yy}(x, y, \sigma) \end{bmatrix} \quad (3.34)$$

Based on Lowe's idea of computing LoG approximations, Bay *et al.* approximate the Gaussian filters using box filters, see Fig. 3.6. These approximated second order Gaussian derivatives can be evaluated very quickly for any filter size by using integral images instead of the original images.

The 9×9 box filters in Fig. 3.6 are approximations for the Gaussian second order derivatives with $\sigma = 1.2$ and represent the lowest scale layer. These

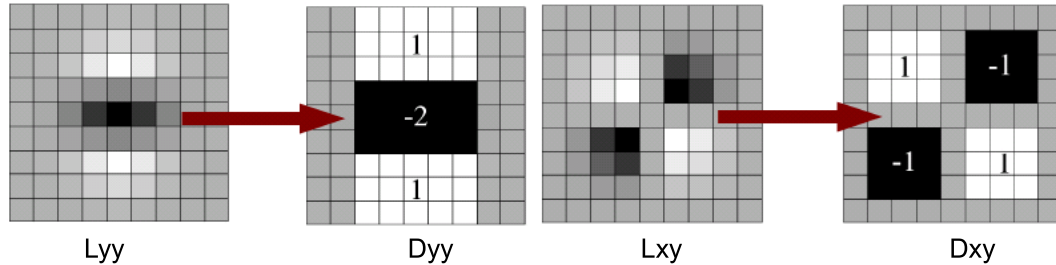


Figure 3.6: SURF: Approximation of the second-order derivatives of Gaussian L_{yy} and L_{xy} with box filters D_{yy} and D_{xy} (mean/average filter).

approximations are denoted as D_{xx} , D_{yy} and D_{xy} . These filter responses are normalized with respect to the mask size, the normalization guarantees a constant Frobenius norm for any filter size as in Eq. 3.35. The SURF scale space instead of iteratively reducing the image size, up-scales the filter size. Such filters could be applied at exactly the same speed directly onto the original image, or even in parallel.

$$\det(H_{approx}) = D_{xx}D_{yy} - (0.9D_{xy})^2 \quad (3.35)$$

The output of the above 9×9 filter is considered as the initial scale layer, referred as $s = 1.2$ (corresponding to Gaussian derivatives with $\sigma = 1.2$). The following layers are obtained by filtering the image with gradually larger masks, taking into account the discrete nature of integral images and the specific structure of the box filters. Specifically, this results in filters of 9×9 , 15×15 , 21×21 , 27×27 , etc. For each octave, the filter size is doubled and the sampling intervals for the extraction of the interest points can be doubled as well. Fig. 3.7 compares the classical approach of multi-scale image pyramid construction and the SURF filter-scaling method.

In order to localize interest points in the image and over scales, the same steps previously explained for SIFT are performed. The comparison to the 26 neighbors in a $3 \times 3 \times 3$ scale space volume is applied for each selected point. The maxima of the determinant of the Hessian matrix are then interpolated in scale with the method proposed by Brown and Lowe. Scale space interpolation is especially important in SURF, because the difference in scale between the first

3. A NOVEL SYSTEM FOR SKIN MOSAICING

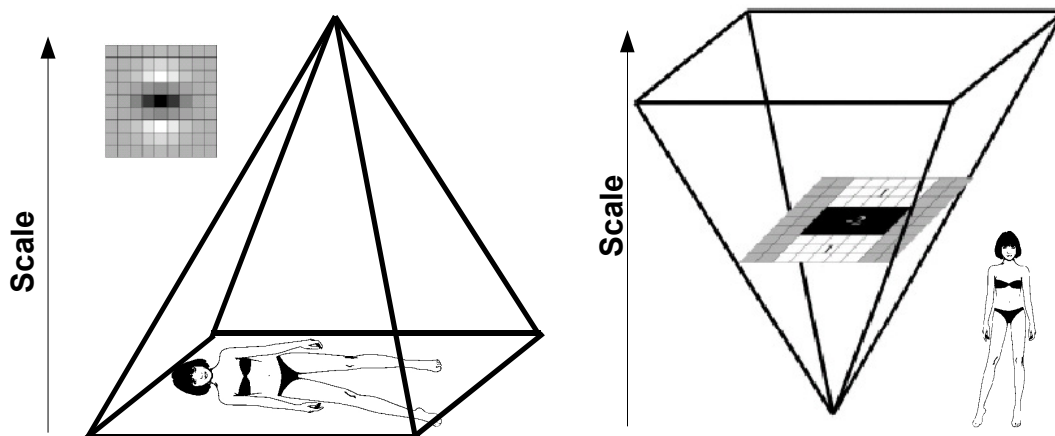


Figure 3.7: SURF: Comparison of the classical approach to image scale space construction and the SURF method using box filters.

layers of every octave is relatively large.

MSER detector. Maximally Stable Extremal Regions (MSER) was proposed by Matas et al.[115] to detect blob regions in wide baseline image registration. As defined by the authors, a maximal region is a connected component of an appropriately thresholded image. In other words, MSER extracts compact regions whose pixels have either higher (bright blobs) or lower (dark blobs) intensity values than all the surrounding pixels.

In order to extract MSER regions, the image is binarized using gradually increasing intensity thresholds. The binarization is used to extract compact dark and bright regions. MSER selects only those regions whose area changes insignificantly over a large range of intensity thresholds. These areas prove to be highly stable in both illumination (linear and non-linear) and affine transformations.

3.1.4 Feature Description

Once a set of keypoints are detected in each image, an invariant descriptor based on the intensity information of the point neighborhood needs to be computed. This descriptor should be invariant to image deformations, change of lighting and robust to the presence of noise. In this subsection, a classical technique based

on correlation [172] is explained as well as descriptor based techniques[27], such as SURF and SIFT will be overviewed for their good results in many different applications.

3.1.4.1 Correlation

Correlation methods are based on template matching. During this process, windows of predefined size taken from two images are tried to be matched using similarity based measures. The cost of this process is very high, since all possible window pairs from the two images should be investigated. This cost could be reduced by opening only windows on previously detected points. This will reduce the number of computations.

Two main groups of techniques are included in correlation based matching techniques, these are similarity measure (NCC - Normalized Cross-Correlation) and difference measure (SAD - Sum of Absolute Differences and SSD - Sum of Squared Differences).

Given a keypoint $p1$ in the first image I_1 , with coordinates (x_1, y_1) , and a keypoint $p2$ in the second image I_2 , with coordinates (x_2, y_2) , the NCC score between two rectangular windows of size $2r \times 2r$ centered on each keypoint is given in Eq. 3.36, where \bar{I} and $\sigma^2(I)$ denotes the average (Eq. 3.37), and the variance (Eq. 3.38) of pixel inside the window.

$$NCC(p1, p2) = \frac{\sum_{i=-r}^{i=r} \sum_{j=-r}^{j=r} (I_1(x_1 + i, y_1 + j) - \bar{I}_1)(I_2(x_2 + i, y_2 + j) - \bar{I}_2)}{r^2 \sqrt{\sigma^2(I_1)\sigma^2(I_2)}} \quad (3.36)$$

$$\bar{I} = \frac{\sum_{i=-r}^{i=r} \sum_{j=-r}^{j=r} I(x + i, y + j)}{r^2} \quad (3.37)$$

$$\sigma^2(I) = \frac{\sum_{i=-r}^{i=r} \sum_{j=-r}^{j=r} (I(x + i, y + j) - \bar{I})^2}{r^2} \quad (3.38)$$

The difference measures, SAD and SSD, are given by Eq. 3.39 and Eq. 3.40 respectively.

3. A NOVEL SYSTEM FOR SKIN MOSAICING

$$SAD(p_1, p_2) = \sum_{i=-r}^{i=r} \sum_{j=-r}^{j=r} |I_1(x_1 + i, y_1 + j) - I_2(x_2 + i, y_2 + j)| \quad (3.39)$$

$$SSD(p_1, p_2) = \sum_{i=-r}^{i=r} \sum_{j=-r}^{j=r} (I_1(x_1 + i, y_1 + j) - I_2(x_2 + i, y_2 + j))^2 \quad (3.40)$$

The limitations of these methods originate in their basic idea. The shape of the windows, usually rectangular or circular, only allows the images to be registered locally by a translation. If the images are deformed by more complex transformations, these window shapes are not able to cover the same regions in both images. Correlation based methods only take into consideration the image intensities and not the structural analysis. Consequently, they are not invariant to intensity changes introduced by noise and illumination. NCC is invariant to linear changes in intensity and quite robust to image distortion, but it fails for more complex lighting changes and is not invariant to rotation. The correlation of these image patches is highly sensitive to changes in the 3D viewpoint and non-rigid deformations. These effects will produce miss registrations.

3.1.4.2 SIFT Descriptor

The SIFT detector provides invariance to translation and scale. Therefore, the invariance to image rotation, changes in lighting and the 3D view point is obtained with the SIFT descriptor. To compute the invariance to image rotation, the descriptor is specially calculated for each keypoint orientation.

To assign a local orientation to each detected keypoint, first the scale of the keypoint is used to select the smoothed image, L , with the closest scale. Then all computations are performed in a scale-invariant manner. For each image sample, $L(x, y)$ in the closest scale, the gradient magnitude, $m(x, y)$ and the orientation, $\theta(x, y)$, are precomputed using pixel differences as shown in Eq. 3.41 and Eq. 3.42.

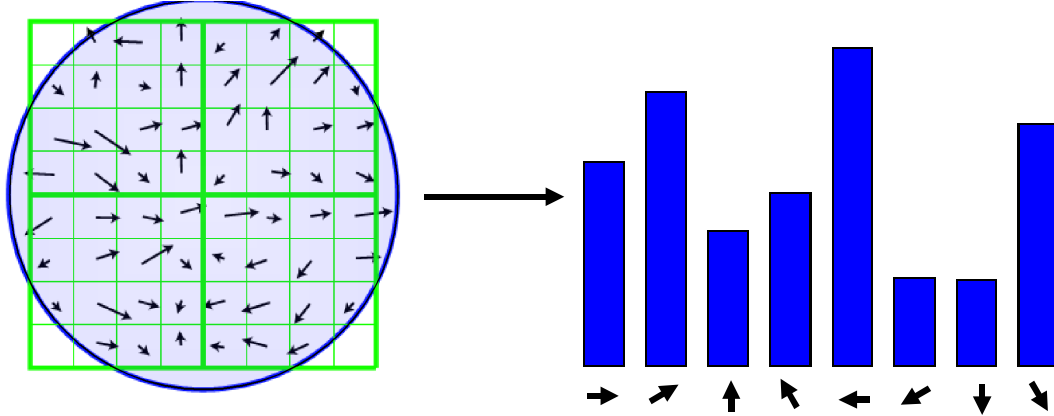


Figure 3.8: SIFT orientation: an orientation histogram formed from the gradient orientations of sample points around the keypoint.

$$m(x, y) = \sqrt{(L(x + 1, y) - L(x - 1, y))^2 + (L(x, y + 1) - L(x, y - 1))^2} \quad (3.41)$$

$$\theta(x, y) = \tan^{-1} \frac{L(x, y + 1) - L(x, y - 1)}{L(x + 1, y) - L(x - 1, y)} \quad (3.42)$$

An orientation histogram is formed from the gradient orientations of sample points within a region around the keypoint. The orientation histogram, shown in Fig. 3.8, has 8 bins covering the 360-degree range of orientation. Each sample added to the histogram corresponds to a dominant direction of local gradients. Once the histogram is constructed, the highest peak and the peaks higher than 80 percent of the maximum are selected to compute the orientation of the keypoint. If the same position has multiple peaks of similar magnitude, multiple keypoints will be created at the same location and scale but with different orientations. Finally, to increase the accuracy, a parabola is fitted to the 3 histogram values closest to each peak to interpolate their position.

Each vector of a detected SIFT keypoint is composed of 5 entries $(x, y, \sigma, m, \theta)$, where x, y and σ represent the spatial coordinate and the scale, and m and θ represent the gradients magnitude and orientation respectively.

Computing the SIFT descriptor is done by first sampling the image's gradient

3. A NOVEL SYSTEM FOR SKIN MOSAICING

magnitude and orientation around the keypoint location, at the selected scale of the keypoint. The gradients are computed in the same way as for computing the orientation assignment. In order to achieve orientation invariance, the coordinates of the descriptor and the gradients orientation are rotated relative to the keypoint orientation, then a Gaussian weighting function with σ equal to half the width of the descriptor window is used to assign a weight to the magnitude of each sample point. This process will avoid sudden changes in the descriptor with small changes in the position of the window, and also has less influence on the gradients farther from the center of the descriptor.

The keypoint descriptor is formed by creating an orientation histogram for each 4×4 sample region of the initial histogram. This orientation histogram is composed of 8 bins covering the 360-degree range of orientations. Each entry within a bin is multiplied by a weight of $1 - d$ for each dimension, where d is the distance of the sample from the central value of the bin as measured in units of the histogram bin spacing.

The final vector that describes the keypoint contains the values of all the orientation histogram entries. The SIFT paper uses a 4×4 array of orientation histograms with 8 orientation bins in each computed from a 16×16 sample array. This will generate an array of $4 \times 4 \times 8 = 128$ entries to define the keypoint. This vector will be normalized to reduce the effects of lighting changes. This process is summarized in Fig. 3.9.

To use SIFT with detectors that do not provide orientation, like Harris, the keypoint descriptors were computed in the same way, but at a single scale.

3.1.4.3 SURF Descriptor

The SURF descriptor is based on the same idea as SIFT. The first step consists of fixing a reproducible orientation, based on the information provided by a circular region around the interest point. Then, a square region is aligned to the selected orientation and the SURF descriptor is extracted from it.

For each detected keypoint, an orientation is assigned to achieve invariance of rotation changes. The orientation is obtained by computing the Haar-wavelet responses in the x and y directions shown in Fig. 3.10. Wavelet responses are

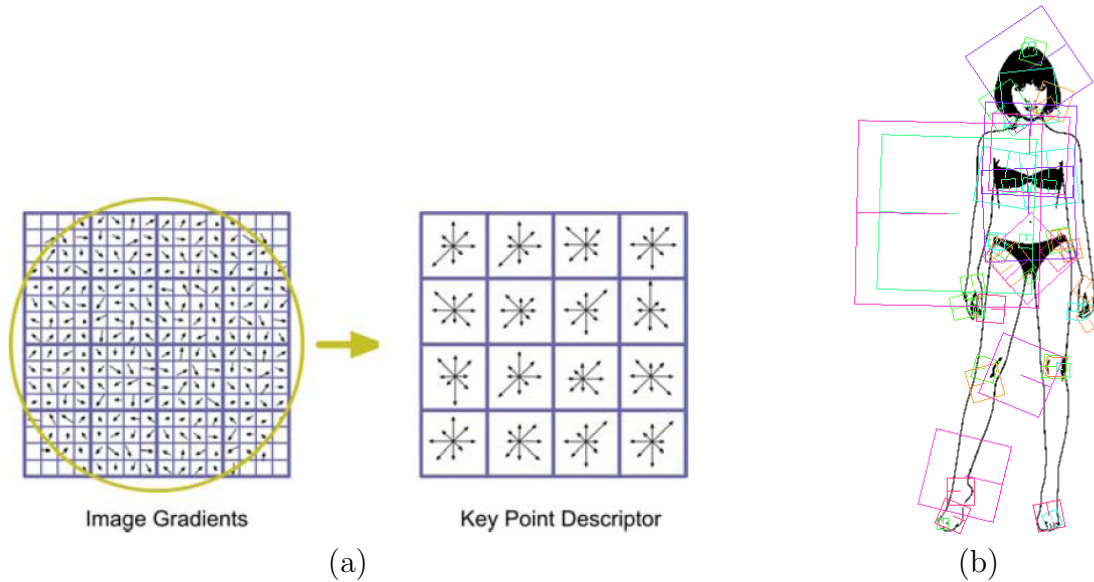


Figure 3.9: The SIFT descriptor: (a) Creation of the orientation histograms, forming the SIFT descriptor of the detected keypoint. The gradient magnitude and orientation of each image sample point is computed in a region around the keypoint location. These are weighted by a Gaussian window, illustrated by the overlaid circle. The samples are then accumulated into orientation histograms summarizing the contents over 4×4 subregions, with the length of each arrow corresponding to the sum of the gradient magnitudes near that direction within the region. This figure shows a 4×4 descriptor array computed from a 16×16 set of samples. (b) SIFT descriptors of several SIFT keypoints detected at different scales.

computed in a circular neighborhood of radius $6s$ around the keypoint, where s corresponds to its scale. The sampling step is also chosen to be s . The size of the wavelet filters is adjusted to the scale, higher scales will define larger filters. Using integral images allows us to compute the response in the x or y direction at any scale with only six operations. The side length of the wavelet is $4s$. Wavelet responses are calculated and weighted with a Gaussian ($\sigma = 2.5s$) centered on the keypoint. These responses are represented as vectors in a space with the horizontal response strength along the abscissa (x axis in Fig. 3.10) and the vertical response strength along the ordinate (y axis in Fig. 3.10). The dominant orientation is computed by the sum of all responses within a sliding orientation

3. A NOVEL SYSTEM FOR SKIN MOSAICING

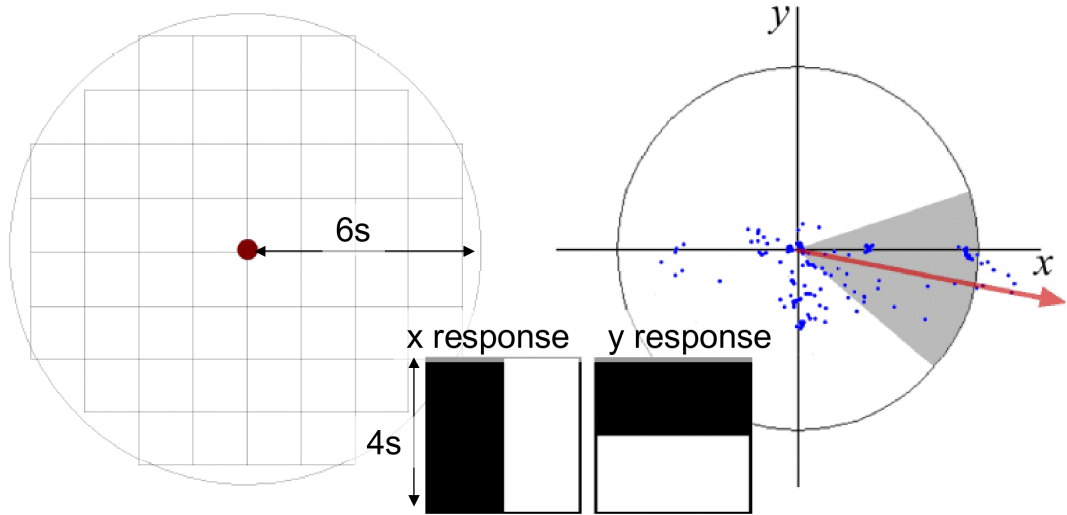


Figure 3.10: A SURF orientation assignment. Left: In the circular neighborhood of radius $6s$ around the keypoint, where wavelet responses are computed. s is the scale of the keypoint. Middle: Haar wavelet 2D filters used for SURF. Right: The space for representation of the wavelet responses as vectors with coordinates x (horizontal response) and y (vertical response). The dominant orientation is estimated by calculating the sum of all vectors within a sliding orientation window covering an angle of $\pi/3$.

window covering an angle of $\pi/3$. The horizontal and vertical responses within the window are summed to form a new vector. The longest vector gives its orientation to the interest point.

To extract the SURF descriptor, a square region centered on the keypoint is constructed and oriented along the selected orientation. The size of this region is $20s$ and is split up regularly into 16 (4×4) smaller square sub-regions, which keeps important spatial information in. For each sub-region, a simple feature of 5×5 regularly spaced sample points is computed.

The Haar wavelet response in the horizontal direction is called d_x and d_y for the vertical (filter size is $2s$). Both are defined in relation to the selected keypoint orientation, as illustrated in Fig. 3.11 (b) The wavelet responses are weighted with a Gaussian ($\sigma = 3.3s$) centered on the keypoint, increasing the robustness against geometric deformations and localization errors. d_x and d_y are summed up independently over each subregion to form a first set of entries for the feature

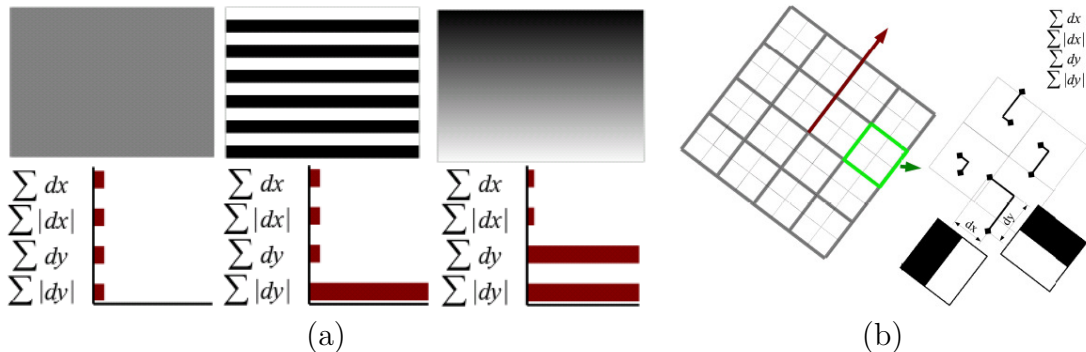


Figure 3.11: SURF Descriptor: (a) The descriptor is constructed with sums of wavelet responses along x and y over the $20s$ keypoint neighborhood splitted regularly into 16 square sub-regions. The descriptor entries of a sub-region represent the nature of the underlying intensity pattern. Left: In case of a homogeneous region, all values are relatively low. Middle: In presence of frequencies in the x direction, the value of $\Sigma|d_x|$ is high, but all others remain low. If the intensity is gradually increasing in the x direction, both values Σd_x and $\Sigma|d_x|$ are high. (b) A square region centered in the keypoint is constructed and aligned to the keypoint orientation. “Horizontal”(x) and “vertical”(y) wavelets responses are defined also in relation to this orientation.

vector, providing information about the polarity of the intensity changes, and the sum of the absolute values responses, $|d_x|$ and $|d_y|$, are also extracted. For all 4×4 regions the descriptor structure $v = (\Sigma d_x, \Sigma d_y, \Sigma|d_x|, \Sigma|d_y|)$ is stored in a final 64 length array. The wavelet responses are invariant to a bias in illumination (offset). Invariance to contrast (a scale factor) is also obtained by turning the descriptor into a unit vector through normalization.

3.1.5 Feature Matching

Once the keypoints and descriptors are computed in two different images, they have to be matched. The simplest method to match two image keypoints is to compute the similarity between their descriptors, computing the Euclidean distance between them. Let $\mathbf{v}_1 = (v_1^1, v_1^2, \dots, v_1^N)$ and $\mathbf{v}_2 = (v_2^1, v_2^2, \dots, v_2^N)$ be descriptors of keypoint p_1 and p_2 in the first and second images respectively, where N is equal to 128 in SIFT and 64 in SURF. The Euclidean distance between v_1 and v_2 is

3. A NOVEL SYSTEM FOR SKIN MOSAICING

given by Eq. 3.45.

$$Dist(p_1, p_2) = \sqrt{\sum_{i=1}^N (v_1^i - v_2^i)^2} \quad (3.43)$$

Each descriptor is highly distinctive. This means that it is highly probable to find a good match taking into account only the Euclidean distance between two descriptors of different images. However there are often many features that will not have a corresponding pair in the other image. This is because some of them could be occluded, do not appear in the second image, or simply, they are not robust enough to be detected in both images. For this reason a threshold describing the maximum distance between the actual descriptor and its closest correspondant is usually set to discard the incorrect matches.

