

**HÉLDER FILIPE PINTO DE OLIVEIRA**

**An Affordable and Practical 3D Solution for the  
Aesthetic Evaluation of Breast Cancer  
Conservative Treatment**

**Ph.D. Thesis**

**FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO  
January 2013**





**FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO**



# **An Affordable and Practical 3D Solution for the Aesthetic Evaluation of Breast Cancer Conservative Treatment**

**Hélder Filipe Pinto de Oliveira**

Programa Doutoral em Engenharia Eletrotécnica e de Computadores

Supervisor: Jaime dos Santos Cardoso (PhD)

Co-supervisor: Maria João Viseu Botelho Cardoso (MD, PhD)

January 2013



# Resumo

O cancro de mama continua a ser uma das doenças malignas mais mediatizadas, não só devido à sua alta incidência e prevalência, mas também devido à implicação de que a mama tem na imagem maternal e sexual da mulher. A introdução do tratamento conservador do cancro da mama (BCCT), em que o tumor é extraído juntamente com uma margem de tecido mamário macroscopicamente livres de doença, tem permitido um controle local da doença, com um nível de sobrevida semelhante à obtida com uma mastectomia, mas com um melhor resultado estético. O resultado estético e a qualidade de vida dos pacientes têm sido um objetivo fundamental do tratamento. No entanto, ao contrário dos resultados oncológicos, o resultado estético depende de vários fatores, muitos deles difíceis de avaliar o que contribui para uma significativa heterogeneidade do resultado.

De modo a superar as limitações de reprodutibilidade e objetividade dos métodos de avaliação do resultado estético do BCCT, um software (BCCT.core) foi recentemente desenvolvido. Este software é baseado na comparação entre a mama tratada e não tratada em fotografias digitais. O sistema permite a medição semiautomática de vários índices relacionados com o resultado estético após cirurgia, tornando a avaliação rápida, fácil e reprodutível. O conjunto de medidas é automaticamente convertido numa avaliação global e objetiva do resultado estético, baseado num algoritmo usando técnicas de aprendizagem automática, especialmente adaptado para classes ordinais. Este algoritmo foi treinado e aperfeiçoado para prever o resultado estético, numa escala de quatro categorias (excelente, bom, razoável e mau). A avaliação objetiva fornecida pelo software é representativa do resultado estético global, permitindo a medição de vários fatores que influenciam significativamente o resultado estético, contribuindo para o melhoramento das técnicas atuais de tratamento e para a comparação efetiva entre diferentes equipas de trabalho.

Embora o BCCT.core apresente resultados satisfatórios, apresenta uma limitação importante. A mama é um objeto tridimensional (3D) complexo e os seus limites não são claramente definidos sendo desta forma difícil identificar os seus pontos anatómicos. Continuamente têm sido desenvolvidas novas técnicas cirúrgicas e equipamentos de modo a otimizar o desempenho dos cirurgiões em procedimentos relacionados com as intervenções mamárias, aumentando a relevância das aplicações utilizadas para avaliar automaticamente o resultado estético através de informação 3D. Vantagens inerentes a este processo incluem a capacidade de visualizar a mama de um diferente número de ângulos, de modo a estimar diferenças de volume e para planear procedimentos cirúrgicos. Vários investigadores têm utilizado tecnologia 3D, contudo estas metodologias são bastante dispendiosas, não são adequados para avaliar automaticamente o resultados estético, e não são uma opção viável para a prática clínica diária.

Este trabalho tem como objetivo desenvolver uma ferramenta bastante prática e fácil de usar, baseada num algoritmo computacional de reconstruções 3D a partir dispositivos de baixo custo, tais como: reconstrução a partir de imagens simples, num ambiente não calibrado, usando geometria epipolar e dispositivos RGB-D. O objetivo é permitir a deteção automática de pontos característicos nas imagens e extrair informação volumétrica, auxiliando na avaliação global do

resultado estético. Os resultados experimentais mostram a adequação da nossa abordagem baseada em mapas de profundidade. O algoritmo proposto provou ser preciso e robusto para um elevado número de pacientes. Além disso, em comparação com os estudos anteriores, o processo para a detecção de pontos característicos foi completamente automatizada.

# Abstract

Breast cancer remains one of the most mediated malignant diseases, not only due to its high incidence and prevalence, but also because of the implication that the breast has on women's maternal and sexual images. The introduction of Breast Cancer Conservative Treatment (BCCT), in which the tumour is excised together with a margin of breast tissue macroscopically free of disease, has allowed a local control of the disease, with a survival similar to that obtained with a mastectomy, but with a better aesthetic result. The cosmetic outcome and Quality of Life (QOL) of patients have been a fundamental objective of such treatment. However, contrarily to oncological results, the cosmetic result depends on several factors, many of them difficult to assess which contributes to a significant heterogeneity in result.

To overcome the limitations of reproducibility and objectivity of the methods in order to evaluate the aesthetic result of BCCT, a recent computer system named Breast Cancer Conservative Treatment.cosmetic result (BCCT.core) was developed. This software used digital images, and is based on the comparison of the treated and non-treated breast. The system introduces the semi-automatic measurement of various indices related to the surgical aesthetic result, making the evaluation fast, easy and reproducible. The set of measures is automatically converted into a comprehensive and objective evaluation of cosmetic results, based on a machine learning algorithm, especially adapted to predict ordinal classes. This algorithm was trained and optimised to predict the overall aesthetic result on a scale of four classes (excellent, good, fair and poor). The objective assessment provided by the system is a summary of the overall aesthetic result, allowing an assessment of several factors that significantly influence the aesthetic result and this contributes to the improvement of current treatment techniques and to the effective comparison between different work teams.

Although the BCCT.core system presents satisfactory results, there is an important limitation on this software. The female breast is a complex three-dimensional (3D) object and its boundaries are rather fuzzily defined, and thus it is difficult to identify anatomical landmarks. New surgical techniques and devices are continuously developed to optimize surgeons' performances in procedures related to breast interventions, which increases the relevance of applications used to automatically evaluate the aesthetic result using 3D information. Potential advantages include the ability to view the breast from a different number of angles, to estimate volume/volume deficit and to plan future surgeries. Several research groups have made attempts with 3D technology, however these methodologies are very expensive, they are not suitable to automatically evaluate aesthetic results and they are not a feasible option for daily clinical practice.

This work aims at producing a practical tool that is easy to use, based on a computational algorithm that performs 3D reconstructions from low-cost systems, namely: reconstruction algorithm from uncalibrated views, using an epipolar geometry approach and RGB-D sensors. The aim is to enable a complete automatic detection of prominent points in BCCT images and extract volumetric information, helping to the overall aesthetic assessment. Experimental results show the suitability of our depth-map based approach. The proposed algorithm has proved to be accurate

and robust for a wide number of patients. Additionally, in comparison with previous research, the procedure for detecting prominent points was completely automated.

# Acknowledgments

I would like to express my sincere gratitude to my advisor Professor Jaime dos Santos Cardoso for the continuous support of my Ph.D study and research, for his patience, motivation, enthusiasm and immense knowledge. His guidance helped me in all the time of research and writing of this thesis. The joy and enthusiasm he has for his research was contagious and motivational for me, even during tough times in the Ph.D pursuit. I am also thankful for the excellent example he has provided me as a successful researcher and Professor. I could not imagine having a better advisor and mentor for my Ph.D study.

Furthermore, I am very grateful to my co-advisor Professor Maria João Cardoso, for the opportunity to develop my Ph.D study in the area of breast cancer, for the insightful comments both in my work and in this thesis, for her support and for many motivating discussions. I appreciate all her contributions of time, ideas, and funding to make my Ph.D experience productive and stimulating.

A very special thank to André Magalhães for his friendship and assistance has meant more to me than I could ever express. I could not complete my work without your invaluable partnership assistance. You were always available when I needed.

I extend my sincere thank to INESC Porto, for all the work conditions, infrastructure and technical support given to my Ph.D thesis and for the other challenges I worked in. I also want to thank the communication service of INESC Porto for the constant English revision assistance of papers submitted during the thesis. A especial thank to Rita Pacheco for the revision of English of my Ph.D thesis.

The members of the VCMi (Visual Computing and Machine Intelligence) group have contributed immensely to my personal and professional time at INESC Porto. The group has been a source of friendship as well as good advice and collaboration. I am especially grateful to: Ana Rebelo, Inês Domingues, A. Filipa Sequeira, Samaneh Khoshrou, Ricardo Sousa, Igor Amaral, Andreas Seufert, Eduardo Marques and Iago Landesa-Vázquez.

I also thank all the members of the Computer Aided Radiation Therapy group of the Laboratory of Biomedical Technologies at the Politecnico di Milano, who received me there for three months, especially to Paolo Patete and Professor Guido Baroni.

Data collection was part of my Ph.D thesis. I thank Hospital São João and Hospital da Trindade, which allowed data acquisition and to all the patients that voluntarily participated.

The work developed was not possible without the kind support provided by FCT – Fundação para a Ciência e a Tecnologia (Portuguese Foundation for Science and Technology), with Ph.D grant (Grant ID - SFRH / BD / 43772 / 2008). This work was also financed by the ERDF - European Regional Development Fund through the COMPETE Programme (operational programme for competitiveness) and by National Funds through the FCT within project PTDC/SAU-ENB/114951/2009. Many thanks to them for supporting and promoting scientific research in Portugal.

Since these four years were not only dedicated to the thesis, I would like to thank Professor Aurélio Campilho for the opportunity and trust to coordinate the BioStar group, making me grow as a researcher. I leave here a word of friendship to Filipe Magalhães, who shared the coordination of BioStar with me. Many were the challenges that we have embraced, it was a pleasure sharing them with you... Thanks Filipe.

I also want to thank the students that contributed to this thesis under my orientation: João Soares and Marco Silva. And also to the other students, who officially and non-officially, I supervised: Hélder Matos, Diogo Dias, Miguel Cova, João P. Monteiro and João C. Monteiro. Thank you all, for helping me to improve.

Lastly, I would like to thank my parents for all their love, encouragement and infinite support throughout everything. You provided me all the opportunities from the very first day of my life, I hope that this work makes you proud. And to Patricia whose love, support, encouragement and (lots and lots) of patience allowed me to finish this journey. Thank you for your support when I have needed it the most. Thank you with all my heart!

Hélder P. Oliveira



*“I have not failed.  
I’ve just found 10,000 ways that won’t work.”*

Thomas A. Edison



# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Motivation . . . . .	2
1.2	Objectives . . . . .	3
1.3	Contributions . . . . .	4
1.4	Document Structure . . . . .	6
<b>2</b>	<b>Literature Review</b>	<b>9</b>
2.1	Breast Cancer . . . . .	9
2.2	Breast Conservative Treatment . . . . .	11
2.3	Determinant Factors to Cosmetic Outcome . . . . .	12
2.4	Cosmetic Assessment Parameters . . . . .	13
2.5	Patient's Evaluation . . . . .	14
2.6	Scale of Assessment . . . . .	14
2.7	Methods of Assessment . . . . .	15
2.7.1	Subjective Methods . . . . .	15
2.7.2	Objective Methods . . . . .	16
2.8	Objective Assessment of Breast Aesthetic Outcome . . . . .	17
2.8.1	Two-dimensional Patient Data . . . . .	17
2.8.2	Three-dimensional Methods . . . . .	22
2.9	Public Databases . . . . .	29
2.10	Conclusion . . . . .	31
<b>3</b>	<b>BCCT.core Software</b>	<b>33</b>
3.1	BCCT.core Design . . . . .	33
3.2	Software Automation . . . . .	33
3.3	Feature Definition . . . . .	34
3.3.1	Asymmetry Features . . . . .	34
3.3.2	Colour Difference Features . . . . .	36
3.3.3	Scar Visibility Features . . . . .	37
3.4	The Aesthetic Classifier . . . . .	39
3.5	Conclusion . . . . .	39
<b>4</b>	<b>Database</b>	<b>41</b>
4.1	Database Requirements . . . . .	41
4.2	Database Construction . . . . .	42
4.3	Conclusion . . . . .	47

<b>5</b>	<b>Closed Shortest Path in Images</b>	<b>49</b>
5.1	Regional Detectors . . . . .	49
5.2	Closed Contour Related Works . . . . .	50
5.3	Conventional Closed Shortest Path Algorithms . . . . .	51
5.3.1	Graph Concepts . . . . .	52
5.3.2	Design of the Weight Function . . . . .	53
5.3.3	(Open) Shortest Path Extraction Algorithm in a DAG . . . . .	54
5.3.4	Closed Shortest Path Extraction Algorithm in Polar Coordinates . . . . .	56
5.4	Closed Shortest Path Algorithm in the Original Coordinates . . . . .	58
5.4.1	Linearization of the DAG . . . . .	58
5.4.2	Edge Cost Modulation . . . . .	60
5.4.3	Distance Between Nodes . . . . .	60
5.5	Experimental Results . . . . .	61
5.5.1	Illustrative Examples . . . . .	61
5.5.2	Mass Segmentation . . . . .	63
5.6	Conclusions . . . . .	67
<b>6</b>	<b>Interpretability Models and Lateral-views for the Aesthetic Result Assessment</b>	<b>69</b>
6.1	Interpretable Models . . . . .	69
6.1.1	Study Population and a Gold Standard . . . . .	70
6.1.2	Feature Selection . . . . .	70
6.1.3	Dealing with Decision Costs . . . . .	71
6.1.4	Results . . . . .	72
6.2	Improving the Model with Lateral Information . . . . .	75
6.2.1	Study Population and a Gold Standard . . . . .	76
6.2.2	Feature Definition . . . . .	77
6.2.3	Classifier . . . . .	79
6.2.4	Results . . . . .	80
6.3	Conclusions . . . . .	81
<b>7</b>	<b>Uncalibrated-based 3D Reconstruction</b>	<b>83</b>
7.1	General Uncalibrated Reconstruction Workflow . . . . .	84
7.2	Camera Model . . . . .	85
7.3	Binocular Disparity and Stereo Vision . . . . .	88
7.4	Feature Detection . . . . .	91
7.4.1	Corner Detectors . . . . .	91
7.4.2	Region Detectors . . . . .	92
7.4.3	Edge Detectors . . . . .	92
7.4.4	Efficient Feature Detectors . . . . .	93
7.5	Feature Descriptors . . . . .	94
7.6	Rectification . . . . .	95
7.6.1	Based on Known Objects . . . . .	95
7.6.2	Uncalibrated Rectification . . . . .	95
7.7	Stereo Matching . . . . .	99
7.8	Experimental Work . . . . .	103
7.8.1	Test images . . . . .	104
7.8.2	Feature Point Detection . . . . .	106
7.8.3	Image Rectification . . . . .	107
7.8.4	Disparity Map Generation . . . . .	107

7.9	Results . . . . .	108
7.9.1	Feature Point Detection and Matching . . . . .	108
7.9.2	Rectification . . . . .	109
7.9.3	Stereo matching . . . . .	112
7.9.4	Real Scenario . . . . .	114
7.9.5	A Simplified Scenario . . . . .	117
7.10	Conclusion . . . . .	122
<b>8</b>	<b>Depth-map Based Methodologies</b>	<b>125</b>
8.1	Depth-map Sensor . . . . .	126
8.1.1	Hardware . . . . .	126
8.1.2	Software . . . . .	127
8.2	The Kinect Sensor as a Tool . . . . .	128
8.3	Validation of the Kinect as a Tool to Assess Aesthetic Outcomes . . . . .	131
8.3.1	Transformation of Raw to Metric Data . . . . .	131
8.3.2	Kinect Sensitivity Assessment . . . . .	133
8.3.3	Kinect Measurements on a Female Phantom Torso . . . . .	136
8.4	Super-Resolution Application . . . . .	138
8.4.1	Super-Resolution Applied to Depth Maps . . . . .	140
8.4.2	Performance Tests . . . . .	142
8.4.3	Performance Evaluation . . . . .	142
8.4.4	Results . . . . .	143
8.4.5	Discussion . . . . .	148
8.5	Joint RGB-D Calibration . . . . .	148
8.5.1	Calibration Method . . . . .	151
8.6	Conclusion . . . . .	153
<b>9</b>	<b>A 3D Low-cost Solution for the Aesthetic Result Assessment</b>	<b>157</b>
9.1	A Kinect Based System for the Aesthetic Evaluation of BCCT . . . . .	157
9.1.1	Overview . . . . .	158
9.2	A Joint RGB-D Based Detection of Prominent Points . . . . .	159
9.2.1	Simultaneous Detection of Breast Contour and Breast Peak Points . . . . .	159
9.2.2	Sternal Notch Detection . . . . .	168
9.2.3	Nipple Detection . . . . .	168
9.3	Volumetric Feature Definition . . . . .	170
9.4	Aesthetic Model . . . . .	171
9.5	Results . . . . .	172
9.5.1	Breast Peak Detection . . . . .	173
9.5.2	Breast Contour Endpoints Detection . . . . .	173
9.5.3	Breast Contour Detection . . . . .	174
9.5.4	Nipple Detection . . . . .	175
9.5.5	Volumetric Feature Definition . . . . .	178
9.5.6	Aesthetic Evaluation . . . . .	179
9.6	Conclusion . . . . .	184
<b>10</b>	<b>Conclusions</b>	<b>187</b>
10.1	Future Work . . . . .	190
	<b>References</b>	<b>193</b>



# List of Figures

2.1	Illustration of the breast surgical procedure. . . . .	10
2.2	Surgical procedure comparison. . . . .	11
2.3	Timeline of some of the most important methodologies for the objective evaluation of the aesthetic outcome. . . . .	17
2.4	BAT software interface. . . . .	20
2.5	BCCT.core software interface. . . . .	21
2.6	Application example from Moyer tool. . . . .	24
2.7	Application example from Farinella tool. . . . .	25
2.8	Application example from Balaniuk tool. . . . .	25
2.9	Application example from Catanuto tool. . . . .	27
2.10	Application example from Henseler tool. . . . .	27
2.11	Patient photograph and generated disparity map. . . . .	28
2.12	Breast Research Group Database - Patient acquisition - Different positions. . . .	30
2.13	Breast Research Group Database - Manually annotation using BCCT.core software.	31
3.1	BCCT.core software interface. . . . .	34
3.2	Illustration of measurements for asymmetry features. . . . .	35
3.3	Breast mask for the extraction of colour features. . . . .	37
3.4	Breast sectors for the extraction of local colour changes induced by the surgical scar.	38
4.1	Data acquired with a portable camera. . . . .	42
4.2	Data acquired using the Kinect device. . . . .	45
4.3	Kinect data ground truth annotation. . . . .	47
4.4	Portable camera data ground truth annotation. . . . .	47
4.5	Nipple height measurement. . . . .	48
5.1	Example of a DAG with several topological orderings. . . . .	53
5.2	Examples of neighbourhoods and corresponding causal neighbours. . . . .	54
5.3	The nodes represent pixels and arcs represent connections between pixels. . . . .	54
5.4	A forbidden path in the digraph. . . . .	55
5.5	Exemplification of shortest paths. . . . .	57
5.6	Two paths intersecting. . . . .	58
5.7	Illustration of the order of the pixels in the linearization approach. . . . .	59
5.8	Illustration of causal neighbours for a few nodes. . . . .	59
5.9	Two closed paths enclosing the centre of the ROI. . . . .	60
5.10	Two movements with different characteristics. . . . .	61
5.11	Illustration of some toy examples. . . . .	62
5.12	Impact of the seed position in final result. . . . .	63

5.13	Illustration of the application of the proposed algorithm in randomly selected real images. . . . .	64
5.14	Illustration of the conventional transform to polar coordinates in randomly selected real images. . . . .	65
5.15	Illustration of some good and bad mass segmentation on Mammograms. . . . .	66
6.1	Typical lateral photographs. . . . .	76
6.2	Illustration showing fiducial points and measurements. . . . .	77
6.3	Classifier flowchart. . . . .	80
7.1	Stereo reconstruction generic workflow. . . . .	85
7.2	Pinhole camera model. . . . .	85
7.3	Simplified pinhole camera model. . . . .	86
7.4	Simplified camera and lens example. . . . .	86
7.5	Geometric model for two-view geometry based on the pinhole camera model. . . . .	88
7.6	Disparity search summary assuming a low baseline. . . . .	90
7.7	Disparity and distance relation. . . . .	90
7.8	Difference of Gaussian space-scale pyramid visual description. . . . .	93
7.9	Results from the rectification algorithm developed by Loop and Zhang. . . . .	96
7.10	Uncalibrated rectification result from Hartley and Zisserman. . . . .	97
7.11	Rectification examples by Fusiello and Irsara. . . . .	98
7.12	Uncertainty in stereo matching. . . . .	99
7.13	Stereo matching results from Kolmogorov and Zabih algorithm. . . . .	100
7.14	Results from the algorithms presented by Šára. . . . .	100
7.15	Experimental results from the algorithm presented by Čech and Šára. . . . .	101
7.16	Results of the stereo matching algorithm by Bhatti and Nahavandi. . . . .	103
7.17	Experimental results from the stereo matching algorithm by Wang and Road. . . . .	104
7.18	Experimental results from the stereo matching algorithm provided by Klaus. . . . .	104
7.19	Classical result set for the stereo matching algorithm by Geiger. . . . .	105
7.20	Initial test images, from the <i>Middlebury</i> dataset. . . . .	105
7.21	Virtual models for large ambiguity tests. . . . .	106
7.22	Dummy model for more real tests, with induced volume difference in the breasts. . . . .	107
7.23	Feature detection performed in the <i>Aloe</i> and <i>Flowerpot</i> . . . . .	109
7.24	Feature detection performed in the three artificial models. . . . .	110
7.25	Best results from the rectification algorithms. . . . .	112
7.26	Results for stereo matching algorithms running over the 3D coloured model. . . . .	113
7.27	Results for stereo matching algorithms running over the Dummy model. . . . .	114
7.28	Original patient pairs, prior to processing. . . . .	115
7.29	Rectification results for the patient pairs. . . . .	116
7.30	Results for stereo matching algorithms running over photographs of Patient A. . . . .	117
7.31	Results for stereo matching algorithms running over photographs of Patient B. . . . .	118
7.32	Female phantom torso. . . . .	119
7.33	Acquisition system from Patete methodology. . . . .	119
7.34	Identification of corresponding points from the female phantom torso. . . . .	120
7.35	Rectified images from the female phantom torso. . . . .	120
7.36	Surface fitting model. . . . .	122
8.1	Prime sense design. . . . .	126
8.2	Avatar reconstruction using the algorithm by Zollhofer. . . . .	128



8.3	Results example of mannequin using the algorithm by Cui and Stricker. . . . .	129
8.4	Widely varying poses from Weiss work. . . . .	130
8.5	Surfel update demonstration by Henry. . . . .	130
8.6	Overview of the reconstruction algorithm using the application by Tong. . . . .	131
8.7	General pipeline using the system by Wang. . . . .	131
8.8	Kinect disparity map from the test scene. . . . .	132
8.9	Kinect capture setup. . . . .	133
8.10	Kinect depth curve <i>vs</i> distance. . . . .	134
8.11	Kinect sensitivity analysis at extreme distances. . . . .	135
8.12	Detailed Kinect sensitivity analysis. . . . .	136
8.13	Detailed Kinect sensitivity analysis of blocks spaced 0.2cm and 0.6cm. . . . .	137
8.14	Different acquisition of phantom with Kinect. . . . .	138
8.15	Aliasing in high frequency zones in the ISO 12233 chart. . . . .	139
8.16	Experimental results using the methodology by Yang. . . . .	141
8.17	Experimental results using the methodology by Garro. . . . .	141
8.18	Results from the application of Super-Resolution algorithms to an example depth image. . . . .	145
8.19	Example of application of Super-Resolution in synthetic images. . . . .	146
8.20	Example of application of Super-Resolution in RGB images. . . . .	147
8.21	Samples from the calibration images. . . . .	150
8.22	Reference frames and transformations presented on the scene. . . . .	150
8.23	Flowchart from the calibration methodology used. . . . .	152
8.24	Measurement uncertainty for varying depths for the two Kinect devices. . . . .	153
8.25	Reprojection error after the Kinect device is calibrated. . . . .	154
8.26	Depth and RGB data overlap. . . . .	154
8.27	A 3D perspective of the scene after correct calibration. . . . .	155
9.1	Microsoft Kinect data for patient#62. . . . .	158
9.2	System flowchart for the joint RGB-D based system. . . . .	158
9.3	Algorithm flowchart for the simultaneous detection of breast contour and breast peak points. . . . .	160
9.4	Adaptive Background Segmentation . . . . .	162
9.5	Toy example of different gradient shapes. . . . .	162
9.6	Breast peak candidate detection. . . . .	163
9.7	Graphical demonstration of correlation approach. . . . .	163
9.8	Brest contour detection. . . . .	166
9.9	External endpoint detection. . . . .	167
9.10	Complete contour detection and internal endpoint definition. . . . .	167
9.11	Neck position detection. . . . .	168
9.12	Nipple detection. a) Harris corner detection inside the breast; b) Areola detection (true corner); c) Areola detection (false corner). . . . .	169
9.13	Volumetric measurements. . . . .	172
9.14	Cascade classifier flowchart. . . . .	172
9.15	Breast peak detection examples. . . . .	174
9.16	Contour endpoints detection examples. . . . .	175
9.17	Complete contour detection examples. . . . .	176
9.18	Nipple detection examples. . . . .	177
9.19	Scatter plot of depth height ratios. . . . .	178
9.20	Patient examples. . . . .	185