SIFT and SURF have more sophisticated techniques to find matches. These methods take into account the distance with respect to the first and the second closest neighbors. Correct matches need to have the closest neighbor significantly closer than the closest incorrect match to achieve reliable matching. The second-closest match estimates the density of false matches within this portion of feature space and, at the same time, identifies specific instances of feature ambiguity.

In the original SIFT [106] paper, the threshold on the ratio of second-closest and first-closest neighbors of each keypoint was set to 1.5. p_1 being a keypoint in image 1 and p_2^{first} and p_2^{second} keypoints detected in the second image, and taking into account the Euclidean distance d_E between the descriptors, the keypoints p_1 and p_2^{first} are matched if Eq. 3.44 holds.

$$\frac{d_E(p_1, p_2^{second})}{d_E(p_1, p_2^{first})} > 1.5 \quad (3.44)$$

3.1.6 Motion Estimation

The number of correspondences from the previous stage is usually quite large, and often half of these initial matches are incorrect[173]. Outliers should be rejected using a robust motion estimation algorithm, which checks if the spatial coordinates of each correspondence agrees with the dominant motion.

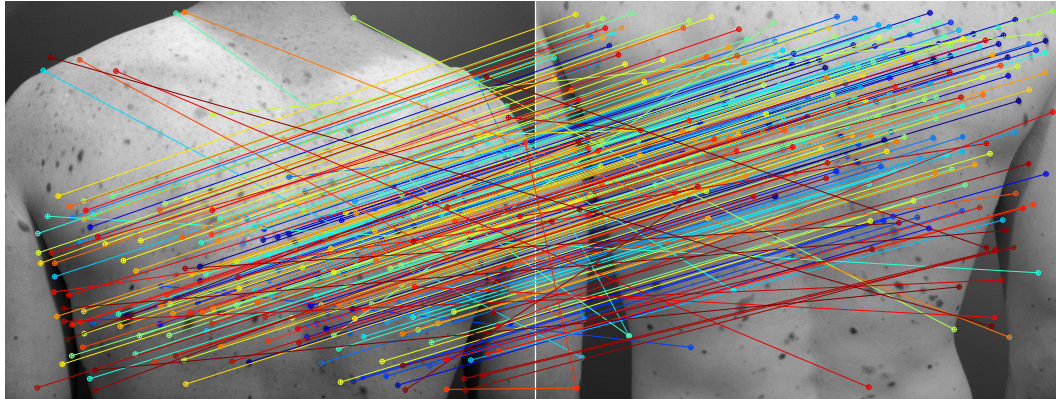


Figure 3.12: Matched images from a patient. SURF was used for detection and description. The number of keypoints detected in the first and second images is 1666 and 1661 respectively. The number of matched keypoints is 256, among which there are 41 false correspondences (*outliers*).

The motion can be represented in 2D by a matrix called Homography [71]. The homograph \mathbf{H} defines a planar transformation between two images. The transformation matrix H relates the points in one image to those in the other. For any point \mathbf{p} in the first image, only one point \mathbf{p}' in the second image is defined, by $\mathbf{p}' = \mathbf{H}\mathbf{p}$, and vice-versa: $\mathbf{p} = \mathbf{H}^{-1}\mathbf{p}'$, where $\mathbf{p} = (x, y, 1)^T$ is a 2D point expressed in homogeneous coordinates.

Several methods have been studied to estimate a Homography according to [71]. These methods can be classified into four different groups according to their degrees of freedom, which are the number of parameters that might vary independently. These models are Euclidean, Similarity, Affine and Projective, in increasing order of degrees of freedom (see Fig. 3.13). It is important to note that a set of image correspondences can always be adjusted better with a projective transformation, because it will minimize the Euclidean norm of the residuals. However, computing a projective transformation does not always provide the expected result, since measurements have errors and the real motion between two images can be much simpler than that estimated by the homograph. Motion estimation using a transformation model with fewer degrees of freedom sometimes produces better results.

3. A NOVEL SYSTEM FOR SKIN MOSAICING

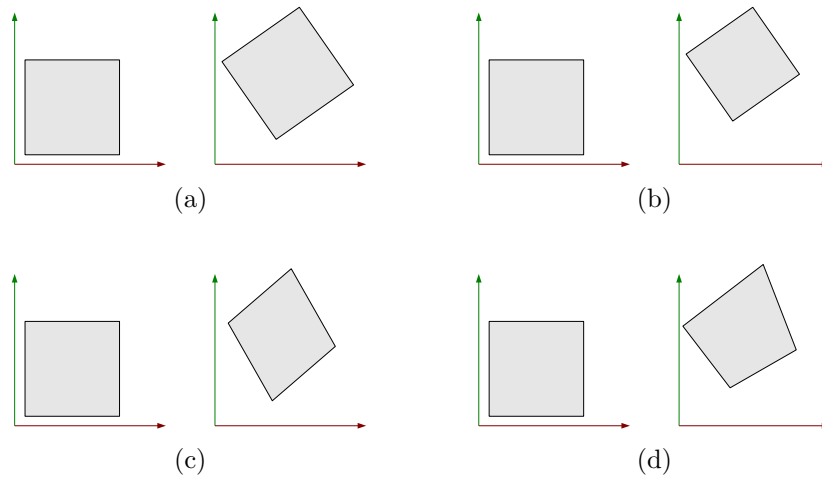


Figure 3.13: Planar transformations: (a) Euclidean (rotation and translation), (b) Similarity (rotation, translation and scale), (c) Affine (translation, rotation, shear and anisotropic scaling), (d) Projective (translation, rotation, shear, anisotropic scaling and perspective distortion).

3.1.6.1 Homography Estimation Methods

Euclidean transformation. Euclidean transformations have three degrees of freedom, two for translation and one for rotation. In order to calculate a Euclidean transformation, a minimum of two correspondences are needed. The homograph is defined by Eq. 3.45.

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta & tx \\ \sin\theta & \cos\theta & ty \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \quad (3.45)$$

Similarity transformation. A similarity transformation is a Euclidean transformation that may suffer a scaling (s). It has four degrees of freedom, one for rotation, two for translation and one for scaling. Two correspondences are enough to calculate similarity transformations.

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} scos\theta & -ssin\theta & tx \\ ssin\theta & scos\theta & ty \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \quad (3.46)$$

Affine transformation. The affine transformation is composed of six unknowns, allowing transformations that include translation, rotation, shear and anisotropic scaling, which is a direction-dependent scaling. The matrix representation of affine transformation is given by Eq. 3.47.

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} h_{11} & h_{12} & tx \\ h_{21} & h_{22} & ty \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \quad (3.47)$$

Affine transformation can be broken down into three rotations and one non-isotropic scaling by using Singular Value Decomposition (SVD) given in Eq. 3.48. First, a rotation by angle ϕ is applied. Next a non-isotropic scaling (l_1, l_2) along the new rotated x and y axes is performed. Then, a back rotation by $-\phi$ and finally, a rotation by θ completes the process.

$$\begin{pmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{pmatrix} = \begin{pmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{pmatrix} \begin{pmatrix} cos(-\phi) & -sin(-\phi) \\ sin(-\phi) & cos(-\phi) \end{pmatrix} \begin{pmatrix} \rho_1 & 0 \\ 0 & \rho_2 \end{pmatrix} \begin{pmatrix} cos\phi & -sin\phi \\ sin\phi & cos\phi \end{pmatrix} \quad (3.48)$$

Projective transformation. The projective transformation includes all the possible deformations in a planar scene; translation, rotation, shear, anisotropic scaling and perspective distortion. It can be represented by Eq. 3.49 where λ is an arbitrary scaling factor.

$$\begin{pmatrix} \lambda x' \\ \lambda y' \\ \lambda \end{pmatrix} = \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \quad (3.49)$$

3. A NOVEL SYSTEM FOR SKIN MOSAICING

$$\begin{aligned}
\lambda x' &= h_{11}x + h_{12}y + h_{13} \\
\lambda y' &= h_{21}x + h_{22}y + h_{23} \\
\lambda &= h_{31}x + h_{32}y + 1
\end{aligned} \tag{3.50}$$

Projective transformation can be calculated from a set of point correspondences in the following way. By substituting the third equation from the linear system 3.50 into the first two, the system can be rewritten as Eq. 3.51.

$$\begin{aligned}
x' &= h_{11}x + h_{12}y + h_{13} - h_{31}x/x - h_{32}x/y \\
y' &= h_{21}x + h_{22}y + h_{23} - h_{31}y/x - h_{32}y/y
\end{aligned} \tag{3.51}$$

In order to find the eight coefficients of the homograph, at least four correspondences are needed, since every correspondence provides two constraints for x and two more for y . For n point correspondences, system 3.51 can be rewritten in the form $b = Ah$ with Eq. 3.52.

$$\begin{bmatrix} x_1' \\ y_1' \\ x_2' \\ y_2' \\ x_3' \\ \vdots \\ x_n' \\ y_n' \end{bmatrix} = \begin{bmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -x_1'/x_1 & -x_1'/y_1 \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -y_1'/x_1 & -y_1'/y_1 \\ x_2 & y_2 & 1 & 0 & 0 & 0 & -x_2'/x_2 & -x_2'/y_2 \\ 0 & 0 & 0 & x_2 & y_2 & 1 & -y_2'/x_2 & -y_2'/y_2 \\ x_3 & y_3 & 1 & 0 & 0 & 0 & -x_3'/x_3 & -x_3'/y_3 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_n & y_n & 1 & 0 & 0 & 0 & -x_n'/x_n & -x_n'/y_n \\ 0 & 0 & 0 & x_n & y_n & 1 & -y_n'/x_n & -y_n'/y_n \end{bmatrix} \begin{bmatrix} h_{11} \\ h_{12} \\ h_{13} \\ h_{21} \\ h_{22} \\ h_{23} \\ h_{31} \\ h_{32} \end{bmatrix} \tag{3.52}$$

The solution for \mathbf{h} which minimizes $\|\mathbf{Ah}\|$ subject to $\|\mathbf{h}\|=1$ is the unit singular vector corresponding to the smallest singular value of \mathbf{A} . The SVD can therefore be used to obtain the solution [71]. The SVD of a given matrix $\mathbf{A}_{m \times n}$ is written as $\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{V}^T$, where $\mathbf{U}_{m \times m}$ and $\mathbf{V}_{n \times n}$ are orthogonal matrices, $\mathbf{U}\mathbf{U}^T = \mathbf{I}$, $\mathbf{V}\mathbf{V}^T = \mathbf{I}$, and \mathbf{D} is a diagonal matrix with non-negative elements.

The elements of \mathbf{D} , d_1, d_2, \dots, d_n , are singular values of \mathbf{A} :

$$\|Ah\| = \|UDV^T h\| = \|DV^T h\| = \|Dz\| \quad (3.53)$$

where $\mathbf{z} = \mathbf{V}^T \mathbf{h}$ and $\|\mathbf{z}\| = 1$ since \mathbf{U} and \mathbf{V} are norm preserving matrices. Eq.3.53 is minimized by setting $\mathbf{z} = (0, 0, 0, \dots, 1)$, as \mathbf{D} is a diagonal rectangular matrix $\mathbf{D}_{m \times n}$ with its elements are sorted in descending order. Finally, the homograph is found by means of the equation $\mathbf{h} = \mathbf{V}\mathbf{z}$, which corresponds to the last column of \mathbf{V} .

3.1.6.2 Rejection of Outliers: RANSAC

The homograph estimation algorithm described above assumes that the only source of error in the set of presented point correspondences is the measurement of the points location, but this assumption is not always true, since in practical situations, mismatched points are also present. These points are considered outliers, and introduce errors in the estimated homograph. The goal of this step is to determine which of the initial matches are inliers and which are outliers. Then, estimate the homograph again through least squares taking into account only the inliers. In order to deal with these correspondences not obeying the dominant motion, two common solutions are used[77]. These solutions imply an outlier rejection algorithm, the most frequently used being: RANSAC (RANdom SAMple Consensus [50]) and LMedS (Least Median of Squares [148]). The RANSAC algorithm needs information on the percentage of outliers, which is not usually available, while the LMedS estimation does not require such information, but is very time consuming. For this reason, RANSAC has been selected for this work.

The RANdom SAMple Consensus (RANSAC) algorithm presented by Fisher and Bolles in 1981 is a robust estimator able to deal with a high proportion of outliers. It has proved extremely effective in computing two-view relations, such as homographies, from point-to-point image correspondences. In general, the RANSAC idea is very simple: two of the dataset points are selected randomly and defines a line. Then the distance from all the other points to the line is computed and those that lie within a distance threshold are selected. This random selection is repeated a number of times and the line with the most accepted points

3. A NOVEL SYSTEM FOR SKIN MOSAICING

is selected as the robust fit. The points that lie within the threshold distance are the inliers while the others are the outliers. If one point is an outlier, the line will not gain much support from it (see Algorithm 2).

Algorithm 2 The RANSAC Robust estimation algorithm

- 1: Randomly select a sample of s data points from S and instantiate the model from this subset.
 - 2: Determine the set of points S_i that are within a threshold distance of the defined threshold.
 - 3: If the size of S_i is greater than any predefined threshold T , re-estimate the model with the points in S_i and terminate.
 - 4: If the size of points of S_i is less than T , select a new subset and repeat the above.
 - 5: After N trials, the largest consensus set S_i is selected and the model is re-estimated.
-

If the model is a planar homograph (\mathbf{H}), and the data a set of 2D point correspondences (x, x') , then the minimal subset consists of four correspondences.

The homograph estimate returned by the algorithm could be used to initialize a non-linear homograph estimator described in the next Section. In this case, the set of point correspondences selected as inliers is used to compute a refinement of the estimated \mathbf{H} .

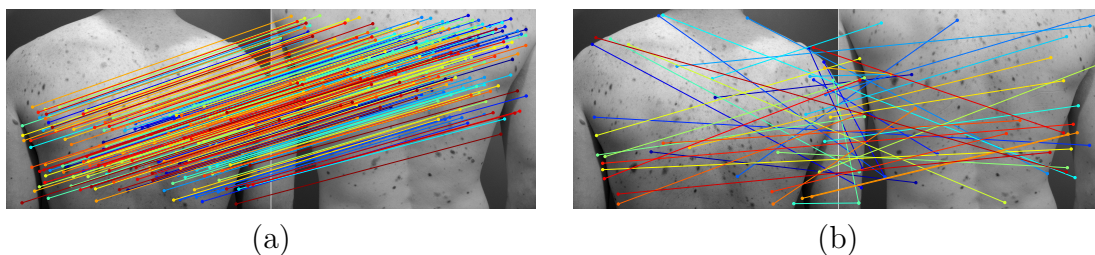


Figure 3.14: RANSAC: The number of initial matched keypoints is 256 from Fig. (3.12), (a) 215 accepted matches (inliers) by RANSAC and (b) 41 false correspondences (outliers).

3.1.6.3 Homography Optimization

A number of nonlinear methods have been proposed to refine homographs [71] [32]. These refinements of the estimation of the homograph basically have the goal of minimizing a defined cost function. This method could be linear or non-linear depending on the function selected and on the homograph model. Minimizing such cost functions requires the use of iterative techniques.

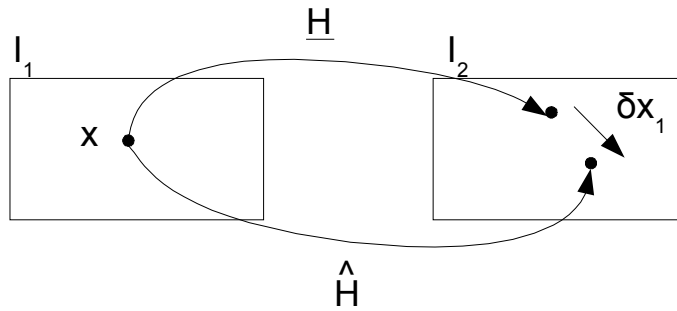


Figure 3.15: The error of the projected point x of image I_1 into I_2 is defined by δx . \hat{H} represents the estimated homograph and \underline{H} is the ground-truth homograph.

From Eq. 3.49, a cost function e could be expressed as the Euclidean squared distance between the detected match and the one estimated by the homograph by Eq. 3.54. This difference is represented in Fig. 3.15.

$$e(h) = \sum_{i=1}^n \left(\left(x_i' - \frac{h_{11}x_i + h_{12}y_i + h_{13}}{h_{31}x_i + h_{32}y_i + 1} \right)^2 + \left(y_i' - \frac{h_{21}x_i + h_{22}y_i + h_{23}}{h_{31}x_i + h_{32}y_i + 1} \right)^2 \right) \quad (3.54)$$

Where n is the number of correspondences and $\mathbf{h} = [h_{11}, h_{12}, h_{13}, h_{21}, h_{22}, h_{23}, h_{31}, h_{32}]^T$. Solving the problem means obtaining a homograph h that minimizes the cost function $e(h)$. This problem is known as a nonlinear least squares problem, which can be solved using iterative methods such as Newton Iteration or Levenberg-Marquardt [98] [114]. Eq. 3.54 can be rewritten in closed form as Eq. 3.55, assuming f as locally linear, a first order Taylor expansion can be formulated around h_0 as Eq. 3.56, where h_0 is the initial estimation of the homograph.

3. A NOVEL SYSTEM FOR SKIN MOSAICING

$$e = |f(h) - x'| \quad (3.55)$$

$$f(h) = f(h_0) + \frac{\partial f}{\partial h}(h - h_0) + r_n \quad (3.56)$$

where r_n is called the *remainder* and is calculated as follows in Eq. 3.57.

$$r_n = \int_{h_0}^h f^{(n+1)}(u) \frac{(x - u)^n}{n!} du \quad (3.57)$$

Consider $J = \frac{\partial f}{\partial h}$ as a linear mapping represented by the jacobian of f with respect to the elements in h and then the function that minimizes $f(h_1) - x'$ can be written as Eq. 3.58

$$f(h_1) = f(h_0) + J\Delta h - x' = e_0 + J\Delta h \quad (3.58)$$

Where $h_1 = h_0 + \Delta h$. Now the function $|e_0 + J\Delta h|$ needed to be minimized over Δh is linear and can be minimized using normal equations like Eq. 3.59

$$\begin{aligned} J^T J \Delta h &= -J^T e_0 \\ \Delta h &= -J^+ e_0 \end{aligned} \quad (3.59)$$

and $h_1 = h_0 - J^+ e_0$. Vector h that minimizes the Eq. 3.54, can be computed iteratively with $h_{i+1} = h_i + \Delta h_i$. In order to start the iteration ($i = 0$), an initial estimation of h_0 should be provided (normally using a linear estimator). This initial estimation plays an important role in achieving converge, since this non-linear approach could be stuck around a local extrema if the initial estimation is not good enough.

3.1.7 Global Registration and Alignment

The objective of this section is to describe two of the last steps in the mosaicing pipeline. First of all, global registration will transform all the images into a global coordinate frame which contains the whole scene.

3.1.7.1 Global Registration

This step will register every image into a global coordinate frame which contains the whole scene. This coordinate frame is usually aligned to one of the images called the *reference frame*. First of all, by using the methods described above to compute the homographies between an image pair, we estimate the transformations between the first image and the second image, between the second and third images, and so on. For each of these transformations, it is possible to obtain the transformation from the reference frame by concatenating (multiplying) the intermediate sequence of homographs as shown in Eq. 3.60.

$$H_{(ref \rightarrow k)} = \prod_{n=ref}^{k-1} H_{(n \rightarrow n+1)} \quad (3.60)$$

where $H_{(n \rightarrow n+1)}$ is the homograph between the reference image n and the next image $n + 1$.

This computation accumulates the error in each homograph, and this effect is usually detected when the image sequence has a loop, which means that the same part of the scene is mapped more than once. To solve this problem, the global alignment methodology presented in the next section is needed.

3.1.7.2 Global Alignment

To create a globally consistent mosaic we have to consider that a point that appears in different images should be mapped to the same point in the mosaic. Some solutions are proposed for this problem in [42] where global consistency of the inter-frame alignment matrices is solved by using a linear system of equations. In [32] the maximum likelihood between all the point matches is computed. Finally, in [150] [84] [145], a graph representation is used.

When consecutive images are registered, a graph-based registration and a simultaneous registration (i.e., bundle adjustment) are performed. The vertices of the graph represent each image and every edge represents the homograph between images see Fig. 3.16. This graph representation allows detection of the edges between non-consecutive images. The last step of simultaneous registration

3. A NOVEL SYSTEM FOR SKIN MOSAICING

uses local constraints to determine the optimal global registration that minimizes an error function.

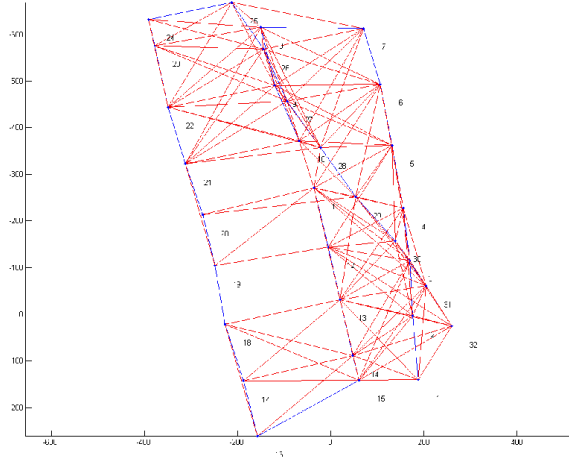


Figure 3.16: Graph-based registration: The vertices represent images and the line edges between them. Blue edges represent consecutive pairs and red edges non-consecutive image overlaps.

The graph can have all the possible links between images, but if all the possible edges are selected, the computational cost of estimating the homographs will be expensive. For this reason, the graph contains only the edges with a higher probability of success in the computation of the homograph. The vertices added are those which have a significant overlap between the image pair. An overlap formula can be computed according to the distance between centroids:

$$\delta_{ij} = \frac{\max(0, |c_i - c_j| - |d_i - d_j|/2)}{\min(d_i, d_j)} \quad (3.61)$$

where c_i, c_j are the centroids of the projection onto the mosaic of images I_i and I_j , respectively. And d_i and d_j are the distances between the image center and the farther image corner when the image is projected onto the mosaic. If in $\delta_{ij} > 1$ there is no overlap, the smaller the value, the bigger the overlap surface between the images.

Once the new edges are detected, the homograph between them is computed with the methods described above. Finally, bundle adjustment[104] is applied to

find the optimal motion parameters, with the aim of globally aligning the mosaic. From these parameters, the absolute homograph of each image can be computed.

For each overlapping pair of images I_i and I_j , with N points-wise correspondences between them, it is possible to define a cost function. These points are mapped to the mosaic according to Eqs. 3.62 and 3.63, respectively.

$${}^M X_i = {}^M H_i \cdot x_i \quad (3.62)$$

$${}^M X_j = {}^M H_j \cdot x_j \quad (3.63)$$

where M represents the mosaic frame, and x_i and x_j are points on the image frame and X_i and X_j are points on the mosaic coordinate system.

The cost function E that minimizes 3.62 and 3.63 can be written as Eq. 3.64, where the sum is calculated over all the pairs of overlapping images I_i and I_j for all the correspondences (n) between them.

$$E = \sum_{i,j} \sum_{k=1}^n (\|{}^i x_k - {}^i H_j \cdot {}^j x_k\|_2 + \|{}^j x_k - {}^i H_j^{-1} \cdot {}^i x_k\|_2) \quad (3.64)$$

For all the image pairs, the function to minimize can be written as Eq. 3.65, where m is the number of image pairs.

$$\min \sum_{k=1}^m E_k^2 \quad (3.65)$$

This approach usually generates a good solution on a Levenberg and Marquardt algorithm and few iterations are needed to get a global minimum.

3.1.7.3 Four Point Transformation

Even when the mosaics are optimized, some problems can still appear. These problems include the scale of the mosaic and image deformations. The reason is that some mosaics do not include loops that allow the refinement of the image correspondences. We added four markers to the scene; two that can be seen in the first image and two other seen in the last image of the sequence. The distance between these points is known in proportion to the algorithm, which

3. A NOVEL SYSTEM FOR SKIN MOSAICING

at the same time detects the points in the acquired images. The four extracted image points are correlated with the 4 reference markers, computing a projective homograph between them, then, the initial mosaic is transformed with respect to this homograph. In this way, the final mosaic has 4 imposed points that correct the size and deformations of the images. Finally, the absolute homograph of each image is corrected and the final mosaic is built.

This approach preserves the local alignment between consecutive images and globally reduces the distortion effect on the image size.

3.1.8 Mosaic Blending

Image blending is the last step to building a mosaic. Blending tries to minimize the visibility of the seams between images in order to achieve a verisimilar appearance of the created mosaic.

Two main approaches to image blending are described in the literature [99]: *transition smoothing* [28] and *optimal seam finding* [45]. The first group of methods tries to minimize the visibility of seams by smoothing the common overlapping regions of stitched images. Optimal seam finding methods place the seam between two images where intensity differences in their common overlapping area are minimal. Furthermore, it is also possible to find techniques that take advantage of the benefits of both approaches [2], in what can be called *combined techniques*. Fig. 3.18 shows an example of image blending.



Figure 3.17: Serratus mountain panorama built using *AutoStitchTM* software developed by M. Brown [27].



Figure 3.18: Blending mosaic corresponding to the mosaic in Figure 3.17 built using *AutoStitchTM* software developed by M. Brown [27].

3.1.8.1 Transition Smoothing Techniques

Transition smoothing is performed in the image intensity and image gradient domains. In this procedure, the images are first broken down into a set of band-pass filtered component images and joined using a weighted average within the transition zone, which is proportional in size to the wavelengths represented in the band. Gradient domain methods reduce the inconsistencies due to illumination changes and variations in the photometric response of the cameras since differences between the gradients in the overlapping regions are invariant to the average image intensity.

3.1.8.2 Optimal Seam Finding Techniques

In contrast to transition smoothing techniques, the optimal seam finding approach uses only image intensity information. Davis [42] described an image blending that computes the relative photometric differences between two images and searches for the dividing boundary along the lowest intensity of the difference image Dijkstra's algorithm.

Based on a similar idea, Efros *et al.* [45] proposed a texture synthesis method that selects the texture blocks to be cut from a given texture sample that agrees with its neighbors and finds a dividing boundary that minimizes intensity inconsistencies.

Other authors, such as Uyttendaele *et al.*, [175] proposed a method based on a set theory that tries to determine the regions of the images belonging to any

3. A NOVEL SYSTEM FOR SKIN MOSAICING

moving object in order to select the images that have to contribute information, rendering the object with the best possible appearance.

The approach of Gracias *et al.* [56] exploited the watershed transform to reduce the complexity of a graph cuts computation when searching for the best placement of the seam between images.

3.1.8.3 Combined Techniques

Other approaches, such as the one proposed by Agarwala *et al.*[2], may be classified as *transition smoothing* and *optimal seam blending* because they use both ideas to perform blending. Firstly, these techniques use graph cuts to find the contribution regions among several images in which each pixel is treated independently, and secondly, a gradient domain fusion is carried out in order to refine and minimize the visibility of the seams in overlapping regions between images.

3.2 A proposal to Skin Image Mosaicing

This section presents a proposal focused on skin mosaicing for skin cancer detection, although it can work on any other kind of image, giving further applications.

The designed algorithm follows the steps described in Algorithm 3. Steps 1 to 4 are applied to all the image sequences with the goal of registering each image in the mosaic. Once the mosaic is created, two further steps can be performed. On one hand, if the 4 markers are found, 4-point warping is applied. On the other hand, if no markers are present but non-consecutive overlaps exist, a global alignment of the image sequence can be performed. Finally, a blending algorithm can be applied for correcting non-uniform lighting.

3.2.1 Marker Detection

This step performs a detection on each image in order to detect a pre-defined marker. These markers are used with the goal of:

1. defining the initial and final image sequence.
2. optimizing the mosaic by applying the 4-point algorithm.

Algorithm 3 Mosaicing Pipeline.

1. Extract and describe keypoints.
 2. Initial matching.
 3. Compute homographies.
 4. Global registration and change on coordinate frame.
 5. Find markers.
 6. **if (Markers present)**
 7. Compute 4-point warping.
 8. **else**
 9. **if (non-consecutive overlaps exists)**
 10. Compute global alignment.
 11. **end if**
 12. **end if**
 13. Mosaic Blending
-

3. determining the relation between pixels in the image and meters in the scene (fixing the zooming factor).

With the use of markers, see Fig. 3.19 (a), it is possible to detect different sequences in the same image set. The first and second images in which two markers are present delimit a sequence. Then, the next steps for computing the final mosaic will be applied for each detected sequence.

The four markers are divided into two pairs. Each pair of markers is composed of two circles separated by 50 cm. One pair will be located near the patient's head and the other pair near to the feet. The circles are formed by alternating quarter black and white quarter-circles, creating an easily distinguishable corner that can be detected using Harris [70]. Harris corners are used since they are very salient.

3. A NOVEL SYSTEM FOR SKIN MOSAICING

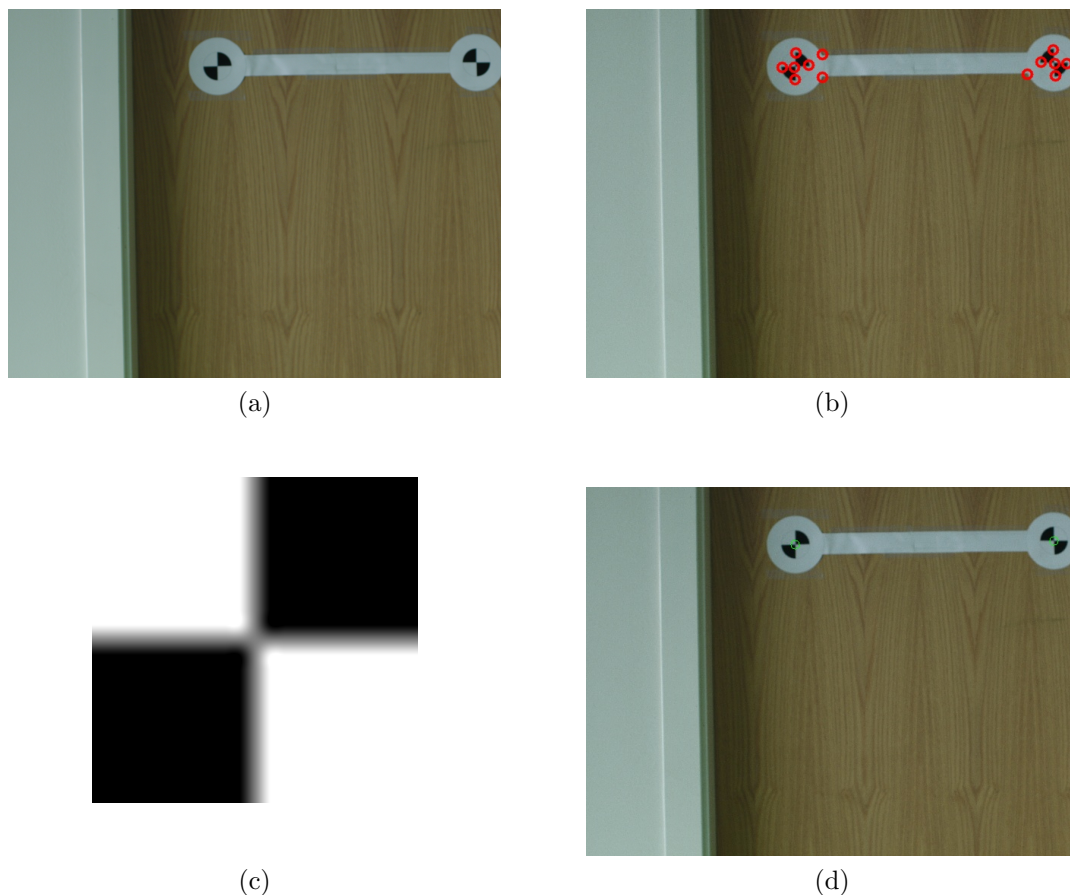


Figure 3.19: Finding Markers: (a) image of a pair of markers. (b) Harris detected keypoints. (c) patch to correlate. (d) Central point for each marker.

Once a set of around 10 to 20 keypoints are detected (see Fig. 3.19 (b)), they are matched applying normalized cross correlation (details at Section 3.1.4.1). The correlation is performed between selected patch centered on the keypoint and a patch for the left marker, Fig. 3.19 (c) and their negatives for the right marker. The corner that gives the best correlation will be selected as the one with the highest probability of being the central point of one of the markers. If only one marker is detected and not its pair, the markers in the image are not used. Fig. 3.19 (c) shows the 2 detected corners in an initial image sequence.

The second reason for using markers is to apply 4-point warping of the algorithm (see Section 3.1.7.3). Its application will be explained in Section 3.2.6.

The final advantage of knowing the position of the markers is to establish a relation between pixels in the image and meters. With this information, it is possible to add more information to the final mosaic. The user will be able to guarantee the size of the nevus allowing him to self-monitor its evolution.

3.2.2 Extraction and description of keypoints.

For each input image, interest points are extracted using SIFT or SURF depending on the detector and descriptor chosen. It is also possible to use a GPU implementation of SIFT [188] that will decrease the computation time with similar results as the non-GPU implementation. Detected points can be different due to the differences on the algorithm implementation. The designed software uses a GPU implementation by default, because it is less demanding in terms of memory and CPU. The next sub-section describes the advantages of GPU. If the computer is not able to use GPU, the program will use SURF or SIFT on the CPU. The final computation time will increase, but the results will not be significantly different.

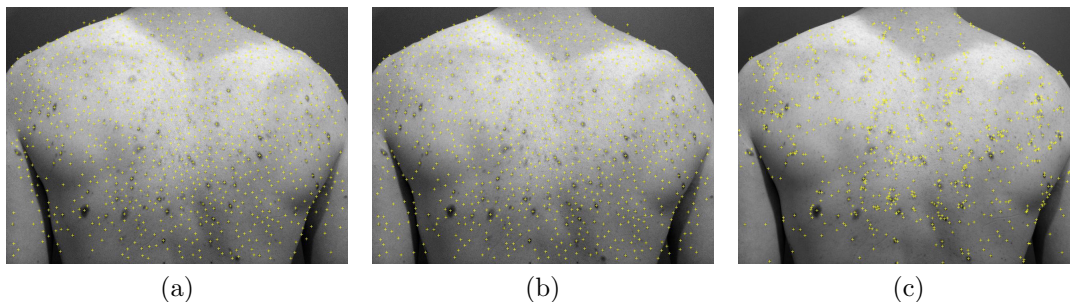


Figure 3.20: Extracted keypoints: (a) Using SURF. (b) Using SIFT. (c) Using GPU-SIFT.

For more details on computing the descriptor for each keypoint see Section 3.1.4.

3. A NOVEL SYSTEM FOR SKIN MOSAICING

3.2.2.1 GPU-SIFT

GPU-based SIFT was designed by Sinha *et al.* [158]. It works on NVIDIA cards and is able to extract about 800 features from an 640 x 480 video at 10Hz, which is approximately 10 times faster than an optimized CPU implementation of SIFT. This improvement is generated by the graphics hardware that contains powerful co-processors (GPUs) with a peak performance of hundreds of GFLOPS, one order of magnitude higher than CPUs. They are designed to independently process streams of vertexes and fragments (pixels) in parallel. However, their data parallel SIMD (Single Instruction Multiple Data) architecture also provides an abstraction for performing general purpose computations on GPUs (GPGPU) and for treating the GPU as a stream processor.

For the implementation of an algorithm on the GPU, some streams are usually needed. For each of these computational steps, the appropriate fragment program is bound to the fragment processor and the render operation is invoked. The rasterization engine generates a stream of fragments and provides a fast way of interpolating numbers in graphics hardware.

To decrease the computational time of the SIFT algorithm, some steps are computed in parallel on the GPU [156]. In fact, the Gaussian scale space pyramid is accelerated on the GPU using fragment programs for separable Gaussian convolution. The values of intensity and DoG are stored in a RGBA texture and computed at the same pass using vector operations in fragment programs. To find the local maxima and minima in the DoG pyramid, blending operations are performed in parallel at each pixel location. Depth and Alpha tests are used to threshold these keypoints. The final results of the keypoint extraction (location and scale) are stored in the RGBA data, which is read back to the CPU and decoded there.

Since translating the gradient pyramid from the texture memory to the CPU is expensive, the next steps performed are also in the GPU. At this point, to compute the orientation of each keypoint, gradient vectors near the keypoint location are Gaussian weighted and accumulated inside an orientation histogram by another fragment program. The orientation histogram is read back to the CPU where its peaks are detected (histogram computation is more expensive on

the GPU, for this reason the data is translated to the CPU).

The final step involves the computation of a 128-element descriptor of SIFT [157]. As this part is not efficiently computable by the GPU, it is divided between the CPU and the GPU. First, each feature's gradient vector patch is weighted by a Gaussian mask using the GPU. The re-sampled and weighted gradient vectors are collected into a tiled texture block which is subsequently transferred back for the CPU and used for compute the descriptor.

Summarizing GPU-SIFT provides a large speed-up in the Gaussian scale-space pyramid construction and keypoint localization steps. The compressed read back of binary images containing feature positions reduces the read back data-size by a factor of 32. The feature orientation and descriptors computation is partitioned between the CPU and GPU in a way that minimizes data transfer. Overall, a 8-10 times speed-up is observed compared to CPU versions.

3.2.3 Initial matching

As discussed in Section 3.1.5, to compute the initial matching between two data sets, it is necessary to compute the nearest neighbor for each point using their descriptor. The distance from each point in the first data set to all the points in the other is computed, and the closest one is selected as the matching point. Another possibility is to take into account the first and second neighbors and compute the ratio between them. Taking into account that the number of points at each data set could be large (more than one thousand points), and the descriptors of these points could be 80 or 128 entries, the library ANN [123], implemented by Mount, will be used. We selected this library because of its advantage of being able to specify a maximum approximation error bound, thus allowing the user to control the trade-off between accuracy and running time.

Efficiently finding the nearest neighbor or the k nearest neighbors, especially in higher dimensions, is a very difficult problem. It is always possible to solve this problem with a simple brute-force process of computing the distances between the query point and each of the data points. However, this may be too slow for many applications that require a large number of queries to be answered in the same data set. Instead of using brute-force, ANN reprocesses a set of data points

3. A NOVEL SYSTEM FOR SKIN MOSAICING

into a data structure from which nearest neighbor queries are then answered. There are a number of data structures that have been proposed for solving this problem [17] [8].

Extracting the exact nearest neighbor has the problem that the running time (or space) grows exponentially as function of data dimensions and usually, these methods are not any better than brute-force on small datasets. If the user is willing to tolerate a small amount of error in the search (returning a point that may not be the nearest neighbor, but is not significantly far from it) then it is possible to achieve significant improvements in running time.

3.2.4 Homography computation

For this proposal, four homograph models are implemented: Euclidean, Similarity, Affine and Projective (moreover, optimization of this last method is implemented and called non-linear projective).

The initial guess of the Homography is obtained using RANSAC [50], based on a *Matlab* implementation provided by [92] where the number of iterations needed to obtain a good result is recalculated at each iteration, taking into account the ratio between inliers and outliers of the best homograph found at the last iteration.

3.2.5 Global registration and coordinate frame change

At this step, the software first computes the absolute homograph, for each image. To compute the absolute homograph one image is chosen as the *reference frame*. In our case, the first frame of the sequence is chosen as reference and the absolute homograph of each image is iteratively computed by multiplying its relative homograph by the absolute homograph of the previous image. By doing this, all the images will be related to the *reference frame*.

By changing the coordinate frame we obtain all the points with positive coordinates in the final mosaic. Top left, top right, bottom left and bottom right have to be known at this stage. Once this information is known, the mosaic is shifted in order to have each corner on a positive coordinate value. Also, these

Element	Attribute	Description
<MOSAIC>		Document root node.
	ver	GMML format version.
	subver	GMML subformat version.
<INIT>		Initialization data of mosaicing system.
	os	Operating System where the mosaic is computed.
	cpu	Processor used to compute the mosaic.
<COMMENT>		Author's comments about the mosaic.
<CONFIGURATION>		Configuration used to compute the mosaic.
	src	Path where configuration file is stored.
<NAV_DATA>		Navigation File information (if exists).
	src	Path where navigation file is stored.
<GEO_DATA>		Geographic localization of the mosaic (if provided).
<MOSAIC_ORIGIN>		Mosaic UTM coordinates.
	x	<i>x</i> coordinate of the mosaic origin.
	y	<i>y</i> coordinate of the mosaic origin.
	z	<i>z</i> coordinate of the mosaic origin.
<PIXEL_SIZE>		Relation between pixel and meters.
	x	Width, in meters, of one pixel.
	y	Height, in meters, of one pixel.
<MOSAIC_SIZE>		Size of the mosaic in pixels.
	x	Width, in meters, of the final mosaic.
	y	Height, in meters, of the final mosaic.
<NODE>		Description data of one mosaic node.
	Time	Acquisition time, for the actual image.
	Index	Order of the image in the sequence.
<IMAGE>		Used Image.
	src	Path where the image is stored.
<HOMO>		Homography that relates the transformation between two images.
	model	Homography model used.
	type	Absolute or relative homograph.
<VALUES>		Matrix containing the homograph values.
<COVARIANCES>		Matrix covariances.
<POINT_MATCH>		List of point and matches.
	src	Path where point-match file is stored.
<POSE>		Position, in meters, of the camera.
	type	Absolute or relative position.
<EDGES>		Set of image overlaps.
<EDGE>		Relation between one frame and the others.
	node	Index of the related node.

Table 3.2: GMML 3.0a specifications: The elements and attributes of GMML 3.0a format.

3. A NOVEL SYSTEM FOR SKIN MOSAICING

steps provide information about the size of the mosaic. All these data will be used in order to render the mosaic.

Finally, all the information retrieved from the last steps will be stored in a GML (Girona Mosaic Markup Language)[93], a format created by the *Underwater Vision Lab* is defined based on an XML (eXtensible Markup Language) and is used to store all the related data generated in the mosaic construction in an ordered and coherent way. Also, the GML format is exportable to other programs like MosaicViewer (also developed by *VICOROB*). Version 3.0a of GML specifications are defined in Table 3.2.

3.2.6 Compute 4-point warping

With the four markers detected, see Section 3.2.1, it is possible to apply a transformation to the mosaic as defined in Section 3.1.7.3. Using this technique, the final mosaic will be corrected in orientation and scale, by imposing a predefined distance between markers.

The four detected markers (see Fig. 3.21) will be multiplied by the absolute homograph of their image. With this operation, the position of the four markers will be known with respect to the mosaic frame. Then, a projective homography representing the transformation between them and the real position of the markers is computed. After that, each absolute homograph will be multiplied by this homograph, and each image will be moved to the correct location. Table 3.3 shows how the 4 detected markers are moved to their correct place. The position of the final position is in pixels, and is equal to the distance introduced by the user multiplied by the relation pixel/meters.

3.2.7 Global alignment

When one region is acquired more than once by non-consecutive photographs, a closing loop may be detected. This allows the software to update the topology of the mosaic, and therefore the position of each image. From the estimated absolute homograph, it is possible to extract the x and y position of each image center. Then, taking into account the projection on the mosaic of each image, it

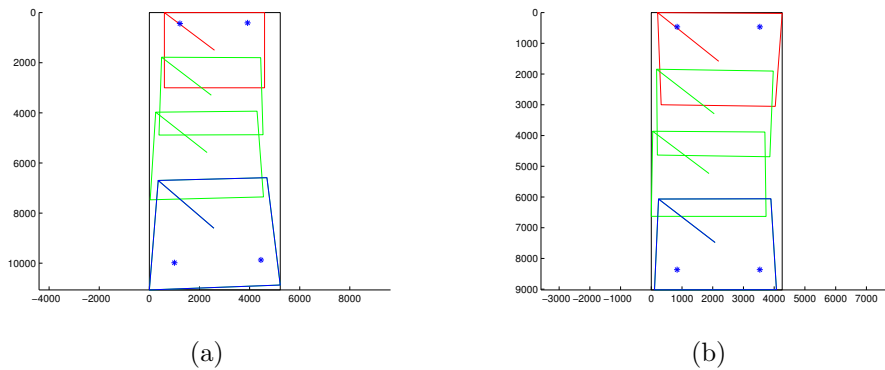


Figure 3.21: 4-point warping: (a) Mosaic preview from a non-warped mosaic, blue stars show the centers of the markers. (b) Mosaic preview from the transformed mosaic.

Marker Position	Non Optimized		4-Point Optimized	
	x	y	x	y
top-left	1219	436	841	464
top-right	3919	412	3532	464
bottom-left	999	9983	841	8366
bottom-right	4454	9871	3532	8366

Table 3.3: 4-point warping: Pixel position of markers before and after the 4-point transformation.

is possible to estimate which images could potentially overlap, see Fig. 3.22 (a) and (b).

After that, the homograph between these image pairs is computed and their correspondences are used to constrain the optimization. This step will try to minimize the error contained in a set of matching points and the error contained on the position of the camera at the moment of acquiring the images (x , y , z , yaw, roll and pitch). Once the camera positions are re-estimated, the absolute homographs are recalculated and stored in a new GMMML (see Fig. 3.22 (c)).

3. A NOVEL SYSTEM FOR SKIN MOSAICING

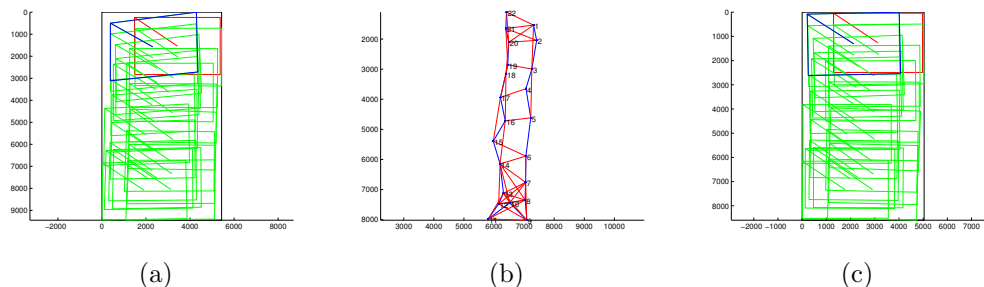


Figure 3.22: Global Alignment: (a) Preview of a non optimized mosaic. (b) Consecutive and non-consecutive edges from last mosaic. (c) Optimized mosaic.