# List of Tables

2.1	Patient, tumour and treatment-related factors . . . . .	13
2.2	Comparison of the most significant 2d technologies involved on breast cancer cosmetic evaluation. . . . .	18
2.3	Comparison of the most significant 3D technologies involved on Breast Cancer cosmetic evaluation. . . . .	23
2.4	Breast Research Group Database - Distribution of the patients over the four classes.	31
4.1	Patients information provided in the database. . . . .	45
4.2	Distribution of the 130 patients over the four classes. . . . .	45
4.3	Disagreement between evaluators. . . . .	46
5.1	Results of mass segmentation on Mammograms. . . . .	66
5.2	Mass segmentation on Mammograms: results for different perturbations on the seed localization. Results are presented in the format “mean (standard deviation)”. . . . .	67
6.1	Distribution of the 143 patients over the four classes. . . . .	70
6.2	Correlation coefficient and average value of each class on feature selection procedure. . . . .	72
6.3	Misclassification error using the SVM classifier. . . . .	73
6.4	Misclassification error using the Decision Tree classifier. . . . .	73
6.5	Confusion matrices for best classifiers, in the test set, using the cost matrix. . . . .	74
6.6	Confusion matrices for best classifiers, in the test set, using the equal misclassification costs. . . . .	75
6.7	Distribution of the 63 patients over the four classes. . . . .	76
6.8	Misclassification error using the lateral features (the preliminary misclassification error for the frontal features was 0.35 and 0.57 for the weighted error). . . . .	80
6.9	Confusion matrices for best classifiers, in the test set, for the different models. (frontal information only and frontal plus lateral information) . . . . .	81
7.1	Feature detection results for the testbed and some evaluation. . . . .	108
7.2	Rectification results and comparison. . . . .	112
7.3	Rectification values for the Patient photographs using the <i>SIFT</i> . . . . .	115
7.4	Mean square error of 3D fitting models. . . . .	121
7.5	Nipples measurements taken from the 3D reference model. . . . .	122
8.1	Comparison of different disparity transformation functions. . . . .	133
8.2	Validation of transformation functions. . . . .	133
8.3	Ratio values obtained with Kinect using the female phantom torso. . . . .	138
8.4	Super-Resolution tested programs. . . . .	142

8.5	Performance Overview . . . . .	144
8.6	Synthetic Image Performance Overview . . . . .	146
8.7	RGB Performance Overview . . . . .	147
8.8	Reprojection error after Kinect device calibration. . . . .	153
9.1	Fitting parameters. . . . .	161
9.2	Breast peak point detection error (in pixels). . . . .	173
9.3	Endpoint detection error (in pixels). . . . .	174
9.4	Breast contour detection error (in pixels). . . . .	174
9.5	Miss classification error (cross-validation). . . . .	176
9.6	Average distance between the manually marked nipple and the predicted nipple in cm (Results are presented in the format "mean (standard deviation)", and miss classification error. . . . .	177
9.7	Volumetric feature values. . . . .	179
9.8	Distribution of the 70 patients over the four classes. . . . .	180
9.9	Distribution of the 70 patients over the three classes. . . . .	180
9.10	Disagreement between evaluators for the subjective evaluation of the 70 patients. . . . .	180
9.11	Distribution of the 70 patients for BCCT.core and linear model. . . . .	181
9.12	Disagreement between output from BCCT.core and linear model. . . . .	181
9.13	Disagreement between output from previous model and the subjective evaluation performed by the physicians with 4 classes. . . . .	181
9.14	Disagreement between output from previous model and the subjective evaluation performed by the physicians with 3 classes. . . . .	181
9.15	Results obtained considering 2D features only. . . . .	183
9.16	Results obtained for the cascade scheme. . . . .	183
9.17	Results obtained for the non-restricted scheme of 2D and 3D features. . . . .	183
9.18	Selected features. . . . .	183
9.19	Impact of 3D features. . . . .	185

# Acronyms

$\chi^2$	Chi-Squared Test
2D	Two-dimensional
3D	Three-dimensional
3DNRA	3D Nipple Retraction Assessment
BAD	Breast Area Difference
BAT	Breast Analysing Tool
BCCT	Breast Cancer Conservative Treatment
BCCT.core	Breast Cancer Conservative Treatment.cosmetic result
BCD	Breast Contour Difference
BCE	Breast Compliance Evaluation
BOD	Breast Overlap Difference
BRA	Breast Retraction Assessment
BSA 0.1	Breast Shape Analyzer 0.1
BSI	Breast Symmetry Index
$c\chi_L^2$	$c\chi^2$ distance between the histograms in the $L$ channel of the CIE L*a*b*
$c\chi_a^2$	$c\chi^2$ distance between the histograms in the $a$ channel of the CIE L*a*b*
$c\chi_b^2$	$c\chi^2$ distance between the histograms in the $b$ channel of the CIE L*a*b*
$c\chi_{Lab}^2$	$c\chi^2$ distance between the histograms in the 3 channels of the CIE L*a*b*
CAD	Computer-Aided Design
$cEMD_L$	EMD distance between the histograms in the $L$ channel of the CIE L*a*b*
$cEMD_a$	EMD distance between the histograms in the $a$ channel of the CIE L*a*b*
$cEMD_b$	EMD distance between the histograms in the $b$ channel of the CIE L*a*b*
$cEMD_{Lab}$	EMD distance between the histograms in the 3 channels of the CIE L*a*b*
CIE L*a*b*	CIE XYZ Color Space

DAG	Directed acyclic graph
DoG	Difference of Gaussian
ELAS	Efficient Large-scale Stereo
EMD	Earth Mover's Distance
FAST	Features from Accelerated Segment Test
GUI	Graphical User Interface
IBP	Iterative Back Projection
IMF	Infra-Mammary Fold
IR	Infra-Red
LBC	Lower Breast Contour
LBCE	Breast Lateral Compliance Evaluation
LBPD	Breast Ptosis Difference
LBSD	Lateral Breast Size Difference
LND <sub>bc</sub>	Nipple Lateral Distance to Breast Contour
LNDE	Nipple Lateral Distances Evaluation
LND <sub>tw</sub>	Nipple Lateral Distance to Thoracic Wall
LOO	Leave One Out
LR	Low-Resolution
MSE	Mean Square Error
MSER	Maximum Stable Extremal Regions
NAC	Nipple-Areola Complex
NHA	Nipple Height Assessment
ORSA	Optimized Random Sampling Algorithm
pBAD	Dimensionless Breast Area Difference
pBCD	Dimensionless Breast Contour Difference
pBCE	Dimensionless Breast Compliance Evaluation
pBOD	Dimensionless Breast Overlap Difference
pBRA	Dimensionless Breast Retraction Assessment
PCA	Principal Component Analysis
PGM	Portable GrayMap Format
pLBC	Dimensionless Lower Breast Contour
PSNR	Peak Signal to Noise Ratio
pUNR	Dimensionless Upward Nipple Retraction

QOL	Quality of Life
$R^2$	Square Correlation Coefficient
RANSAC	Random Sample Consensus
rLBCE	Breast Lateral Compliance Evaluation Ratio
rLBDP	Breast Ptois Difference Ratio
rLBSD	Lateral Breast Size Difference Ratio
rLNDbc	Nipple Lateral Distance to Breast Contour Ratio
rLNDE	Nipple Lateral Distances Evaluation Ratio
rLNDtw	Nipple Lateral Distance to Thoracic Wall Ratio
RBF	Radial Basis Function
ROI	Region-of-Interest
$s\chi_a^2$	surgical scar visibility measured with $c\chi^2$ in the $a$ channel of the CIE $L^*a^*b^*$
$s\chi_b^2$	$c\chi^2$ surgical scar visibility measured with $c\chi^2$ in the $b$ channel of the CIE $L^*a^*b^*$
$s\chi_L^2$	surgical scar visibility measured with $c\chi^2$ in the $L$ channel of the CIE $L^*a^*b^*$
$s\chi_{Lab3D}^2$	$c\chi^2$ surgical scar visibility measured with $c\chi^2$ in the 3 channels of the CIE $L^*a^*b^*$
SAA	Surface Area Assessment
$sEMD_a$	surgical scar visibility measured with EMD in the $a$ channel of the CIE $L^*a^*b^*$
$sEMD_b$	surgical scar visibility measured with EMD in the $b$ channel of the CIE $L^*a^*b^*$
$sEMD_L$	surgical scar visibility measured with EMD in the $L$ channel of the CIE $L^*a^*b^*$
$sEMD_{Lab}$	surgical scar visibility measured with EMD in the 3 channels of the CIE $L^*a^*b^*$
SIFT	Scale-Invariant Feature Transform
SR	Super-Resolution
SURF	Speeded Up Robust Feature
SUSAN	Smallest Univaluse Segment Assimilating Nucleus
SVM	Support Vector Machine
ToF	Time-of-Flight
UNR	Upward Nipple Retraction
VDA	Volume Difference Assessment





# Chapter 1

## Introduction

Breast cancer is considered a public health problem, and it is the most frequently diagnosed cancer and the leading cause of death in women worldwide. One in ten women will develop breast cancer at some point in their lifetime. According to recent worldwide data, this disease accounted for 23% (1.38 million) of the total new cancer cases and for 14% (458 400) of the total cancer deaths in 2008 [115]. The incidence of breast cancer has increased in most countries worldwide over the last decades due to reproductive behaviour and the use of exogenous hormones, as well as differences in weight, exercise, diet and alcohol consumption [25].

Approximately 90% of breast cancers are curable if detected in their initial stage and treated properly. It is a very frequent disease and remains one of the most publicised malignancies not only because of its high incidence and prevalence, but also because of the impact that the breast has on women's bodies, sexual and maternal images. Routine screening has been taken in order to promote early diagnosis with regular medical prevention consultations depending on the presence or absence of risk factors [173].

As far as treatment is concerned, it has progressed to the establishment of the BCCT as an alternative to the classic mutilating mastectomy. Contrary to the mastectomy where the entire breast is removed, in the conservative treatment the tumour is excised macroscopically together with a margin of cancer free breast tissue. Subsequently the patient undergoes radiotherapy to the remaining breast tissue. This conservative approach, in its various forms, has made local control of the disease possible, with survival rates similar to those obtained with mastectomy, but with better cosmetic results [78,245].

The aesthetic result and patients' Quality of Life (QOL) have become fundamental objectives in this type of treatment. However, results depend on several factors (patient, tumour, surgical technique and factors related to treatment, and of course radiotherapy) which contribute to very different outcomes [49,50].

The assessment of the cosmetic result, as a means to evaluate one of the aspects of treatment quality, has become essential to any institution that performs breast cancer treatment [38]. This evaluation makes it possible to refine the current techniques and to identify the variables with a significant impact on the final aesthetic result and with the potential for improvement [50]. With

the development of new oncoplastic techniques in breast conservation, it is even more important to have means to compare cosmetic results. With the quantification of aesthetic results, it is possible to tailor the spectrum of techniques available to individual cases without compromising the oncologic and cosmetic results [10]. However, while the oncological outcome of breast conservation procedures can easily be assessed objectively, the cosmetic outcome does not yet have an evaluation standard [38].

## 1.1 Motivation

Until recently, the most common technique used to assess the aesthetic results was the subjective assessment by one or more observers, who focus directly on the patients or on photographic representations of them. The opinion on the final aesthetic result is graded using one of several existing scales that rank the results, usually by comparing the operated breast with the untreated breast. The most frequently used evaluation scale was introduced by Harris in 1979 [96]. However, there are some problems regarding the interpretation of the results of the studies which use this type of assessment: for example, exemption is not always guaranteed since it is often performed by professionals who are involved in the treatment. Therefore, its reproducibility is not usually high [36] and when measured, the level of agreement between observers is low or moderate [40].

Objective methods were introduced in an attempt to overcome the lack of objectivity and reproducibility. These methods consisted of comparing the two breasts with simple measurements marked directly on patients or on photographs of them [50, 149]. Almost all of the measurements suggested in the literature capture the breast's asymmetry and were subjected to significant intra-observer and inter-observer variability.

The current methodologies in studies that evaluate the cosmetic outcome of BCCT continue to show a significant lack of standardisation, not only in the type of assessment used, but also in the factors included in this evaluation and the instruments used for this analysis. There was a need to replace or enhance the expert human evaluation of the aesthetic results of BCCT with a validated objective tool. This tool should be easy to use and highly reproducible and acceptable to those who would be evaluated. One of those cases is the Breast Analysing Tool (BAT) that was introduced by Fitzal *et al.* [80]. However, the most prominent tool to objectively perform the aesthetic evaluation of BCCT was introduced by Cardoso and Cardoso [28]. This computer-aided medical system, called BCCT.core, aims to overcome the acute shortage of such software systems and exploit the unique ability of computational methods to provide an effective and easy to use tool to improve the outcome of breast cancer patient care. The BCCT.core is a semi-automatic system capable of objectively evaluating the overall aesthetic results of the BCCT. The development of the BCCT.core consisted of automatically extracting several features from frontal patient's photographs, capturing some of the factors that are considered to have an impact on the overall cosmetic results: breast asymmetry, skin colour changes due to the radiotherapy treatment and the appearance of the surgical scar. In a second phase, a Machine Learning algorithm is applied

to predict the overall cosmetic result using the recorded features [31]. The objective classification outputted by this module constitutes a valuable summary of the aesthetical result, enabling an effective comparison between different medical teams and centres all over the world. While innovative and reproducible, the approach of this tool has several points that need to be addressed as is often suggested by the users. The first case is the interpretability of the model that relates the aesthetical result with the input measures; the second is the fact that only face-view photographs of the patients are used not including the information from side or oblique views; and finally the lack of the third dimension.

Now, more than ever, it is important not only to compare results after treatment, but also to predict cosmetic results before the procedure has taken place. Therefore, there is a need for a tool capable of simulating surgical options and outcomes in terms of volume, gain or loss, and of better educating patients on a more informed choice of breast surgery procedures. A more accurate and objective tool to predict surgical outcomes to guide the patient and surgeon in the decision-making or planning process is feasible using 3D imaging and surgical simulation. A simulation model also allows patients to visualise the possible outcomes of different surgical options. It is generally accepted that 3D imaging has great potential in a clinical environment, although there are factors that may influence its use in the near future. Several research groups have made efforts with 3D technology [67, 106]. However, current 3D technologies are very costly and difficult to operate, thus requiring specialised staff, and are not a feasible choice for daily clinical practice. Therefore, their widespread use in the near future is not predictable. Consequently, the search for low cost and easy to operate equipment is highly desirable.

## 1.2 Objectives

The impact of the research to be conducted in this work is reflected in its ability to improve the quality of the health care in the area of Breast Cancer Treatment.

In this work, the aim is to produce an affordable and practical tool for the aesthetic evaluation after breast cancer surgeries, developing a computational algorithm that reconstructs 3D objects using low-cost methodologies. Preliminary studies will be based on the analysis of the current limitations and on higher expectations related to the diversity of new breast interventions. Firstly, a single photograph of the patient is insufficient for an accurate pre-operation decision. As an extension to the method currently used, features extracted from side-views photographs will be used, combined with frontal features, aiming to improve the global assessment result. The work will focus on the application of 3D methods for the aesthetical evaluation of the surgical intervention result, based on 3D data reconstruction from one or more uncalibrated views of the patient and using RGB-D sensors. Secondly, the aim is to investigate the complete automatic detection of prominent points in BCCT images, making it easier to extraction of volumetric information, providing a better quantitative and reliable assessment of the aesthetic outcomes of breast surgeries, which is fundamental for the reproducibility of the method.

All the research conducted depends on the availability of training and testing examples that can be used to develop the models. Therefore, this research will be preceded by the collection of training data manually annotated with the help of an expert, creating a multi-modal database.

### 1.3 Contributions

We summarise below the contributions of the thesis for the aesthetic evaluation after breast cancer surgeries. In this thesis we have

1. created a multi-modal database, which could be used in the future by breast cancer researchers in order to compare different methodologies. The database created is a unique tool and can lead to new developments in the area;
2. proposed a new closed contour computation in the original coordinates, in contrast to the typically approach which transforms the image into polar coordinates and the closed contour is transformed into an open contour between two opposite margins;
3. investigated the application of interpretable models to BCCT.core software, in contrast with the non-linear models currently deployed in the software;
4. proposed new features extracted from side-views photographs to perform the aesthetic evaluation of the cosmetic result;
5. introduced the use of low-cost systems to acquire 3D information to the breast cancer area; and proposed new algorithms to detect fiducial points and breast contour without user intervention and new models for the aesthetic evaluation of the cosmetic result incorporating volumetric information.

#### List of Publications Related to the thesis

The work related with this thesis resulted in the publication of the followings journal papers:

- H. P. Oliveira, J. S. Cardoso, A. Magalhães, and M. J. Cardoso. A 3D Low-cost Solution for the Aesthetic Evaluation of Breast Cancer Conservative Treatment, 2013 (Submitted)
- J. S. Cardoso, H. P. Oliveira, and Inês Domingues. Closed shortest path in images: Avoiding unnecessary transformation, 2013 (Submitted)
- H. P. Oliveira, J. S. Cardoso, A. Magalhães, and M. J. Cardoso. Methods for the aesthetic evaluation of breast cancer conservation treatment: A technological review. *Current Medical Imaging Reviews*, 9(1):32–46, 2013

From this thesis resulted the following international conference papers:

- H. P. Oliveira, M. D. Silva, A. Magalhães, M. J. Cardoso, and J. S. Cardoso. Is kinect depth data accurate for the aesthetic evaluation after breast cancer surgeries? In *Proceedings of Iberian Conference on Pattern Recognition and Image Analysis*, pages 261 – 268, 2013
- H. P. Oliveira, J. S. Cardoso, A. Magalhães, and M. J. Cardoso. Simultaneous detection of prominent points on breast cancer conservative treatment images. In *Proceedings of The IEEE International Conference on Image Processing*, pages 2841 – 2844, 2012
- H. P. Oliveira, P. Patete, G. Baroni, and J. S. Cardoso. Development of a bcct quantitative 3d evaluation system through low-cost solutions. In *Proceedings of the 2<sup>nd</sup> International Conference on 3D Body Scanning Technologies*, pages 16 – 27, 2011
- R. Sousa, H. P. Oliveira, and J. S. Cardoso. Feature selection with complexity measure in a quadratic programming setting. In *Proceedings of Iberian Conference on Pattern Recognition and Image Analysis*, pages 524 – 531, 2011
- H. P. Oliveira, A. Magalhaes, M. J. Cardoso, and J. S. Cardoso. Improving the bcct.core model with lateral information. In *Proceedings of the 10<sup>th</sup> IEEE International Conference on Information Technology and Applications in Biomedicine*, pages 1 – 4, 2010
- H. P. Oliveira, A. Magalhaes, M. J. Cardoso, and J. S. Cardoso. An accurate and interpretable model for bcct.core. In *Proceedings of the 32<sup>nd</sup> Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 6158 – 6161, 2010

The work was also validated by the medical community, with the publication of the following international medical conferences papers:

- A. T. Magalhães, H. P. Oliveira, J. Soares, A. J. Moura, J. S. Cardoso, and M. J. Cardoso. 3D Model for Aesthetic Objective Evaluation after Breast Cancer Surgery using Infrared Laser Projector. In *Proceedings of the 3<sup>rd</sup> International Meeting of Oncoplastic and Reconstructive Breast Surgery*, 2011
- A. T. Magalhães, H. P. Oliveira, A. J. Moura, J. S. Cardoso, and M. J. Cardoso. Aesthetic Evaluation after Mastectomy and Immediate Breast Reconstruction using the BCCT.core Software. In *Proceedings of the 3<sup>rd</sup> International Meeting of Oncoplastic and Reconstructive Breast Surgery*, 2011

- A. T. Magalhães, H. P. Oliveira, S. Costa, J. S. Cardoso, and M. J. Cardoso. Value of Photographic Side-views in the Objective Evaluation of the Aesthetic Results of Breast Cancer Conservative Treatment. In Proceedings of the 33<sup>rd</sup> Annual San Antonio Breast Cancer Symposium, 2010

I also participated in national conferences with the following papers:

- H. P. Oliveira, and J. S. Cardoso. Gradient Flow Based Detection of Prominent Points on BCCT Depth-Map Images. In Proceedings of the 18<sup>th</sup> Portuguese Conference on Pattern Recognition, pages 1 – 2, 2012
- H. P. Oliveira, J. S. Cardoso, A. T. Magalhães, and M. J. Cardoso. Depth-Map Images for the Aesthetic Evaluation of Breast Cancer Treatment. In Proceedings of the 1<sup>st</sup> PhD. Students Conference in Electrical and Computer Engineering, pages 1 – 2, 2012
- M. D. Silva, H. P. Oliveira, and J. S. Cardoso. Super-Resolution on Depth-Map Images. In Proceedings of the 1<sup>st</sup> PhD. Students Conference in Electrical and Computer Engineering, pages 1 – 2, 2012
- H. P. Oliveira, A. T. Magalhães, and J. S. Cardoso. Kinect Based Method for the BCCT Quantitative 3D Evaluation. In Proceedings of the 17<sup>th</sup> Portuguese Conference on Pattern Recognition, pages 1 – 2, 2011
- H. P. Oliveira, and J. S. Cardoso. BCCT.core Model Enhancement With Interpretability and Lateral Information. In Proceedings of the 16<sup>th</sup> Portuguese Conference on Pattern Recognition, pages 1 – 2, 2010

## 1.4 Document Structure

This thesis is organised into ten Chapters, each one describing the work conducted during the last four years. The remainder of this Section presents the main motivations, objective and thesis contributions.

Afterwards, in Chapter 2 a literature review is provided on all the topics related with aesthetic evaluation of breast cancer surgeries, with an emphasis on the objective methods and corresponding technologies used. The Chapter starts by presenting subjective methods and their disadvantages comparatively to objective methodologies. Then, focusing on objective methods, both the traditional methods based on patient photographs and the approaches based on 3D data are discussed, providing a critical analysis and put in a chronological perspective. Finally, trends related with aesthetic evaluation are discussed, with a focus on the automation of the evaluation process.

Chapter 3 presents the main characteristics of the BCCT.core software, demonstrating the importance of the software for the breast cancer related research. The information reported in this

Chapter provides a general idea of the tool, helping understand the capabilities and drawbacks which motivated this thesis.

During this thesis a multi-modal database was created and its characteristics are presented in Chapter 4. The database is based on depth and colour information, with ground truth annotation both for fiducial point position and subjective aesthetic evaluation.

A new closed contour computation is presented in Chapter 5. Here, is introduced a formulation in the original coordinates, in contrast to the typical approach using polar coordinates.

In Chapter 6 improvements on the BCCT.core are investigated, comparing the performance of different interpretable methods with the current BCCT.core model, and with the introduction of knowledge extracted from patient side-views, complementing the currently used frontal information.

A study of 3D reconstruction based on uncalibrated view is presented in Chapter 7, including the state-of-the-art algorithms for feature detection, matching, rectification and stereo matching, applied to our data.

Chapter 8 introduces the main characteristics of the RGB-D devices, demonstrating their great potential. Due to the limitation related with the Kinect resolution, methods to improve it are presented here. Afterwards, calibration methodologies to align information from depth and colour images are also discussed.

The combination of RGB and depth data acquired with the Kinect device is studied in Chapter 9. Here, is demonstrated how the automation of fiducial points can be assisted with the fusion of RGB and depth data, helping define 3D features and the overall aesthetic evaluation.

Finally, conclusions and lines for future research on this topic are provided in Chapter 10.





## Chapter 2

# Literature Review\*

In this chapter we try to discuss all the questions related with aesthetic evaluation of breast cancer surgeries, with an emphasis in the objective methods and corresponding technologies used. We start by presenting subjective methods and their disadvantages compared to objective methods. As part of the objective methods, we discuss both the traditional methods based on patients' photographs and the more recent approaches based on 3D data, with a critically analysis and put in chronological perspective. Trends related with evaluation of the BCCT aesthetic result are discussed, namely objective methods using two-dimensional (2D) photographs and 3D data, with a focus also on the automation of the evaluation process.

### 2.1 Breast Cancer

In the past, breast cancer surgical treatment consisted essentially of only one of two procedures: mastectomy or lumpectomy (see Figure 2.1).

The mastectomy is the procedure where the entire breast is removed. There are five different types of mastectomy: "simple" or "total" mastectomy, modified radical mastectomy, radical mastectomy, partial mastectomy, and subcutaneous (nipple-sparing) mastectomy. Simple or total mastectomy concentrates on the breast tissue itself. Beyond the removal of the entire breast, occasionally lymph nodes are also removed because they happen to be located within the breast tissue taken during surgery. In this surgical procedure no muscles are removed from beneath the breast. In the modified radical mastectomy both breast tissue and axillary lymph node dissection are performed, during which levels I and II of underarm lymph nodes are removed (B and C in Figure 2.1(a)). The radical mastectomy is the most extensive type of mastectomy, where levels I, II, and III of the underarm lymph nodes are removed (B, C, and D in Figure 2.1(a)). The surgeon also removes the chest wall muscles under the breast. This type of intervention is recommended only when the breast cancer has spread to the chest muscles under the breast. Although common in the past, radical mastectomy is now rarely performed because in most cases modified radical mastectomy has proved to be just as effective and less disfiguring. In the partial mastectomy, the

---

\*Some portions of this Chapter appeared in [179]

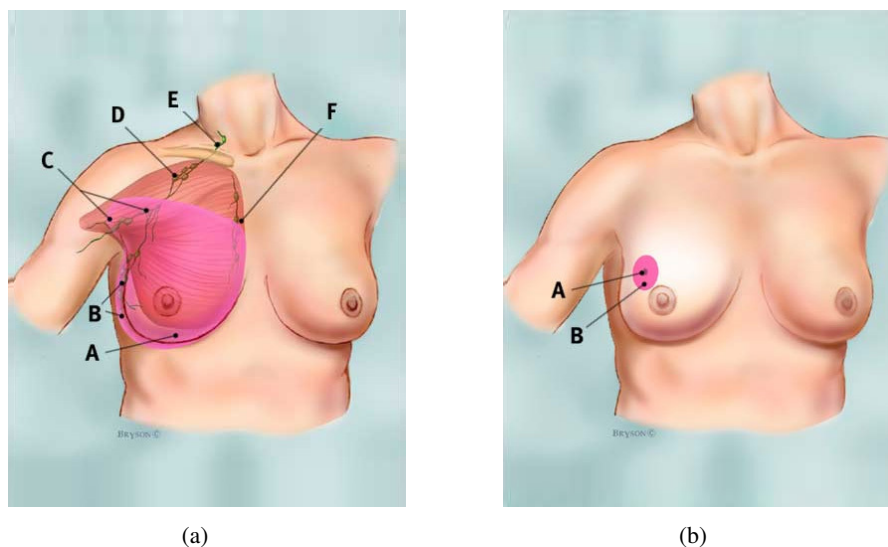


Figure 2.1: Illustration of breast surgical procedure <sup>1</sup>. a) Mastectomy; b) Lumpectomy.

cancerous part of the breast tissue and some normal tissue around it are removed. In the subcutaneous ("nipple-sparing") mastectomy, all of the breast tissue is removed; however, the nipple remains intact. Subcutaneous mastectomy is performed less frequently than simple or total mastectomy because more breast tissue is left behind, and afterwards that could develop cancer. Some physicians have also reported that breast reconstruction after subcutaneous mastectomy can result in distortion and possibly numbness of the nipple. Because subcutaneous mastectomy is still an area of controversy among some physicians, doctors may recommend simple or total mastectomy instead <sup>1</sup>.