3.2.8 Mosaic Blending

Mosaic blending is the last step of the mosaicing pipeline and will correct the photo-metric errors appearing in the final mosaic.

The first step is rendering the mosaic, where the value of each mosaic pixel is computed by searching the images containing this point. This step can be done using various techniques:

1. The simplest one takes into account only the first or the last image that contributes to each position.
2. More sophisticated techniques choose the closest image to each pixel. For every pixel on the mosaic the distance between this point and the center of the contributing images is computed. The image with an smaller distance is used and the intensity value of the pixel image is applied on the mosaic.

The second technique was implemented in the software as it provides the best results. Fig. 3.23 (a) shows the contributing image at each pixel, called closest map. Figure 3.23 (b) shows the intensity value of the closest map. This second image is called the closest mosaic.

From the closest mosaic, x and y gradient images are formed and the seams between the images are forced to be equal to zero. Then, the data is fitted into a Poisson-Neumann [48][136] solver in order to recover the intensity image of each channel.

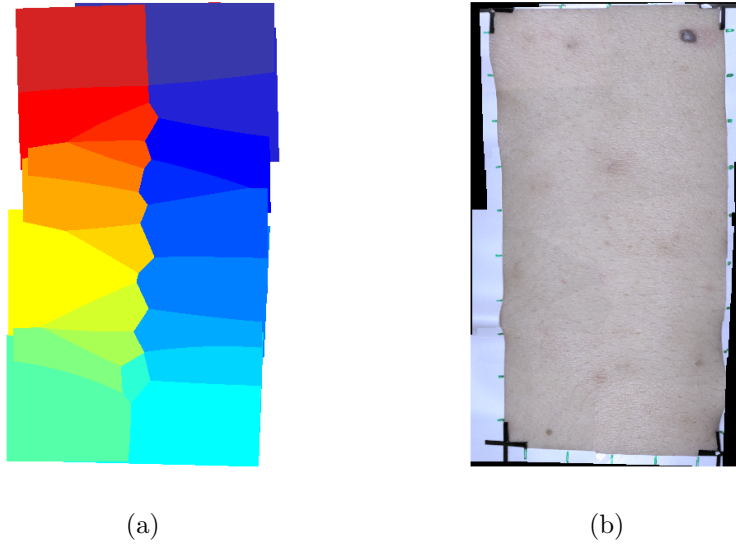


Figure 3.23: Mosaic Blending: (a) Closest Map. (b) Closest Mosaic.

3.3 Results

This section illustrates some of the results obtained from applying the presented mosaicing software to different data sets. Some of these datasets are related to whole body maps, but others are completely independent, showing the different possibilities provided by the software. Finally, some problems that can appear in the registration of images are also discussed.

3.3.1 Test Data Set Description and Results

The image data sets used in the tests are shown below. The particularities of their stitching and the properties that want to be highlighted in each case will also be exposed.

Data Set 1

The following images show the building process of a mosaic with a rigid object. The movement of the camera between the 4 images, can be described by a z translation, and a rotation on the axis defined by the image plane and the z axis. The images are acquired with a *Canon A 650 IS*.

3. A NOVEL SYSTEM FOR SKIN MOSAICING

Fig. 3.24(a) shows the results of the mosaic computed by only applying consecutive homographs. This mosaic presents good registration between images, but the shape of the full mosaic is not correct. This problem is due to the motion described by the camera. It is possible to see that the top two markers have a different horizontal distance than the bottom ones. This problem can be corrected by applying a four point transformation.

The four detected markers are translated to their correspondent positions introduced by the user on the interface. By doing this, the vertical and horizontal distances between the markers are corrected, and also any supposing that the markers are oriented in the same way a it is possible to correct rotational problems. Fig. 3.24(b) shows the results after 4-point wrapping.

Finally, image blending is used in order to correct the illumination change between images. Fig. 3.24(c) shows the final result were the seams between images are not visible.

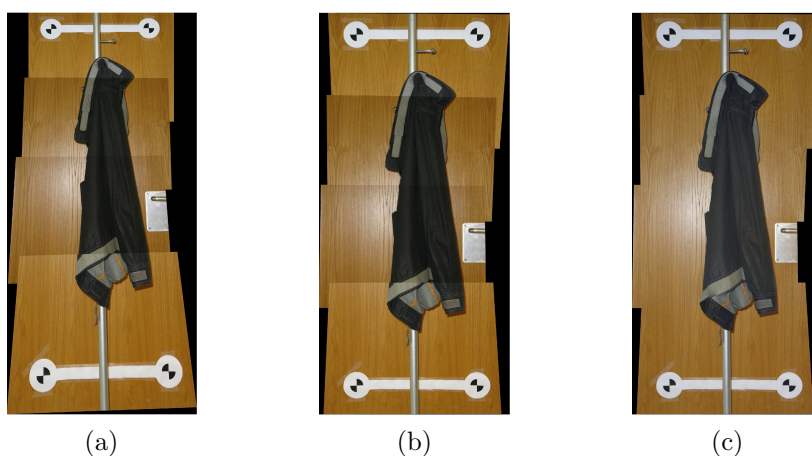
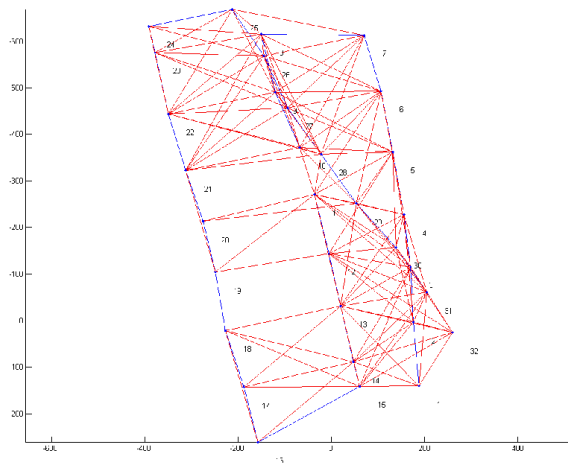


Figure 3.24: Rigid Object Mosaic: (a) Not optimized, (b) Optimized and (c) blended.

Data Set 2

These images correspond to a 10 x 5 cm. mosaic of the back of a patient. The acquisition procedure is done by adding a led ring to a Canon EOS 40D reflex camera. The final mosaic is composed of 22 images. The sequence 3.25 (a) is acquired by first taking one vertical strip from the top to the bottom and then

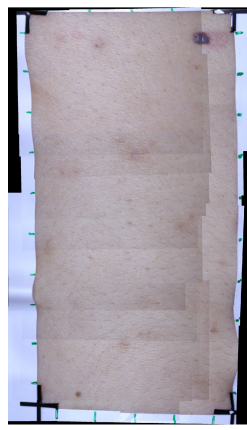
the second strip is acquired from the bottom to the top. This acquisition was done without undergoing any rotation.



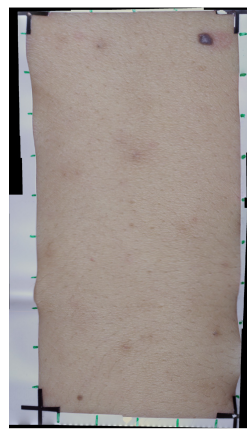
(a)



(b)



(c)



(d)

Figure 3.25: Back Mosaic: (a) Topological Map, (b) Non-optimized, (c) Optimized and (d) Blended.

The mosaic construction seems to be easily solved by applying an Euclidean transformation between images, but due to changes in the relative depth between

3. A NOVEL SYSTEM FOR SKIN MOSAICING

the camera and the patient, and some small rotations caused by small movements of the camera, a projective homograph is needed.

The mosaic obtained from computing the homographs from the consecutive images is shown in Fig. 3.25(a). From this result, it is possible to see some misalignments due to the accumulative error. This error is more evident when comparing the first and last images of the sequence, in which the same zone is mapped into different pixels. As this sequence is closing a loop, it is possible to apply a global alignment. Fig. 3.25(b) shows the results after optimization, with the result of a reduced misalignment.

Finally, the mosaic is blended correcting in all the photometric differences produced by the illumination changes produced in the image acquisitions. Fig. 3.25(b) shows the blended mosaic.

Data Set 3

Sequence number 3 is provided by the *Melanoma Unit at the Hospital Clinic of Barcelona*. It shows the back of a patient with a high density of nevi. It is formed by of only 3 images. Initial mosaicing results are shown in Fig. 3.26(a). In this sequence, it is not possible to compute a global optimization of a 4-point transformation, since it does not use markers and there are not any loops in the sequence.

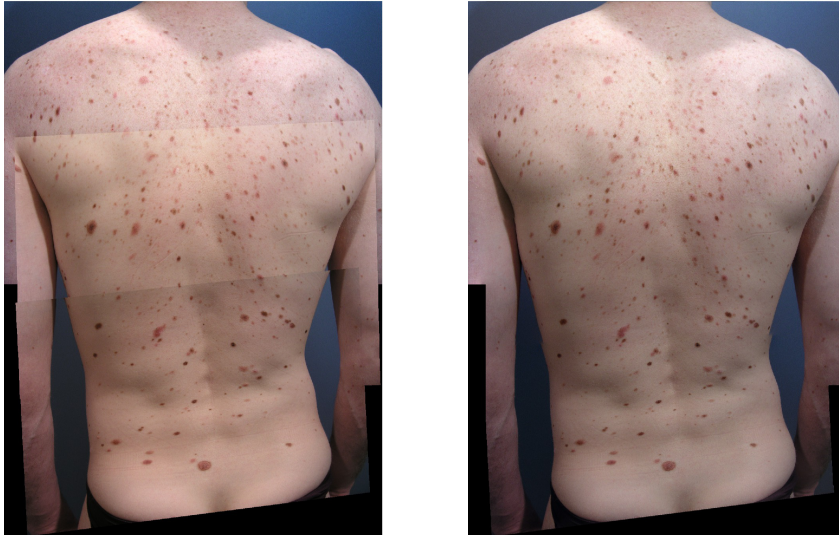
The initial mosaic is blended minimizing the small differences between the images caused by some illumination changes. Fig.3.26(b) shows the blended mosaic.

Data Set 4

The last data set is formed of 4 different sequences, describing the front, sides and back of a patient. The first three mosaics contain the 4 markers, but the last one is cut in the middle. For this reason, it is only possible to apply a 4 point transformation in the first 3 mosaics.

Each mosaic is blended reducing the photometric errors. In this way, it is possible to correct a highly noticeable intensity difference between the first and last images of each sequence.

Some errors were found once these sequences were registered were registered. Further details are provided in the next section.



(a)

(b)

Figure 3.26: Patient A mosaic: (a) No optimized and (b) blended.

3.4 Registration Problems

The last mosaics show some problems in registration for two different reasons:

- The non-rigidness of the body provokes problems in different parts of the mosaic, for example, the arms and other parts of the body that move due to breathing. Fig. 3.28 (a).
- Parallax problems occur from the translations from the camera. Where 3D parts of the scene could not be registered by a 2D homograph. Fig. 3.28 (b).

3. A NOVEL SYSTEM FOR SKIN MOSAICING

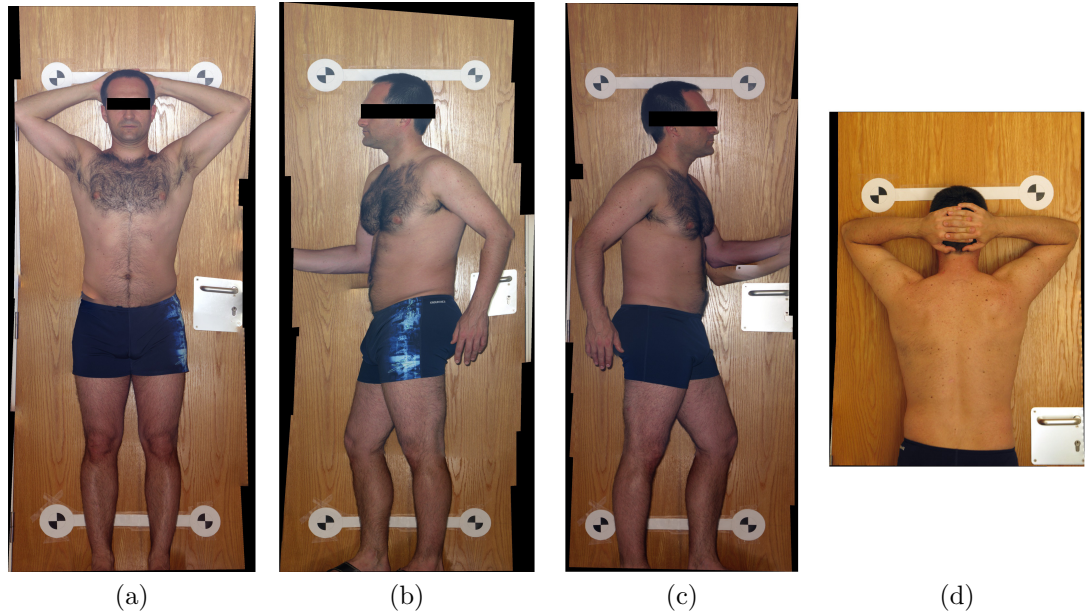


Figure 3.27: Mosaic of B patient: (a) Frontal mosaic (b)(c) lateral mosaics and (d) rear mosaic



Figure 3.28: Registration Problems: (a) Non-rigidity: breathing movement. (b) Parallax

The first problem can be solved or at least minimized by applying non-rigid transformation techniques [20]. These techniques involve elastic registration to cope with the deformation of the subject (due to breathing, anatomical changes, etc.). These deformations can be produced between images or between different explorations that need to be compared.

The homography matrix is only able to describe planar transformations, implying that the images involved have to belong to a planar scene. This constraint is too restrictive for a wide range of possible mosaicing applications as shown in Fig. 3.28 (b). The regions of the images belonging to an object that presents parallax [89] may be ignored in the motion estimation procedure only if they are not significant with respect to the size of the images. Nevertheless, misalignment will still exist. For this reason it will be necessary to determine which image will contribute to the mosaic with the information relative to the conflict object.

Registration is also possible when *moving objects* are present in the scene if the regions that belong to the moving objects are not considered in the estimation of the motion. Nevertheless, when registration into a mosaic is performed under these conditions, moving objects may appear deformed in the overlapping regions. Specifically, the objects can be cut in the borders of the images and do not correctly match in the seams between consecutive images. In the worst case, the objects may appear duplicated in more than one location, if the movement has placed it outside the overlapping region between the images involved.

3.4.1 Computational Cost

Table 3.4 shows the computational time of each sequence at a common workstation (1.80GHz Intel[®] Pentium[®] Dual CPU E2160). The graphic card used in order to compute the GPU-SIFT is a NVIDIA[®] GeForce 7300 GT.

3.5 Conclusions

We have presented a tool for the medical community, specially addressed to patients with skin cancer, in order to enable doctors to create optical maps of the whole body. These maps will allow them to monitor unusual changes in the skin such as a sore that does not heal, a nevus that has grown or changed color or the detection of new nevi.

Firstly, the basic concepts required to proceed with a deep study of skin mosaicing have been explained. These techniques focus on the detection and description of interest points, matching of images, homograph computation, mosaic

3. A NOVEL SYSTEM FOR SKIN MOSAICING

Sequence number:	1	2	3	4(a)	4(b)	4(c)	4(d)
Marker Detection	14"			8"	7"	8"	
Keypoint	36"	2' 25"	12"	2' 16"	1' 59"	1' 42"	26"
Homogrpahy	40"	55"	5"	1' 44"	1' 31"	1' 18"	8"
Global Alignment		3' 30"					
4 Point Warping	10 "			5"	5"	5"	
Blending	1' 10 "	4' 00 "	30"	1' 40"	1' 35"	1' 37"	26"
Final time	2' 50 "	10' 50 "	47"	5' 53"	5' 17"	4' 50"	1' 00"

Table 3.4: Computational Time: Time required for each of the 4 sequences.

optimization and image blending.

The photomosaicing system is fully designed with respect to the knowledge acquired in the study of the main steps of the mosaicing procedure.

Also, a graphical user interface is designed to enable the interaction between the dermatologist or the patient and the image processing software, see Fig 3.29.

Finally, some results from different image sequences have been presented, showing the performance of these algorithms. Some errors produced due to poor registration have been presented.

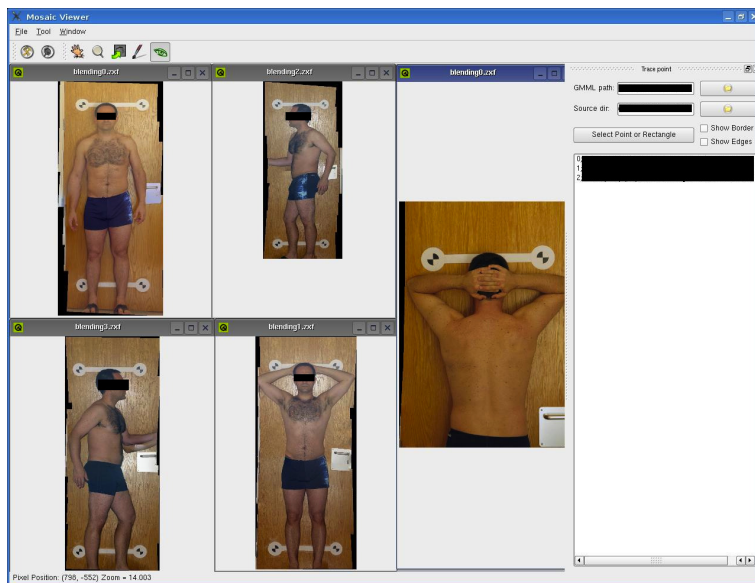


Figure 3.29: Mosaic Viewer interface.

3. A NOVEL SYSTEM FOR SKIN MOSAICING

Chapter 4

Design and development of a new 3D Body Scanner

This chapter describes the current technology for 3D body scanning and its applications to skin cancer detection. It also presents two new devices specifically designed for skin cancer detection.

Although dermoscopy is nowadays a well-established practice for dermatologists in order to diagnose melanocytic tumors, it involves a long time-consuming examination for each patient (30-50 minutes), and images from different explorations have to be compared every 3-6-12 months in order to detect changes in already existing nevi or to detect new nevi. The utilization of this technique has demonstrated its utility in reducing the number of biopsies and allowing the early diagnosis of melanoma[177]. Unfortunately, only a small group of patients can benefit from this methodology due to its complexity and cost.

The main drawbacks of this approach are:

- It is a tedious, slow and costly process, since each exploration needs a significant amount of time of an expert.
- Visual inspection is prone to errors due to inattention on the part of the physician.
- A new nevus may not be detected without the aid of a previous image with enough resolution of the affected zone.

4. DESIGN AND DEVELOPMENT OF A NEW 3D BODY SCANNER

- Only a limited number of dermatoscopic images are acquired for each patient (from 10 to 200) but never for all the lesions on patients with a high number of nevi.

Dermatologic images can be stored in systems such as PhotoMAX, where a technician or a dermatologist links each image with the current location of the mole on the body by hand. Later on, dermatologists can visually compare images of the same lesion in different explorations, and decide whether a lesion is benign or malignant. In this process, three steps are needed: first, acquire the image; then, select where its correspondence lays on the body; and, finally, diagnose. All these steps are not automatized and for this reason, human intervention is always needed.

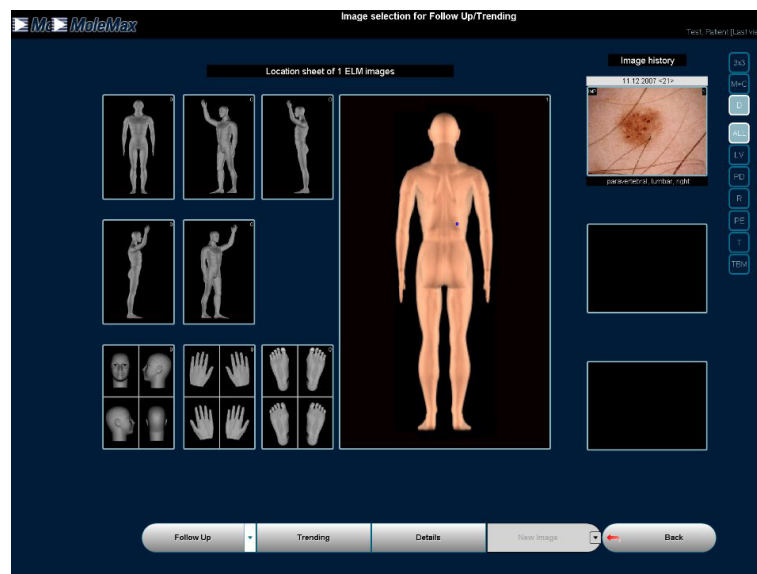


Figure 4.1: PhotoMAX software for managing patient images and database. Images are linked to an avatar by hand. From Derma Medical Systems (<http://www.dermamedicalsystems.com>)

This thesis proposes the development of an optical scanner, which allows the acquisition of dermatoscopic images of the whole skin surface of the patient, with the aim of building a high-resolution map in a fully automatic way. This photomosaic will be compared with photomosaics acquired in previous explorations,

allowing the detection of any type of suspicious change. In summary, we will initially solve the acquisition and positioning of each image in an automatic way, and only the diagnosis will remain manual.

Two scanner prototypes have been implemented on this thesis, namely horizontal and vertical scanners. In both cases, the scanner performs a full exploration of the patient, taking images of the whole cutaneous surface. Both systems use cross-polarized lighting.

The next section reviews the current state-of-the-art in 3D body scanning. This review will guide us in finding the best solution for implementing our proposed 3D scanner.

4.1 State-of-the-art in 3D Body Scanning

By means of 3D body scanning, the acquisition of images can be performed automatically. Recently, some new devices have appeared on the market. These systems have been used in different fields[4][105] such as the medical sciences, the textile industry, cosmetics, ergonomics, anthropology, fitness, security, animation, sculpture and communications, among others.

The acquisition technologies of these devices can be divided into laser scanning, coded structured light scanning, photogrammetry and others (see Fig. 4.2).

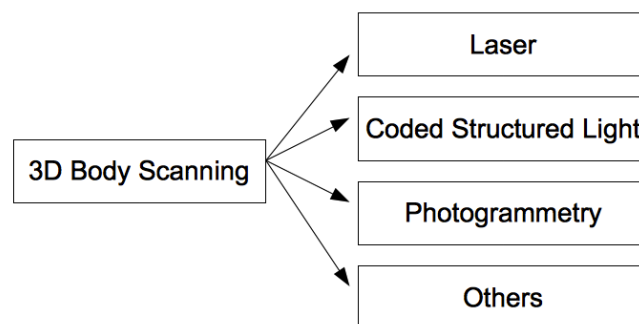


Figure 4.2: 3D Body Scanning Classification

4. DESIGN AND DEVELOPMENT OF A NEW 3D BODY SCANNER

4.1.1 Laser Scanning

Laser scanning systems[97] use one or more strips of lasers projected onto the body while a camera simultaneously acquiring images of the patient. The surface of the body is measured applying simple geometrical rules.