Lumpectomy is a form of "breast-conserving" or "breast preservation" surgery. Technically, a lumpectomy is a partial mastectomy because part of the breast tissue is removed, as well as the breast tumour (the "lump") and some of the normal tissue that surrounds it (see Figure 2.1(b)). However, more tissue is removed in partial mastectomy than in lumpectomy. The amount of tissue removed can vary greatly, which is related to several factors, such as: patient characteristics, tumour characteristics and stage of the tumour. The patient then undergoes radiotherapy on the remainder of the breast <sup>1</sup>.

Therefore, a good aesthetic outcome is an important endpoint of breast cancer treatment, being closely related to psychosocial recovery and QOL. The importance of a good aesthetic outcome is well recognised by experts in this field although it is known that this is often not achieved. The conservative approach has made it possible to locally control the disease and it has a similar survival rate to that obtained with the mastectomy, but with better cosmetic results [78, 245]. A basic concept of the conservative treatment is that there must be an amount of healthy breast tissue removed, enough to obtain a surgical margin free of neoplasia. The greater the amount of healthy tissue excised, the lower the probability of incomplete removal of the neoplasia. At

<sup>1</sup><http://www.breastcancer.org>

the same time, the probability of local recurrence due to the growth of the remaining neoplasia is also lower. However, the larger the amount of breast tissue removed, the lower the chances of obtaining satisfactory cosmetic result. This comparison can be clearly demonstrated in the diagram in Figure 2.2.

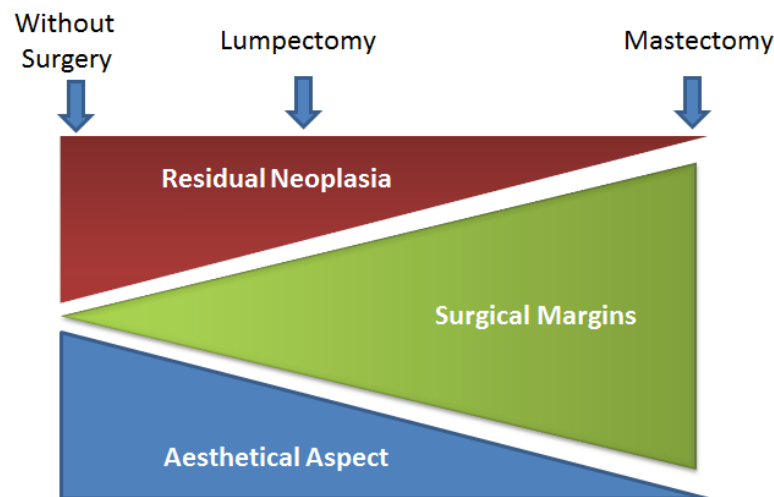


Figure 2.2: Surgical procedure comparison.

Surgical outcome depends on several factors, many of them difficult to assess, thus leading to a significantly heterogeneous results. For this reason, it is fundamental to evaluate specific surgical procedures on the basis of their aesthetic outcome through specific quantitative tools.

Although breast conservation techniques have been widely studied, different forms of evaluation and heterogeneous working practices have contributed to different aesthetic results. As this scenario suggests, the evaluation of aesthetic results should be mandatory in any institution performing breast cancer treatment, contributing to the improvement of current strategies by enabling the identification of variables which have a significant impact on the final aesthetic result.

In the process of assessing cosmetic outcomes there are several important issues that should be considered: which factors have a crucial impact on the cosmetic outcome of BCCT; which parameters or features should be evaluated in the cosmetic assessment of BCCT; how patients are evaluated; which scales are used in this evaluation; which methods and technological solutions are available for the evaluation of cosmetic results of BCCT.

## 2.2 Breast Conservative Treatment

The breast conservative treatment was introduced by Umberto Veronesi during the 1970s [244], and is well accepted by the international medical professionals. In this technique, the neoplastic tissue is ablated with a portion of healthy surrounding parenchyma. During the last 40 years this technique suffered profound alterations and evolutions [158]. While the oncologic outcomes of breast conservation procedures are easily estimated objectively by disease-free and overall survival

rates, cosmetic outcomes had no universally accepted auditing method [38]. For that reason, very few surgeons audit their results. Existing methods were time-consuming, there was no defined time after surgery to perform this assessment and even if surgeons wanted to evaluate cosmetic results there were no standards for comparison. Many works have been carried out in order to overcome all previously stated obstacles.

The BCCT has led patients to prevention regarding the very important factor related to the early diagnosis since this is directly related to the preservation (or not) of the breast. The aesthetic problem related to the absence of one breast or to the ablation of only a portion in conservative treatment, was considered a minor issue and aesthetic interventions to correct defects were delayed and performed voluntarily when women, only a minority, felt the need to do so. During the last ten years the concept of QOL raised the interest on oncoplastic breast surgery, the best of both sides, the ideal oncologic treatment with a good aesthetic outcome. As a result, a myriad of new techniques appeared either in immediate breast reconstruction in mastectomy patients and in oncoplastic procedures in breast conservation. However, the growing interest in this new field of oncoplastic surgery brought all sorts of aesthetic results of breast procedures, and the lack of adequate training of surgeons in this area brought additional concern regarding the need for a reproducible method to evaluate results.

Although considerable research has been put on breast-conservation techniques, different forms of performing them, as well as incorrect working practices, contribute to different aesthetic results [49]. As this scenario suggests, the evaluation of aesthetic results should be mandatory in any institution performing breast cancer treatment. This can improve current strategies by enabling the identification of variables, with a significant effect on the final aesthetic result and amenable to improvement [50].

## 2.3 Determinant Factors to Cosmetic Outcome

Breast conservation surgery, radiotherapy included, has not a definite standard for every patient. There are technical variations related both to surgery and radiotherapy and even though BCCT has an identical meaning worldwide, it is not a uniform procedure [49]. To obtain a good aesthetic outcome, it is essential to have a clear notion of the factors that can influence results [10]. Historically, factors with an impact on aesthetic outcomes are divided into patient, tumour and treatment-related factors [37] (see Table 2.1).

Factors most frequently reported as being important to the final aesthetic result are patient weight [51], breast size [51, 149, 200, 249], tumour location [148, 213, 225], tumour size [51, 61, 200, 206, 229], specimen weight or volume [50, 51, 148, 184, 193, 200, 229, 249], placement of incisions [50, 51, 229, 249], chemotherapy [18, 118, 201, 206] and radiotherapy to the breast [96, 148, 184].

Today with upcoming oncoplastic techniques, new factors related essentially to surgery will arise. If the specimen size and weight are seen as the most consensual of the factors, and if this concept is applied to the therapeutic reduction mammoplasty technique, where large volumes are

Patient-related	Tumour and treatment-related
Age	Incision size (cm), tumour size (mm)
Height (cm)	Incision type (radial, sulcus, periareolar, curved)
Weight (kg)	Specimen weight (g)
Thorax perimeter (cm)	Side (left, right)
Body mass index (BMI)	Tumour location (UOQ, UIQ, LOQ, LIQ, OQT, SQT, IQT, LQT)
Breast size (radius cm)	Type of intervention (tumorectomy or tum, with sentinel node biopsy or with axillary dissection)
Bra cup size (A, B, C, D)	Scar visibility (not, slightly, very visible)
Age of menarche	Surgeon (general, breast → 150 cases / year)
No. pregnancies	Chemotherapy (yes, no)
Age of first pregnancy	Hormone therapy (yes, no)
Oral contraceptives (yes, no)	Follow-up (months)
Menopause (yes, no)	

UOQ, upper outer quadrant; UIQ, upper inner quadrant; LOQ, lower outer quadrant; LIQ, lower inner quadrant; OQT, outer quadrant transition; SQT, superior quadrant transition; IQT, inner quadrant transition; LQT lower quadrant transition.

Table 2.1: Patient, tumour and treatment-related factors (From [37]).

removed along with the tumour, there will be a paradoxical effect of a better aesthetic outcome because both breasts are intervened simultaneously [121]. Nevertheless, having a clear idea of which factors influence the cosmetic result is essential to achieve a reproducible way of measuring these outcomes, regardless of the technique used. Only then can we correlate results with practice and subsequently improve results.

## 2.4 Cosmetic Assessment Parameters

To evaluate the aesthetic result of BCCT, an observer usually identifies and evaluates colour, shape, geometry, irregularity and roughness of the treated breast, compared to the untreated breast. Ideally, the comparison of the aesthetic result of BCCT would be performed using pre-treatment images of the patient. Unfortunately, the habit of capturing images before initiating treatment is not a rule in the majority of centres treating breast cancer. As a consequence, published works about the aesthetic results of BCCT use the comparison between both breasts, assuming that better results correspond to more similar breasts. This is indeed a much more practical approach and, in the era of new emerging oncoplastic techniques where both breasts are frequently submitted to surgical procedures, comparison is even more demanding [10]. The ultimate objective is to obtain identical breasts taking into account not only the surgical aspect, but also the adjuvant radiotherapy treatment applied in the majority of patients.

There are several factors contributing to global aesthetic result. Asymmetry in size is probably the most important factor which impacts the global cosmetic result and primarily depends on the

amount of excised tissues, when surgery is unilateral [4,50,52,137,149,193,210,236,249]. Surgery and radiotherapy derived fibrosis, with upward retraction of the inferior mammary sulcus and the nipple-areola complex (NAC) can also have a consequence in the evaluated symmetry without impairing the size of the breast [149, 193]. There are other factors with great impact on aesthetic result. Scar visibility and length can also influence the aesthetic result [4, 50, 236, 248]. Finally, other aspects that need to be considered, generally attributed to radiotherapy, are differences in colour, both the hyperpigmentation of the treated breast or the hypopigmentation of the NAC complex, and to a lesser extent telangiectasias [50, 138, 218, 248].

## 2.5 Patient's Evaluation

Patients are usually evaluated either by direct observation [1, 2, 8, 44, 59, 61, 73, 196, 213], using conventional photographs or slides [4, 19, 50, 96, 149, 190, 192, 248], special cameras - telecameras [175, 210] or more recently using digital images [28, 66, 80].

Direct observation of patients is promoted by the vast majority of personnel involved in BCCT as the most complete form of cosmetic evaluation, allowing physicians not only to get a better global appreciation of the results but also of other factors, which are not visualised in a photograph, such as skin atrophy and oedema of the breast and arm [50, 192]. However, photographs are more practical. Images can be saved permanently and visualised when necessary, and can be easily observed by different physicians [66, 248]. More recently, digital photographs have replaced paper prints and slides, making the entire process even easier and less expensive, while maintaining picture quality [28, 66, 80].

Eadie *et al.* [66] compared the aesthetic evaluation of breast conserving treatment in 10 patients by direct observation, printed analogical photographs, printed digital photographs and digital images on screen. Although printed analogical photographs presented better results, the correlation of results between direct observation and images did not depend on the type of images used [66].

## 2.6 Scale of Assessment

The most widespread scale used since the beginning of conservation breast procedures, until today, is the Harvard scale introduced by Jay Harris in 1979 [96]. It classifies the overall cosmetic result in four classes: Excellent (treated breast nearly identical to the untreated breast), Good (treated breast slightly different from the untreated one), Fair (treated breast clearly different from the untreated breast but not seriously distorted), and Poor (treated breast seriously distorted). Harris also introduced a scoring system to classify both the results of the surgical procedure (0-scar unapparent, 1-scar apparent, 2-major tissue loss) and of radiotherapy (0-none, 1-slight, 2-moderate, 3-severe).

Several other authors have tried to obtain a more complete and discriminative classification which included other factors such as skin oedema and thickening, mammary fibrosis, retraction and telangiectasias [11, 91, 137]. They used scores for each of the considered parameters and added

the values to obtain a final score that would allocate the patients in the four classes described by Harris [50, 91, 150] or into other similar scales, some using more than four categories [149] and others using less [12].

More sophisticated scores have been described with the objective of measuring BCCT aesthetic result [12, 137]. However, none of these have been established as a preferred alternative to the Harvard scale. In the absence of more valuable alternatives, these scales, while subjective and without discrimination power, provide useful information on the cosmetic result of BCCT [2, 59, 147].

## 2.7 Methods of Assessment

Generally there are two groups of methods used on the cosmetic evaluation of BCCT: subjective and objective methods. Subjective methods include patients' self-evaluation, evaluation by a single observer or by a panel of observers. Objective evaluation depends essentially on the type of quantitative measurements used. For both groups evaluation can be made directly on the patient or using photographs (prints, slides or digital images).

### 2.7.1 Subjective Methods

#### 2.7.1.1 Self-Evaluation

Patient self-evaluation is probably the easiest way to analyse the cosmetic outcome of BCCT. It is surely the one that best translates the psychosocial adaptation of patients to the result [3]. However, its reproducibility is low because it depends on several factors not amenable to quantification, such as age and socioeconomic status [137, 210], each one of them having a direct impact on how women see themselves after treatment. For those who support this method, the self-evaluation process and satisfaction with the result are the most important factors [5]. In works where self-evaluations are compared with external observer evaluations, patients invariably evaluate themselves more favourably [18, 23, 50, 51, 96, 120, 147, 210]. One plausible reason for this better evaluation is related to the fear that patients have to express criticism towards treatment or responsible caregivers [147]. This feeling is even stronger when a self-evaluation questionnaire is answered during a follow-up visit [190]. However, this fear can exist even with questionnaires sent by e-mail [5]. Another important point is the fact that BCCT patients tend to consider this as an alternative to mastectomy and even when results are worse, they are still seen as a better option than having the entire breast removed [147, 190].

It is undoubtedly important to have feedback from patients on aesthetic results; however, the low reproducibility of this method makes it impossible to use it for comparison purposes [23, 50, 120, 147, 210, 248].



### 2.7.1.2 Observer Evaluation

Currently, the most used evaluation of aesthetic results is the subjective assessment made by one or more observers, who focus directly on patients or on photographic representations of them [18, 196]. The final aesthetic result is assessed using, usually, the Harvard scale [96], where results are ranked according to comparisons between the operated breast and the untreated breast. This evaluation is undertaken by one [2, 8, 11, 18, 56, 58–61, 91, 120, 150] or several observers [4, 44, 50, 51, 57, 66, 73, 92, 147, 205, 206, 210, 218, 235, 248].

Even though this subjective evaluation has been criticised, it is still the most frequently used method [2, 59, 147, 150, 235, 253]. However, there are some problems regarding the interpretation of results from studies which use this type of assessment [40]: for example, impartiality is not always guaranteed since it is often performed by professionals involved in the treatment [11, 18, 56, 58, 59, 61, 81, 110, 184]. Moreover, its reproducibility is usually not high and agreement, when measured between observers, is only low or moderate [50, 192, 248]. The impracticality and the invasion of patients' privacy inherent to this type of evaluation is not negligible: if the evaluation is carried out by one or more observers involved in the treatment process and at different monitoring stages, the discomfort caused to patients is high, especially when the evaluation is performed directly and not through the use of photographs.

Another negative aspect related to this type of evaluation is the impractical side of the method, especially if we think that patients should be evaluated by multiple observers in different phases of treatment and follow-up. Even using digital images and electronic mail evaluation, this process can take months or even years, when performed in a large number of patients with multiple observers [36].

### 2.7.2 Objective Methods

Pezner *et al.* [192] published an evaluation of a series of frontal views from 14 patients performed by 44 observers that challenged the value of subjective evaluations in terms of reproducibility. They also proposed the first measure to objectively assess the cosmetic outcome of BCCT: the Breast Retraction Assessment (BRA) [193], which determines the treated breast retraction comparatively to the untreated breast. Using a marked acrylic sheet over the thorax of the patients, they calculated upward and inward retraction of the treated versus the non-treated breast. Higher BRA values corresponded to worse cosmetic results. BRA was subsequently correlated with tumour size, chemotherapy and radiation fields [193].

BRA and the measures suggested afterwards in the literature (solely capturing the breast asymmetry) were often complemented by a subjective assessment made by observers [4, 23, 50, 137, 149, 175, 210, 218, 238, 248]. Even with these additions, all available methods were subject to significant intra-observer and inter-observer variability. More than trying to evaluate the aesthetic result from the measurements, these methods correlate them with subjective evaluation. Current methodologies in use continue to show a significant lack of standardisation, not only in the type of



assessment used, but also in terms of the factors that should be considered in this evaluation and even in the instruments used for this analysis.

Next, we elaborate on the objective assessment of BCCT, detailing most important methods based on singular photographs and 3D information.

## 2.8 Objective Assessment of Breast Aesthetic Outcome

Objective methods were introduced in an attempt to overcome the lack of objectivity and reproducibility. There are two different methodologies to objectively evaluate the aesthetic result: patients' 2D photographs and 3D information. Figure 2.3 shows some of the most significant works related to the objective assessment of breast aesthetic outcome in breast cancer treatment.

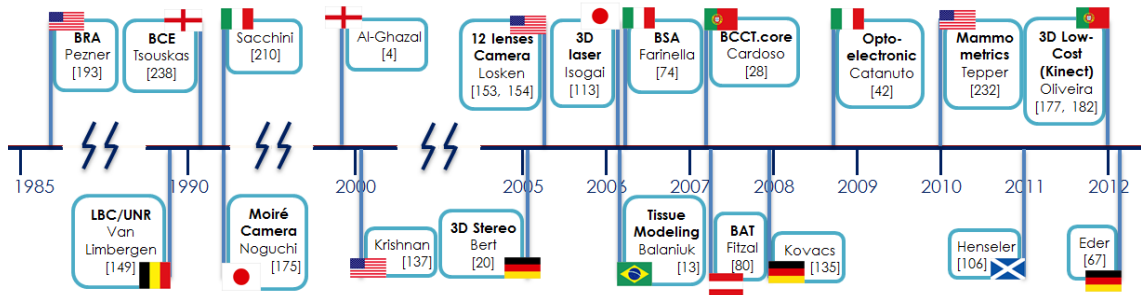


Figure 2.3: Timeline of some of the most important methodologies for the objective evaluation of the aesthetic outcome.

### 2.8.1 Two-dimensional Patient Data

These methods consist of comparing the two breasts with simple measurements marked on patients' photographs [50, 149, 193]. These methods are essentially based on asymmetries between treated and nontreated breasts that improved the assessment reproducibility (see Table 2.2). However, it has been argued that they do not take into account the global appearance of the aesthetic results, due to the lack of the third dimension [50, 137, 210, 248].

#### 2.8.1.1 First-generation Methods

As stated before, the first measure proposed for the objective assessment of the cosmetic outcome of BCCT was the BRA [193] to determine retraction in the treated breast compared to the untreated breast. Using identical reasoning applied to patient photographs, Van Limbergen *et al.* [149] came up with two new asymmetry measurements, besides the BRA: the Lower Breast Contour (LBC) and the Upward Nipple Retraction (UNR). They also found a strong correlation between the obtained values and a subjective classification performed by observers.

Tsouskas and Fentiman [238] described the Breast Compliance Evaluation (BCE) as a measurement of objective cosmetic outcome. Compliance is the difference between the distance from

Authors	Papers	year	Description
Turesson and Notter	[239,240]	1984	Turesson Skin damage classification graded as the area of telangiectasias and skin necrosis, due to radiotherapy
Pezner <i>et al.</i>	[193]	1985	Breast Retraction Assessment (BRA), which determines the treated breast retraction comparatively to the untreated breast
Limbergen <i>et al.</i>	[149]	1989	Lower Breast Contour (LBC) and the Upward Nipple Retraction, which determines the nipple displacement and breast contour retraction of the treated breast respectively, comparatively to the untreated breast
Tsousskar and Fentiman	[238]	1990	Breast Compliance Evaluation (BCE), which determines the difference between the distance from the nipple to the Infra-Mammary Fold (IMF) of the treated breast, comparatively to the untreated breast
Noguchi <i>et al.</i>	[175]	1991	Moiré topographic camera to compute the differences between the displayed curves in both breasts
Cardoso and Cardoso	[28]	2007	BCCT.core software for the aesthetic outcome assessment entailed a semi-automatic extraction of features considered to have an impact on the overall cosmetic result (asymmetry, colour differences and scar visibility features)
Fitzal <i>et al.</i>	[80]	2007	BAT software and Breast Symmetry Index (BSI) definition to evaluate the cosmetic outcome of BCCT
Kim <i>et al.</i>	[126]	2007	Breast ptosis measurement based on ratios of distances between fiducial points manually identified in oblique and lateral patient photographs
Oliveira <i>et al.</i>	[181]	2010	Features extracted from lateral patients photographs for the aesthetic outcome assessment
Lee <i>et al.</i>	[139]	2012	Quantitative measure of breast curvature based on catenary

Table 2.2: Comparison of the most significant 2d technologies involved on breast cancer cosmetic evaluation.

the nipple to the Infra-Mammary Fold (IMF) with the patient in the erect and supine positions. The same line of thought was followed by several other authors in more recent work [23, 50, 210, 218, 248].

Noguchi *et al.* [175] introduced an innovative sum of objective and subjective evaluations. The objective part was undertaken with a Moiré topographic camera, which solved the differences between the displayed curves in both breasts. Other parameters were evaluated subjectively by observers (skin changes and scar) and the final result was the sum of both evaluations.

The same path was followed by others who used both types of method and the final result was always the sum of the two obtained scores [4, 137]. The idea of adding the results of both subjective

and objective evaluations resides in the fact that, for the majority of authors, neither method is successful in translating the complete evaluation of aesthetic results. When a bad aesthetic result is obtained, for example, due to a very visible and badly placed scar, the correlation between asymmetry measurements and subjective classification is obviously lost [149].

A study to examine the convergence between objective and subjective indices of cosmetic and functional status after patients undergo BCCT was developed by Krishnan *et al.* [137], where the relations between the objective indicators and QOL were tested. They found that these scores were well correlated with the patients' self-reported cosmesis questionnaires. The findings revealed positive cosmetic and functional treatment outcomes, showing good or excellent cosmetic results.

The fifth edition of the EORTC manual for clinical research in breast cancer published in 2005 [72] also stressed the idea of adding qualitative and quantitative evaluations. The subjective evaluation should be made by a panel of at least five observers classifying results according to the Harris Scale [96] and an objective evaluation using the measurements of asymmetry described by Pezner [193], and finally adding Turesson Skin damage classification [239, 240], graded as the area of telangiectasias and skin necrosis, due to radiotherapy.

The first-generation of models introduced objective quantification of individual aspects of the overall cosmetic outcome. These developments represent significant advances from a complete reliance on expert judgement; however, we have yet to see these capabilities substantially differentiate operational evaluation methodology. The simple sum of individual indices, objective and subjective, fails to acknowledge that different aspects have different contributions to the overall result; the reliance on the subjective assessment of individual aspects makes the final decision subjective and irreproducible; even for the objective measures, the manual measurement makes them partially subjective; as a consequence the process is still time consuming and hard to implement under routine clinic conditions. In many ways, the development of these models has been a step in the right direction, but we are still far from an acceptable system.

### 2.8.1.2 Second-generation Methods

Two approaches emerged in 2007 in an attempt to bring something new to the evaluation of aesthetic results in breast conservation [28, 80]. These methods were introduced to make it possible to predict the global aesthetical result. These software tools are based on different individual characteristics automatically and objectively extracted from patient photographs. These approaches aimed to overcome the acute shortage of such tools and to exploit the unique ability that computational methods have to provide an effective, easy, fast, reliable and reproducible tool to evaluate the outcomes of breast cancer patient care. In both works, the team started by creating a database of patient images evaluated subjectively by a panel of experts, obtaining classifications which served as a ground truth the developing new methods and comparing results.

Fitzal *et al.* [80] described a Breast Symmetry Index (BSI) to evaluate the cosmetic outcome of BCCT. They created software (Breast Analysing Tool - BAT) to measure differences between left and right breast sizes from a patient's digital picture (see Figure 2.4).

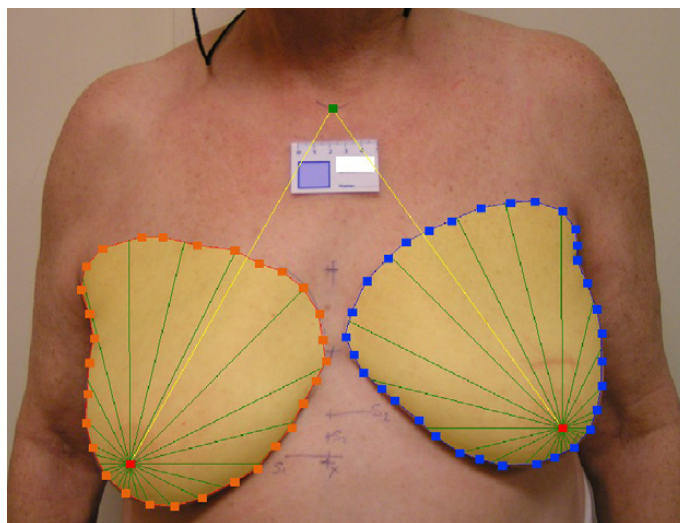


Figure 2.4: BAT software interface (From [80]).

The sum of all differences results in the breast symmetry index (BSI) score, which is capable of measuring differences in size between breasts. They were able to demonstrate BAT© accuracy and additionally showed that BSI is reproducible and may be used by experts and non-experts with the same results. The BSI index measured from frontal pictures correlates highly with subjective votes from experts and significantly differentiates between good and fair cosmetic outcomes. However, there was no correlation between the BSI index and the patients' self-evaluation. Moreover, the BSI did not further differentiate between excellent and good or fair and poor cosmetic outcomes. These data show that the BSI measured with the BAT© may be used to some extent to evaluate cosmetic results of BCCT.

Cardoso and Cardoso [28] introduced a computer-aided medical system named BCCT.core (see Figure 2.5). The development of BCCT.core entailed the semi-automatic extraction of several features from patient photographs, capturing some of the factors which are considered to have an impact on the overall cosmetic result. To accomplish cosmetic categorisation, a concise representation of a BCCT image is first obtained based on asymmetry, colour differences and scar visibility features.