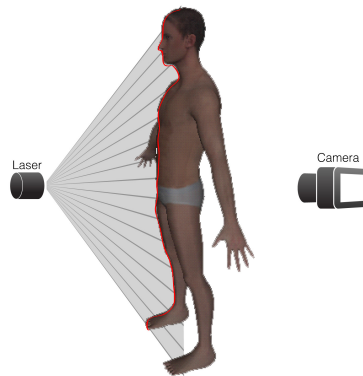


Figure 4.3: Schema of 3D body scanning based on laser scanning.

Some systems use special optical systems and mirrors in order to generate multiple strips from a single laser[51]. Two kind of systems are available: those that rotate the object[149], and those that rotate the optical system (light sensor and laser)[33].

Many different types of laser scanners have been proposed ranging from foot scanners such as the one proposed by I-Ware Laboratory[78], to head scanners, such as in the one proposed by Cyberware[39], or even full-body scanners such as Vitronic[178], Cyberware[39], Hamamatsu Photonics[68] and Human-Solutions[162].

The main disadvantage of these systems is the high cost of producing the required hardware. The correct calibration of these systems is crucial in order to reconstruct the object. The time needed to explore the entire body is a huge problem because the whole body must be kept immobile during the acquisition time. Uncontrollable movements such as breathing can induce errors. Lasers present a risk to the eye as well as to the body in skin cancer patients. Laser

scanning only provides a point cloud, representing the range of the 3D volume scanned; no information on texture or color is provided.

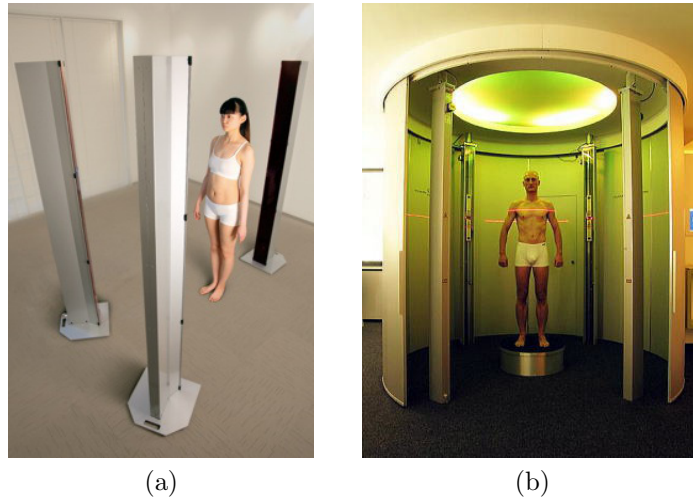


Figure 4.4: 3D body scanning based on laser scanning: (a) Cartesia[®] 3D Body Scanner (Spacevision) and (b) 3D Vitus[®] (Vitronic)

4.1.2 Coded Structured Light Scanners

Coded structured light scanners[133] are based on the projection of a light pattern (usually in the form of strips) onto the human body. A digital camera measures the strips on the body surface and, singularly, recovers their positions by using triangulation. Usually, binary coding systems are used to determine the origin of each strip. Also, to increase their resolution, the projected strips are shifted in space and time.

This system presents the advantage of reducing the digitalization time compared with laser systems. An entire body image can be acquired in a short period (seconds) so human body images can be digitalized without problems caused by movement.

If the area to be scanned is large, systems with multiple projectors and cameras are used.

Examples of this kind are the scanners developed by Beuckmann GmbH[19]

4. DESIGN AND DEVELOPMENT OF A NEW 3D BODY SCANNER

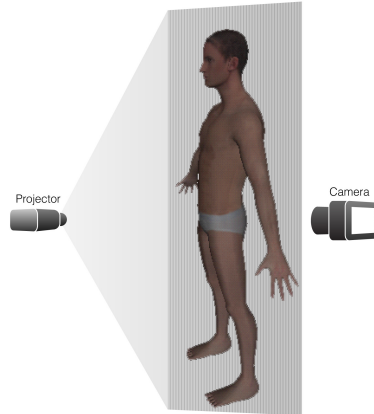


Figure 4.5: Scheme of 3D body scanning based on Coded Structured Light

and InSpeck[79] for head scans or Textile and Clothing Technology Corporation[171] or Telmat[170] and inSpeck[79] for a full body scan. These systems present good results in the recovery of the 3D volume. They are also robust on the occluded parts of the body. Their main drawback is the poor texture recovery due to the pattern projected on it. This problem is also present when segmentation methods are applied to the 3D volume due to the discontinuity of the texture and color through the volume.

4.1.3 Photogrammetry

Photogrammetry systems use one or multiple cameras for data acquisition. Once the images are acquired, computer vision and image processing techniques are employed to obtain 3D data from the images[144]. Three different systems are considered within photogrammetry: the multi-image photogrammetry, visual hull and silhouette analysis.

Multi-image photogrammetry uses multiple images from one or more calibrated cameras. Once the images are acquired, matching algorithms determine corresponding points in the images. By intersecting rays between camera positions, it is possible to compute the 3D points in the scene, creating a dense 3D point cloud.

The visual hull method[95] employs a set of images acquired from different



(a)



(b)



(c)



(d)

Figure 4.6: 3D body scanning based on coded structured light: (a) bodySCAN[®] (Breuckmann), (b) NX-16 3D Body Scanner ([TC]²_®), (c) Mephisto 3D Body Scanner (4DDynamics[®]) and (d) Symcad[®] (Telmat Industrie)

directions. The 3D reconstruction is performed using the volume intersection approach. The bounding geometry of the object scanned can be obtained by intersecting the cones formed by their projection onto the image planes and the focal points of the cameras. The 3D volume scanned is called a visual hull[96][34].

The silhouette analysis approach acquires two images of the human body; one from the front and another from the side. By using the body's symmetry, some relevant measurements with sufficient accuracy can be determined from the silhouettes[21].

Examples of this kind are found in the scanners developed by Human-Solutions[162] that use the silhouette analysis, 4D View Solutions[161] which employ a visual

4. DESIGN AND DEVELOPMENT OF A NEW 3D BODY SCANNER

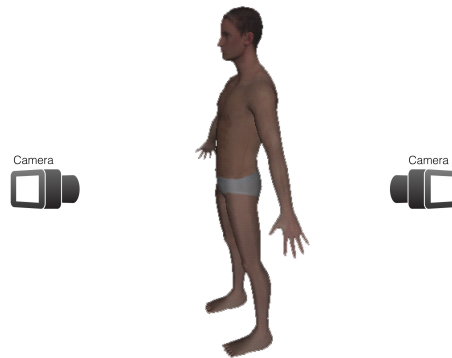


Figure 4.7: Scheme of 3D body scanning based on photogrammetry

hull method and Orscan[132] that commercialize a multi-image photogrammetry 3D scanner.



Figure 4.8: A 3D body scanner based on Photogrammetry: Skin-Vision360TM by Orscan Technologies

4.1.4 Other

Other 3D sensors use ultra-high frequency radio waves, time-of-light (TOF) cameras and contact measurements.

Ultra-high frequency radio waves allow a whole body scan while the person remains fully clothed. A transceiver illuminates the body with extremely low-powered millimeter waves; the radiation penetrates clothing and reflects off the body. The reflected signals are then collected by an array of transceivers and processed for measurement[72][57]. Intellfit Corporation commercializes ultra-high frequency radio wave scanners.

TOF cameras are based on an array of diodes that emit a near-infrared wave front that is intensity-modulated[131]. The light is reflected by the object and imaged by an optical lens onto the dedicated 3D-sensor. By measuring the phase delay between the emitted and received light, it is possible to measure the distances between the objects[94]. CSEM (Centre Suisse d'Electronique et de Microtechnique)[38] commercialize TOF cameras.

Contact measurement devices, such as Elinvision[46] and Amfit Inc.[3], provide digital measurements of the body, either with electronic tape measurements or with scanners that have a touch probe matrix.

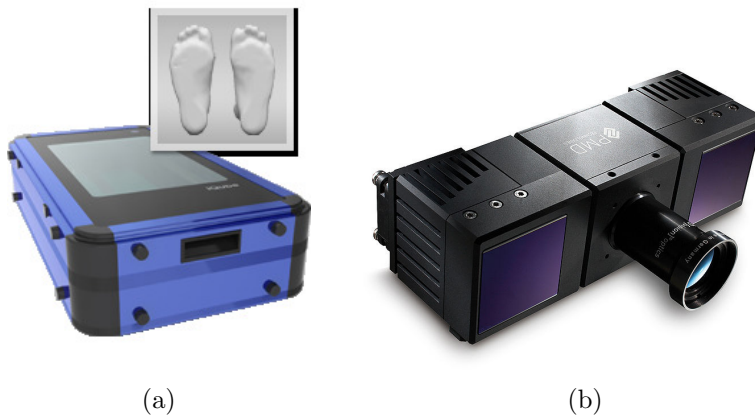


Figure 4.9: Other 3D body scanning systems: (a) Icube[®] (Elinvision) and (b) CamCube[®] (PMDTech)

4. DESIGN AND DEVELOPMENT OF A NEW 3D BODY SCANNER

4.1.5 Scanner Proposal

After reviewing the different systems, it is noticeable that most of them are not appropriate for medical applications such as skin cancer detection. Skin cancer detection first needs to have high-resolution images in order to detect lesions and classify them, either by a dermatologists or in an automatic way. Secondly, texture and color are needed for most diagnosis systems, characteristics not present in most of the 3D scanners presented. Above, we show the characteristics of each group of devices.

Laser scanning can scan the entire body, from head to foot, but presents problems in the high cost of the device, the need of a precise calibration to render 3D reproductions and a long acquisition time that can introduce errors due to small movements.

Coded Structured Light scanners need less time to perform the acquisition. Their main problem is the need of different devices (cameras and projectors) to cover the whole body. Due to the interference between the strips of the different projectors, it is not possible to shoot them simultaneously, but it should be done in a sequential way, increasing the final acquisition time.

Photogrammetry systems already on the market suffer from either not being able to cover the entire body (Multi-image photogrammetry) or not having enough resolution (visual hull and silhouette analysis). In general, they have low resolution due to the high computational cost of the 3D reconstruction algorithms.

Other sensors such as ultra-high frequency radio waves, present a perfect 3D scan but without any texture or color information. TOF cameras do not have high resolution or good color quality, but they provide 3D scans without any extra computation. Contact measurements only provide distances from different parts of the body for a posteriori to compute a 3D avatar.

All these systems have been mainly developed for the textile, ergonomy and fitness industries, where a clear reconstruction of the body is preferred over a high resolution of any specific part of it. Also, most of these systems will not provide more than an avatar or a set of measurements once they finish the acquisition. For these reasons, we decided to use photogrammetry with multiple cameras covering the full body in order to achieve good resolution of the body and a 3D

reproduction of the person as well.

4.2 Design and development of a cutaneous scanner

This section describes our two cutaneous scanners (see Figs. 4.11 and 4.16) developed for the acquisition of high quality images of the whole body. The two proposals were developed in order to cover different kinds of patients with different characteristics. The horizontal scanner is designed to be used with older people or people with difficulties in standing steady for some minutes. The advantage of the horizontal scanner is that it is easy to maintain the same position if the patient is lying on a table. The main disadvantage is the time required to perform a full exploration with the horizontal scanner. This time is double the time needed for the vertical scanner. Furthermore, the vertical scanner does not require the use of a dark room so that ambient light does not mix with the polarized light of the scanner. On the other hand, the cabin of the vertical scanner can cause claustrophobia in the patient.

4.2.1 Horizontal Scanner

The horizontal scanner was developed between 2008 and 2009 and is composed of a small bed with a mattress and an acquisition header. The patient lies on the table either on his/her back or on his/her stomach. The acquisition head is composed of CCD cameras and a lighting system. This head moves over the patient, acquiring images every 10 cm over the entire body. Once the scanner arrives at its last position, the patient turns over and a new data acquisition is performed to complete the exploration.

4.2.1.1 Mechanical Aspects

The scanner is constructed on the basis of modular aluminium profiles, which allowed design modifications in the prototyping phase of development. The final external dimensions of this scanner are 1700×2500 mm by 1400 mm in height

4. DESIGN AND DEVELOPMENT OF A NEW 3D BODY SCANNER

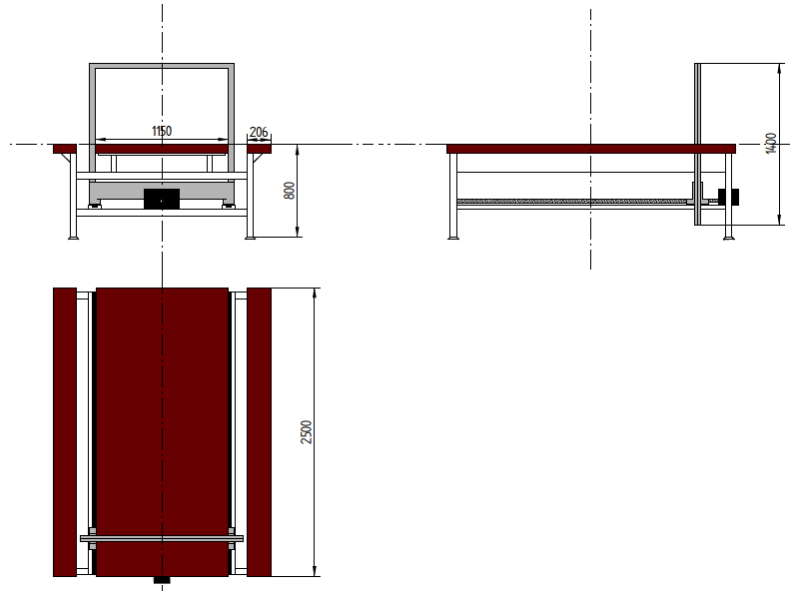


Figure 4.10: Horizontal Scanner Schema



Figure 4.11: Horizontal Scanner.

(see 4.10). The dimensions of the scanner in width and length are primarily defined by the size of the table for the patient and the height of the acquisition

head, which is placed as close as possible to the patient, yet still allowing the cameras to focus.

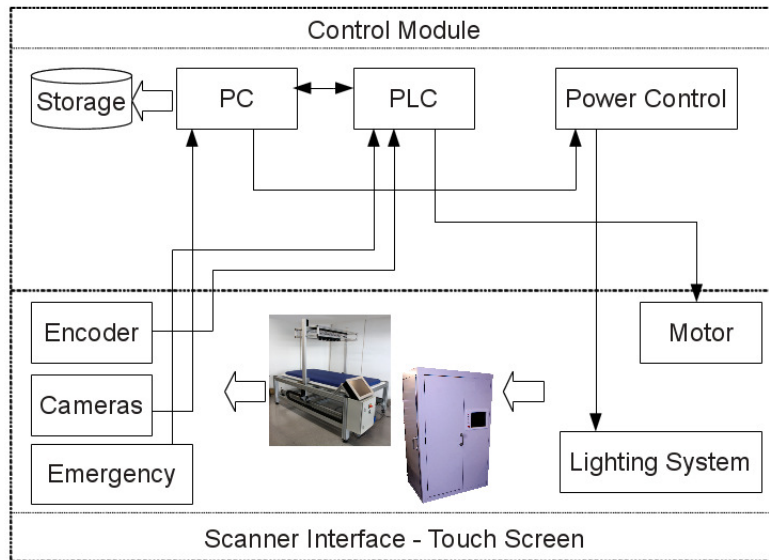


Figure 4.12: Bloc diagram of the proposed scanner.

4.2.1.2 Computer Module

The scanner uses an industrial panel PC that has a touchscreen as an input device, where a technician can operate the scanner. The PC has connections to the main parts of the scanner such as the Power Control motherboard, the Programmable Logic Controller (PLC) and the acquisition module.

The Power Control motherboard is connected to the scanner through a serial port, and is used to switch on/off the lighting system and to start/stop the acquisition device explained below.

The PLC module is in charge of controlling their actuators, which will be described in the next section. The PLC is connected to the computer module and uses a serial port for sending information. Its main functions are sending start/stop signals to the actuators and receiving the current position from the encoder. Also, other options such as configuring the velocity and acceleration of the motor are enabled.

4. DESIGN AND DEVELOPMENT OF A NEW 3D BODY SCANNER

The PC is linked to the acquisition module, connecting all the cameras and sending configuration and acquisition messages to them through a USB port.

4.2.1.3 Actuators

The actuator suite is based on a motor for moving the acquisition head. This motor has an encoder in order to control the position. The system is connected to the PLC in order to control them. The PLC used is a Danfoss VLT Automation Drive FC-300 which allows the use of three-phase current motors.

The PLC on the scanner communicates directly to the PC by sending serial messages that contain the position to be reached by the scanner.

4.2.1.4 Lighting System

The lighting system is designed to maximize the amount of light on the scene, while keeping it within comfortable levels for the patients. For this reason, 8 fluorescent tubes are used together with a customized reflector in order to direct all the light intensity towards the patient (see Fig. 4.13).

Two linear polarizers are used to generate cross-polarized images. The first is placed in front of the lighting system, while the second is placed in front of the cameras and rotated 90 degrees with respect to the first one. Following this method, it is possible to eliminate any reflections on the skin. The scanner should be used in a dark room to avoid interference from any other light source that may disturb the cross-polarized system.

Having a high quantity of light allows the cameras to use a smaller aperture, thus giving a better depth of field and image quality. Also, having more light reduces the exposition time, minimizing motion blur problems in the images due to small movements of the patient.

4.2.1.5 Image Acquisition Module

The scanner has 11 cameras with the goal of maximizing the overlap between images (see Fig. 4.14) and covering the entire body. All the cameras are Canon PowerShot G9 with 12 MegaPixels (4000×3000 pixels) on a sensor of $1/1,7''$



Figure 4.13: Detail image from the acquisition head, showing the acquisition system (11 cameras) and the lighting system (8 fluorescents, reflectant system and polarizer film).

and an optical zoom of $6\times$. We chose these cameras as they made the best quality/price compromise when we started the project. Also, as opposed to most industrial cameras, these cameras have the possibility of setting up different parameters such as focal point, zoom, etc, which will be needed to focus on different parts of the body. With cameras without auto-focus, the various distances from the camera to the body would not allow a single focus for all the images, requiring a smaller lense aperture and even more light.

Software The first prototype was developed under a Windows operating system and used the *Canon Software Development Kit (SDK)*. The interfaces were implemented in *qt*, providing a user-friendly communication with the scanner technician. Further work will allow the system to manage patients' databases. Also, an input/output data module will be implemented to send images to servers in order to post-process them.

The acquisition module allows as to configure most of the camera's parameters, such as the acquisition mode (Auto, Manual, Aperture Value and Time Value), the exposition time (seconds), aperture (F-value), ISO, focus point (central or multiple focusing points), zoom (\times factor), white balance (Daylight, Tungsten, etc), etc.

The acquisition system also allows us to synchronize image acquisition, trig-

4. DESIGN AND DEVELOPMENT OF A NEW 3D BODY SCANNER

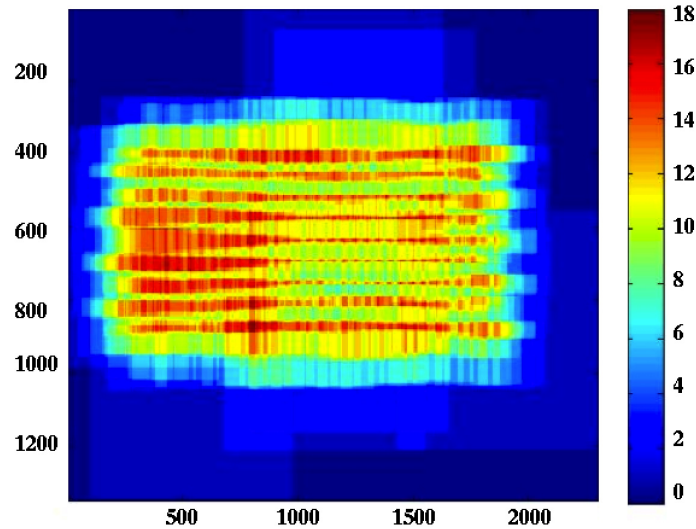


Figure 4.14: Overlapping images on the Horizontal Scanner: With a person lying on the table, the numbers shown on the right of the figure are the number of images that cover each position of the table, having at least 6 views of the same position. This will minimize the problems due to occlusions. Moreover, only the closer images will be used for mapping in order to select those that present the best focus in each position.

gering them simultaneously, and storing the images either on the SD cards or directly in the PC.

Control. Lights and power are controlled by a Programmable Interrupt Control (PIC) connected to the PC serial port. The objective of this module consists of sending start/stop messages to either the camera power or the acquisition light.

4.2.2 Vertical Scanner

The second prototype was conceived to reduce the size of the system and the acquisition time. Moreover, the patient's position was also modified (the patient stands up, as opposed to lying on a bed). In this version of the scanner, the patient is placed in the center of a rotating platform. The cameras and the lighting system

are placed on one side of the cabin, and the scanner acquires images while the patient rotates. Once half of the entire acquisition is done (after a rotation of 180 degrees), the patient changes his position to cover the remaining half of the body not yet covered in the first acquisition. Once the patient has changed position, the scanner acquires images of the patient's back.

This scanner has a dimension of 990×1400 and 2272 mm in height, approximately half the size of the horizontal one, allowing installation in smaller rooms (see 4.15). We also worked on the external aspect of the scanner making it more presentable in a medical ambient (see 4.16). The use of a cabin to cover the scanner allows the system to be used in any room without the need of switching off the ambient lights to prevent undesired effects on the cross-polarization configuration of the scanner's lighting system.

The computer module is approximately the same, but the industrial PC has been changed to a standard PC because of its lower cost. However, a PC allows us to add extra USB-PCI cards, thus reducing the need for hubs to connect each camera.

However, the cabin covering the scanner generates a new problem: the technician operating the scanner can not see what is happening during the acquisition. For this reason, a webcam is connected and visualized during the acquisition application in order to control the state of the patient when the cabin is closed.

The control of the actuators is performed by a different PLC; the one used in this version is a Danfoss VLT Micro Drive FC-51. This PLC is smaller than the one used in the horizontal scanner since it does not have to control large motors, like the one used on our first version of the scanner. The communication between the PLC and the PC has also been improved in this version. This new PLC allows the use of the *libnodave* library. This standard library allows the exchange of data with Siemens PLCs. In our case, we use it for sending/receiving position values.

The lighting system presents two modifications with respect to the horizontal scanner. First, we increased the number of fluorescents to 12 to maximize the amount of light, and secondly, we added what we call the ambient light, that is a small LED light placed on the roof of the scanner. This light goes on when the patient enters the scanner before starting the acquisition and helps the technician

4. DESIGN AND DEVELOPMENT OF A NEW 3D BODY SCANNER

to put the patient in place. The light is controlled by the Power Control module connected to the PC.

Also, the number of cameras is increased from 11 to 19. The difference in the number of cameras is due to a number of factors, like the different movements of each scanner, the zone to be covered at each step and the final acquisition time.

The linear movement on the horizontal scanner allows the system to take images of the body's side, while the vertical scanner only acquires images from either the top or the bottom of the body at each step. Therefore, the area of the body covered by the vertical scanner is larger than on the previous scanner. On the other hand, the disadvantage of having more cameras is compensated by the reduction in the acquisition time on the vertical scanner, since it is possible to cover the whole body with the same resolution as with the horizontal scanner but with fewer steps.

The second scanner is developed under the Linux operating system and uses open source libraries (*libgphoto*) to control the cameras. This change made the system more stable as well as adding extra options to it. The webcam on the vertical scanner is controlled using the *OpenCV* libraries.