The asymmetry between both breasts was conveyed under a large set of indices. To extract the colour features, the colour content of each breast is first characterised by a colour histogram; the dissimilarity between both histograms serves as the colour features. The scar visibility was translated into local colour dissimilarity, by comparing corresponding sectors of the breasts. The measurements are preceded by the automatic localisation of fiducial points (nipple complex, breast contour and jugular notch of the sternum) [34] on the photographs. Measurements for asymmetry, skin colour change and surgical scar appearance are then supported on these fiducial points. The representation is then further analysed by a pattern classifier performing the categorisation [31]. The pattern classifier finds the combination of individual features that approximate best the overall classification as given by the panel of experts. A correct classification rate of about 70% was

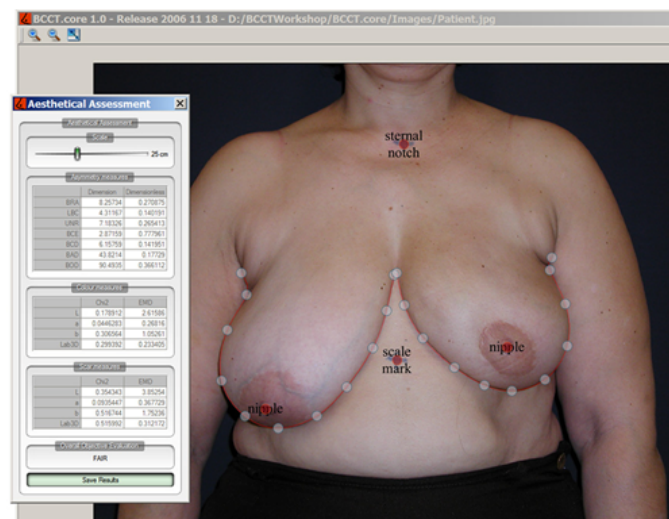


Figure 2.5: BCCT.core software interface (From [28]).

obtained, when comparing with the subjective consensual classification from the panel [36].

### 2.8.1.3 Summary

The BAT and the BCCT.core have been recently compared [260] regarding the same set of cases. The conclusion of this study, was that the BAT software maintains a moderate performance regardless of picture quality using only asymmetry measurements. On the other hand, the BCCT.core, using other parameters such as colour differences and scar appearance, obtains different results with good performance with better photographs but only moderate performance in lower quality pictures [39]. Moreover, the BCCT.core has been used in many works related with BCCT cosmetic evaluation [101, 103, 104].

These tools, while innovative and reproducible, have several drawbacks that need to be addressed. Both use only frontal photographs of the patients, which do not include lateral or oblique views. This approach is rare in the history of the aesthetic outcome of BCCT [181] and only a small number of researchers have used lateral features in their work. The exceptions include Kim *et al.* [126], who compare objective measurements based on breast ptosis with ratings on a subjective scale made by experienced clinical observers and more recently, Lee *et al.* [139] who introduces a novel quantitative measure of breast curvature based on catenary. They compared the length, the area enclosed by the curve, and the curvature measure from the catenary curve to those from manual tracings of the breast contour. Likewise, the same procedure was applied to frontal photographs. This can be related with the difficulty of extracting robust features from lateral views. The question is: “is it important to include this information?” Fitzal *et al.* [80] stated that lateral and oblique views are difficult to standardise and the solution will be to move into a 3D version of the software.

## 2.8.2 Three-dimensional Methods

Potential advantages of 3D imaging as a tool for objective cosmetic evaluation include the ability to view the breast from a significant different number of angles, to estimate volume/volume deficit and to plan future surgeries. There are several variations of this tool, from relatively simple volumetric analyses to more sophisticated programmes, which make it possible to perform quantitative measurements and software that makes it possible to simulate the post-operative outcome. Although these benefits are known, some authors still doubt the ability of current 3D image analysis methods [35].

### 2.8.2.1 Existing Systems

Several research groups have recently made attempts with 3D approaches (see Table 2.3).

3D imaging in breast surgery can be used for several clinical purposes. The most practical applications are evaluating breast asymmetries, both congenital and acquired, and factors affecting breast shape in augmentation mammoplasty. Other uses of 3D imaging for clinical purposes include the evaluation of patients who wished to have a reduction mammoplasty and the evaluation of patients undergoing unilateral breast reconstruction to determine the expander and permanent implant size that gives the best symmetry with the contralateral breast.

Galdino *et al.* [86] investigated the use of clinical 3D imaging to determine quantitative information about the breast, such as volume or projection. They applied this approach to real cases, providing objective data on the breast and surgical mammoplasty (especially augmentation mammoplasty). This helped surgeons better understand the factors that contribute to breast shape and influence surgical outcomes. A couple of limitations were found, highlighted by patients with significant ptosis or suffering from obesity, which may introduce errors into the 3D data, making them unreliable.

Losken and colleagues [153, 154, 174] developed an objective technique based on a 3D camera and software to quantify the cosmetic results of BCCT. This software package enables the comparison between the treated and untreated breast by analysing the surface area and volume differences. The camera includes 12 individual digital lenses arranged in 3 planes with a single focal point at the manubrium. Images are captured with the patients' arms in two different positions: at their sides and on their hips. The software is then used to determine the level of asymmetry between the breasts by overlapping the 3D meshes from the two breasts (see Figure 2.6).

Bert *et al.* [20] described a 3D stereo-photogrammetry system which captured 3D surface anatomy data, making it possible to perform a volumetric assessment of the breasts. In order to reproduce the position of the patient in fractionated radiotherapy, the patient surface model is registered to a previously recorded reference surface. This is performed using surface registration, where the system determines the rigid-body transformation that minimises the distance between the treated and the reference surface models in a region-of-interest (ROI). The implemented system is accurate and sufficiently stable to be used in clinics. Errors were computed by comparing



Type	Authors	Papers	year	Technology	Type of vision	Compl.
Multi - camera	Losken <i>et al.</i>	[153,154,174]	2005 - 2008	Camera with 12 individual digital lenses arranged in 3 planes with a single focal point	Passive	High
Stereo - photogrammetry	Bert <i>et al.</i>	[20]	2005	3D stereo - photogrammetry and surface registration	Passive	Average
3D Laser Scanner	Isogai <i>et al.</i>	[113]	2006	3D laser scanning and quantitative comparison with Moiré patterns projection	Active	Average
3D Laser Scanner	Farinella <i>et al.</i>	[74,75]	2006	3D laser scanning combined with anatomical landmarks	Active	High
3D Laser Scanner	Tepper <i>et al.</i>	[230,231]	2006 - 2008	3D laser scanning and surface registration	Active	High
3D Laser Scanner	Balaniuk <i>et al.</i>	[13]	2006	3D laser scanning in combination with a generic 3D model of the internal structure of the breast using virtual reality approaches combined with soft tissue modeling	Active	High
3D Laser Scanner	Kovacs <i>et al.</i>	[134–136]	2006 - 2007	3D laser scanning and comparison with reference measurements obtained nuclear magnetic resonance imaging	Active	High
3D Laser Scanner	Catanuto <i>et al.</i>	[41,42,189]	2008 - 2009	3D laser scanning and stereo opto-electronic tracking system with breathing artifact correction	Active	High
Multi - camera	Henseler <i>et al.</i>	[106]	2011	3D model generated with a 4 pods and 8 cameras, 2 cameras on each pod and comparison with volume measured with BAT© software [80]	Passive	Average
3D Low Cost RGB-D camera	Oliveira <i>et al.</i>	[177,182]	2011 - 2012	Disparity map generation with Kinect device. Automation of prominent points and breast contour, and comparison with results obtained with BCCT.core [28]. Volumetric detection using a low cost solution.	Active (Infra-red)	Low
3D Laser Scanner	Eder <i>et al.</i>	[67]	2012	3D laser scanning by overlapping the mirrored left breast with the right breast. 3D contour definition and comparison with contour obtained with BCCT.core software [28]	Active	Average

Table 2.3: Comparison of the most significant 3D technologies involved on Breast Cancer cosmetic evaluation.

the surface model with the computed tomography (CT) geometry, and were in the order of sub-millimeters. Alignment deviation and gating-signal tests were of the same magnitude.

A quantitative analysis using 3D imaging was proposed by Isogai *et al.* [113], who used laser

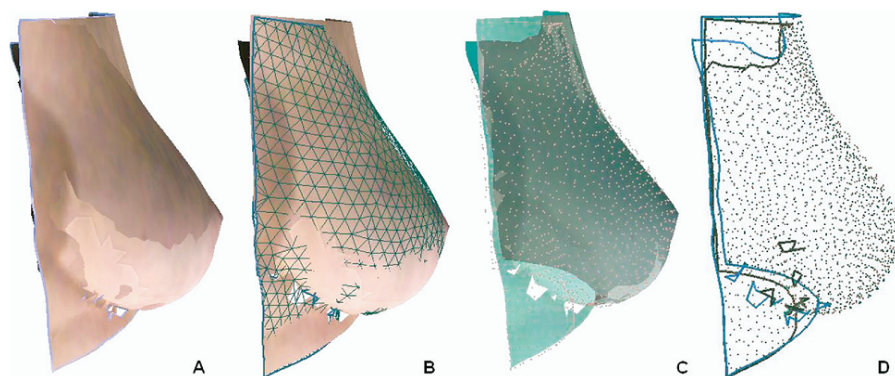


Figure 2.6: Application example from Moyer *et al.* [174]. (A) represents the lateral image of the treated breast overlapped on the healthy breast; (B) presents the comparison between the healthy breast (natural colour) with the treated breast (meshed). The software analyses the surface through the (B) mesh interpolation and the (C) division into random individual points. The distance between the two surfaces meshes (D) is computed by the software and given as the RMS, representing the asymmetry score. All individual distances between corresponding points are summed to give a total volume difference, designated as the square root of each squared difference.

light scanning to generate 3D images, which were then subjected to Moiré patterns for a quantitative comparison of the reconstructed breast with the normal breast. The evaluation is based on breast symmetry, volume and shape after breast reconstruction using the 3D laser light scanner system. Captured images of the normal breast were mirror-reversed and overlapped on images of the reconstructed breast.

Farinella *et al.* [74,75] used 3D laser scanning combined with anatomical landmarks identified by surgeons and developed the Breast Shape Analyzer 0.1 (BSA 0.1) software, which provided useful objective quantitative measurements to surgeons. This technique uses only well-defined anatomical points (Figure 2.7(a)), which have been identified and selected by surgeons. A simple sequence of geometric operations is performed to divide the breast surface into four anatomic subunits, according to the clinically derived breast meridian and equator lines. Using this breast shape partition, useful measurements can be extrapolated on a 3D model data set. The models used in this work were collected following the most common criteria for cosmetic evaluation and reconstruction (such as age, menopausal status, ptosis classification). The acquisition was made using a commercial laser scanner, applied on volunteers sitting on a chair with their back at 45 degrees (Figure 2.7(b)). Each volunteer was scanned three times: facing the camera and rotating the chair at 45 degrees to the left and to the right. This system has an important limitation related to some of the patients' uncontrollable physiological movements. During the acquisition, the breathing and imperceptible body torsions influenced the data, and thus the authors decided not to merge the different scans. The authors suggest that this can be overcome by adopting a faster but less accurate shape-from-video technique.

Tepper *et al.* [230,231] used 3D imaging to enable an objective volumetric analysis of breasts prior to breast reconstruction. Surgeons are then capable of visualising the size, shape, contour



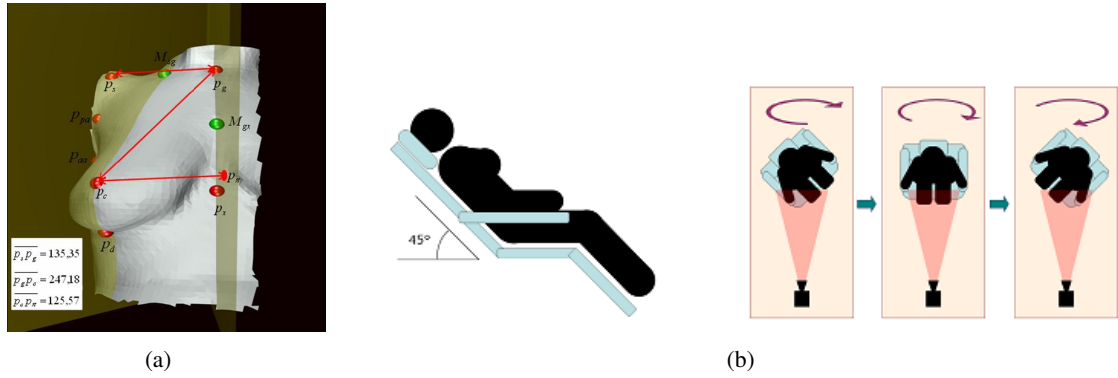


Figure 2.7: Application example from Farinella *et al.* [75]. a) Anatomical landmarks; b) Volunteers sitting positions (Patients' posture and rotation of the chair with respect to the camera).

and symmetry of the breast with 3D breast models, as well as obtaining quantitative breast measurements and volumetric calculations. This application represents a significant advance from traditional approaches to aesthetic and reconstructive breast surgery.

A 3D imaging tool was developed by Balaniuk *et al.* [13], using virtual reality approaches combined with soft tissue modelling methods which allow surgical simulations to predict cosmetic outcomes and options for breast augmentation and reconstructive procedures. They discuss the adequacy of the different medical procedures to the patient's choice, and state that it is very important, for patients and for surgeons, to have a tool to visualise the potential outcomes of the surgery and make decisions on their surgical options. In Figure 2.8, we can see an example of the model created with this application for a breast with and without an implant.

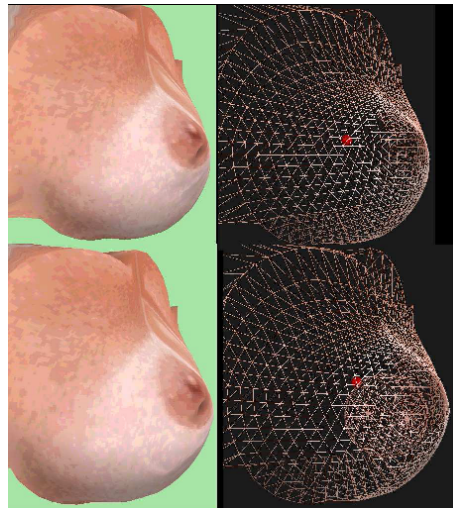


Figure 2.8: Application example from Balaniuk *et al.* [13]. The breast with and without an implant.

In the work carried out by Kovacs *et al.* [134–136], it is stated that the 3D scanner represents a simple and promising method to measure the breast volume. In their work, they present a research

study on the optimisation of the imaging technology for the mammary region with a 3D laser scanner, which is used to evaluate the method's precision and accuracy, and to allow optimum data reproducibility. The standardisation of the imaging and scanning system is decisive in order to guarantee that the method is reproducible. They have tested the most favourable imaging technology on dummy models for scanner-related factors, such as the scanner position in comparison with the torso and the number of scanners and single shots. This application showed the correct settings for 3D imaging of the breast region with a laser scanner with an acceptable degree of accuracy and reproducibility.

Catanuto *et al.* [41,42] present a set of parameters to unambiguously estimate the shape of the natural and of the reconstructed breast, using an optoelectronic tracking system in seven female volunteers. This study allows a real-time breathing artifact correction [189] and a surface patch fusion with no intervention from the operator. With these parameters, it is possible to describe several anatomical geometrical properties, such as: distances (linear measurements between relevant anatomical landmarks); surface measurements (total surface area, anatomical subunits area); angles (divergence between the nipples from the patient's point of view, inclination of the breast mound on the chest wall); curvature (curved surface properties); symmetry (symmetry between the two sides of the chest wall); and IMF shape (natural characterisation and reconstructed breast shape). With this technique, the most interesting result obtained was a graphic depiction of the curvature of the thoracic surface. While the first, the "divergence angle", gives a perspective on the point of view of the patients' own breasts (Figure 2.9(a)), the latter introduces a representation of the breast based on a colour map which expresses flat regions and curvature (Figure 2.9(b)). The authors conclude that the colour based curvature representation can replace volume terminology, as curvature can be expressed either with colours or its degree by coefficients. The breast surface was segmented into four quadrants using reproducible landmarks. The proposed methodology relies on geometric planes (bilateral symmetry plane, meridian plane and equatorial plane) that make it possible to calculate significant clinical angles.

A 3D multiple stereo camera system was introduced by Henseler *et al.* [106] to objectively assess breast shape. The system is composed of 4 pods and 8 cameras, 2 cameras on each pod. Several dummy patient breast models and breast surfaces were captured. Then a 3D model was constructed and the volume was measured with BAT software [80]. The researchers stated that the obtained results were very satisfactory, on 3D data obtained from the dummy models. In the real breast model, all measurements of total breast volume were consistently higher compared to the results obtained with stereophotogrammetry (see Figure 2.10). Moreover, the arbitrary identification of the chest wall during the immersion of the breast was the cause of the poor reproducibility of this method. This highlights the technical difficulties of the method. The measurements of volume were more reproducible with the 3D imaging method. Nevertheless, the true volumes of the breasts remained uncertain due to undefined breast boundaries.

Eder *et al.* [67] conducted a study based on a 3D evaluation protocol to analyse breast symmetry according to 3D breast contour differences between the left and the right breasts using surface imaging. In this study, researchers objectively compared breast symmetry using a 3D scanner by

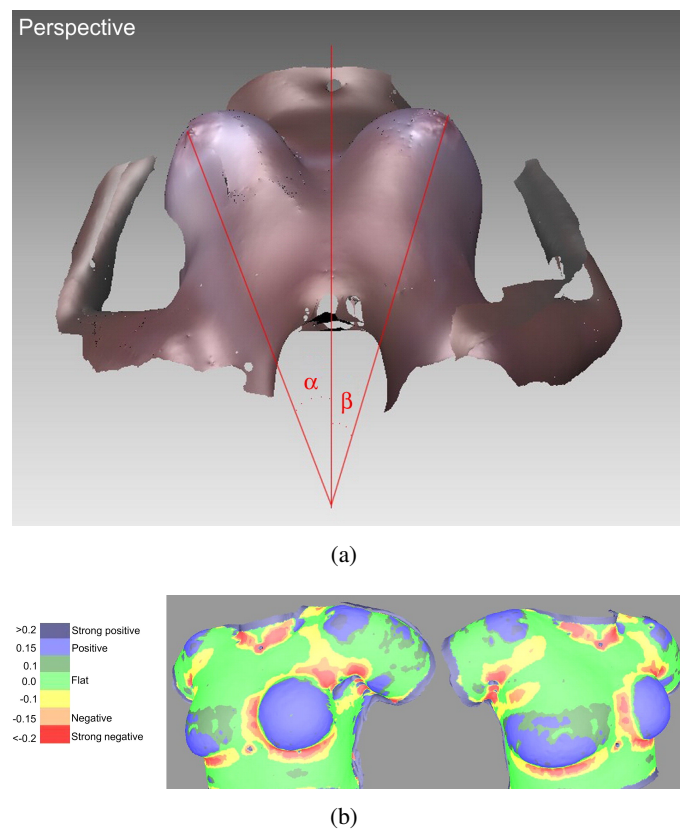


Figure 2.9: Application example from Catanuto *et al.* [42]. a) Example of divergence angles on a 3D model; b) Example of colour based map. Colour scale associates values to a curvature colour scale.

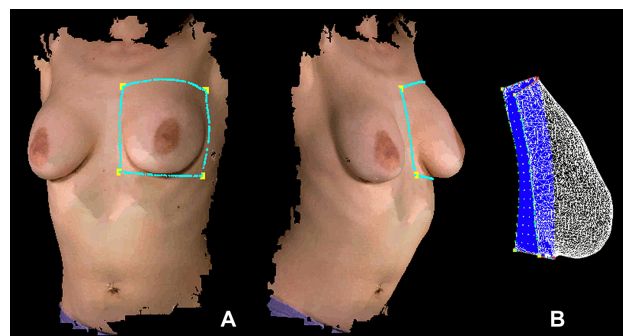


Figure 2.10: Application example from Henseler *et al.* [106]. Segmented breast on 3D model, breast segment with chest wall.

overlaying the mirrored left breast with the right breast and determining the mean 3D contour difference between the two breast surfaces. Three observers analysed the evaluation protocol precision using two dummy models. They also evaluated the potential of this approach in clinical applications comparing it with BCCT.core [28], on 23 breast reconstruction patients. The authors concluded that this approach could assist surgeons in the pre-operative planning and optimisation

of breast corrections after reconstruction and breast conservative approaches.

In 2010 Tepper *et al.* [232] provided an overview of 3D breast photography, with an emphasis on its potential role to establish a standardised system for breast analysis, by introducing a new concept entitled "mammometrics". In this concept, 3D-based breast measurements can be used to help guide operative planning and objectively analyse surgical results. In their work, the researchers validated the use of 3D breast photography in various clinical arenas, including autologous breast reconstruction, prosthetic breast reconstruction, reduction mammoplasty, and augmentation mammoplasty. They state that 3D imaging technology has the ability not only to obtain well-established breast measurements accurately, but also to find new measurements that were not previously possible with conventional tools, such as total breast volume, volumetric distribution and breast projection.

Recently the Kinect (Microsoft Corp. Redmond, WA) [177, 182] was introduced as promising low cost and easy to use equipment for BCCT cosmetic evaluation. This tool can not only facilitate automation, but also provide volumetric information (see Figure 2.11).

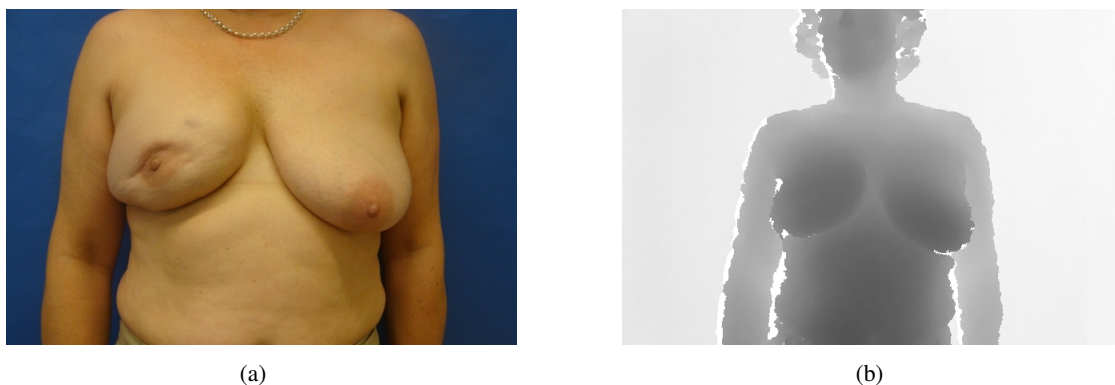


Figure 2.11: Patient photograph and generated disparity map (From [182]).

With this solution authors improved automation of prominent points and the breast contour [177], comparatively to those obtained with the original BCCT.core [28]. Furthermore, other promising results were achieved, including the detection of volumetric differences of the breasts using the disparity map generated from the Kinect [182]. It was shown that depth-map images facilitate the automation of BCCT.core and thus this software remains low cost and an easy to use solution. Obtained results also showed an excellent performance and robustness for a wide variety of patients. Authors believe that this kind of systems could be a feasible solution to use in the future for 3D assessments of BCCT results.

### 2.8.2.2 Summary

All research work presented was particularly related to BCCT and more recently other types of breast cancer locoregional treatment, and also with plastic surgery. Some of the approaches can

only be applied to a specific procedure because the characteristics to consider in each situation may be different and the precision level can also differ according to each situation.

Currently, the approaches are mainly based on 3D laser scanners using active and passive lighting. However, these methods do not fill completely physicians' expectations. The most recent works [67, 106], are based on methodologies introduced previously, and compared the performance with results obtained with 2D tools [28, 80], which continue to be used as a reference. Over the last few years unfortunately there was not a technological evolution of these methodologies, foreseeing a radical change, which may well pass through the use of simpler approaches based on low-cost systems. The main drawback of 3D techniques is the demand for specialised hardware, software and personnel. The high cost and the difficulty of using these methods on a daily basis prevent their widespread use in the near future. Additionally, almost all currently used techniques based on 3D models do not try to predict the aesthetic result for a more informed choice of treatment, nor are they suitable for the automatic evaluation of the aesthetic result after the surgery [42, 136, 153]. Other issue is related with the long acquisition time needed to obtain the 3D model, which can result in discomfort for the patient. During the period of acquisition any movement made by the patient can result in unreliable 3D models. Apart from the 3D scanners based systems, researchers are still using methodologies based on stereo vision [140] or multi-camera systems [122]. However, 3D reconstruction of breasts is a very difficult task since it is featureless.

Another important issue is related with automation of the software. All solutions need total or partial input from the user to obtain features that are used to achieve the expected result. The automatic identification of fiducial points can be difficult to localise, like the manubrium and sternal notch, especially with poor lighting and among overweight and obese patients. This makes a fair comparison between different researchers and institutions difficult. Moreover, reproducibility of the results is very difficult to obtain. The comparison between different methodologies is also very hard to achieve, since databases used in the different studies are not shared.

It would be desirable to have completely automated software, low cost and easy to operate hardware and the ability to compare results with public databases

## 2.9 Public Databases

The demand for more capable and accessible databases is welcome and necessary. Public datasets provide common standards to measure and compare accuracies of proposed approaches. Databases should be available, preferentially over the Internet, and continuous user support is also indispensable. The importance of having public datasets is recognized, both to evaluate the performance of existing systems and to encourage the development of new algorithms. One of the trends is related to the database size. In this context, a larger data set will give more reliable results. Another trend is that databases have to become more diversified, with respect to patient characteristics, or surgical interventions. Annotation of the data, ground truth, is also mandatory in order to give an independent comparison of the results obtained from different methodologies. In particular,



BCCT authors are not accustomed to provide access to the data used in their works, in contrast to radiology researchers [102, 172, 226]. This fact makes it difficult to compare results, metrics and the different methodologies employed.

The only exception is the database provide by the Breast Research Group <sup>1</sup>. This dataset is composed of data acquired from three different institutions in Portugal, comprising photographs from 120 patients. Data acquired was collected from patients treated with conservative breast surgery, with or without auxiliary surgery and radiotherapy, with treatment completed at least one year before acquisition. All patients signed an informed consent to participate. Breast images were acquired employing a 4 Megapixel digital camera. Photographs were taken in four different positions with the patient standing on floor marks: facing, arms down; facing, arms up; operated side, arms up; contralateral side, arms up. Figure 2.12 presents a typical set of images for a patient. A mark was made on the skin at the suprasternal notch and at the midline 25 cm below the first mark. These two marks create a correspondence between pixels measured on the digital photograph and the length in centimetres on the patient.

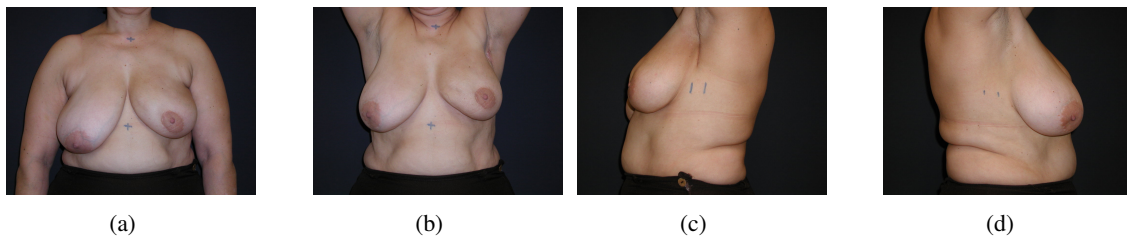


Figure 2.12: Breast Research Group Database - Patient acquisition - Different positions (From [28]).

This database has two different kinds of annotation: subjective evaluation of the aesthetic result and the definition of breast contour and fiducial points (nipple position, suprasternal notch position and the 25cm mark). The subjective evaluation used the opinion of several experts with experience in BCCT [40]. A set of 24 clinicians working in 13 different countries were selected, based on their experience in BCCT (number of cases seen per year and/or participation in published work on evaluation of aesthetic results). They individually evaluated a series of 120×4 photographs taken from 120 patients. They were asked to evaluate overall aesthetic results, classifying each case using the Harris scale [96]: excellent, good, fair and poor. In order to obtain a consensus among observers, the Delphi process was used [99, 119]. In the end, the consensus on aesthetic results was obtained in 113 of the 120 evaluated patients, equivalent to 94% of the cases. The distribution of cases is presented in Table 2.4.

With respect to breast contour and fiducial points annotation, 8 independent users manually marked these features, in frontal patient views, using the BCCT.core software tool [28] (see Figure 2.13).

<sup>1</sup><http://medicalresearch.inescporto.pt/breastresearch/>

Class	#cases
Excellent	14
Good	64
Fair	24
Poor	11

Table 2.4: Breast Research Group Database - Distribution of the patients over the four classes (From [28]).



Figure 2.13: Breast Research Group Database - Manually annotation using BCCT.core software (From [28]).

Both patient data and ground truth annotation files can be accessed at the website of the Breast Research Group <sup>1</sup>.

## 2.10 Conclusion

The widespread method to evaluate aesthetic results in breast cancer treatment consisted of observer evaluations, known as subjective assessments. Harris *et al.* [96] introduced an overall cosmetic score: excellent, good, fair and poor. In this method, the aesthetic of the breast patient is evaluated directly or using a photograph. However, the global aesthetical results are the combination of several features, estimated and subjectively combined by the observers by visual inspection. Subjective evaluation is poorly reproducible, even when performed by specialists, which significantly limits comparison between centres dedicated to breast cancer.

Initial objective methods, which appear to overcome the setbacks of subjective methodologies, consisted of comparing the two breasts with simple measurements marked directly on the patients or in photographs. Trying to overcome the sense that objective asymmetry measurements were insufficient, other groups proposed the sum of the individual scores of subjective and objective indices. However, adding subjective evaluation may result in important setbacks related to the complete automation of the approaches, fundamental for high reproducibility. There remains a

<sup>1</sup><http://medicalresearch.inescporto.pt/breastresearch/>

need to extract volumetric information in order to improve the overall cosmetic evaluation. Furthermore, the solutions should remain low cost.

Potentially, a more accurate and objective tool to predict surgical outcomes to guide the patient and surgeon in the decision-making or planning process is using 3D imaging and surgical simulations. A simulation model also allows patients to visualise the possible outcomes of different treatment options. It is generally accepted that 3D imaging has great potential in a clinical environment, although there are factors that may influence its use in the near future. The high cost of the equipment and the need for specialised people to operate are undesirable circumstances. Consequently, the search for low cost and easy to use equipment is highly desirable.

Two important conclusions are drawn from the concepts addressed previously. The first one is that cosmesis is an important objective of breast conserving treatment and it should be assessed along with oncological outcome. The second one is that it is essential to have easy and reproducible methods of assessment that are consistent worldwide. The immediate consequence of these two assumptions is that subjective methods must be replaced by objective ones to obtain the desired reproducibility and easy manipulation.

A good method of evaluation should be consistent and simple to use with all kinds of images. For example, standards for photograph quality (definition, backlight, background) are required if one is to expect discriminative power in evaluation of aesthetic results. Are we there yet? We do not think so but the way is now open to the development these new techniques. With the upcoming new breast oncoplastic surgeries there will be an even greater demand for evaluation of cosmetic results. Radiotherapy also with the use of the partial breast irradiation techniques must now pave the way for a comparison between the new cosmetic results and the ones obtained with an entire irradiation of the breast.

In breast-conserving surgery, there is evidence that approximately 30% of women receive a suboptimal or poor aesthetic outcome; however there is currently no standardised method of identifying these women. We agree that a new generation of tools will see considerable growth in the near future; these tools, tailored to the individual patient, should also target the prediction (rather than just the evaluation of the treatment) of the aesthetic outcome of breast conserving surgery. The combination of 3D photography and routinely acquired radiological images (i.e. mammography, ultrasound and MRI, when available), together with information about the tumour (size, location, shape etc.) will enable evaluation of alternative therapeutic strategies and the consequences of the available options. This will aid communication with patients and empower patients to take an active role in a shared decision making process.



## Chapter 3

# BCCT.core Software

The BCCT.core is a semi-automatic system capable of objectively evaluating the overall aesthetic result of BCCT. A brief introduction was already made in Section 2.8.1.2.

This chapter presents the main characteristics of the BCCT.core software. Here, the features extracted from the input data are presented, as well the quantification of the aesthetic results based on an optimised objective score obtained from the extracted features. The methodology relies in the cross-section of medical expert systems, soft computing and machine learning.

### 3.1 BCCT.core Design

BCCT.core [28] introduces the semi-automatic measurement of various indices related to aesthetic results of surgical procedure, making a quick assessment that is easy and reproducible, categorizing the aesthetic result of BCCT on the Harris scale [96]. To accomplish the categorization, first, a concise representation of a BCCT image is obtained based on asymmetry, skin colour changes and surgical scar appearance. These measurements are preceded by the semi-automatic localization of fiducial points (nipple complex, breast contour and jugular notch of sternum) on the digital photographs [33,34]; measurements are then based on these fiducial points. All measurements are immediately visible to the user (Figure 3.1) and recorded on a database. Simultaneously, the set of measures is automatically converted into an objective assessment of the aesthetic result, based on a machine learning algorithm.

### 3.2 Software Automation

Many works tried to automate the detection above mentioned fiducial points. In [221] authors introduced a semi-automatic method to detect the breast contour in frontal colour images. The contour is found based on previously known start and end points. Subsequently, the algorithm automatically finds the contour in-between using graph theory. As a result, the algorithm finds the minimum cost path between previously known end points provided manually by the user. This line of work was improved in [34], where both the start and end points are found automatically.

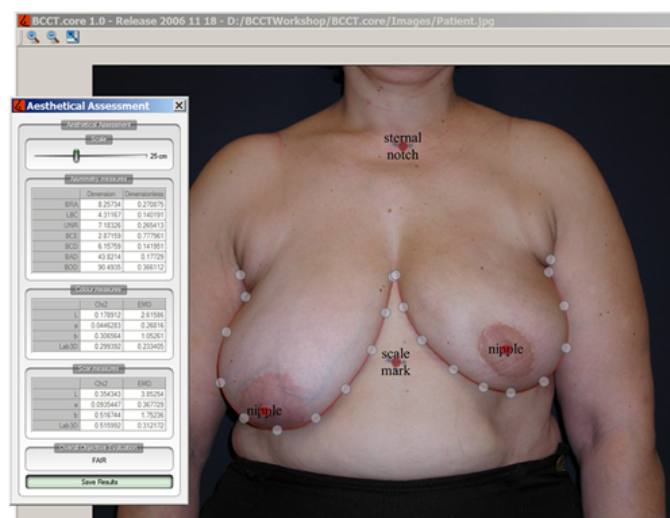


Figure 3.1: BCCT.core software interface.

External points of the contour are detected, considered the area of the body where the arm contour intersects the trunk contour. However, and depending on the patient's arm position, the contour of the arm and trunk can be overlapped, and thus it can be impossible to distinguish when patients put their upper limbs down. This problem is overcome by searching for the strongest vertical gradient lines. However, this does not work if patients assume different arm positions.

### 3.3 Feature Definition

When an observer evaluates the cosmetic result of BCCT, he identifies and assesses aspects, such as shape, colour, geometry, roughness and irregularity of the visual appearance of the treated breast, compared with the untreated breast. All these factors are subjectively combined into an overall assessment.

In the process of designing a software that mimics the human evaluation, one of the first steps is the generation of features. These features must capture all factors with impact on the overall aesthetic result and should be measurable in digital patient photographs after treatment.

As possible sets of discriminative features, were used those already identified by experts as relevant to the human evaluation of surgery procedures. It is universally accepted that the aesthetic outcome is essentially determined by visible skin changes, radiotherapy consequences such as hyperpigmentation or telangiectasia, the scar caused by the surgery and by changes in breast size or shape [149].

#### 3.3.1 Asymmetry Features

The most disclosed measures used on the objective evaluation of aesthetic results are based on the asymmetry between the treated and non-treated breast. Having different levels of asymmetry, it

was chosen to measure all indices presented in the literature, without limiting the performance of subsequent tasks. Some asymmetry measurements were also introduced, which seemed to complement the existing ones [28]. The inserted measurements capture aspects, such as asymmetry on the nipple position, differences on the lowest vertical position of the breast contour, breast area differences, breast contour length differences, and others. All these measurements require a conversion to the metric scale, using marks made on the patient's skin at a known distance.

The indices, recorded to assess breast asymmetry were the following (see Figure 3.2):

- Breast retraction assessment (BRA) =  $\sqrt{(X_1 - X_2)^2 + (Y_1 - Y_2)^2}$ : quantifies the difference in nipple position in both breasts and reflects the degree of breast retraction;
- Lower breast contour (LBC): difference between levels of inferior breast contour;
- Upward nipple retraction (UNR):  $= |Y_1 - Y_2|$ : difference between nipple levels;
- Breast compliance evaluation (BCE):  $|NI_1 - NI_2|$ : difference between left and right nipple to IMF distance;
- Breast contour difference (BCD): the difference between lengths of left and right breast contours;
- Breast area difference (BAD): the difference between areas of left and right breasts;
- Breast overlap difference (BOD): the non-overlapping area of the two breast after flipping one of them along a vertical line and coinciding both points of junction with the thorax.

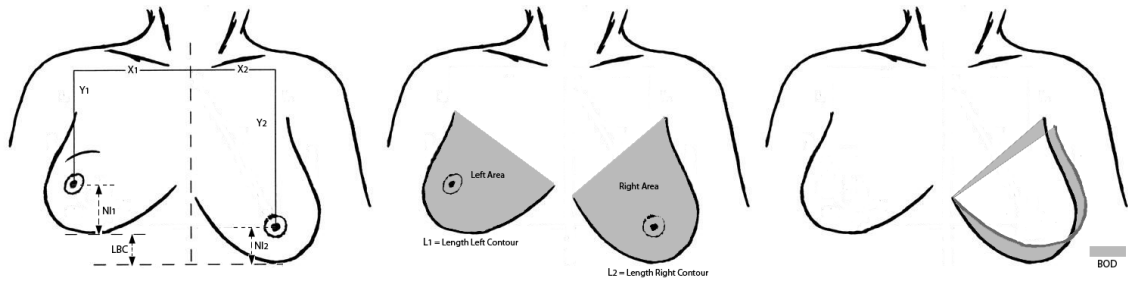


Figure 3.2: Illustration of measurements for asymmetry features.

The need to use additional marks conditions the automation of the process and, for this reason, the adequacy of relative measures was studied. These measurements are the result of a difference ratio between values taken from the treated and non-treated breast:

$$\frac{|V_1 - V_2|}{0.5 \times (V_1 + V_2)} \quad (3.1)$$

where  $V_1$  and  $V_2$  are values measured on the treated and non-treated breast. Using relative measurements, it becomes unnecessary to use the scale mark on the patient.

The dimensionless asymmetry measurements, with respect to dimension ones, are:

- $pBRA = \frac{BRA}{(\sqrt{X_1^2+Y_1^2}+\sqrt{X_2^2+Y_2^2})/2}$ : quantifies the dimensionless difference in nipple position in both breasts and reflects the degree of breast retraction;
- $pLBC = \frac{LBC}{(Y_1+NI_1+Y_2+NI_2)/2}$ : dimensionless difference between levels of inferior breast contour;
- $pUNR = \frac{UNR}{(Y_1+Y_2)/2}$ : dimensionless difference between nipple levels;
- $pBCE = \frac{BCE}{(NI_1+NI_2)/2}$ : dimensionless difference between left and right nipple to IMF distance;
- $pBCD = \frac{BCD}{(L_1+L_2)/2}$ : the dimensionless difference between lengths of left and right breast contours;
- $pBAD = \frac{BAD}{(right\ area+left\ area)/2}$ : the dimensionless difference between areas of left and right breasts;
- $pBOD = \frac{BOD}{(right\ area+left\ area)/2}$ : the dimensionless non-overlapping area of the two breast after flipping one of them along a vertical line and coinciding both points of junction with the thorax.

### 3.3.2 Colour Difference Features

Others measures were also considered. The skin colour changes were introduced for the first time by Cardoso and Cardoso [28], measuring the difference of colour between the treated and non-treated breast, an aspect that is highly valued in the subjective evaluation because hyperpigmentation is a typical consequence of radiotherapy, and it depends largely on the dose of radiation that is used.

The task of measuring the colour difference between images or between parts of the same image takes a central role in many areas of engineering and is performed as follows: first, compute the colour distribution in the important regions to compare (in this case histograms are used as an estimate of the distribution); second, choose a comparative measurement for the histograms, which adequately reflects the difference between the distributions.

The selection of the two regions to compare only involves the selection of two breast contours (Figure 3.3).

Several measurements were made, without limiting the subsequent operations: the colour difference on each channel and on the three channels were computed simultaneously. Each of these comparisons were made using two different metrics: the  $\chi^2$  rate and the Earth Mover's Distance (EMD). The several experiments made with various colour spaces did not show important differences, so only the CIE  $L^*a^*b$  colour space was kept to be more perceptually uniform. A total of eight colour changes measures are collected [28]. The list of features is:

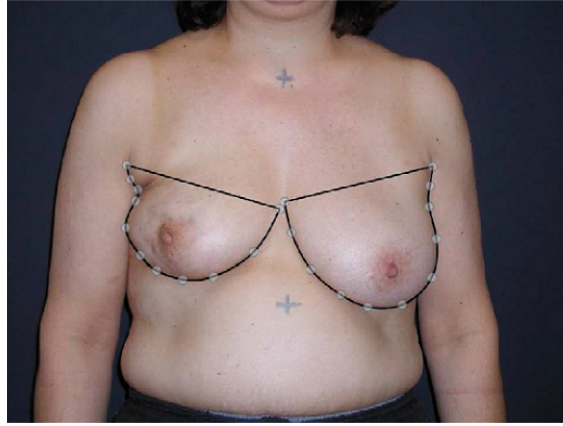


Figure 3.3: Breast mask for the extraction of colour features.

- $c\chi_L^2$ : distance between the histograms of the left and right breasts measured in the  $L$  channel of the CIE  $L^*a^*b^*$  colour space, using the  $\chi^2$  measure;
- $c\chi_a^2$ : distance between the histograms of the left and right breasts measured in the  $a$  channel of the CIE  $L^*a^*b^*$  colour space, using the  $\chi^2$  measure;
- $c\chi_b^2$ : distance between the histograms of the left and right breasts measured in the  $b$  channel of the CIE  $L^*a^*b^*$  colour space, using the  $\chi^2$  measure;
- $c\chi_{Lab3D}^2$ : distance between the histograms of the left and right breasts measured in the 3 channels of the CIE  $L^*a^*b^*$  colour space, using the  $\chi^2$  measure;
- $cEMD_L$ : distance between the histograms of the left and right breasts measured in the  $L$  channel of the CIE  $L^*a^*b^*$  colour space, using the EMD;
- $cEMD_a$ : distance between the histograms of the left and right breasts measured in the  $a$  channel of the CIE  $L^*a^*b^*$  colour space, using the EMD;
- $cEMD_b$ : distance between the histograms of the left and right breasts measured in the  $b$  channel of the CIE  $L^*a^*b^*$  colour space, using the EMD;
- $cEMD_{Lab}$ : distance between the histograms of the left and right breasts measured in the 3 channels of the CIE  $L^*a^*b^*$  colour space, using the EMD.

### 3.3.3 Scar Visibility Features

The contribution of surgical scars is the most difficult to quantify due to their enormous variability. Aspects such as size, position and contrast with the surroundings contribute to that perception. In the BCCT.core software, the visibility of the scar was translated as local difference of colour. To calculate this difference, the breast was divided into 12 angular sectors ( $30^\circ$ ) with the vertex at

the nipple (Figure 3.4) and the subsequent calculation of colour difference between corresponding sectors, following the same procedure as that described for skin colour changes features; the visibility of the scar is translated as the maximum of these differences.

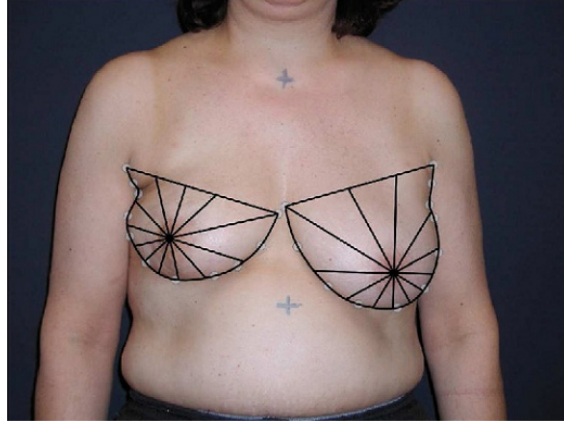


Figure 3.4: Breast sectors for the extraction of local colour changes induced by the surgical scar.

The list of features related with surgical scar is:

- $s\chi_L^2$ : visibility of the surgical scar captured in the  $L$  channel of the CIE  $L^*a^*b^*$  colour space, using the  $\chi^2$  measure;
- $s\chi_a^2$ : visibility of the surgical scar captured in the  $a$  channel of the CIE  $L^*a^*b^*$  colour space, using the  $\chi^2$  measure;
- $s\chi_b^2$ : visibility of the surgical scar captured in the  $b$  channel of the CIE  $L^*a^*b^*$  colour space, using the  $\chi^2$  measure;
- $s\chi_{Lab3D}^2$ : visibility of the surgical scar captured in the 3 channels of the CIE  $L^*a^*b^*$  colour space, using the  $\chi^2$  measure;
- $sEMD_L$ : visibility of the surgical scar captured in the  $L$  channel of the CIE  $L^*a^*b^*$  colour space, using the EMD;
- $sEMD_a$ : visibility of the surgical scar captured in the  $a$  channel of the CIE  $L^*a^*b^*$  colour space, using the EMD;
- $sEMD_b$ : visibility of the surgical scar captured in the  $b$  channel of the CIE  $L^*a^*b^*$  colour space, using the EMD;
- $sEMD_{Lab}$ : visibility of the surgical scar captured in the 3 channels of the CIE  $L^*a^*b^*$  colour space, using the EMD.

### 3.4 The Aesthetic Classifier

Development an objective method for evaluating the aesthetic result of BCCT using digital images calls for a standard for result comparison (gold standard). As a solution to this problem, a consensus evaluation was used performed by a panel of observers, experts in the field of Breast Cancer Treatment [28,40]. The model was built over a database of digital images of patients submitted to the BCCT (see Section 2.9) and an evaluation of the aesthetic result is provided for each image resulting from the consensus of the panel. The method attempts to mimic the ground truth evaluation.

The set of measures is automatically converted into an overall objective classification of the aesthetical result, using a Support Vector Machine (SVM) classifier trained to predict the overall aesthetical classification on the aforementioned scale of four classes.

SVMs have proved to be capable of representing complex classification or mapping functions. They discover the representations using powerful learning algorithms. In the problem addressed here, there is an inherent ordering between the classes, which motivated the use of models specially targeted for this kind of ordinal data. Using a mapping of the data replication method for ordinal data in SVMs [29], and performing a simplified feature selection, was reached an SVM classifier through a Radial Basis Function (RBF) kernel and requiring 4 features only: two to capture asymmetry, one for skin colour changes and one for scar visibility [28].

### 3.5 Conclusion

BCCT.core is currently being used by many international groups in prospective studies, including: Nottingham Breast Institute, UK; Leiden University Medical Centre, The Netherlands; Cancer Care Center, Sydney, Australia; University of Heidelberg, Breast Center, Heidelberg, Germany; Medical University, Vienna, Austria. Moreover, comprehensive reports can be generated and the results can be stored in a database to facilitate trend and statistical analyses. The hassle-free installation and the intuitive windows-based interface of the BCCT.core allow users to start working on cosmetic quantification immediately.

Although it is innovative and reproducible, this software has some issues that require improvements, as is often suggested by the users. Suggested improvements include the model's interpretability relating the aesthetic result with input measurements and the fact that only frontal view photographs are used, disregarding information from side or oblique views. However, there are important setbacks related to the complete automation of the software – which is fundamental for high reproducibility – and the capability to extract volumetric information to improve the overall cosmetic evaluation.





## Chapter 4

# Database

As stated in Section 2.9, the BCCT researchers usually do not provide access to the data used in their work. That fact creates additional difficulties when researchers need to compare their results. Some parts of the work conducted in this thesis used the Breast Research Group Database (see Section 2.9). However, this database is only composed of frontal and lateral photographs of patients, which is not enough to perform volumetric measurements. In order to perform the proposal work, it was necessary to create a database meeting the requirements of this project.

### 4.1 Database Requirements

The process of creating a database for aesthetic evaluation in BCCT should meet specific requirements:

1. The collection of data should be performed by a Breast Cancer specialist;
2. Patients who participate in the study have to be subjected to a surgical intervention and, as far as possible, include a significant number of cases of each type of intervention;
3. Ground truth annotation should also be available concerning fiducial points position (nipples position, suprasternal notch position) and breast contour definition;
4. Subjective aesthetic evaluation is fundamental for researchers to develop their algorithms. This subjective evaluation must preferably be performed by a panel of specialists;
5. All confidential medical information should not be included in the data;
6. When possible, the database should be available together with ground truth annotation on the Internet so that other researchers can use it. The continuous user support is also indispensable.

## 4.2 Database Construction

The acquisition of pictures in medical environment is not always easy to perform. Patients must agree with the process, and sometimes doing that on a voluntary basis makes it difficult to increase the number of cases in the database. Currently, the database includes data from 135 patients acquired during several sessions throughout more than one year, performed by a breast cancer expert physician. The complete information of each acquired patient can be found in Table 4.1.

For each patient the acquired data include colour images (from different angles) acquired with a standard 5MP portable camera (See Figure 4.1), as well as several (up to six) depth-map and colour frontal images acquired with the Kinect<sup>1</sup> (640 × 480 px) (See Figure 4.2). However, some patients do not include all modalities data. This was caused mainly by three issues: technical problems with one of the acquisition devices; logistical problems related to the patient; and the adaptation of thesis requirements, throughout the 4 years.

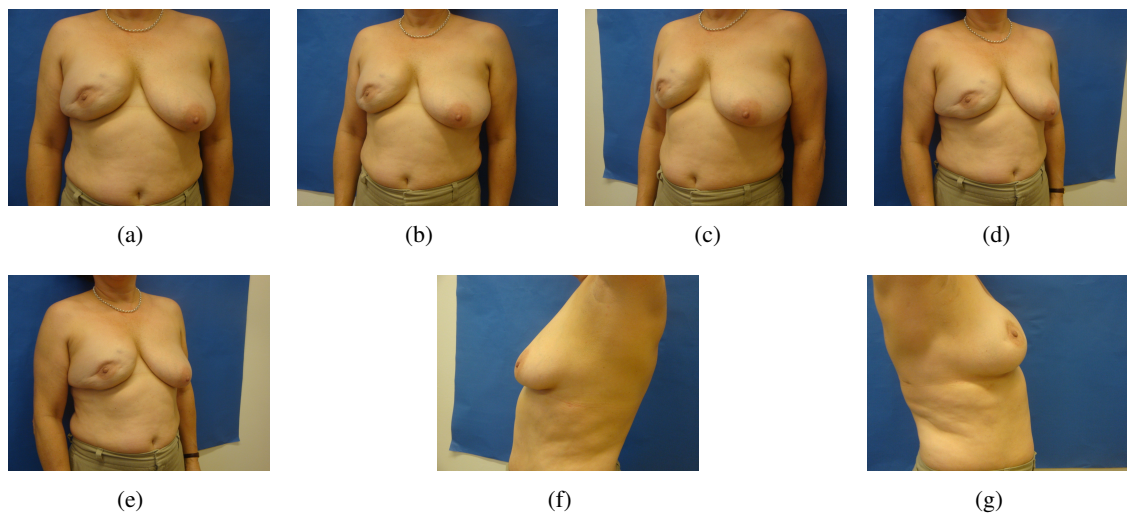


Figure 4.1: Data acquired with a portable camera (Patient #15).