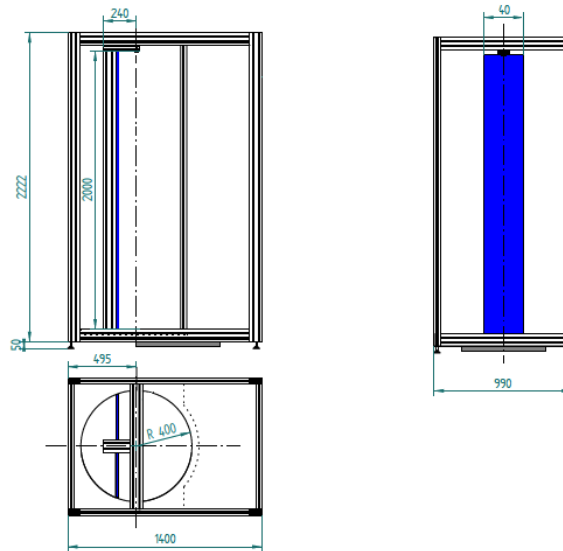


Figure 4.15: Vertical Scanner Schema

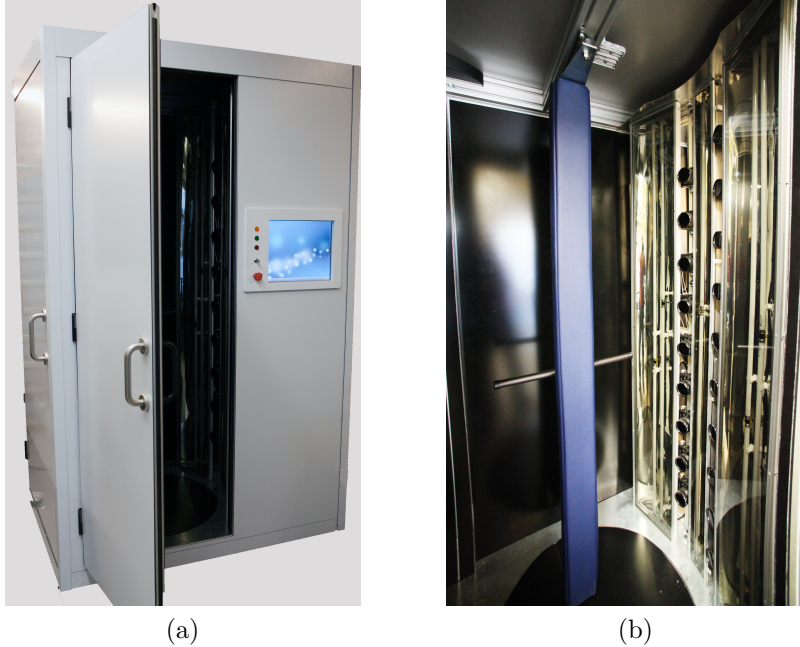


Figure 4.16: Vertical Scanner: (a) Exterior part of the vertical scanner and (b) interior of the scanner with the patient mattress, the lighting system and the acquisition.

4.3 3D in vivo skin mapping

Both scanners are equipped with a software developed in our lab and able to create 3D body maps. A common shortcoming of the existing systems outlined above is the limited image and texture resolution they provide on the reconstructed 3D models. For this reason, our proposed 3D skin mapping system was designed to have enough resolution to enable automatic detection of temporal changes. In contrast with the state-of-the-art devices currently employed by dermatologists, our approach can map the entire body, thus going beyond the standard procedure of manually linking images of lesions on top of avatars. Our solution not only allows the detection of new lesions, but also allows recording the state of the nevi at every exploration, thus enabling the detection of changes over time. In our system, images could be automatically located and mapped onto an avatar. This automatic location enables the comparison of the image containing the lesion and

4. DESIGN AND DEVELOPMENT OF A NEW 3D BODY SCANNER

the image of the same lesion but from another exploration.

4.3.1 Method

4.3.1.1 Image preprocessing

The proposed 3D body mapping system is not only designed to be used with our scanner, but is able to perform 3D reconstruction using images acquired by other means, ranging from low-resolution images provided by cell phone cameras to high-resolution images such as those acquired by professional digital cameras. However, the higher the number of images, the more detailed 3D reconstruction can be obtained.

The only pre-requisite is the calibration of the intrinsic camera parameters[22] in order to model the image formation process and to correct radial and tangential image distortions. Once the images are corrected, the proposed image processing pipeline is based on a series of successive steps detailed in the following subsections, namely: estimation of the position and orientation of the cameras, dense reconstruction, surface reconstruction and texture mapping.

In comparison with other products already available on the market, our solution significantly increases the resolution and provides the required texture for change detection. This higher resolution will allow the development of automatic change detection software to assist dermatologists in the detection of new lesions and/or to detect small changes in an already existing lesion.

4.3.1.2 Estimation of Camera positions

First, the position of the images is estimated by a Structure From Motion (SFM) procedure similar to the one described in[130]. This algorithm performs an incremental reconstruction based on the direct registration of the images to a set of already reconstructed 3D points. It comprises two main steps: first, an initial model is obtained by using a standard motion estimation technique[71], and then this initial model is used to directly register new images.

Fixing the first image in the sequence as the reference frame initializes the model. Then, a second image is selected from the set of available images as the one

that maximizes the overlap with respect to the reference image. In typical SFM algorithms, the fundamental matrix is usually used for camera motion estimation. This is the correct approach when the scene has a high 3D relief or has a noticeable parallax. However, some of the images in our intended application may present very low parallax or nearly planar content, thus generating an ill-conditioned fundamental matrix[71]. In this case, a more robust approach consists of modeling that motion using a homography[71]. The SFM algorithm selects one or another model depending on the scene observed.

Once the initial model has been built, positions of new images can be computed by finding correspondences between imaged 2D points and reprojected 3D points. To perform this matching, each 3D point has an associated description vector, which is the mean of the different descriptors[106] of each of the 2D points that have generated it. By having the correspondence between 2D and 3D points, the procedure to compute the position also follows a dual approach of using a projection matrix or a homography based on the planarity of the scene. The estimated final position is refined through Bundle Adjustment[104] to further reduce the reprojection error.

4.3.1.3 Dense Reconstruction

In this step we do not focus on finding correspondences between the most distinctive features in the images, like in the SFM procedure. Instead, the goal here is to obtain a comparatively larger set of correspondences by using a less constrained approach based on the distinctive measure of each feature. For this purpose, we use the method described in[53], which creates a model in the form of a dense set of points with associated normals. This method consists of three main blocks: matching, expansion and filtering.

4.3.1.4 Matching

The matching step first detects a set of corners and blob features using the Harris detector[70] and Difference-of-Gaussians (DoG). The selected features are forced to be uniformly spread to cover the entire image. The features are then matched across multiple images to reconstruct a sparse set of patches. For each feature f

4. DESIGN AND DEVELOPMENT OF A NEW 3D BODY SCANNER

detected in image I , a set of features F is selected from the other images, using epipolar constraints to restrict the search. Once the set of correspondences has been found, the computation of the patch is performed by constraining its position along a line of sight of a reference image, and then finding the position and orientation parameters of the patch by minimizing the photometric consistency of its projection across images.

4.3.1.5 Expansion and Filtering

Once the matching step has been performed, the expansion and filtering steps are iterated. The expansion step aims to add new neighbors to existing patches until they cover the surfaces visible in the scene. The filtering step accounts for visibility consistency of the recovered patches, i.e., it removes outliers taking into account self-occlusions among patches in 3D. This filtering is carried out in two steps. The first filter focuses on removing patches that lie outside the surface described by the points, while the second filter focuses on outliers lying inside this surface.

4.3.1.6 Surface Reconstruction

The points obtained from the dense reconstruction described above can be seen as discrete samples of the scanned surface. This representation is not sufficient to get clear visibility, i.e., an observer cannot tell whether a point is closer to or farther away from another without moving the point of view around the model. For this reason, a triangle mesh describing the surface is needed. Furthermore, this representation will allow the projection of the texture from the original images onto it, thus obtaining a useful model for 3D body inspection.

The set of points resulting from the dense reconstruction step are called unorganized, since they are not assumed to follow any underlying structure. There are several methods in the literature to deal with the ill-posed problem of finding a surface from a set of unorganized points. One of the more widely used techniques is the Poisson method[85], which discretizes the working space into a voxel grid, and splats the samples on it. The volume of the object will then be recovered implicitly in this grid as an indicator function, i.e., the value of a given voxel will

be zero if it is part of the outside of the object, and will be one if it is part of the inside. From this representation, the surface is extracted using the Marching Cubes algorithm[102].

The splatted points and normals are seen as a discretization of the gradient of the indicator function described above. The problem then is to find the inverse of this gradient, i.e., to find the indicator function whose gradient best approximates the vector field defined by the oriented samples. By means of applying the divergence operator, this problem can be cast as a Poisson problem, consisting of finding the scalar function whose Laplacian (divergence of gradients) equals the divergence of the vector field.

4.3.1.7 Texture Mapping

Once the surface mesh is built, the original images are used to obtain the texture corresponding to each of its triangles. The set of views compatible with a given triangle (i.e., the views where it is visible) are computed using line-of-sight constraints, and then the triangle is back-projected into these images. The image containing the reprojected triangle with largest area is the one selected to provide the texture.

4.4 Results

4.4.1 3D reconstruction results

In this section we illustrate a 3D reconstruction process using only an arm and not the entire body to clearly show the full process. The sequence has 23 images, more than those used in the scanner to describe the arm once the full body is acquired. The arm is one of the more challenging parts of the body due to the difficulty of reconstructing occluded zones such as some parts of the middle fingers. Fig. 4.17 shows the pipeline of matching, expanding and filtering of the 3D model.

Fig. 4.18 presents a final view of the reconstructed arm, with an example of the obtained resolution.

4. DESIGN AND DEVELOPMENT OF A NEW 3D BODY SCANNER

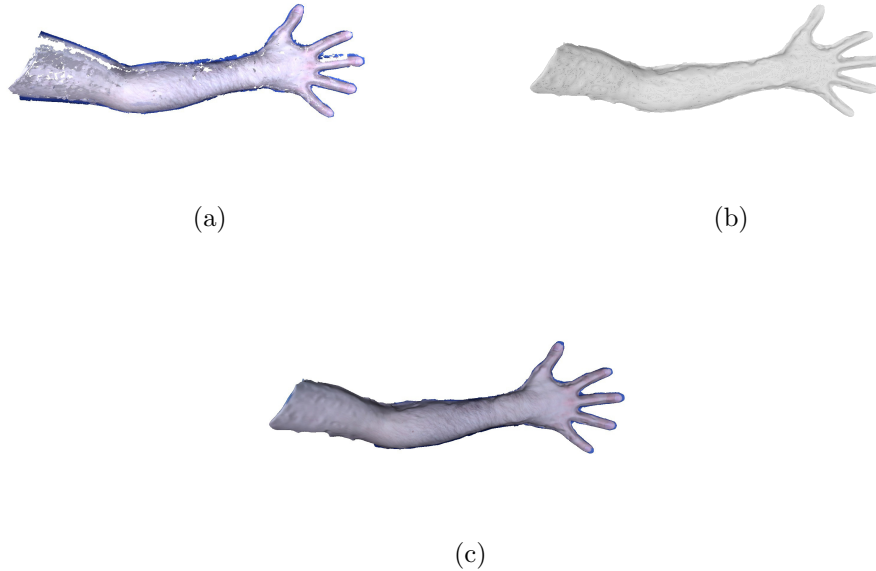


Figure 4.17: Sample reconstruction pipeline: (top) dense 3D points; (middle) mesh of triangles forming the surface as returned by the Poisson method; and (bottom) textured surface.

4.4.2 2D - 3D results on the scanner

In this section we present some results from the images acquired by the horizontal scanner. The image sequence contains 352 images from 11 cameras. All the cameras are setup with the same parameters. The average resolution of the images is 25 pixels per millimeter. This means that small changes between sequences can be returned as a big difference in the acquired image. These results represent some of the possibilities the scanner can obtain, but the final goal of our system is to place each lesion of a patient to a standardized 3D map of the patient in order to compare the same image from different sequence acquisitions.

First, we present the mosaic in Figure 4.19 built by using the techniques and software described in chapter 3. The mosaic presents some inaccuracies due to the presence of 3D on the scene and therefore unable to be considered planar for a 2D representation.

Although 2D mosaicing allows the estimation of an approximate location of

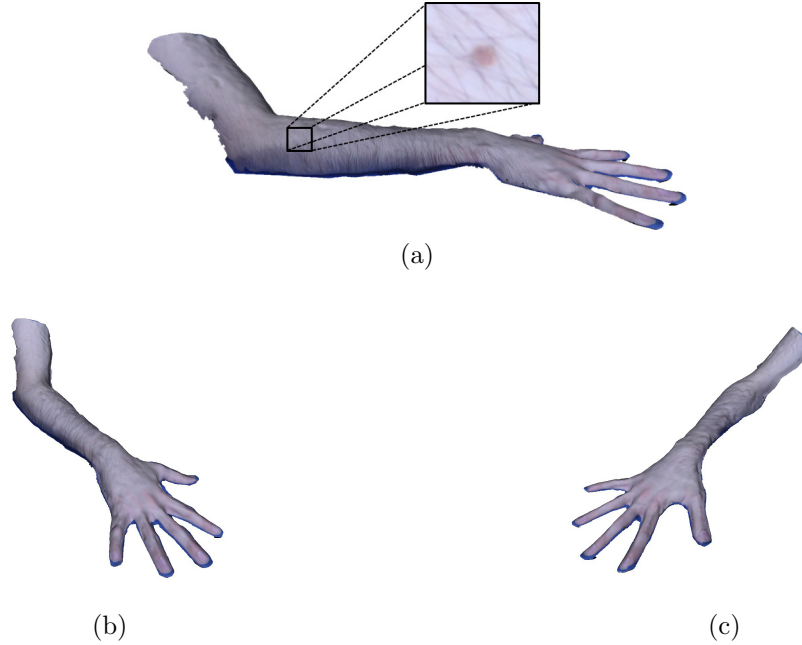


Figure 4.18: Three views of the final reconstruction after applying the described pipeline.

every image with respect to the patient body, accurate location of the patient's lesions (needed for analysis of their temporal evolution) requires a 3D approach, which is not affected by parallax. Unfortunately, a 3D reconstruction approach would pay the price of reducing the resolution of the texture patch associated to the lesion. For this reason, we propose a mixed model where the 3D mesh allows the proper mapping of lesions within the patient body, and the original 2D image will be used for diagnosis and change detection.

Fig. 4.20 illustrates the 3D point cloud recovered from the intersection of interest points from among the cameras.

The expansion of these points through all the images produces 3D representations such as those illustrated in Figures 4.21 and 4.22. The first one presents the mesh of triangles forming the surface as returned by the Poisson method and the second figure presents the resulting texturized surface.

Fig. 4.23 represents a small part for the patient's back. A detailed image of the 3D texturized view of three lesions and the original image are illustrated. It

4. DESIGN AND DEVELOPMENT OF A NEW 3D BODY SCANNER

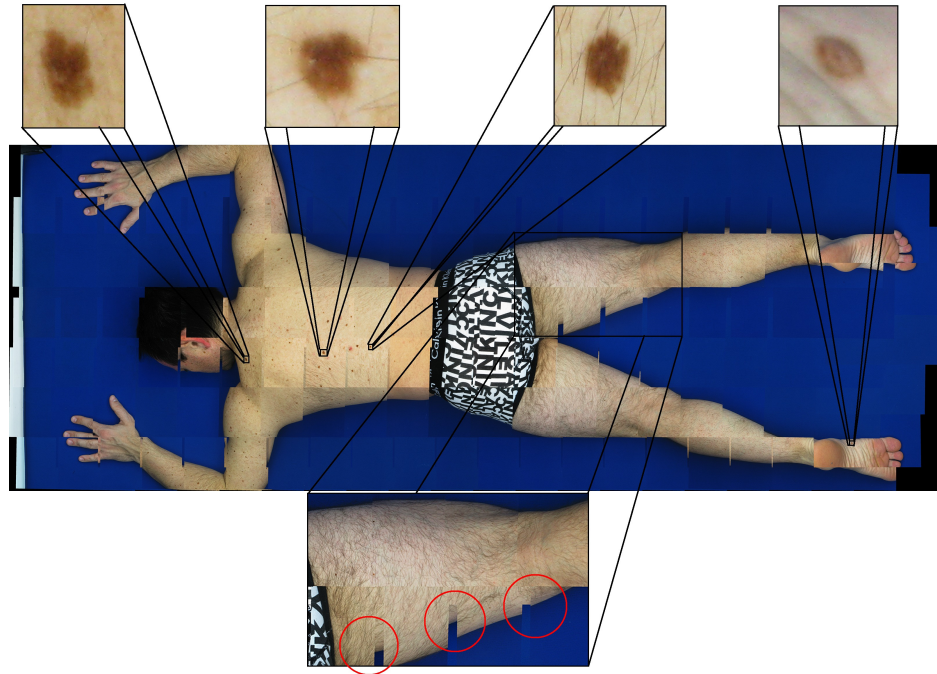


Figure 4.19: 2D Mosaic of the full body. The small images on the top shows the real resolution of the scanner, every millimeter on the human body is represented by at least 20 pixels on the mosaic. Bottom: incorrect alignment due to the use of a 2D approach under the presence of strong parallax.

should be noted that the 3D model has a lower resolution than the original 2D image, since it is not intended for diagnosis purposes, but only to correctly localize every lesion within the patient's body. Its lower resolution is due to the different processing steps performed, avoiding extremely large 3D meshes, but still allowing the 3D model to encode information about the normal vector of each lesion, i.e., the orientation of the patch where the lesion is located, avoiding slanted views of the lesions. These normals would improve local alignment between images of the same lesion, acquired in different explorations. Therefore, the original 2D image is more suitable for diagnosis, and the 3D model enables the correction of the orientation of the lesion with respect to the camera, making 2D comparison more reliable.

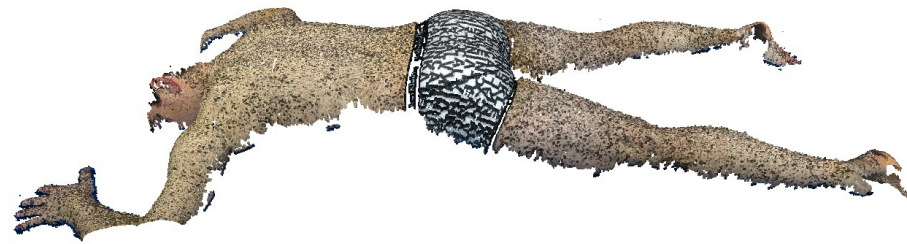


Figure 4.20: 3D Point Cloud representation of the patient's back. Each point has correspondence in at least 3 cameras.

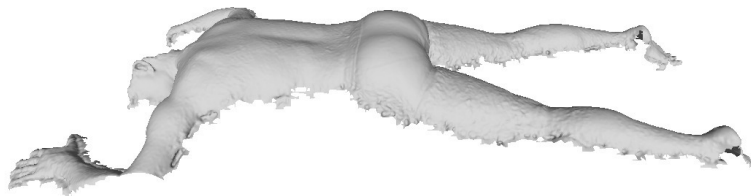


Figure 4.21: Mesh of triangles forming the surface as returned by the Poisson method for the patient's back.

4.5 Conclusions

This chapter presented a proposal for 2D/3D mapping with the aim of providing the basis for automatic detection over time. First, we reviewed most of the 3D scanners available on the market, showing their characteristics, advantages and

4. DESIGN AND DEVELOPMENT OF A NEW 3D BODY SCANNER



Figure 4.22: Texturized Surface for the patient's back.

3D reconstruction Info.	
N° of images	209
N° of matched images	155
N° of 3D Dense Points	11,002,962
N° of vertex	189,484
N° of triangles	373,276

Table 4.1: 3D Reconstruction Results.

drawbacks. From these approaches, we decided to proceed with a photogrammetric approach because it is the only one that provides texture and color information needed for skin cancer detection.

We developed two prototypes: the first one is called the horizontal scanner and the second one is a vertical scanner. Now both are in the last stages of development and close to being ready for the test phase in a real environment at the Hospital Clinic in Barcelona.

Finally, we developed and tested a 3D reconstruction software in order to obtain a model of the patient for locating moles. This approach has some problems due to the inaccuracies of the procedure. For this reason, we will implement a simpler function in the future that will relate any acquired image with a 3D posi-

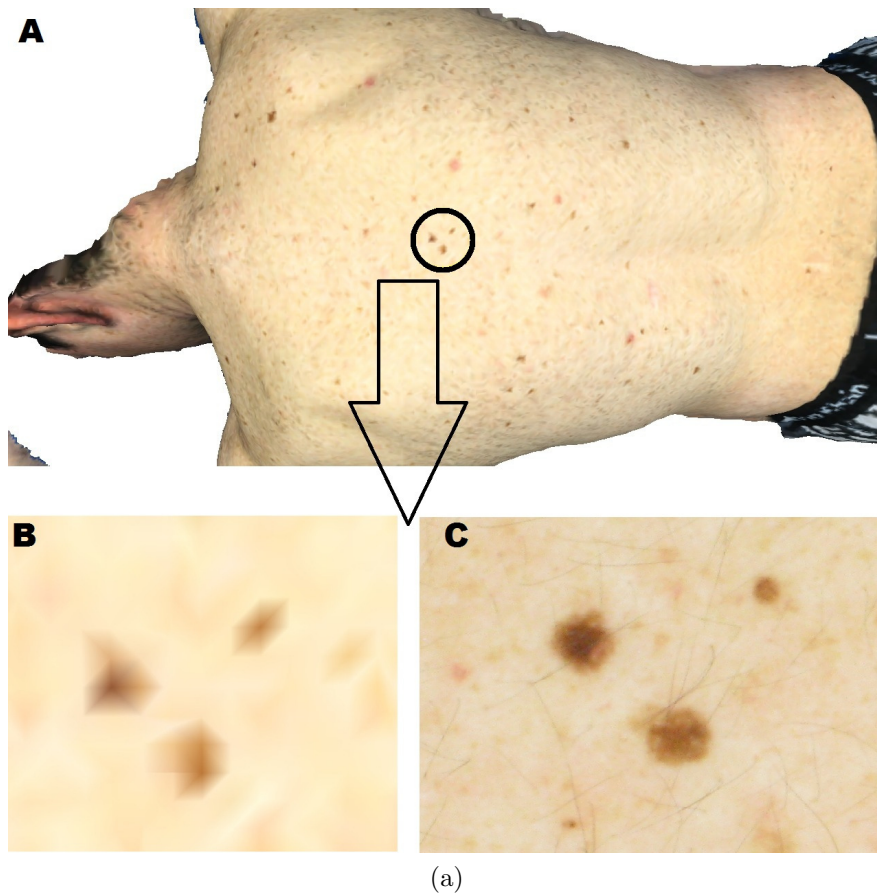


Figure 4.23: (a) Texturized Surface for the patient's back. (b) Detail of a small part of the 3D surface. (c) same part on the original image.

tion on an avatar. For comparing images over time, the images lying in the same position on the avatar will be compared individually, looking for new or changed lesions.

4. DESIGN AND DEVELOPMENT OF A NEW 3D BODY SCANNER

Chapter 5

Conclusions

This chapter concludes the work presented through this document. It first summarizes the thesis by reviewing the contents described in each chapter and some aspects that, although not accomplished during the thesis, point out some interesting future research issues. The chapter then points out the research contributions extracted from the proposals and experiments. Finally, the publications related to this work are listed.