Subjective evaluations of the aesthetic results are also provided based on the Harris scale [96]. Ideally the overall aesthetic assessment should correlate coherently with a significant number of expert assessments, from different areas of the world, which probably provide the desired reference classification [40]. This evaluation panel has been planned; however, final assessment decision was not achieved until this work was concluded. In order to provide the subjective evaluation on the images acquired from patients, two experts in the field of breast cancer have voluntarily provided their evaluation. In Table 4.2, it is possible to visualise the distribution of patients over the four classes for the two evaluators.

In Table 4.2 it is possible to confirm that there are some discrepancies when comparing the evaluations of the two evaluators. For example, evaluator #2 gives a better aesthetic assessment

<sup>1</sup><http://www.xbox.com/en-US/KINECT>

Pat. #	Acqui. date	Nipple Height		HD cam. # views	Kinect (# photos)		Sub. Eval.	
		Left	Right		Depth	Colour	Phy. #1	Phy. #2
1	22-03-2011	n/a	n/a	7	3	n/a	n/a	n/a
2	22-03-2011	n/a	n/a	7	4	n/a	n/a	n/a
3	22-03-2011	n/a	n/a	7	3	n/a	n/a	n/a
4	22-03-2011	n/a	n/a	7	3	n/a	n/a	n/a
5	22-03-2011	n/a	n/a	7	6	n/a	n/a	n/a
6	29-03-2011	5	6.5	7	3	n/a	2	1
7	29-03-2011	2	2.8	7	4	n/a	3	2
8	29-03-2011	3.2	4.5	7	4	n/a	2	2
9	29-03-2011	4	6.5	7	3	n/a	2	1
10	29-03-2011	3.5	4	7	3	n/a	1	2
11	29-03-2011	3.5	5.5	7	4	n/a	3	2
12	31-03-2011	4.5	4.5	7	6	n/a	1	3
13	31-03-2011	3	3	1	3	n/a	1	2
14	31-03-2011	6	4.5	7	3	n/a	4	2
15	05-04-2011	6	5	7	4	n/a	4	4
16	05-04-2011	2.5	3	7	5	n/a	4	1
17	05-04-2011	2.5	4	7	2	n/a	1	1
18	15-04-2011	5.8	2.3	7	2	n/a	4	3
19	15-04-2011	3.6	4.2	7	3	n/a	1	2
20	15-04-2011	3.8	2.7	7	3	n/a	4	2
21	15-04-2011	2.4	4	7	3	n/a	4	2
22	15-04-2011	3.4	3.5	7	3	n/a	3	1
23	15-04-2011	1.3	1.3	7	3	n/a	2	1
24	15-04-2011	2.9	3.2	7	3	n/a	1	1
25	15-04-2011	3.4	2.4	7	3	n/a	3	3
26	17-05-2011	5	5.9	7	3	n/a	2	2
27	17-05-2011	5.6	7	7	3	n/a	3	3
28	17-05-2011	2.9	3.4	7	4	n/a	2	1
29	17-05-2011	3.1	2.5	7	3	n/a	1	1
30	17-05-2011	2.6	2.5	7	3	n/a	2	2
31	17-05-2011	4.5	4.1	7	3	n/a	2	1
32	17-05-2011	5	4.5	7	3	n/a	1	2
33	24-05-2011	5.1	4	7	3	n/a	3	2
34	24-05-2011	5	5.8	7	3	n/a	4	2
35	24-05-2011	5	3.4	7	3	n/a	4	4
36	24-05-2011	2.5	1.3	7	3	n/a	2	1
37	24-05-2011	2.7	3.5	7	3	n/a	1	2
38	24-05-2011	3.6	3.3	7	1	n/a	1	2
39	24-05-2011	4.2	4.6	7	3	n/a	4	4
40	24-05-2011	3.3	3.7	7	3	n/a	1	3
41	24-05-2011	2.8	2.4	7	3	n/a	4	3
42	24-05-2011	3.4	3.9	7	3	n/a	1	2
43	24-05-2011	2.1	1.3	7	3	n/a	3	3
44	24-05-2011	2.4	2.5	7	3	n/a	4	4
45	24-05-2011	2.5	1.4	7	3	n/a	2	2
46	24-05-2011	4.7	3	7	3	n/a	1	1
47	24-05-2011	2.9	3.1	7	2	n/a	4	3
48	05-04-2011	5	6	7	n/a	n/a	2	1
49	15-04-2011	0.9	1.4	7	n/a	n/a	2	1
50	24-05-2011	5.9	4.5	7	n/a	n/a	1	1
51	22-11-2011	4.4	4.8	3	3	3	4	4
52	22-11-2011	4.2	3.9	3	3	3	3	2
53	22-11-2011	2.5	2.1	3	2	2	3	2
54	22-11-2011	2.3	1.7	3	3	3	2	1
55	22-11-2011	3.7	1.9	3	2	2	2	2
56	16-03-2012	n/a	n/a	3	2	n/a	1	1
57	16-03-2012	n/a	n/a	3	2	n/a	4	3
58	16-03-2012	n/a	n/a	3	3	n/a	3	2
59	16-03-2012	n/a	n/a	3	3	n/a	2	1

Pat. #	Acqui. date	Nipple Height		HD cam. # views	Kinect (# photos)		Sub. Eval.	
		Left	Right		Depth	Colour	Phy. #1	Phy. #2
60	16-03-2012	n/a	n/a	3	3	n/a	2	1
61	21-03-2012	n/a	n/a	3	2	2	1	1
62	21-03-2012	n/a	n/a	3	4	4	2	1
63	21-03-2012	n/a	n/a	3	2	2	4	3
64	21-03-2012	n/a	n/a	3	1	1	2	1
65	21-03-2012	n/a	n/a	3	n/a	n/a	1	1
66	23-03-2012	2.2	3	3	5	5	1	2
67	23-03-2012	5.4	5.4	3	5	5	1	2
68	23-03-2012	6	8.5	3	5	5	4	2
69	23-03-2012	n/a	n/a	3	5	5	2	1
70	28-03-2012	3.6	4	3	3	3	4	2
71	28-03-2012	1.3	0.8	3	3	3	1	1
72	28-03-2012	1.8	2.6	3	4	4	2	2
73	28-03-2012	4.8	4.4	3	5	5	1	1
74	28-03-2012	6.1	4.9	3	3	3	2	1
75	28-03-2012	3.3	3.8	3	4	4	1	1
76	28-03-2012	8	8	3	3	3	2	1
77	30-03-2012	4.7	5.8	3	2	2	2	1
78	30-03-2012	2.3	3.4	3	3	3	2	2
79	30-03-2012	3.8	3.6	3	4	4	1	1
80	04-04-2012	4.6	4.3	3	3	3	2	1
81	04-04-2012	3.9	3.7	3	3	3	3	1
82	04-04-2012	4.1	3.8	3	4	4	3	3
83	04-04-2012	2.5	2.4	3	4	4	1	1
84	04-04-2012	4.1	4.7	3	4	4	2	1
85	04-04-2012	3.2	3.8	3	5	5	2	2
86	04-04-2012	2.7	3.7	3	4	4	1	1
87	04-04-2012	3.2	4.2	3	5	5	2	1
88	11-04-2012	3.7	4.7	3	4	4	2	1
89	11-04-2012	3	4	3	4	4	2	1
90	11-04-2012	1.5	0.9	3	2	2	3	2
91	11-04-2012	2.9	4.4	3	4	4	2	1
92	11-04-2012	6	5.3	3	5	5	4	2
93	11-04-2012	1.5	2.5	3	4	4	4	4
94	23-05-2012	3.6	3.6	3	5	5	1	1
95	18-04-2012	2.3	2.3	3	4	4	1	1
96	18-04-2012	5.7	5.3	3	4	4	3	2
97	18-04-2012	2.3	2.8	3	5	5	1	1
98	18-04-2012	4	4	3	5	5	1	1
99	18-04-2012	4.5	3.5	3	4	4	3	1
100	18-04-2012	4.4	4.1	3	5	5	2	1
101	20-04-2012	4.3	4.6	3	5	5	2	1
102	20-04-2012	3	3.3	3	2	2	4	2
103	27-04-2012	4	3.6	3	5	5	3	1
104	27-04-2012	3.1	3.5	3	5	5	2	1
105	27-04-2012	4.2	4.2	3	5	5	2	2
106	27-04-2012	2.2	3.1	3	5	5	1	1
107	27-04-2012	3.6	3.8	3	4	4	1	1
108	27-04-2012	3	3.5	3	5	5	1	2
109	27-04-2012	3.8	3.8	3	5	5	1	1
110	02-05-2012	3.5	3.5	3	5	5	3	3
111	11-05-2012	2	3.4	3	5	5	3	2
112	11-05-2012	1.9	4.2	3	4	4	3	3
113	16-05-2012	1.8	2.3	3	5	5	1	1
114	16-05-2012	2.6	2.6	3	5	5	1	1
115	16-05-2012	2.5	3.7	3	4	4	1	1
116	16-05-2012	3	4	3	5	5	2	1
117	18-05-2012	4.2	3.7	3	5	5	1	1
118	18-05-2012	3.9	3.3	3	5	5	3	2
119	23-05-2012	2.1	1.9	3	5	5	2	1

Pat. #	Acqui. date	Nipple Height		HD cam. # views	Kinect (# photos)		Sub. Eval.	
		Left	Right		Depth	Colour	Phy. #1	Phy. #2
120	23-05-2012	3.1	4.7	3	5	5	3	3
121	23-05-2012	3	3	3	5	5	1	1
122	25-05-2012	4.3	4	3	5	5	4	3
123	25-05-2012	4	2.6	3	4	4	2	2
124	25-05-2012	2.8	3.6	3	4	4	3	2
125	25-05-2012	2.8	2.2	3	4	4	1	1
126	25-05-2012	3.3	4.3	3	4	4	1	1
127	25-05-2012	3	2.6	3	5	5	2	1
128	06-06-2012	3.3	3.4	3	5	5	3	2
129	06-06-2012	4.2	4.2	3	4	4	2	1
130	06-06-2012	3.8	3.3	3	4	4	2	1
131	13-06-2012	5	4.1	3	3	3	1	1
132	13-06-2012	4.9	4.4	3	5	5	3	1
133	13-06-2012	3	2.5	3	5	5	3	2
134	13-06-2012	2.5	3.2	3	5	5	2	2
135	13-06-2012	4.1	4.6	3	5	5	3	1

Table 4.1: Patients information provided in the database.

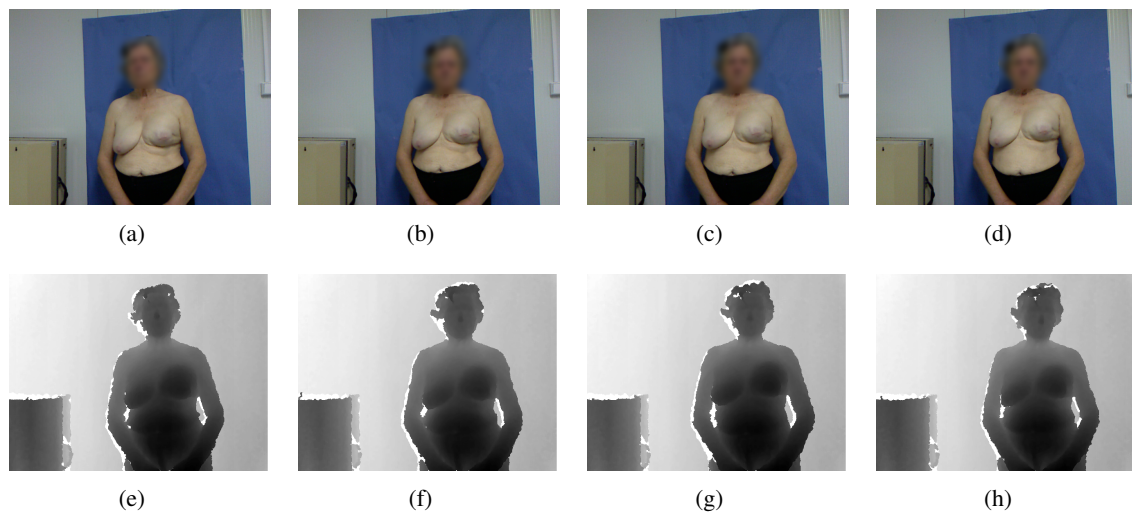


Figure 4.2: Data acquired using the Kinect device (Patient #93). Colour and corresponding depth-map image.

Class		Excellent	Good	Fair	Poor	Total
# cases	Phy. #1	41	42	26	21	130
	Phy. #2	67	42	15	6	

Table 4.2: Distribution of the 130 patients over the four classes.

comparatively to evaluator #1. To better understand the disagreement between evaluators, the correlation coefficient and Cohen's kappa coefficient are shown in Table 4.3.

By looking at Table 4.3, it is easy to observe that agreement is not very significant, corroborating the need for an evaluation performed by a panel of experts. Particularly, the Weighted

Disagreement	Correlation Coefficient	Kappa	Weighted Kappa
0.59	0.60	0.16	0.34

Table 4.3: Disagreement between evaluators.

Kappa (**wk**) coefficient presents a very low value, considered "fair", which is normally assessed as: 0-0.25 - poor; 0.26-0.50 - fair; 0.51-0.75 - moderate; 0.76-1 - good.

The Kappa statistics was designed to evaluate the real agreement between two different criteria to measure the same characteristic. This coefficient measures inter-rater agreement for qualitative characteristics. It is generally more robust than simple percent agreement calculation since Kappa statistics looks to the agreement occurring by chance. As a result, this measurement is now more and more used in medical environments [48]. Since aesthetic classification has a particularly order, Weighted Kappa can be used as it was especially designed for this kind of data, counting disagreements differently. In this case, the weighted matrix was as follows:

$$\mathbf{wk} = \begin{bmatrix} 1 & 2/3 & 1/3 & 0 \\ 2/3 & 1 & 2/3 & 1/3 \\ 1/3 & 2/3 & 1 & 2/3 \\ 0 & 1/3 & 2/3 & 1 \end{bmatrix}$$

Manual ground truth annotation was performed both in Kinect data and for the data acquired with a portable camera (frontal data only). With Kinect data the positioning of the breast peaks, defining breast contour and locating endpoints were annotated in depth-map images, as well as nipple position and suprasternal notch position in the Kinect colour image (see Figure 4.3).

Kinect data (depth and colour) are not aligned due to construction characteristics. This particularity will be further addressed in Section 8.5. In Figure 4.3, it is possible to visualise the ground truth annotation on some patient data. Breast contour and breast peak point can be visualised in the depth-map images, represented as cyan lines and yellow crosses respectively. While breast contour was manually annotated, the breast peak points (the area in the breast closer to the camera or further away from the chest wall, not necessarily the nipple) were semi-automatic defined, based on depth-intensity information in an area delimited manually. The nipple and suprasternal notch positions were annotated in the Kinect colour image.

Ground truth annotation for the data acquired with a portable camera was performed using the BCCT.core software for frontal patient data only, defining the breast contour and fiducial points (nipple position and suprasternal notch position) (see Figure 4.4).

Additionally, the height of both nipples (distance between the medial projection of the nipple and the sternum measured with 2 rulers) were manually obtained by the physician during the acquisition period (see Figure 4.5).

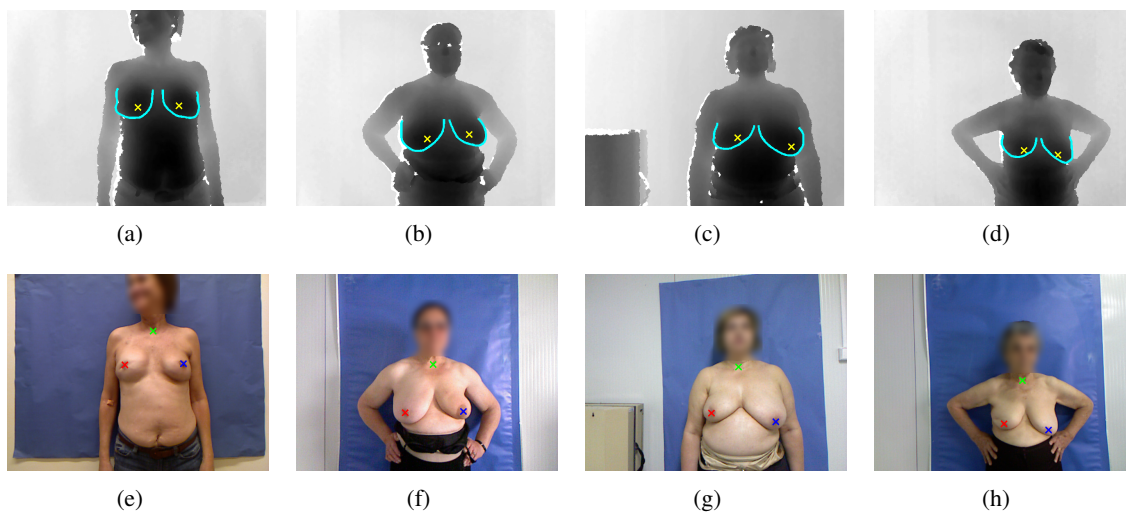


Figure 4.3: Kinect data ground truth annotation (breast contour definition, breast peak point position, nipple position and suprasternal notch position) (for Patient #55, #85, #92 and #122).

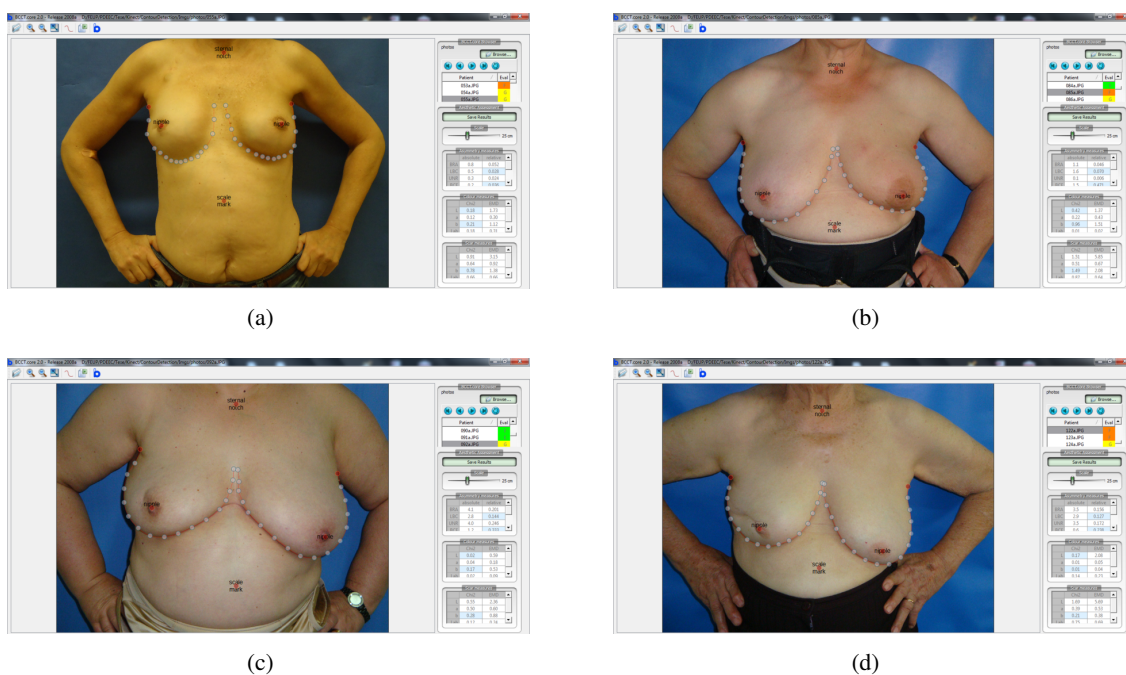


Figure 4.4: Portable camera data ground truth annotation (breast contour definition, nipple position and suprasternal notch position) (for Patient #55, #85, #92 and #122).

### 4.3 Conclusion

It is well recognised that breast cancer databases are not publicly available, particularly in the area of aesthetic result assessment, which difficult the comparison between researchers. In this thesis, we addressed this gap by proposing a multi-modal database based on colour and depth-





Figure 4.5: Nipple height measurement.

map images. We do acknowledge that not all images in the database respect all quality assessment criteria. Nevertheless, the database reflects a wide variability of cases and conveys the reality on patients which performed a breast surgical procedure.

Despite the fact that our database has a limited number of images, we strongly believe that it is enough to conduct the proposal objectives on this thesis. In any case, the number of cases can be increased in the future. Annotation is a subjective, tedious, and extremely time-consuming task, which should be preferentially performed by specialists, which can turn into an extremely difficult and costly task. In this work, there was a big concern in making precise annotations. However, while fiducial points annotation presents satisfactory results, the subjective evaluation for the aesthetic result was only performed by two specialists. For that reason, in future, will be integrated in the database a subjective evaluation performed by a panel of specialist with the indispensable agreement.

We consider that this database has the potential to be a greet value in the breast cancer field. With this database, we aim at increasing available resources in the breast imaging, facilitating the comparison algorithms of different investigators. This can motivate computer vision researchers to develop methodologies that take advantage of a multi-modality database. The high image diversity of this database will provide new challenges, but extremely useful to design more robust algorithms to evaluate the aesthetic result. For these reasons, we believe that this database can be a reference for future works in the breast cancer imaging area. Both patient data and ground truth annotation will be available at the website of the Breast Research Group <sup>1</sup>.

---

<sup>1</sup><http://medicalresearch.inescporto.pt/breastresearch/>



## Chapter 5

# Closed Shortest Path in Images\*

Shortest path algorithms are extensively used in image processing applications, ranging from staff line detection in music scores to detection and tracking of vibrating structures, such as lines and cables in civil engineering, encompassing fiducial point detection. There are applications where the starting and ending positions of the shortest path need to be constrained, defining a closed contour enclosing a previously detected seed. The detection of the external contour of the iris can be addressed as such for biometric applications. Usually the closed contour computation is addressed by transforming the image into polar coordinates, where the closed contour is transformed in an open contour between two opposite margins.

In this Chapter, after calling the attention to some limitations of this process, we show how to compute the closed contour in the original coordinates. After defining a directed acyclic graph appropriate for this problem, we address the main difficulty in operating in the original coordinates. Since small paths collapsing in the seed point are naturally favored, we modulate the cost of the edges to counterbalance that bias. The method is illustrated with examples from The Berkeley Segmentation Dataset and a thorough evaluation is conducted with datasets from the breast cancer field. During this thesis, this algorithm will be applied to compute the areola contour in RGB images (see Chapter 9).

### 5.1 Regional Detectors

The image interpretation has and still is being hampered by the lack of an efficient representation (or abstraction) for the visual data. One standard approach is to represent the image on the basis of the low-level features of images, such as colour, texture, and shape. Almost all these systems are founded on the premise that images can be characterised by global descriptors. The global descriptor consists of features computed on the whole image. For example in Rubner *et al.* [209], the authors proposed a Histogram based algorithm to characterise an image by its color distribution or histogram, using the EMD for matching histograms. Another option is based on local descriptors [241], building on the idea that some points in the image are more relevant than

---

\*Some portions of this Chapter appeared in [32]

others for operations such as alignment or establishment of a dense set of correspondences so that a 3D model can be constructed or an in-between view can be generated. During the feature detection (extraction) stage, each image is searched for locations that are likely to match well in other images. At the feature description stage, each region around detected keypoint locations is converted into a more compact and stable (invariant) descriptor that can be matched against other descriptors [228].

Regional descriptors are usually not considered since it is assumed that they require a segmentation step, typically named as a medium level operation, in opposition to the local and global descriptors, which are obtained by low level operations. These type of descriptors are typically divided into three groups: intensity-based regions, maximally stable extremal regions, and superpixels. These regions are provided by different methodologies but focus on similar image structures and share similar properties. Segmentation-based methods present limited robustness to changes in viewing conditions but are currently more and more used in the context of image recognition.

We argue that regional properties, by themselves or in conjugation with local and global descriptors, around key or candidate points can be relevant for the high level interpretation of the visual content. These type of detectors are typically used in a larger scale region to compute the descriptor, such that it also contains part of the surrounding image structures and captures the shape of the region boundary. This is usually enough to increase the discriminative power and match regions between images. Our work is oriented to relatively large image segments, which are too large for local features methodologies. For this reason, our approach focus on segmentation-based methods.

Here, the computation of a closed contour enclosing a seed point. Usually that is addressed by transforming the image to polar coordinates, where the closed contour is transformed in an open contour between two opposite margins. After calling the attention to some of limitations of this process, we show how to compute the closed contour in the original coordinates. Closed contours (segmented regions) will enclose a seed previously detected. This seed can be manually or automatically defined depending of the application.

## 5.2 Closed Contour Related Works

The detection of closed contours in an image is a common task in different problems. There are some methods which make assumptions about the characteristics of the content present in the scene, which can incorporate low-, mid-, and high-level shape priors [203]. Other methods are based on contour completion [204], which are a powerful tool to detect closed regions.

Other approaches based on Gestalt cues of parallelism, collinearity, proximity and symmetry to group contours have been also addressed [223, 259]. These methodologies faced an important challenge related with the complexity of pairwise contour grouping to detect symmetry-related contour pairs. Levinstein *et al.* [142, 143] attempt to overcome this computational complexity limitation by constraining the symmetric parts to be collections of superpixels.

Other methods use simple methodologies based on continuity and proximity, such as the notion of boundary gap, used to measure the absence of edges along the closed contour. The probability of the connection between two adjacent contours was modeled by Elder and Zucker [70], in order to find contours using a shortest path algorithm. Wang *et al.* [251] created a measure to evaluate the gap of a contour using the ratio cut approach, only based on the total boundary gap. The problem of perceptual completion of occluded surfaces was addressed by Williams and Hanson [255], defined by set of occluded surfaces from image contours computed with a labeled knot-diagram.

The stated methodologies find the appropriate closed contour with high complexity approaches, mostly, using heuristics to prune the range of search. In Zhu *et al.* [267] this problem is solved by transforming the edge into polar coordinates. The method of Jermyn and Ishikawa [116] computed the closed contour using ratio cuts working directly with pixels using a 4-connected image grid and by minimizing many different closure costs in a simply connected planar graph. However, this is not an optimal solution, since individual pixels representation difficult the gap computation.

The idea of transforming the problem of detecting edge segments in a problem of finding the minimum-cost path in a weighted and directed graph was initially proposed by Martelli [163, 164].

When comparing to other traditional methods, minimum cost segmentation has several advantages. Watershed techniques produce closed contours, but are known to oversegment the images. Active contours (or snakes) also produce a closed contour. But initialization needs to be close to the real contour. Live-wire methods use the minimum cost idea, but needs seed points (usually given by the user) that must belong to the final contour. Riverbed contributes to reduce the number of these seed points along complex shapes, but live wire is still better in dealing with badly defined borders [169]. Graph cuts have the disadvantage of being biased toward producing small contours.

For a survey on graph-related methods, we refer to [258]. If your interest is on medical-imaging applications, we point to [68, 194].

Despite the several advantages of the shortest path calculation, its traditional use to compute a closed contour needs a transformation to the polar domain (and back to the Cartesian domain) [227]. In this transformation resolution may be lost, discontinuities may be introduced and extra computational power and time are required.

### 5.3 Conventional Closed Shortest Path Algorithms

In order to understand the method here proposed, several background notions are necessary. In this Section graph terminology is introduced in Subsection 5.3.1, followed by the description of some traditional neighbourhoods and the design of the weight function in subsection 5.3.2. Subsection 5.3.3 shows how the computation of the shortest path is made and in subsection 5.3.4 methodologies to use the shortest path algorithm in the extraction of closed contours are detailed.

### 5.3.1 Graph Concepts\*

A graph  $G = (V, A)$  is composed of two sets  $V$  and  $A$ .  $V$  is the set of nodes, and  $A$  the set of arcs  $(p, q)$ ,  $p, q \in V$ . The graph is *directed* (digraph) if the edges have a direction associated with them (the set  $A$  is a set of ordered pairs of nodes). An arc  $a = (p, q)$  is considered to be directed from  $p$  to  $q$ ;  $q$  is called the *head* and  $p$  is called the *tail* of the arc;  $q$  is said to be a direct successor of  $p$ , and  $p$  is said to be a direct predecessor of  $q$ .

The graph is *weighted* if a weight  $w(p, q)$  is associated to each arc. On images, the weight of each arc,  $w(p, q)$ , is usually a function of pixels values and pixels relative positions. A path from vertex (pixel)  $v_1$  to vertex (pixel)  $v_n$  is a list of unique vertices  $v_1, v_2, \dots, v_n$ , with  $v_i$  and  $v_{i+1}$  corresponding to neighbour pixels. The total cost of a path is the sum of each arc weight in the path  $\sum_{i=2}^n w(v_{i-1}, v_i)$ .

A path from a source vertex  $v$  to a target vertex  $u$  is said to be the *shortest path* if its total cost is minimum among all  $v$ -to- $u$  paths. The distance between a source vertex  $v$  and a target vertex  $u$  on a graph,  $d(v, u)$ , is the total cost of the shortest path between  $v$  and  $u$ .

A path from a source vertex  $v$  to a sub-graph  $\Omega$  is said to be the shortest path between  $v$  and  $\Omega$  if its total cost is minimum among all  $v$ -to- $u \in \Omega$  paths. The distance from a node  $v$  to a sub-graph  $\Omega$ ,  $d(v, \Omega)$ , is the total cost of the shortest path between  $v$  and  $\Omega$ :

$$d(v, \Omega) = \min_{u \in \Omega} d(v, u). \quad (5.1)$$

A path from a sub-graph  $\Omega_1$  to a sub-graph  $\Omega_2$  is said to be the shortest path between  $\Omega_1$  and  $\Omega_2$  if its total cost is minimum among all  $v \in \Omega_1$ -to- $u \in \Omega_2$  paths. The distance from a sub-graph  $\Omega_1$  to a sub-graph  $\Omega_2$ ,  $d(\Omega_1, \Omega_2)$ , is the total cost of the shortest path between  $\Omega_1$  and  $\Omega_2$ :

$$d(\Omega_1, \Omega_2) = \min_{v \in \Omega_1, u \in \Omega_2} d(v, u). \quad (5.2)$$

A directed acyclic graph (DAG) is a directed graph with no directed cycles. That is, it is formed by a collection of vertices and directed edges, each edge connecting one vertex to another, such that there is no way to start at some vertex  $v$  and follow a sequence of edges that eventually loops back to  $v$  again.

A topological sort or topological ordering of a DAG is a linear ordering of its vertices such that, for every edge  $(u, v)$ ,  $u$  comes before  $v$  in the ordering. Equivalently, each DAG gives rise to, at least, a partial order  $\leq$  on its vertices, where  $u \leq v$  exactly when there exists a directed path from  $u$  to  $v$  in the DAG. The DAG shown in Figure 5.1 has many valid topological sorts, including  $\{7, 5, 3, 11, 8, 2, 9, 10\}$ ,  $\{3, 5, 7, 8, 11, 2, 9, 10\}$  and  $\{3, 7, 8, 5, 11, 10, 2, 9\}$ .

Some algorithms become simpler when used on DAGs instead of general graphs, based on the principle of topological ordering. For example, it is possible to find shortest paths and longest paths from a given starting vertex in DAGs in linear time by processing the vertices in a topological

---

\*Some parts of this Section will appear again in Chapter 9. Since this Chapter is mostly based on a journal paper, we decided to replicate the information to maintain the coherence of the text

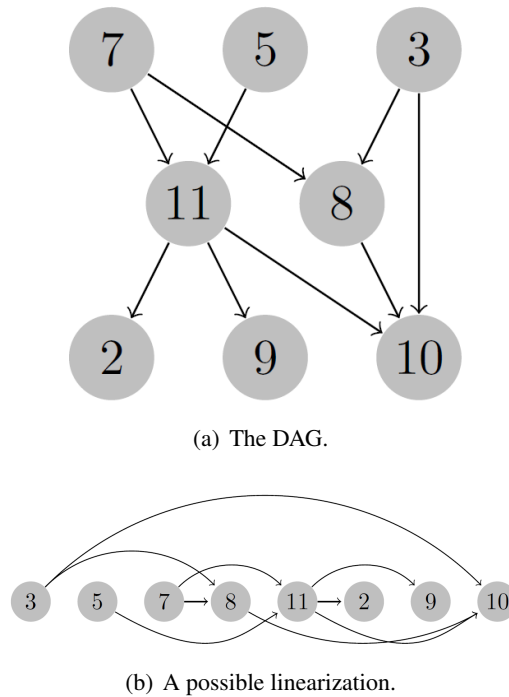


Figure 5.1: Example of a DAG with several topological orderings.

order, and calculating the path length for each vertex to be the minimum or maximum length obtained via any of its incoming edges. In contrast, for arbitrary graphs the shortest path may require slower algorithms such as Dijkstra's algorithm or the Bellman-Ford algorithm, and longest paths in arbitrary graphs are NP-hard to find.

### 5.3.2 Design of the Weight Function

The  $d$ -neighbourhood of a node (pixel)  $p$  is the set of the  $d$  closest nodes of  $p$ , for some adopted distance (usually the Euclidean distance). We defined the causal neighbourhood of a node  $p$  as the subset of neighbours  $v_i$  such that  $v_i \leq p$ , that is,  $v_i$  precedes  $p$  in the topological ordering. Figure 5.2 illustrates some typical neighbourhoods in images and the corresponding causal neighbourhoods, assuming the linear ordering of the vertices is given by column number.

The weight of an edge in the graph is usually set as a function of the values of the incident nodes (pixels). One starts by computing some feature(s) (e.g., gradient or directional derivatives, texture, etc.) and the distance between the node and each of its causal neighbours. Finally the cost is set as a monotonic function of these. Then the shortest path should minimize the sum of the costs over the path. Like [227], we will assume that the cost is a function of the features computed at the head of the arc and the distance to the causal neighbours.

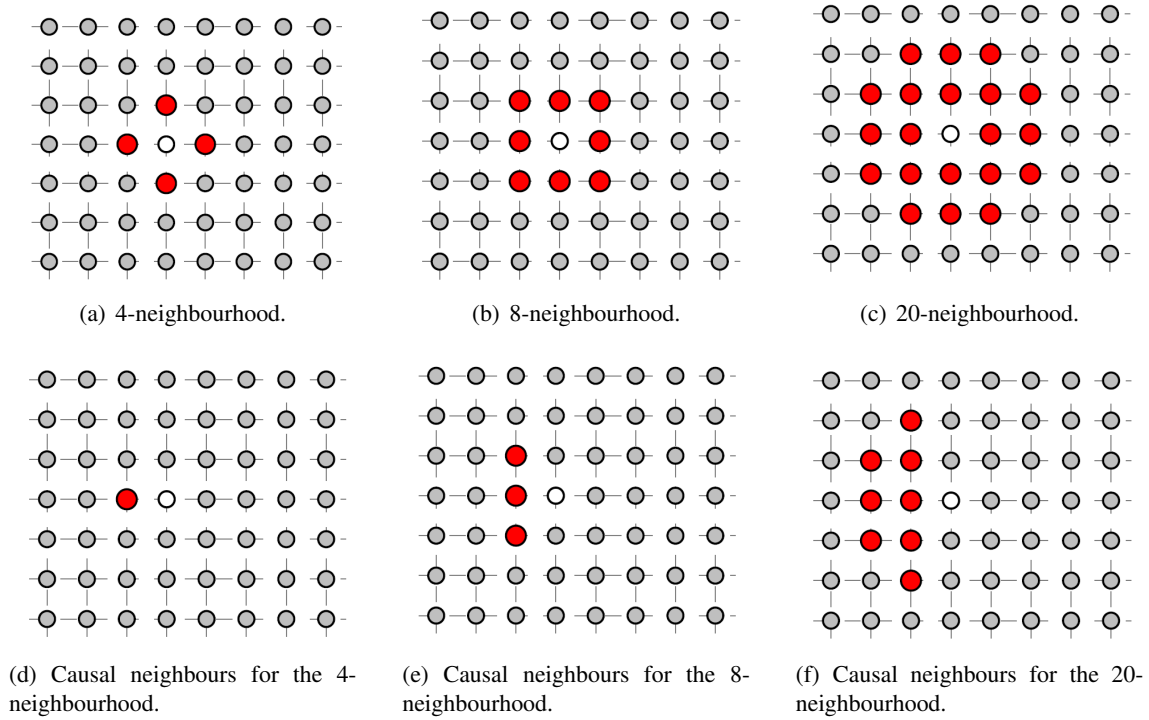


Figure 5.2: Examples of neighbourhoods and corresponding causal neighbours.

### 5.3.3 (Open) Shortest Path Extraction Algorithm in a DAG

Figure 5.3 represents a DAG appropriate to find a path from the left side to the right side of an image or grid such that the cost of the path is minimum. Note that: a) we have assumed the 8-neighbourhood; b) this DAG does not allow paths that zigzag back and forth, since at least one arc would be traversed in the wrong direction, an example of which is presented in Figure 5.4; c) the DAG to compute the shortest path from right to left is obtained by inverting the direction of the arcs.

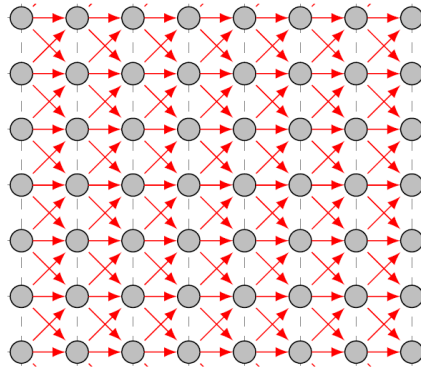


Figure 5.3: The nodes represent pixels and arcs represent connections between pixels.

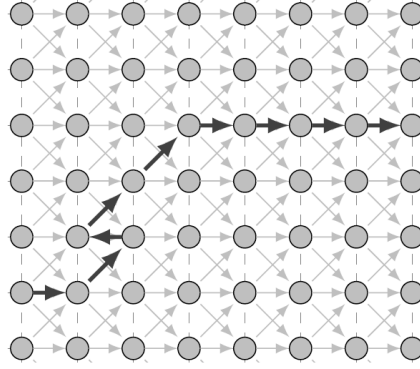


Figure 5.4: A forbidden path in the digraph of Figure 5.3.

In the case of a DAG, the shortest path is easily solved by examining the nodes according to the natural order and for each of them applying the following typical iteration (see Algorithm 1):

---

**Algorithm 1:** Shortest path in a DAG.

---

```

1 for node  $j = 2, 3, \dots, N$  do
2    $d_j \leftarrow \min(d_i + c_{ij}) : (i, j) \text{ enters } j,$ 
3 end
```

---

where  $d_i$  is the accumulated cost of the path until  $i$ . Although it is also possible to follow a forward form, we have adopted backwards computation, leading to the classic dynamic programming algorithm [227].

For the grid shown in Figure 5.3, the ordering is given by the column position. The linearisation step can be removed since it amounts to scan the image column by column, from top to bottom (although inside a column the scanning could be different). The computation of the shortest path can be rephrased as follows, directly in the image grid. Let  $I$  be an  $N_1 \times N_2$  window (after polar coordinate transform) with  $N_1$  columns and  $N_2$  rows; define an admissible path to be

$$\mathbf{s} = \{(x, y(x))\}_{x=1}^{N_1}, \text{ s.t. } \forall x \ y(x-1) \in \text{Neighbourhood}(x, y),$$

where  $y$  is a mapping  $y : [1, \dots, N_1] \rightarrow [1, \dots, N_2]$ . That is, an admissible path is a path of pixels in the image from left to right, containing at most one pixel in each column of the image and consecutive pixels in the path are neighbours.

The first step is to traverse the image from the second column to the last column and compute the cumulative minimum cost  $d$  for each entry  $(i, j)$ :

$$d(i, j) = \min(d(\ell, k) + c(p_{\ell, k}; p_{i, j})) : (\ell, k) \text{ enters pixel } (i, j)$$

where  $c(p_{\ell, k}; p_{i, j})$  represents the weight of the edge incident with pixels at positions  $(i, j)$  and

$(\ell, k)$ . At the end of this process,

$$\min_{j \in \{1, \dots, N_2\}} d(N_1, j)$$

indicates the end of the minimal connected path. (Note that here we used a double index notation to represent pixels (nodes).) Hence, in the second step, one backtrack from this minimum entry on  $d$  to find the optimal path.

### 5.3.4 Closed Shortest Path Extraction Algorithm in Polar Coordinates

The method discussed in the previous subsection can be applied to the computation of closed shortest paths in images. Given a point that should be enclosed by the path, one starts by performing a transformation to polar coordinates using the given point as origin. With this transformation the closed shortest path results in the shortest path between opposite margins, with the constraint that it should start and end in the “same” line. (Note that the line does not need to be exactly the same; the start and end points only need to be neighbours, assuming that the image wraps around, e.g. the leftmost and the rightmost columns are actually neighbouring columns.)

In spite of the efficiency of the computation of the shortest path between the whole left and right margins, or between two pre-defined points in the margins, or between one of the margins and a pre-defined point in the other margin, the search for the shortest path between the left and right margins with the constraint that the path should start and end in the same row seems to increase the complexity of the procedure. As typical, optimisation with constraints is more difficult than without.

Had one been interested in the simple shortest path between the left and right margin and the computation would be very efficiently performed using dynamic programming, as seen above. Note that the above procedure gives not only the shortest path between the left and right margins but also yields the shortest path between any point in the right margin and the whole left margin: for any point  $j$  in the right margin,  $C(N_1, j)$  indicates the cost of the shortest path between  $j$  and the whole left margin, see Figure 5.5. Finally, it should be clear how to change the initial conditions of the above procedure to yield the shortest path between two pre-defined points in the opposite margins.

Unfortunately, the computation of the shortest path constrained to start and end in the same row (corresponding to closed contours in the original window) does not seem amenable to such an efficient procedure. The brute force solution of computing the shortest path between the  $i$ -point in the left margin and the  $i$ -point in the right margin, for  $i = 1 \dots N_2$ , and taking the minimum, is not compatible with requirements of near real-time in many applications.

Noting that if  $j$  and  $\ell$  are two distinct points in the right margin, then the shortest paths between each of these points and the whole left margin do not intersect, it is trivial to conclude that there is a point  $m$  in the right margin for which the shortest path between  $m$  and the whole left margin starts also at row  $m$ . Note that this path  $\mathcal{P}_1$  corresponds to a closed path in the original coordinates (not necessarily the shortest one). Similarly, interchanging the role of the left and right margin, it is possible to obtain a point  $n$  in the left margin for which the shortest path  $\mathcal{P}_2$  to the whole right

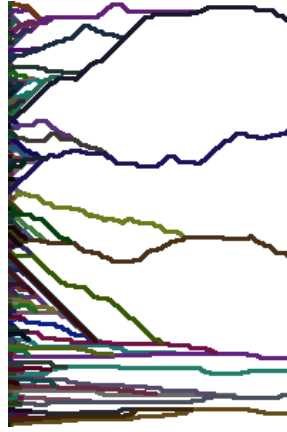


margin is closed. Therefore, we chose to approximate the (optimal) shortest path between the left and right margins starting and ending on the same row by  $\min(\mathcal{P}_1, \mathcal{P}_2)$ .

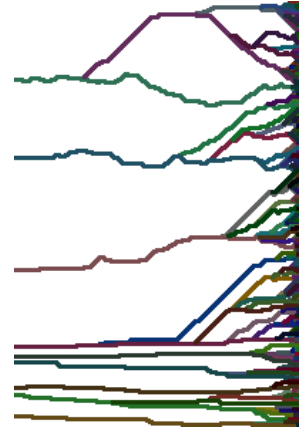
Note that this algorithm corresponds to the multiple back-tracking algorithm (MBTA) of [227] but we execute it both from left to right and from right to left. In Figure 5.5(b) the shortest paths between *each point* on the left margin and the *whole* right margin are traced. Likewise, Figure 5.5(c) shows the shortest paths between *each point* on the right margin and the *whole* left margin.



(a) Original image.



(b) Shortest paths from each pixel in the left column and the whole right column.



(c) Shortest paths from each pixel in the right column and the whole left column.

Figure 5.5: Exemplification of shortest paths.

The authors in [227] failed to realize that there is always at least one closed path among the paths obtained with MBTA (“We have found that on an image or regular grid one can *almost always* find a circular path although this circular path may not be the circular shortest path.”). Here we prove that it is indeed possible to always find at least one.

*Proof.* A key argument in the demonstration is that the shortest path between a whole margin and a point  $p$  in the opposite margin does not intersect the shortest path between the whole margin and a different point  $q$ . Let’s prove by contradiction: if the two paths do intersect in a point  $a$  (see Figure 5.6) then one of the paths can be improved by taking the other sub-path between  $a$  and the whole margin. In the special case that both sub-paths between  $a$  and the whole margin have the same cost then we will always make the same decision (e.g., always choosing the option with the starting point in the highest row). Let’s now consider the path ending in the first line. If the starting point of the path is also on the first line we are done; otherwise the starting line is higher than 1 ( $\geq 2$ ). Let’s now consider the path ending in the second line. If the starting point of the path is also on second line (note that it cannot be on the first line by the argument before) we are done; otherwise the starting line is higher than 2 ( $\geq 3$ ). Proceeding until the path ending in last line, the path cannot start in the line before that (nor after!), and therefore the path is closed.  $\square$

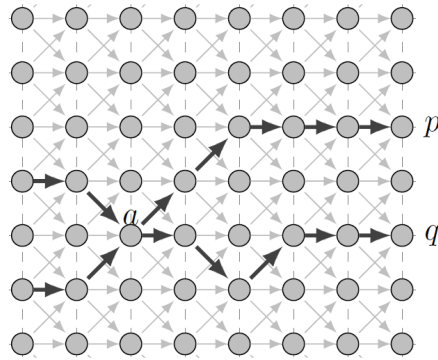


Figure 5.6: Two paths intersecting in  $a$ . One path is the shortest path between  $p$  and the whole left margin; the other path is the shortest path between  $q$  and the whole left margin.

Although this method is quite interesting in practice, it also has its weaknesses. A first concern is the transformation of the input image (or input ROI) to polar coordinates. The transformation itself requires the definition of a resolution for theta and the radius, which is not obvious. A high resolution will result in a bigger polar image and more computational effort; a low resolution may result in a coarse estimation of the path. Moreover, there is the danger that closed contours in polar coordinates become open when performing the inverse transformation for the original coordinates. In fact, when transforming back to original coordinates, the interpolation step will originate grey-level values (which result in a path fuzzily defined except in the case of the nearest-neighbour interpolation, which can produce disconnected paths if the theta resolution is not over-dimensioned for the maximum possible radius.

## 5.4 Closed Shortest Path Algorithm in the Original Coordinates

The computation of the closed shortest path in the original coordinates has some initial difficulties. A first task involves creating a DAG from the grid and an appropriate linearisation of the DAG. Secondly, since paths closer to the centre have less pixels, they will be naturally selected even if the cost of the weight of the edges is slightly higher than in paths further away from the centre. Thirdly, one must consider if the Euclidean distance between nodes (pixels) is appropriate or if another notion of distance is more effective to capture the distance in the context of closed paths enclosing a given node.

### 5.4.1 Linearization of the DAG

While working in polar coordinates, the linear ordering of the vertices is given by column number, since one wants to go from left to right. The causal neighbours of a given node are the neighbours with lower column number and edges are oriented from the causal neighbours to the given point.

Now, working in the original coordinates, one wants to go around a given seed point  $\mathcal{O}$ , from 0 to 360 degrees. It is therefore natural to order the nodes by the angle  $\theta$  of the node relative to  $\mathcal{O}$  – see Figure 5.7.

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 21 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 31 & 28 & 26 & 22 & 20 & 18 & 15 & 0 & 0 \\ 0 & 35 & 32 & 30 & 27 & 23 & 19 & 16 & 14 & 11 & 0 \\ 0 & 37 & 36 & 33 & 29 & 24 & 17 & 13 & 10 & 9 & 0 \\ 0 & 40 & 39 & 38 & 34 & 25 & 12 & 8 & 7 & 6 & 0 \\ 41 & 42 & 43 & 44 & 45 & 0 & 1 & 2 & 3 & 4 & 5 \\ 0 & 46 & 47 & 49 & 54 & 61 & 72 & 77 & 79 & 80 & 0 \\ 0 & 48 & 50 & 53 & 58 & 62 & 68 & 73 & 76 & 78 & 0 \\ 0 & 51 & 52 & 56 & 59 & 63 & 67 & 70 & 74 & 75 & 0 \\ 0 & 0 & 55 & 57 & 60 & 64 & 66 & 69 & 71 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 65 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Figure 5.7: For a ROI with a radius of 5, this image illustrates the order of the pixels in the linearization starting at  $\theta = 0$ . The zero entries are points that do not belong to the ROI.

The causal neighbours of a given node are the neighbours with lower  $\theta$  and edges are oriented from the causal neighbours to the seed point. Note that now the number of causal neighbours varies from node to node, depending on  $\theta$  and the position of the node in the ROI. Edges are oriented from the causal neighbours to the given point – see Figure 5.8.

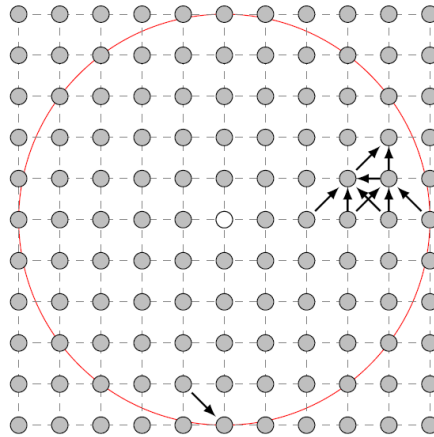


Figure 5.8: For a ROI with a radius of 5 (red) and an 8-neighbourhood, this figure illustrates the causal neighbours for a few nodes. The number of causal neighbours varies from 1 to 4.

It is important to stress that the linearization step does not depend on the content of the image and for a fixed size of the ROI, it can be pre-computed or computed only once or even hardcoded in the software (or hardware).

### 5.4.2 Edge Cost Modulation

A difficulty with searching for the closed shortest path enclosing a given point  $C$  (shortest in the sense of minimizing the cost of the path) is that small paths, collapsing in the point  $C$ , are naturally favoured. We overcome that difficulty by manipulating the cost of an edge appropriately. In Figure 5.9 we depict two closed paths enclosing a point  $C$ . The pixels assume the same value in both contours. If the cost of an edge is solely based on features extracted from the edge or a neighbourhood of the pixel, the cost of the edge takes the same value in both contours and the smallest contour is preferred since it has less edges and therefore smaller overall cost. Assuming that, in general, one does not want to give preference to any of them, costs should be adapted to correct the increase of the number of edges in a path with the distance. As such we weight the cost of an edge by  $1/r$ , where  $r$  is the distance of the head node of the edge. Roughly speaking, the perimeter (and therefore, the number of edges in the contour) grows proportionally to  $r$ ; the normalization makes the overall cost approximately independent of  $r$ .

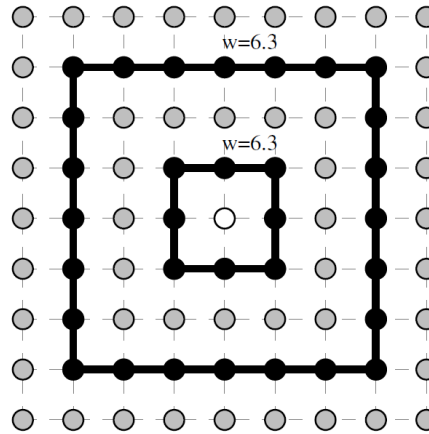


Figure 5.9: Two closed paths enclosing the centre of the ROI. Without a proper modulation, the inner path presents a smaller overall cost.

Note that this same behaviour is true with the conventional polar coordinates transformation: after transformation, all circular paths centred in the seed point have the same cost; in here this is achieved implicitly with resampling, transforming every circular path to a line with the same number of points.

### 5.4.3 Distance Between Nodes

As previously mentioned, the cost of an edge will include a factor related with the feature(s) computed at the head of the edge and a factor related with the distance between the head and the tail of the edge. Let us concentrate in the second factor.