5.1 Summary of the Thesis

The aim of this thesis has been the development of a reliable set of tools to assist dermatologists in the early detection of skin cancer. We began studying and analyzing the proposals available in the literature. We made an extensive literature review of all the systems which are commercially available and can be potentially associated for the diagnosis of skin cancer. Pros and cons of each technique/instrument are detailed and compared leading to the conclusion that some diagnosis methods can be improved and new technologies can be applied to dermatology. Mainly increasing the resolution of dermatological images and improving their color consistency.

We developed a new framework which takes three objectives into account; the quality of the dermatological images, the registration of full body exploration, and the automatic acquisition of high quality full body images.

Over the last few decades color correction has been widely investigated, but

5. CONCLUSIONS

some improvements can be still made, such as the ones proposed in this thesis. Our algorithm minimizes the problems produced by the use of different dermatoscopes and cameras with the goal of having more similar colors between these different devices. Correcting the colors prior to any further computations is a mandatory preprocessing step for automatic skin cancer detection. Also, inexperienced dermatologists or general practitioners would be able to increase their diagnostic accuracy by applying algorithms such as the 7-Point checklist or ABCDE rule that uses color as one of their main criteria.

Chapter 2 presents our algorithm for color calibration of dermatological images, comparing them with the state-of-the-art methods. We established a new procedure for acquiring dermatoscopical images as well as, new procedures for color correcting them. Based on the prior knowledge of the acquisition lighting, the spectral distribution of illuminant is introduced to the calibration algorithm, improving the results presented in the literature. The performance of our algorithm is tested on different cameras and image formats. RAW images and images acquired with high quality cameras produces the best results which are better, in terms of color differences that the state-of-the-art technique. Also a Custom White Balance (CWB) technique is presented for low cost cameras that only produces JPEG images.

Our second objective in this work, described in chapter 3, addresses the problem of how to automatically combine images from the same patient in a single image, generating a single frame composed from different images. This software will allow storage of some views of the patient instead of saving all the individual images. Every detected nevus will be easy to point out on the map, and the patient will be able to look at the same lesion in different explorations, lessening the time needed to go image by image to find the same mole. Once the same mole is detected, the patient can see if any change has taken place, and decide to visit a dermatologist for an accurate diagnosis. The Chapter starts presenting an state-of-the art of the latest keypoint detectors and descriptors. The matching process and image registration is modelled by an homography. Optimization and blending methods are presented in order to improve the final results. With full body image sequences 4 markers are placed on the scene that help calculating a global homography from the first to the last frame. Finally some examples

of different mosaic are presented highlighting some strengths and weakness of the method. Hardware optimization have also been considered by means of the use of the GPU-SIFT detector and descriptor running over a CUDA-adapted graphical card.

Chapter 4 presents a new and innovative proposal to acquire full body explorations from patients. Some devices which obtain 3D full body representations of the human body are described in the literature, most of them are being used in the textile, ergonomics, and fitness industries, among others but not in medical usage. The other systems reviewed in Chapter 4 present accurate 3D representations of the body useful for defining the size and shape of the body. Their main limitation is the lack of image quality, texture and color information, which does not allow dermatologists to see the moles and form a diagnosis from their 3D representation. Our system produces high-quality images that provide texture and color information. Two options are presented for viewing our images. The first is the construction of a 3D model using all the images, and the second is to automatically locate each image on to the avatar of a person.

5.2 Contributions

The main contributions of this thesis are:

- An extensive survey of dermatological devices for skin cancer detection, which are classified according to the technology used. The different methods are compared and the main advantages and disadvantages are shown.
- A color correction scheme to improve the quality of dermatoscopic images. The differences produced by using different dermoscopes and cameras are corrected after applying the algorithm rendering the colors from different images closer to the real ones. The approach is validated with synthetic and real experiments. The results are consistent and present a major improvement with respect to the state-of-the-art.
- A new method based on mosaicking for registering various explorations of the whole body is presented. Multiple images from the same patient can be

5. CONCLUSIONS

combined in a single one, allowing self-exploration by the patient himself and easily enabling the detection of changes over time.

- A new device for acquiring full body explorations is proposed. This device allows the generation of 3D body models with the aim of increasing the resolution and quality of the state-of-the-art.
- 2D and 3D body maps are obtained, presenting the techniques and software required to generate them.

5.3 Further Work

Short term perspectives

- The color calibration algorithm has been probed to work in synthetic images and on few real skin images. Our future plan is to apply this algorithm to a huge number of dermatoscopical images and compare the diagnosis results if dermatologists uses corrected or uncorrected images.
- The approach presented in this work to acquire and generate 3D full body models has been tested on a reduced number of sequences acquired in our research lab. The next step would be to acquire new sequences in a clinical environment, mapping the nevi on a 3D model of the body. This remains as immediate future work.
- It is a must to continue collaborating with dermatologists in order to improve the algorithms developed, and incorporating them in the daily practice.

Long term perspectives

- This thesis presented some algorithms that will help dermatologists to automatically detect skin lesions. Our further goal is to automatize this process.
- In terms of change detection, the use of existing approaches, as well as the performance of new techniques will be tested, both using the proposed 2D and 3D approaches.

-
- It would be interesting to validate the 3D scanner in a clinical study, enabling the future commercialization of this device.
 - Finally, a very ambitious aim is to complete a set of tools in order to acquire, improve the quality of dermatological images, segment skin lesions and finally diagnose according to the changes discovered in each lesion in a fully automated way.

5.4 Related Publications

A list of publications by the author for the PhD candidacy is given below, ordered according to their topic.

Publications related to camera color calibration:

- [CMIG11] J. Quintana, R. Garcia, L. Neumann. A novel method for color correction in epilluminescence microscopy. *Computerized Medical Imaging and Graphics*. Volume 35. Numbers 7-8. October–December 2011. Pag. 646–652.

Publications related to the novel system for skin mosaicking:

- [MSC08] J. Quintana. Assessment of cutaneous lesions through mosaicing. Thesis Submitted for the Degree of Msc Automatic Control and Computer Engineering. Girona. July 2008.

Publications related to the design and development of a new 3D body scanner:

- [MICCAT11] J. Quintana, R. Campos, N. Gracias, J. Freixenet, R. Garcia. 3D Skin Mapping for Melanoma Detection Conference. Barcelona. December 2011.
- [WCD12] J. Quintana, R. Campos, R. Garcia, J. Freixenet, N. Gracias, S. Puig, J. Malvehy. A novel acquisition device for Total Body Photography. World Congress of Dermoscopy. Brisbane. May 2012.

5. CONCLUSIONS

References

- [1] ABERG, P., GELADI, P., NICANDER, I., HANSSON, J., HOLMGREN, U., AND OLLMAR, S. Non-invasive and microinvasive electrical impedance spectra of skin cancer - a comparison between two techniques. *Skin Research and Technology* 11, 4 (2005), 281–286. 29
- [2] AGARWALA, A., DONTCHEVA, M., AGRAWALA, M., DRUCKER, S., COLBURN, A., CURLESS, B., SALESIN, D., AND COHEN, M. Interactive digital photomontage. In *SIGGRAPH '04: ACM SIGGRAPH 2004 Papers* (New York, NY, USA, 2004), ACM, pp. 294–302. 102, 104
- [3] AMFIT. Amfit inc. <http://www.amfit.com>, feb 2012. 133
- [4] APUZZO, N. Recent advances in 3d full body scanning with applications to fashion and apparel. 127
- [5] ARGENZIANO, G., CATRICALA, C., ARDIGO, M., BUCCINI, P., SIMONE, P. D., EIBENSCHUTZ, L., FERRARI, A., MARIANI, C., SILIPO, V., SPERDUTI, I., AND ZALAUDEK, I. Seven-point checklist of dermoscopy revisited. *The British journal of dermatology* 164, 4 (2011), 785–790. 17
- [6] ARGENZIANO, G., FABBROCINI, G., CARLI, P., DE GIORGI, V., SAMMARCO, E., AND DELFINO, M. Epiluminescence microscopy for the diagnosis of doubtful melanocytic skin lesions. comparison of the abcd rule of dermoscopy and a new 7-point checklist based on pattern analysis. *Arch Dermatol* 134, 12 (1998), 1563–70. 16

REFERENCES

- [7] ARGENZIANO, G., SOYER, H., GIORGI, V., PICCOLO, D., CARLI, P., AND DELFINO, M. Dermoscopy: A tutorial. *EDRA Medical Publishing* (2002). 12
- [8] ARYA, S., MOUNT, D., NETANYAHU, N., SILVERMAN, R., AND WU, A. An optimal algorithm for approximate nearest neighbor searching fixed dimensions. *J. ACM* 45, 6 (1998), 891–923. 110
- [9] ASAWANONDA, P., AND TAYLOR, C. Wood’s light in dermatology. *Int J Dermatol* 38, 11 (1999), 801–7. 9
- [10] BALSEIRO, S., AND CORREIA, H. Is olfactory detection of human cancer by dogs based on major histocompatibility complex-dependent odour components?—a possible cure and a precocious diagnosis of cancer. *Med Hypotheses* 66, 2 (2006), 270–2. 10
- [11] BANKY, J., KELLY, J., ENGLISH, D., YEATMAN, J., AND DOWLING, J. Incidence of new and changed nevi and melanomas detected using baseline images and dermoscopy in patients at high risk for melanoma. *Arch. Dermatol* 141, 8 (2005), 998–1006. 10
- [12] BARNARD, K., AND FUNT, B. Camera characterization for color research. *Color Research and Application* (Jan 2002). 37
- [13] BARNHILL, R., FANDREY, K., LEVY, M., MIHM, M., AND HYMAN, B. Angiogenesis and tumor progression of melanoma. quantification of vascularity in melanocytic nevi and cutaneous malignant melanoma. *Lab Invest* 67, 3 (1992), 331–7. 28
- [14] BAY, H., ESS, A., TUYTELAARS, T., AND GOOL, L. V. Speeded-up robust features (surf). *Comput. Vis. Image Underst.* 110 (June 2008), 346–359. 80
- [15] BAY, H., TUYTELAARS, T., AND GOOL, L. J. V. SURF: Speeded up robust features. In *European Conference on Computer Vision* (2006). 80
- [16] BEAUDET, P. *Rotationally invariant image operators*, vol. 579. 1978, pp. 579–583. 74

REFERENCES

- [17] BENTLEY, J. K-d trees for semidynamic point sets. In *SCG '90: Proceedings of the sixth annual symposium on Computational geometry* (New York, NY, USA, 1990), ACM, pp. 187–197. 110
- [18] BERARDI, P., AND ARCANGELI, F. The tape stripping toluidine blue (tstb) method in the diagnosis of malignant melanoma: an investigator-blind study. *Melanoma Res* 2, 2 (1992), 93–9. 27
- [19] BEUCKMANN. Beuckmann gmbh. <http://www.beuckmann.com>, feb 2012. 129
- [20] BOOKSTEIN, F. Principal warps: thin-plate splines and the decomposition of deformations. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 11, 6 (Jun 1989), 567–585. 120
- [21] BOTTINO, A., AND LAURENTINI, A. A silhouette based technique for the reconstruction of human movement. *Computer Vision and Image Understanding* 83, 1 (2001), 79–95. 131
- [22] BOUGUET, J. Camera calibration toolbox for Matlab, 2008. 144
- [23] BOYLE, P., DOR, J., AUTIER, P., AND RINGBORG, U. Cancer of the skin: a forgotten problem in europe. *Ann Oncol* 15, 1 (jan 2004), 5–6. 1
- [24] BRAINARD, D. H. Color appearance and color difference specification. *The Science of Color* (2003), 191–216. 48
- [25] BRANZAN, A., LANDTHALER, M., AND SZEIMIES, R. In vivo confocal scanning laser microscopy in dermatology. *Lasers in Medical Science* 22, 2 (2007), 73–82. 22
- [26] BROWN, M., HARTLEY, R., AND NISTER, D. Minimal solutions for panoramic stitching. In *IEEE Conference on Computer Vision and Pattern Recognition* (2007). 65, 67
- [27] BROWN, M., AND LOWE, D. Recognising panoramas. In *ICCV '03: Proceedings of the Ninth IEEE International Conference on Computer Vision*

REFERENCES

- (Washington, DC, USA, 2003), IEEE Computer Society, p. 1218. 65, 83, 102, 103
- [28] BURT, P., AND ADELSON, E. A multiresolution spline with application to image mosaics. *ACM Trans. Graph.* 2, 4 (1983), 217–236. 102
- [29] BUSAM, K., CHARLES, C., LEE, G., AND HALPERN, A. Morphologic features of melanocytes, pigmented keratinocytes, and melanophages by in vivo confocal scanning laser microscopy. *Mod Pathol* 14, 9 (2001), 862–8. 22
- [30] CALDEMEYER, K., AND BUCKWALTER, K. The basic principles of computed tomography and magnetic resonance imaging. *Journal of the American Academy of Dermatology* 41, 5 Pt 1 (1999), 768–771. 30
- [31] CANNY, J. A computational approach to edge detection. *Pattern Analysis and Machine Intelligence, IEEE Transactions on PAMI-8*, 6 (nov 1986), 679–698. 72
- [32] CAPEL, D. *Image Mosaicing and Super-Resolution (Cphc/Bcs Distinguished Dissertations.)*. SpringerVerlag, 2004. 97, 99
- [33] CHEN, C., AND KAK, A. Modelling and calibration of a structured light scanner for 3d robot vision. In *in IEEE Conference on Robotics and Automation* (1987), pp. 807–815. 128
- [34] CHEUNG, K., BAKER, S., AND KANADE, T. Visual hull alignment and refinement across time: A 3d reconstruction algorithm combining shape-from-silhouette with stereo. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (June 2003). 131
- [35] CHURCH, J., AND WILLIAMS, H. Another sniffer dog for the clinic? *Lancet.* 358, 9285 (Sep 2001), 930. 10
- [36] CHWIROT, B., CHWIROT, S., REDZISKI, J., AND MICHNIEWICZ, Z. Detection of melanomas by digital imaging of spectrally resolved ultraviolet light-induced autofluorescence of human skin. *Eur J Cancer* 34, 11 (1998), 1730–4. 9

REFERENCES

- [37] CORBALAN, M., MILLAN, M., AND YZUEL, M. Color measurement in standard cielab coordinates using a 3ccd camera: correction for the influence of the light source. *Optical Engineering* 6, 39 (2000), 1470–1476. 48, 63
- [38] CSEM. Centre suisse d’électronique et de microtechnique. <http://www.csem.ch>, feb 2012. 133
- [39] CYBERWARE. Cyberware inc. <http://www.cyberware.com>, feb 2012. 128
- [40] DANI, P., AND CHAUDHURI, S. Automated assembling of images: image montage preparation. *Pattern Recognition* 28, 3 (1995), 431–445. 65
- [41] DANI, P., AND CHAUDHURI, S. Automated assembling of images: Image montage preparation. *Pattern Recognition* 28, 3 (1995), 431–445. 67
- [42] DAVIS, J. Mosaics of scenes with moving objects. In *CVPR '98: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (Washington, DC, USA, 1998), IEEE Computer Society, p. 354. 99, 103
- [43] DEBEVEC, P., AND MALIK, J. Recovering High Dynamic Range Radiance Maps from Photographs. *Computer Graphics* 31 (1997), 369–378. 45
- [44] DELANEY, P., HARRIS, M., AND KING, R. Fiber-optic laser scanning confocal microscope suitable for fluorescence imaging. *Appl. Opt.* 33, 4 (Feb 1994), 573–577. 22
- [45] EFROS, A., AND FREEMAN, W. Image quilting for texture synthesis and transfer. In *SIGGRAPH '01: Proceedings of the 28th annual conference on Computer graphics and interactive techniques* (New York, NY, USA, 2001), ACM, pp. 341–346. 102, 103
- [46] ELINVISION. Elinvision - intelligent e-vision. <http://www.elinvision.lt>, feb 2012. 133
- [47] EPSTEIN, E. Magnifiers in dermatology: a personal survey. *J Am Acad Dermatol* 13, 4 (1985), 687–98. 8

REFERENCES

- [48] FATTAL, R., LISCHINSKI, D., AND WERMAN, M. Gradient domain high dynamic range compression. 114
- [49] FEIT, N., DUSZA, S., AND MARGHOUB, A. Melanomas detected with the aid of total cutaneous photography. *The British journal of dermatology* 150, 4 (2004), 706–714. 11
- [50] FISCHLER, M., AND BOLLES, R. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM* 24, 6 (1981), 381–395. 95, 110
- [51] FOREST, J., AND SALVI, J. A review of laser scanning three-dimensional digitisers. In *IEEE/RJS Int'l Conf. Intelligent Robots and Systems*, pp. 73–78. 128
- [52] FRIEDMAN, K., AND WAHL, R. Clinical use of positron emission tomography in the management of cutaneous melanoma. *Semin Nucl Med* 34, 4 (2004), 242–53. 30
- [53] FURUKAWA, Y., AND PONCE, J. Accurate, Dense, and Robust Multi-View Stereopsis. In *Computer Vision and Pattern Recognition, 2007. CVPR '07. IEEE Conference on* (jun 2007), pp. 1–8. 145
- [54] GLICKMAN, Y., FILO, O., DAVID, M., YAYON, A., TOPAZ, M., ZAMIR, B., GINZBURG, A., ROZENMAN, D., AND KENAN, G. Electrical impedance scanning: a new approach to skin cancer diagnosis. *Skin Research and Technology* 9, 3 (2003), 262–268. 29
- [55] GONZALEZ, S., AND GILABERTE-CALZADA, Y. In vivo reflectance-mode confocal microscopy in clinical dermatology and cosmetology. *Int J Cosmet Sci* 30, 1 (2008), 1–17. 22
- [56] GRACIAS, N., GLEASON, A., NEGAHDARIPOUR, S., AND MAHOOR, M. Fast image blending using watersheds and graph cuts. In *Proc. of the British Machine Vision Conference (BMVC06)* (Edinburgh, U.K., September 2006). 104

REFERENCES

- [57] GRAFULLAGONZALEZ, B., LEBART, K., AND HARVEY, A. Physical optics modelling of millimetre-wave personnel scanners. *Pattern Recognition Letters* 27, 15 (2006), 1852–1862. 133
- [58] GRANA, C., PELLACANI, G., AND SEIDENARI, S. Practical color calibration for dermoscopy, applied to a digital epiluminescence microscope. *Skin research and technology* 11, 4 (2005), 242–247. 38
- [59] GRANA, C., PELLACANI, G., SEIDENARI, S., AND CUCCHIARA, R. Color calibration for a dermatological video camera system. *17th International Conference on Pattern Recognition (ICPR) 3* (Jul 2004), 798 – 801 Vol.3. 38, 49
- [60] GUFLER, H., FRANKE, F., AND RAU, W. High-resolution mri of basal cell carcinomas of the face using a microscopy coil. *AJR Am J Roentgenol* 188, 5 (2007), W480–4. 30
- [61] GUPTA, D., DEAVERS, M., SILVA, E., AND MALPICA, A. Malignant melanoma involving the ovary: a clinicopathologic and immunohistochemical study of 23 cases. *Am J Surg Pathol.*, 6 (2004), 771–80. 30
- [62] GUPTA, L., AND SINGHI, M. Wood’s lamp. *Indian J Dermatol Venereol Leprol* (2004), 131–5. 9
- [63] HAEGHEN, Y., NAEYAERT, J., LEMAHIEU, I., AND PHILIPS, W. An imaging system with calibrated color image acquisition for use in dermatology. *IEEE transactions on medical imaging* 19, 7 (Jul 2000), 722–30. 37, 62, 63
- [64] HAEGHEN, Y. V., AND NAEYAERT, J. Consistent cutaneous imaging with commercial digital cameras. *Archives of Dermatology* 142, 1 (2006), 42–46. 37
- [65] HALPERN, A. The use of whole body photography in a pigmented lesion clinic. *Dermatologic Surgery* 26, 12 (2000), 1175–1180. 11

REFERENCES

- [66] HALPERN, A., MARGHOOB, A., BIALOGLOW, T., WITMER, W., AND SLUE, W. Standardized positioning of patients (poses) for whole body cutaneous photography. *J Am Acad Dermatol* 49, 4 (2003), 593–8. 10
- [67] HALPRIN, K. Diagnosis with wood’s light. tinea capitis and erythrasma. *JAMA* 199, 11 (1967), 841. 9
- [68] HAMAMATSU. Hamamatsu photonics. <http://hamamatsu.com>, feb 2012. 128
- [69] HAR-SHAI, Y., GLICKMAN, Y., SILLER, G., MCLEOD, R., TOPAZ, M., HOWE, C., GINZBURG, A., ZAMIR, B., FILO, O., KENAN, G., AND ULLMANN, Y. Electrical impedance scanning for melanoma diagnosis: a validation study. *Plastic and Reconstructive Surgery* 116, 3 (2005), 782–790. 29
- [70] HARRIS, C., AND STEPHENS, M. A combined corner and edge detector. In *Proceedings Alvey Conference* (Manchester, UK, August 1988), pp. 189–192. 72, 105, 145
- [71] HARTLEY, R., AND ZISSERMAN, A. *Multiple View Geometry in Computer Vision*, second ed. Cambridge University Press, ISBN: 0521540518, 2004. 91, 94, 97, 144, 145
- [72] HAWORTH, C., SAINT-PERN, Y. D., CLARK, D., TRUCCO, E., AND PETILLOT, Y. Detection and tracking of multiple metallic objects in millimetre-wave images. *International Journal of Computer Vision* 71, 2 (2007), 183–196. 133
- [73] HENNING, J., DUSZA, S., WANG, S., MARGHOOB, A., RABINOVITZ, H., POLSKY, D., AND KOPF, A. The cash (color, architecture, symmetry, and homogeneity) algorithm for dermoscopy. *J Am Acad Dermatol* 56, 1 (2007), 45–52. 19
- [74] HENNING, J., STEIN, J., YEUNG, J., DUSZA, S., MARGHOOB, A., RABINOVITZ, H., POLSKY, D., AND KOPF, A. Cash algorithm for dermoscopy revisited. *Arch Dermatol* 144, 4 (2008), 554–5. 19

REFERENCES

- [75] HONG, G., LUO, M., AND RHODES, P. A study of digital camera colorimetric characterization based on polynomial modeling. *Color Research & Application*, 26 (2001), 76–84. 43
- [76] HORN, B., AND SCHUNCK, B. Determining optical flow, 1981. 68
- [77] HUANG, J., LAI, S., AND CHENG, C. Robust fundamental matrix estimation with accurate outlier detection. *J. Inf. Sci. Eng.* 23, 4 (2007), 1213–1225. 95
- [78] I-WARE. I-ware laboratory co. <http://www.iwl.jp>, feb 2012. 128
- [79] INSPECK. Inspeck inc. <http://inspeck.com>, feb 2012. 130
- [80] IYATOMI, H., CELEBI, M., SCHAEFER, G., AND TANAKA., M. Automated color normalization for dermoscopy images. *IEEE 17th International Conference on Image Processing (ICIP)* (2010), 4357–4360. 38
- [81] IYATOMI, H., CELEBI, M., SCHAEFER, G., AND TANAKA, M. Automated color calibration method for dermoscopy images. *Computerized medical imaging and graphics* 35, 2 (2011), 89–98. 38
- [82] JEMEC, G., GNIADOCKA, M., AND ULRICH, J. Ultrasound in dermatology. part i. high frequency ultrasound. *Eur J Dermatol* 10, 6 (2000), 492–7. 24
- [83] KALKHORAN, S., MILNE, O., ZALAUDEK, I., PUIG, S., MALVEHY, J., KELLY, J., AND MARGHOUB, A. Historical, clinical, and dermoscopic characteristics of thin nodular melanoma. *Arch Dermatol* 146, 3 (2010), 311–8. 8
- [84] KANG, E., COHEN, I., AND MEDIONI, G. A graph-based global registration for 2d mosaics. In *ICPR '00: Proceedings of the International Conference on Pattern Recognition* (Washington, DC, USA, 2000), IEEE Computer Society, p. 1257. 99

REFERENCES

- [85] KAZHDAN, M., BOLITHO, M., AND HOPPE, H. Poisson surface reconstruction. In *Proceedings of the fourth Eurographics symposium on Geometry processing* (Aire-la-Ville, Switzerland, Switzerland, 2006), SGP '06, Eurographics Association, pp. 61–70. 146
- [86] KELLY, J., YEATMAN, J., REGALIA, C., MASON, G., AND HENHAM, A. A high incidence of melanoma found in patients with multiple dysplastic naevi by photographic surveillance. *Med J Aust* 167, 4 (1997), 191–4. 11
- [87] KITCHEN, L., AND ROSENFELD, A. Gray level corner detection. *PRL* 1, 2 (1982), 95–102. 72
- [88] KITTLER, H. Dermatoscopy: introduction of a new algorithmic method based on pattern analysis for diagnosis of pigmented skin lesions. *Dermatopathology: Practical & Conceptual*, 13 (2007), 1. 20
- [89] KOENDERINK, J., AND DOORN, A. V. Invariant properties of the motion parallax field due to the movement of rigid bodies relative to an observer. *Optica Acta* 22 (1975), 773–791. 121
- [90] KOENDERINK, J., AND RICHARDS, W. Two-dimensional curvature operators. *Journal of the Optical Society of America A* 5, 7 (1988), 1136–1141. 72
- [91] KOPF, A., SALOPEK, T., SLADE, J., MARGHOOB, A., AND BART, R. Techniques of cutaneous examination for the detection of skin cancer. *Cancer* 75, 2 Suppl (1995), 684–90. 9
- [92] KOVESI, P. Model fitting and robust estimation, 1997. 110
- [93] LAB., U. V. Girona mosaic markup language (gmml). <http://porcsenglar.udg.edu/svn/GMML>, feb 2012. 112
- [94] LANGE, R., AND SEITZ, P. Solid-state time-of-flight range camera. *IEEE Journal of Quantum Electronics* 37, 3 (mar 2001), 390–397. 133

REFERENCES

- [95] LAURENTINI, A. The visual hull concept for silhouette-based image understanding. *IEEE Trans. Pattern Anal. Mach. Intell.* 16 (February 1994), 150–162. 130
- [96] LAURENTINI, A. How far 3d shapes can be understood from 2d silhouettes. *IEEE Trans. Pattern Anal. Mach. Intell.* 17 (February 1995), 188–195. 131
- [97] LERCH, T., MACGILLIVRAY, M., AND DOMINA, T. 3d laser scanning: A model of multidisciplinary research. *Journal Of Textile And Apparel Technology And Management* 5, 4 (2006), 1–22. 128
- [98] LEVENBERG, K. A method for the solution of certain problems in least squares. In *QAM* (1944), vol. 2. 97
- [99] LEVIN, A., ZOMET, A., PELEG, S., AND WEISS, Y. Seamless image stitching in the gradient domain. In *ECCV04* (2004), pp. 377–389. 65, 102
- [100] LINDBERG, T. Detecting salient blob-like image structures and their scales with a scale-space primal sketch: A method for focus-of-attention. *International Journal of Computer Vision* 11 (1993), 283–318. 72
- [101] LINDBERG, T. Feature detection with automatic scale selection. *International Journal of Computer Vision* 30 (1998), 79–116. 72
- [102] LORENSEN, W., AND CLINE, H. Marching cubes: A high resolution 3d surface construction algorithm. *COMPUTER GRAPHICS* 21, 4 (1987), 163–169. 147
- [103] LORENTZEN, H., WEISMANN, K., SECHER, L., PETERSEN, C., AND LARSEN, F. The dermatoscopic abcd rule does not improve diagnostic accuracy of malignant melanoma. *Acta Derm Venereol* 79, 6 (1999), 469–72. 15
- [104] LOURAKIS, M., AND ARGYROS, A. The design and implementation of a generic sparse bundle adjustment software package based on the levenberg-marquardt algorithm. Tech. Rep. 340, Institute of Computer Science - FORTH, Heraklion, Crete, Greece, Aug. 2004. 100, 145

REFERENCES

- [105] LOVATO, C., CASTELLANI, U., FANTONI, S., MILANESE, C., ZANCANARO, C., AND GIACHETTI, A. Computer assisted estimation of anthropometric parameters from whole body scanner data. In *Proceedings of 3DPH* (Nov-Dec 2009), pp. 71–83. 127
- [106] LOWE, D. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vision* 60, 2 (2004), 91–110. 72, 76, 90, 145
- [107] LUCAS, B. *Generalized Image Matching by the Method of Differences*. PhD thesis, Robotics Institute, Carnegie Mellon University, July 1984. 68
- [108] LUCAS, B., AND KANADE, T. An iterative image registration technique with an application to stereo vision (ijcai). In *Proceedings of the 7th International Joint Conference on Artificial Intelligence (IJCAI '81)* (April 1981), pp. 674–679. 68
- [109] MADJIDI, H., AND NEGAHDARIPOUR, S. On robustness and localization accuracy of optical flow computation from color imagery. In *Proceedings of the 3D Data Processing, Visualization, and Transmission, 2nd International Symposium* (Washington, DC, USA, 2004), 3DPVT '04, IEEE Computer Society, pp. 317–324. 70
- [110] MAGLOGIANNIS, I., AND KOSMOPOULOS, D. A system for the acquisition of reproducible digital skin lesions images. *Technology and health care* 11, 6 (2003), 425–441. 38
- [111] MARCHESINI, R., TOMATIS, S., BARTOLI, C., BONO, A., CLEMENTE, C., CUPETA, C., PRATO, I., SICHIROLLO, E. P. A., AND CASCINELLI, N. In vivo spectrophotometric evaluation of neoplastic and nonneoplastic skin pigmented lesions. iii. ccd camera-based reflectance imaging. *Photochemistry and Photobiology* 62 (1995), 151–154. 25
- [112] MARGHOOB, A., CHARLES, C., BUSAM, K., RAJADHYAKSHA, M., LEE, G., CLARK-LOESER, L., AND HALPERN, A. In vivo confocal scanning laser microscopy of a series of congenital melanocytic nevi suggestive of having developed malignant melanoma. *Arch Dermatol* 141, 11 (2005), 1401–12. 22

REFERENCES

- [113] MARGHOOB, A., SWINDLE, L., MORICZ, C., NEGRON, F., SLUE, B., HALPERN, A., AND KOPF, A. Instruments and new technologies for the in vivo diagnosis of melanoma. *Journal of the American Academy of Dermatology* 49, 5 (2003), 777–797. 12, 25
- [114] MARQUARDT, D. An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial and Applied Mathematics* 11, 2 (1963), 431–441. 97
- [115] MATAS, J. Robust wide-baseline stereo from maximally stable extremal regions. *Image and Vision Computing* 22, 10 (2004), 761–767. 72, 82
- [116] MCCAMY, C., MARCUS, H., AND DAVIDSON, J. A color-rendition chart. *Journal of Applied Photographic Engineering* 2, 3 (Summer 1976), 95–99. 37
- [117] MENZIES, S. W., INGVAR, C., CROTTY, K., AND MCCARTHY, W. Frequency and morphologic characteristics of invasive melanomas lacking specific surface microscopic features. *Arch Dermatol.*, 132 (1996), 1178–82. 17, 19
- [118] MIKOLAJCZYK, K. Scale & affine invariant interest point detectors. *International Journal of Computer Vision* 60, 1 (2004), 63–86. 75
- [119] MIKOLAJCZYK, K., TUYTELAARS, T., SCHMID, C., ZISSERMAN, A., MATAS, J., SCHAFFALITZKY, F., KADIR, T., AND GOOL, L. V. A comparison of affine region detectors. *International Journal of Computer Vision* 65, 1-2 (2005), 43–72. 75
- [120] MONCRIEFF, M., COTTON, S., CLARIDGE, E., AND HALL, P. Spectrophotometric intracutaneous analysis: a new technique for imaging pigmented skin lesions. *British Journal of Dermatology* 146, 3 (March 2002), 448–457. 24
- [121] MONHEIT, G., COGNETTA, A., FERRIS, L., RABINOVITZ, H., GROSS, K., MARTINI, M., GRICHNIK, J., MIHM, M., PRIETO, V., GOOGE, P., KING, R., TOLEDANO, A., KABELEV, N., WOJTON, M., AND

REFERENCES

- GUTKOWICZ-KRUSIN, D. The performance of melafind: A prospective multicenter study. *Arch Dermatol* 147, 2 (2011), 188–94. 24
- [122] MORHENN, V., CHANG, E., AND RHEINS, L. A noninvasive method for quantifying and distinguishing inflammatory skin reactions. *J Am Acad Dermatol* 41, 5 Pt 1 (1999), 687–92. 26
- [123] MOUNT, D., AND ARYA, S. Ann: A library for approximate nearest neighbor searching, 1997. 109
- [124] MURTAGH, J. Wood’s light examination. *Aust Fam Physician*. 14, 6 (Jun 1985), 525. 9
- [125] NAKAJIMA, N., IKETANI, A., AND SATO, T. Video mosaicing for document imaging. Tech. rep., Common Platform Software Research Laboratories, NEC Corporation, 2004. 65
- [126] NEGAHDARIPOUR, S. Revised definition of optical flow: Integration of radiometric and geometric cues for dynamic scene analysis. *IEEE Trans. Pattern Anal. Mach. Intell.* 20 (September 1998), 961–979. 70
- [127] NEGAHDARIPOUR, S., AND MADJIDI, H. Direct estimation of motion from sea floor images for automatic station-keeping of submersible platforms. *IEEE Journal of Oceanic Engineering* 24 (1999), 370–382. 70
- [128] NEGAHDARIPOUR, S., AND MADJIDI, H. Robust optical flow estimation using underwater color images. In *OCEANS* (2003). 70
- [129] NEGAHDARIPOUR, S., XU, X., KHAMENE, A., AND AWAN, Z. 3-d motion and depth estimation from sea-floor images for mosaic-based station-keeping and navigation of rovs/auvs and high-resolution sea-floor mapping. In *In Workshop on Autonomous Underwater Vehicles AUV* (1998), pp. 191–200. 70
- [130] NICOSEVICI, T., GRACIAS, N., NEGAHDARIPOUR, S., AND GARCIA, R. Efficient three-dimensional scene modeling and mosaicing. *J. Field Robot.* 26 (October 2009), 759–788. 144

REFERENCES

- [131] OGGIER, T. An all-solid-state optical range camera for 3d real-time imaging with sub-centimeter depth resolution (swissranger). *Proceedings of SPIE 5249*, 65 (2004), 534–545. 133
- [132] ORSCAN. Orscan technologies ltd. <http://www.orscantec.com>, feb 2012. 132
- [133] PAAR, G., AND SMOLLE, J. Stereoscopic skin mapping for dermatology. In *CEUR - Bildverarbeitung fr die Medizin, Springer, Erlangen, Germany, March 9-11, 2003* (2003), vol. 80, pp. 211–215. 129
- [134] PEHAMBERGER, H., STEINER, A., AND WOLFF, K. In vivo epiluminescence microscopy of pigmented skin lesions. i. pattern analysis of pigmented skin lesions. *J Am Acad Dermatol 17*, 4 (1987), 571–83. 14
- [135] PELLACANI, G., AND SEIDENARI, S. Comparison between morphological parameters in pigmented skin lesion images acquired by means of epiluminescence surface microscopy and polarized-light videomicroscopy. *Clinical Dermatology 20*, 3 (2002), 222–7. 36
- [136] PEREZ, P., GANGNET, M., AND BLAKE, A. Poisson image editing. *ACM Transactions on Graphics (SIGGRAPH'03) 22*, 3 (2003), 313–318. 114
- [137] PINGLE, K. K. Visual perception by a computer. In *Automatic Interpretation and Classification of Images* (1969), 277–284. 72
- [138] POUCKE, S. V., HAEGHEN, Y., VISSERS, K., MEERT, T., AND P.JORENS. Automatic colorimetric calibration of human wounds. *BMC Medical Imaging 10*, 7 (2010), 7. 37
- [139] PSATY, E., SCOPE, A., HALPERN, A., AND MARGHOOB, A. Defining the patient at high risk for melanoma. *Int J Dermatol 49*, 4 (2010), 362–76. 4
- [140] RAJADHYAKSHA, M., GONZALEZ, S., ZAVISLAN, J., ANDERSON, R., AND WEBB, R. In vivo confocal scanning laser microscopy of human skin ii: advances in instrumentation and comparison with histology. *The Journal of investigative dermatology 113*, 3 (1999), 293–303. 22

REFERENCES

- [141] RALLAN, D., BUSH, N., BAMBER, J., AND HARLAND, C. Quantitative discrimination of pigmented lesions using three-dimensional high-resolution ultrasound reflex transmission imaging. *J Invest Dermatol* (2006). 31
- [142] RALLAN, D., DICKSON, M., BUSH, N., HARLAND, C., MORTIMER, P., AND BAMBER, J. High-resolution ultrasound reflex transmission imaging and digital photography: potential tools for the quantitative assessment of pigmented lesions. *Skin Res Technol* 12, 1 (2006), 50–9. 31
- [143] REDDY, B., AND CHATTERJI, B. An fft-based technique for translation, rotation, and scale-invariant image registration. *IEEE Transactions on Image Processing* 5 (1996), 1266–1271. 68
- [144] REMONDINO, F. Human body reconstruction from image sequences. In *Pattern Recognition (DAGM 2002), Lecture Notes in Computer Science 2449* (2002), Springer, pp. 50–57. 130
- [145] ROBERTO, M., FUSIELLO, A., AND MURINO, V. High resolution video mosaicing with global alignment. *cvpr 01* (2004), 692–698. 99
- [146] ROSTEN, E., AND DRUMMOND, T. Machine learning for high-speed corner detection. In *European Conference on Computer Vision* (May 2006), vol. 1, pp. 430–443. 72
- [147] ROULLOT, E. A unifying framework for color image calibration. *15th International Conference on Systems, Signals and Image Processing (IWSSIP)* (May 2008), 97–100. 36
- [148] ROUSSEEUW, P. Least median of squares regression. 871–880. 95
- [149] SATO, Y., KITAGAWAS, H., AND FUJITA, H. Shape measurement of curved objects using multiple slit-ray projections. *IEEE Transactions on Pattern Analysis and Machine Intelligence PAMI* 4, 6 (1982), 641–646. 128
- [150] SAWHNEY, H., HSU, S., AND KUMAR, R. Robust video mosaicing through topology inference and local to global alignment. In *ECCV '98: Proceedings of the 5th European Conference on Computer Vision-Volume II* (London, UK, 1998), Springer-Verlag, pp. 103–119. 99

REFERENCES

- [151] SCHAEFER, G., RAJAB, M., CELEBI, M., AND IYATOMI, H. *Skin Lesion Extraction in Dermoscopic Images Based on Colour Enhancement and Iterative Segmentation*. 2009. 39
- [152] SCHAEFER, G., RAJAB, M., CELEBI, M., AND IYATOMI, H. Colour and contrast enhancement for improved skin lesion segmentation. *Computerized medical imaging and graphics* 35, 2 (2011), 99–104. 39
- [153] SERUP, J. High tech dermatology. *Eur J Dermatol* 17, 2 (2007), 178–180. 10
- [154] SHI, J., AND TOMASI, C. Good features to track. *Computer Vision and Pattern Recognition, 1994. Proceedings CVPR '94., 1994 IEEE Computer Society Conference on* (jun 1994), 593–600. 72
- [155] SILVA, N., KURAGO, Z., POLVERINI, P., HANKS, C., AND PAULINO, A. Malignant melanoma of the oral mucosa in a 17-year-old adolescent girl. *Arch Pathol Lab Med* (2002), 1110–3. 30
- [156] SINHA, S., FRAHM, J., AND GENÇ, Y. Gpu-based video feature tracking and matching. In *EDGE 2006: workshop on Edge Computing Using New Commodity Architectures* (Chapel Hill, 2006). 108
- [157] SINHA, S., FRAHM, J., AND GENÇ, Y. Feature tracking and matching in video using programmable graphics hardware. 1–11. 109
- [158] SINHA, S., FRAHM, J., AND POLLEFEYS, M. Gpu-based video feature tracking and matching. Tech. rep., University of North Carolina at Chapel Hill, 2006. 108
- [159] SMITH, S. M., AND BRADY, J. M. Susan - a new approach to low level image processing. *International Journal of Computer Vision* 23 (1995), 45–78. 72
- [160] SOCIETY, T. A. C. Cancer facts & figures, 2010. 1
- [161] SOLUTIONS, D. V. 4d view solutions. <http://r24085.ovh.net>, feb 2012. 131

REFERENCES

- [162] SOLUTIONS, H. Human solutions. <http://www.human-solutions.com>, feb 2012. 128, 131
- [163] STOLZ, W., BRAUN-FALCO, O., BILEK, P., LANDTHALER, M., BURGFORF, W., AND COGNETTA, A. *Color atlas of dermoscopy*. Blackwell Science, 2002. 12
- [164] STOLZ, W., RIEMAN, A., AND COGNETTA, A. Abcd rule of dermoscopy: a new practical method for early recognition of malignant melanoma. *European Journal of Dermatology* 4 (1994), 521–527. 15
- [165] STUCKER, M., ESSER, M., HOFFMANN, M., MEMMEL, U., HIRSCHMULLER, A., BORMANN, C. V., HOFFMANN, K., AND ALTMAYER, P. High-resolution laser doppler perfusion imaging aids in differentiating between benign and malignant melanocytic skin tumours. *Acta Derm Venereol* 82, 1 (2002), 25–9. 28
- [166] STUCKER, M., HORSTMANN, I., NUCHEL, C., ROCHLING, A., HOFFMANN, K., AND ALTMAYER, P. Blood flow compared in benign melanocytic naevi, malignant melanomas and basal cell carcinomas. *Clin Exp Dermatol* 24, 2 (1999), 107–11. 28
- [167] SZELISKI, R. Image mosaicing for tele-reality applications. In *IEEE Workshop on Applications of Computer Vision* (1994), pp. 44–53. 65, 67
- [168] SZELISKI, R., AND SHUM, H. Creating full view panoramic image mosaics and environment maps. In *International Conference on Computer Graphics and Interactive Techniques* (1997), pp. 251–258. 65
- [169] TADROUS, P. Methods for imaging the structure and function of living tissues and cells: 1. optical coherence tomography. *The Journal of Pathology* 191, 2 (June 200), 115–119. 26
- [170] TELMAT. Telmat industrie. <http://telmat.com>, feb 2012. 130
- [171] TEXTILE, AND CORPORATION, C. T. Textile and clothing technology corporation. <http://www.tc2.com>, feb 2012. 130

REFERENCES

- [172] TOMMASINI, T., FUSIELLO, A., TRUCCO, E., AND ROBERTO, V. Making good features track better. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '98)* (Santa Barbara, CA, June 1998), IEEE Computer Society Press, pp. 178–183. 83
- [173] TORR, P., AND ZISSERMAN, A. Feature based methods for structure and motion estimation. In *Vision Algorithms: Theory and Practice, number 1883 in LNCS* (2000), Springer-Verlag, pp. 278–295. 90
- [174] ULRICH, M., STOCKFLETH, E., ROEWERT-HUBER, J., AND ASTNER, S. Noninvasive diagnostic tools for nonmelanoma skin cancer. *The British journal of dermatology* 157 Suppl 2 (2007), 56–58. 24, 25
- [175] UYTTENDAELE, M., EDEN, A., AND SZELISKI, R. Eliminating ghosting and exposure artifacts in image mosaics. In *Proc. Int. Conf. on Comp. Vision and Patt. Recog. (CVPR2001)* (2001), pp. 509–516. 103
- [176] VEIRO, J., AND CUMMINS, P. Imaging of skin epidermis from various origins using confocal laser scanning microscopy. *Dermatology* 189 (1994), 16–22. 26
- [177] VESTERGAARD, M., MACASKILL, P., HOLT, P., AND MENZIES, S. Dermoscopy compared with naked eye examination for the diagnosis of primary melanoma: a meta-analysis of studies performed in a clinical setting. *British Journal of Dermatology* 9999, 9999 (2008). 125
- [178] VITRONIC. Vitronic inc. <http://www.vitronic.de>, feb 2012. 128
- [179] VOIGT, H., AND CLASSEN, R. Topodermatographic image analysis for melanoma screening and the quantitative assessment of tumor dimension parameters of the skin. *Cancer* 75, 4 (1995), 981–8. 11
- [180] WELZEL, J., LANKENAU, E., HTTMANN, G., AND BIRNGRUBER, R. Oct in dermatology. In *Optical Coherence Tomography*, W. Drexler and J. G. Fujimoto, Eds., Biological and Medical Physics, Biomedical Engineering. Springer Berlin Heidelberg, 2008, pp. 1103–1122. 25

REFERENCES

- [181] WIGGER-ALBERTI, W., AND ELSNER, P. [fluorescence with wood's light. current applications in dermatologic diagnosis, therapy follow-up and prevention]. *Hautarzt* 48, 8 (1997), 523–7. 9
- [182] WIGHTON, P., LEE, T., LUI, H., MCLEAN, D., AND ATKINS, M. Chromatic aberration correction: an enhancement to the calibration of low-cost digital dermoscopes. *Skin research and technology* (2011). 38, 62, 63
- [183] WILLIAMS, H., AND A., P. Sniffer dogs in the melanoma clinic? *Lancet*. 1, 8640 (Apr 1989), 734. 9
- [184] WILLIS, C., CHURCH, S., GUEST, C., COOK, W., MCCARTHY, N., BRANSBURY, A., CHURCH, M., AND CHURCH, J. Olfactory detection of human bladder cancer by dogs: proof of principle study. *BMJ* 329, 7468 (2004), 712. 10
- [185] WOLBERG, G., AND ZOKAI, S. Robust image registration using log-polar transform. In *In Proc. IEEE Int. Conf. image processing* (2000), pp. 493–496. 68
- [186] WONG, C., STRANGE, R., AND LEAR, J. Basal cell carcinoma. *BMJ* 327, 7418 (2003), 794–8. 4
- [187] WOOD, R. Secret communications concerning light rays. *Journal of Physiology* (1919), 5e serie: t IX. 9
- [188] WU, C. SiftGPU: A GPU implementation of scale invariant feature transform (SIFT). <http://cs.unc.edu/~ccwu/siftgpu>, 2007. 107
- [189] YOUNG, C. Solar ultraviolet radiation and skin cancer. *Occupational Medicine* 59, 2 (2009), 82–88. 4
- [190] ZITOVA, B., AND FLUSSER, J. Image registration methods: a survey. *Image and Vision Computing* 21, 11 (October 2003), 977–1000. 71