The Euclidean distance between the two nodes seems at first a reasonable solution. Note however that in this case the (position of the) seed point does not impact the distance. Attend to Figure 5.10. In one case the contour is moving directly to the centre, while in the other case the

contour is almost orthogonal to the radius of the current node. Note that with the conventional polar coordinates transformation not only all circular paths centred in the centre of the ROI have the same cost but they are also the shortest paths when all the pixels have the same information. In fact, they are transformed into straight lines between opposite margins.

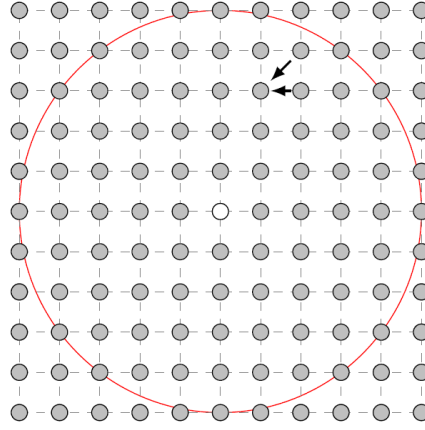


Figure 5.10: Two movements with different characteristics.

Trying to mimic this behaviour, we also experimented with two other measures. The first one is exactly what would be obtained in the polar domain, using a resolution of one degree per pixel and radius unit per pixel:

$$d_{polar} = \sqrt{(\Delta r)^2 + (\Delta \theta)^2}$$

The second measure is the Euclidean distance modulated by a function of  $\alpha$ , the angle between the direction orthogonal to the radius at the current node and the vector from the causal neighbour to the current point:

$$d_{cos} = \frac{d_{Euclidean}}{\cos(\alpha)}$$

Note that in this way we penalize steps that depart from the expected circular path.

## 5.5 Experimental Results

In this Section we evaluate the behaviour of the proposed algorithm in multiple scenarios. We start by illustrating the behaviour in some toy examples. Next, we apply the method to real images; finally, we evaluate the performance in two computer vision challenges: mass segmentation in mammograms and nipple detection for aesthetic evaluation of breast cancer treatments.

### 5.5.1 Illustrative Examples

We start by explaining the behaviour of the method in synthetic images. The results in Figure 5.11 were obtained using the 8-neighbourhood and the polar distance. In Figure 5.11(a) -5.11(c) the weight of the edge is directly the magnitude of the derivative in the radial direction (centred in

the seed point). In Figure 5.11(d) the weight was set as a nonlinear function of the derivative (motivated by previous work on similar settings [177]):

$$w = f_l + (f_h - f_l) \frac{\exp(\beta(255 - d)) - 1}{\exp(\beta 255) - 1} \quad (5.3)$$

with  $f_l = 2$ ,  $f_h = 128$ ,  $\beta = 0.0208$ , and  $d$  equal to the magnitude of the derivative.

Note that in this case the detected contour follows the expected path, while previously it had a tendency to smooth the contour. Nonetheless, in general, the method is quite robust to different cost functions.

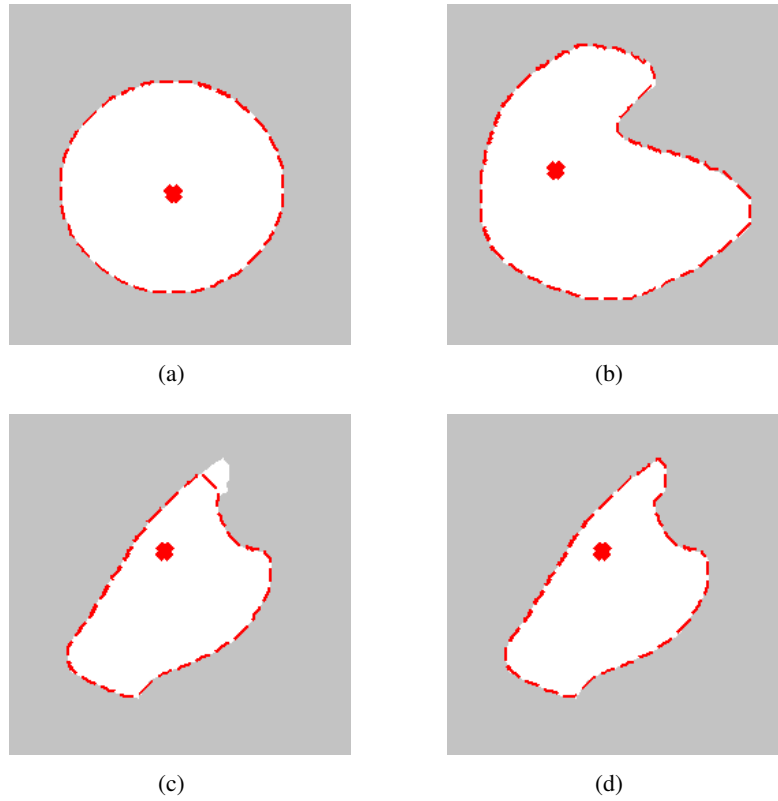


Figure 5.11: Illustration of some toy examples. In a), b) and c) the weight of the edge is directly the magnitude of the derivative in the radial direction. In d) the weight was set as a nonlinear function of the derivative.

Next, we illustrate the influence of the seed position in the final result. In Figure 5.12 we can visualise that impact in one of the previous toy examples. It is important to highlight that for any angle centred in the seed, the output contour will intersect the radial direction only once. Therefore, for the seed position in Figure 5.12(b), it will be impossible to perfectly delineate the true boundary.

In Figure 5.13 we show the application of the proposed algorithm in some images randomly selected from The Berkeley Segmentation Dataset [165]. The seed was manually placed; the polar

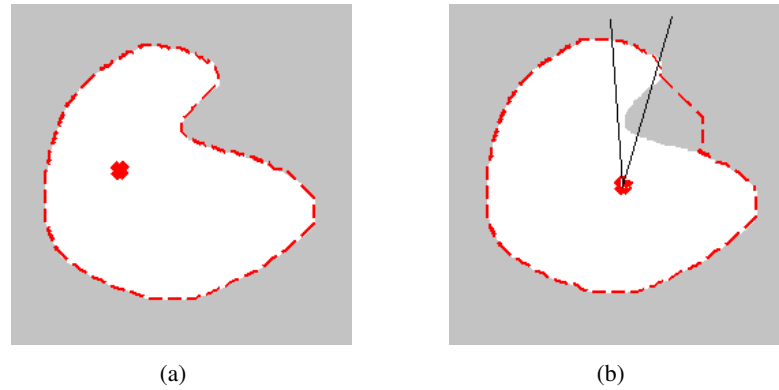


Figure 5.12: Impact of the seed position in final result. In the highlighted circle sector, for the selected seed, it is impossible to perfectly follow the boundary.

distance and the nonlinear weighting were again selected. The results are quite intuitive and within the expected behaviour.

Finally, in Figure 5.14, we show the results obtained with the conventional transform to polar coordinates on the same subset of images from the Berkeley Segmentation Dataset. As expected, for some images the differences are indiscernible. When differences exist, they are small, and without a clear advantage of one method over the other.

In the next subsection we depart from the subjective analysis in a small number of images to an objective evaluation in mass segmentation approach on mammography images <sup>1</sup>.

### 5.5.2 Mass Segmentation

The most common breast cancer screening test is mammography. When radiologists examine mammograms, they look for specific abnormalities. A breast can have: only mass(es), only calcifications or both. The first stage in computer-aided detection (CAD) in screening mammography usually involves candidate generation which identifies suspicious unhealthy candidate regions of interest (candidate ROIs, or simply candidates) from a medical image. These ROIs are then analysed to differentiate malignant cancers from other candidates. In here, we focus on finding the mass inside the ROI.

Ninety nine images from the INBreast database [172] were selected. The selection criterion was that the image contained one and only one mass. The ROI was generated from the ground truth contour of each mass, centred in the centre of the mass and enclosing all the mass. The costs were based on the nonlinear function of the derivative introduced in 5.3.

Evaluation was made both with Region and Contour-based metrics. The firsts include the area overlap measure (AOM), and a combined measure (CM) of under-segmentation, over-segmentation

<sup>1</sup>The validation of the algorithm with the data from this thesis will be presented in Chapter 9

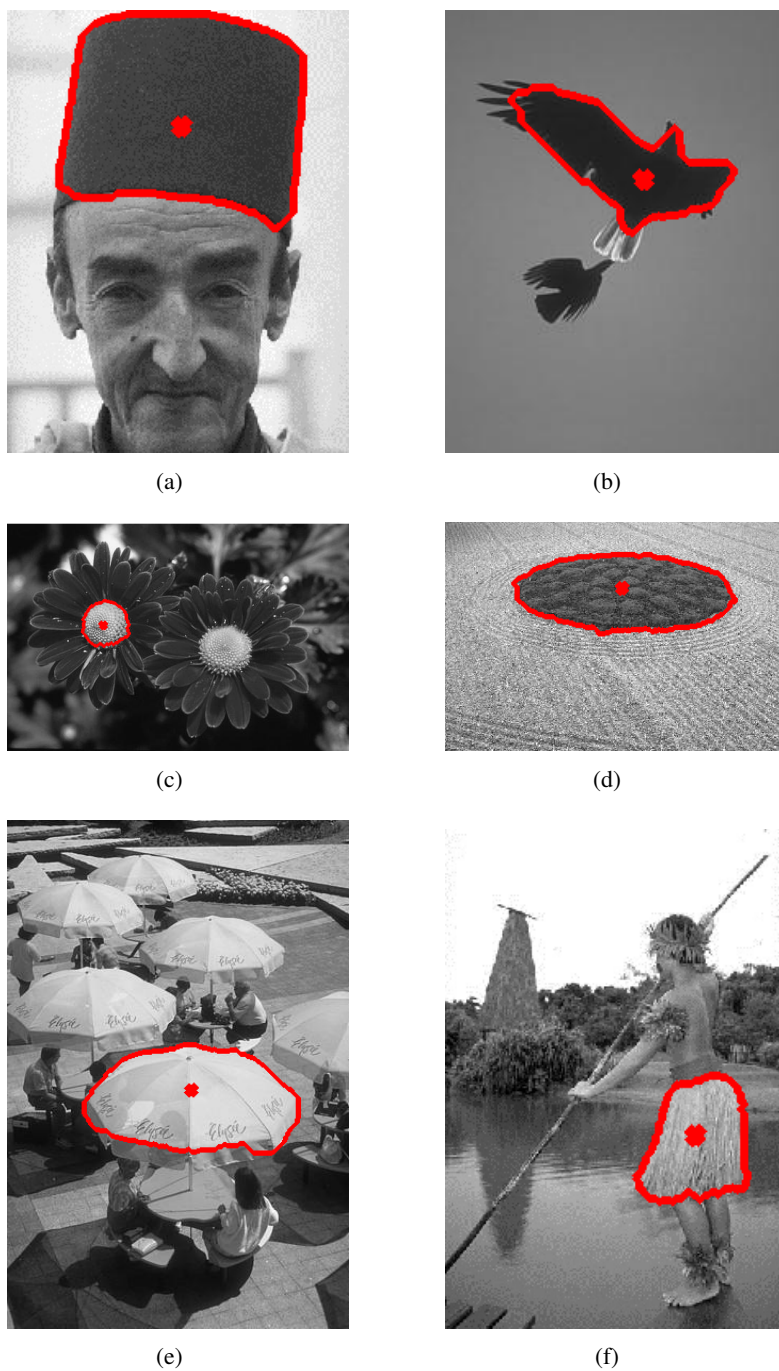


Figure 5.13: Illustration of the application of the proposed algorithm in randomly selected real images.

and AOM [71]. Contour-based metrics include average distance (AD), average minimum Euclidean distance (AMED) and Hausdorff distance (HD) [220]. These last measures were normalized by the ROI diagonal. Note that the selected region based measures are measures of accuracy (the higher the better), while contour based measures are measures of error (the lower the better).



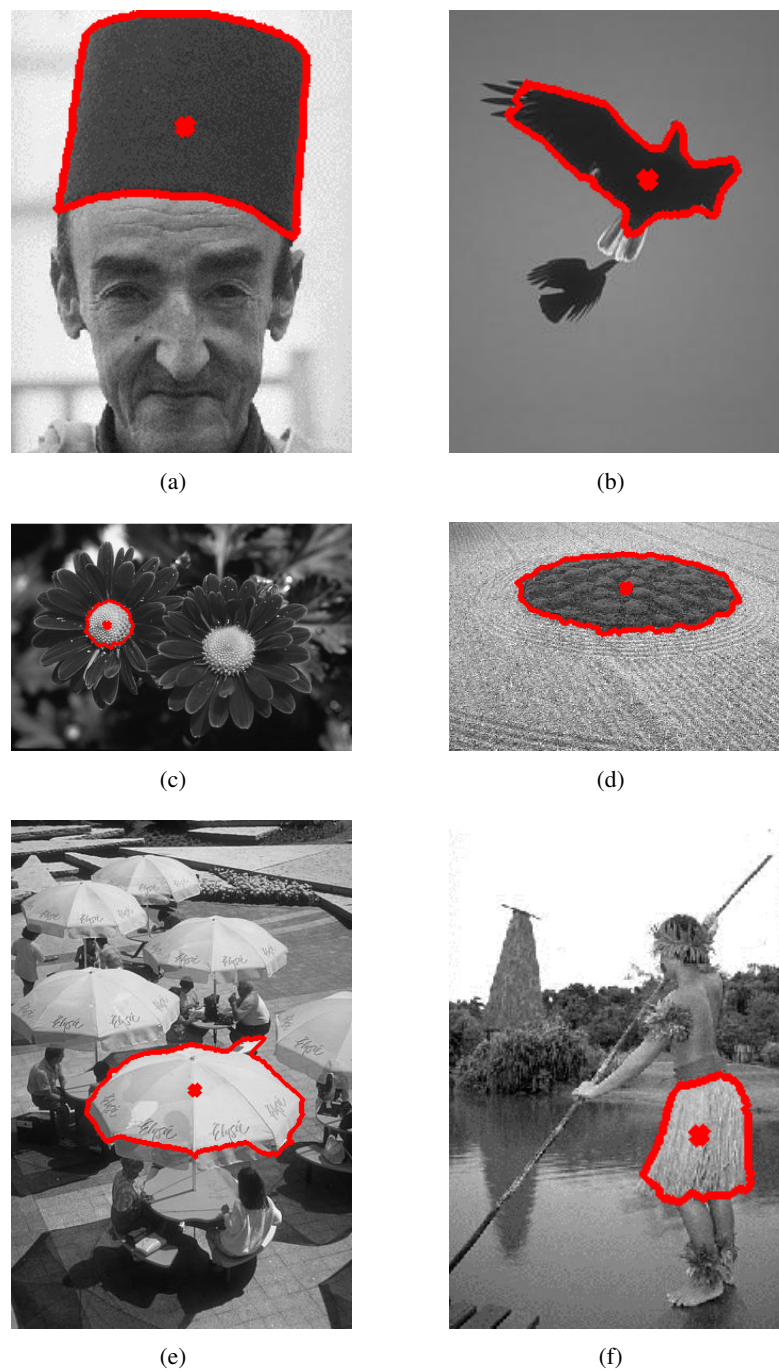


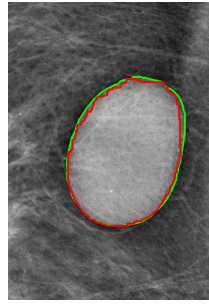
Figure 5.14: Illustration of the conventional transform to polar coordinates in randomly selected real images.

Results for different values of the neighbourhood and distance can be seen in Table 5.1.

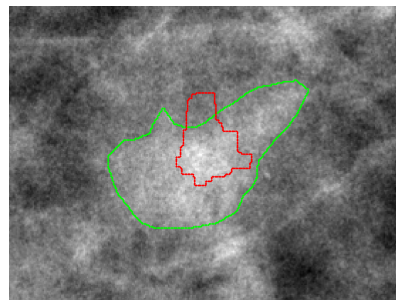
The polar distance consistently outperformed the other alternatives, for all neighbourhoods and according to all metrics. The best size for the neighbourhood is 4 for region based metrics and 8 for most of the contour equivalents. Some selected results are shown in Figure 5.15.

neighbourhood	distance	AOM	CM	AD	AMED	HD
4	Euclidean	0.702 (0.159)	0.794 (0.108)	0.024 (0.017)	0.027 (0.019)	0.088 (0.036)
	Polar	0.707 (0.160)	0.797 (0.111)	0.024 (0.017)	0.027 (0.019)	0.087 (0.037)
	Cos	0.702 (0.159)	0.794 (0.108)	0.024 (0.017)	0.027 (0.019)	0.088 (0.036)
8	Euclidean	0.700 (0.161)	0.791 (0.113)	0.024 (0.018)	0.027 (0.020)	0.084 (0.034)
	Polar	0.707 (0.150)	0.795 (0.107)	0.024 (0.016)	0.026 (0.017)	0.085 (0.034)
	Cos	0.700 (0.161)	0.791 (0.113)	0.024 (0.018)	0.027 (0.020)	0.084 (0.034)
12	Euclidean	0.688 (0.165)	0.784 (0.114)	0.025 (0.019)	0.028 (0.020)	0.087 (0.035)
	Polar	0.692 (0.166)	0.786 (0.116)	0.025 (0.019)	0.028 (0.020)	0.086 (0.036)
	Cos	0.688 (0.164)	0.784 (0.114)	0.025 (0.019)	0.028 (0.020)	0.087 (0.035)
20	Euclidean	0.663 (0.188)	0.766 (0.128)	0.029 (0.025)	0.032 (0.026)	0.092 (0.039)
	Polar	0.668 (0.180)	0.769 (0.124)	0.028 (0.024)	0.030 (0.025)	0.090 (0.036)
	Cos	0.663 (0.188)	0.766 (0.128)	0.029 (0.025)	0.032 (0.026)	0.092 (0.039)

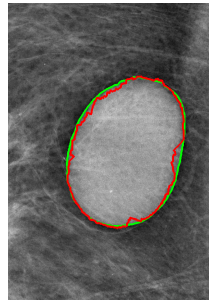
Table 5.1: Mass segmentation on Mammograms: results for different neighbourhoods and distances. Results are presented in the format “average (standard deviation)”.



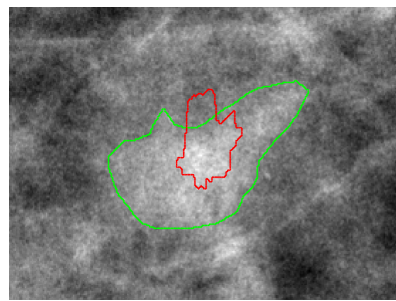
(a) Good result: neighbourhood 4, polar distance.



(b) Bad result: neighbourhood 4, polar distance.



(c) Good result: neighbourhood 8, polar distance.



(d) Bad result: neighbourhood 8, polar distance.

Figure 5.15: Illustration of some good and bad mass segmentation on Mammograms.

Some other works have computed mass contours based on the shortest path algorithm in polar coordinates.

Timp and Karssemeijer [233] had an AOM of 0.69, while Dominguez and Nandi [62] reached an AOM of 0.72. More recently, Song *et al.* [220] obtained an AOM of 0.73.

In [71], the CM metric was used, having the dynamic programming method reached an CM of less than 0.70.

Since the used databases are not the same, a fair comparison cannot be made. Moreover, all the papers found use region-based evaluation metrics and no contour-based metrics are available. However, our segmentation method seems to have a competitive accuracy when comparing to previous implementations.

The next experiment was to study the degradation of the contour quality in relation to the seed position. The seed was randomly selected inside a window centred in the ground truth seed and with size  $\alpha$  times the size of the ROI. The only restriction was that the seed was kept inside the mass.  $\alpha$  was varied from 0.02 to 0.14. For each  $\alpha$  we repeated the experiment 10 times to average results. Results using the 4-neighbourhood and the polar distance are presented in Table 5.2. As expected, the performance decreases monotonously; nevertheless the decrease is smooth without abrupt changes.

noise percentage	AOM	CM	AD	AMED	HD
2%	0.657 (0.188)	0.762 (0.128)	0.030 (0.027)	0.032 (0.027)	0.093 (0.040)
4%	0.656 (0.186)	0.761 (0.127)	0.029 (0.025)	0.032 (0.026)	0.093 (0.040)
6%	0.648 (0.194)	0.756 (0.131)	0.030 (0.027)	0.033 (0.028)	0.095 (0.042)
8%	0.648 (0.194)	0.756 (0.132)	0.031 (0.027)	0.033 (0.028)	0.095 (0.042)
10%	0.639 (0.196)	0.749 (0.134)	0.032 (0.028)	0.035 (0.029)	0.098 (0.044)
12%	0.636 (0.195)	0.747 (0.133)	0.032 (0.027)	0.035 (0.029)	0.099 (0.045)
14%	0.631 (0.197)	0.743 (0.135)	0.032 (0.028)	0.036 (0.029)	0.100 (0.045)

Table 5.2: Mass segmentation on Mammograms: results for different perturbations on the seed localization. Results are presented in the format “mean (standard deviation)”.

## 5.6 Conclusions

We have proposed a novel method to find a closed shortest path in an image. The closed shortest path obtained in the image ensures that the starting and ending positions are connected. This algorithm has applications in object boundary detection, borehole image analysis, fiducial point detection, among others.

Contrarily to conventional methods, the proposed method works on the original space, instead of requiring a mapping into polar transformation. Besides efficiency advantages (avoiding unnecessary steps), there is no loss in resolution and the path is always connected. The algorithm was shown to be fast and reliable in tests on several different types of real images.

We have explored the application of this method in the description of a region enclosing a given point. Instead of just capturing local information in the seed point, for instance, in the nipple detection problem, we describe the region around a candidate since it should correspond to the areola complex (see Chapter 9). The overall quality of the candidate integrates information both from the point and the region around. We believe that this reasoning will find applications in several other domains.



## Chapter 6

# Interpretability Models and Lateral-views for the Aesthetic Result Assessment\*

As stated in Chapter 3, the development of the BCCT.core involved the semi-automatic extraction of several features from frontal patient' photographs. Some of the factors that are considered to have greater impact on the overall cosmetic result are captured: breast asymmetry, skin colour changes due to radiotherapy treatment and surgical scar appearance. After that, an SVM classifier was trained in order to predict the overall cosmetic result from the recorded features. However, improvements on the interpretability of the model used to assess the overall aesthetic result and the inclusion of different views from the patient (side, oblique) were often suggested.

In this chapter we investigate the improvement of BCCT.core, by comparing the performance of different interpretable methods with the currently deployed in BCCT.core, without sacrificing the estimated performance and the introduction of knowledge extracted from patients side-views. This information will be used as a complement for the currently used frontal data trying to improve the accuracy of the software.

### 6.1 Interpretable Models

Although SVMs have proven to be very useful in machine learning, there is a significant drop of understandability of the learned hypothesis, especially when using nonlinear kernels, as is the case of the RBF kernel in BCCT.core. The RBF kernel is a non-linear kernel, with an implicit mapping to a higher dimensional feature space. The relationships uncovered by the SVM in this implicit feature space are not easily portrayed in the initial space. This nonlinear model is often used as a black box, to which data are presented and the predicted class is outputted. This represents a strong limitation for the caregivers, aiming to understand how the overall classification is being attained.

---

\*Some portions of this Chapter appeared in [180, 181]

In this study we consider decision trees and a linear SVM to replace the current RBF SVM. Tree-based models are simple, but widely used, models that work by partitioning the input space into cuboid regions, whose edges are aligned with the axes, and then assigning a simple model (for example, a constant) to each region [22]. A key property of tree based models, which makes them popular in fields such as medical diagnosis, is that they are readily interpretable by humans because they correspond to a sequence of binary decisions applied to the individual input variables. For instance, to predict a patient's disease, we might first ask "is their temperature greater than some threshold?". If the answer is yes, then we might ask next "is their blood pressure lower than some threshold?". Each leaf of the tree is then associated with a specific diagnosis [22].

Linear models for classification create decision surfaces that are linear functions of the input feature vector and hence are defined by hyperplanes within the input space. The decision surface is therefore a simple weighted sum of the input features, with higher weights being assigned to features that are more important for the decision process.

### 6.1.1 Study Population and a Gold Standard

We use a set of 143 frontal photographs<sup>1</sup> taken at different breast centers. To train a classifier to perform an automated analysis, a gold standard or ground truth is needed. We use the gold standard taken from the evaluation of patients by an international panel of experts, following a Delphi methodology [99, 119]. A first set of 113 photographs was already available from the initial study [28] (see Section 2.9); an additional set of 30 photographs was taken since then. The first set of cases was used on the training phase, while the second was used on the testing phase. The distribution of the patients over the four different classes is summarised in Table 6.1.

Class	Excellent	Good	Fair	Poor	Total
# cases	20	74	34	15	143

Table 6.1: Distribution of the 143 patients over the four classes.

Since the extraction of the features from the image is still not completely automated, we asked 8 users to manually place the necessary fiducial points; all  $143 \times 8$  cases were later used to develop the classification models.

### 6.1.2 Feature Selection

In supervised learning, variable selection is used to find a subset of the available inputs that accurately predict the output. The objectives of variable selection are to improve prediction performance of the model, provide faster and more cost-effective predictors, and provide a better understanding of the underlying process that generated the data [93].

<sup>1</sup>In this study, we used the Breast Research Group Database (see Section 2.9), instead of the database created during this thesis.

Feature selection algorithms can be divided into 'filter' and 'wrapper' [222], according to the nature of the methods used to evaluate features. Wrapper algorithms evaluate features with the classification accuracy provided by a target classification algorithm. A target classifier is included in the feature selection process of the wrapper approach. This leads to the improvement of classification accuracy of the wrapper classifier, but it also contributes to the increase in computation times. Furthermore, because the derived feature subset is biased to the wrapped classification algorithm, good performance may not occur when the feature subset is used to build models with other classification algorithms.

Filter algorithms are independent of any learning algorithms; feature evaluation is carried out according to a measurement that evaluates the goodness of individual features, such as correlation. Features are ranked according to their values on the selected measure. The first ranked features are chosen from the entire set of available characteristics, according to prior domain knowledge or a user-specified threshold value.

Evaluating all possible subsets is unfeasible; remember that we recorded 30 features, leading to  $2^{30}$  different subsets of features (see Chapter 3). Therefore, we performed a simplified selection, also benefiting from our previous work [28]. We had already concluded that the 7 dimensionless asymmetry features perform as good as or even better than the dimensional asymmetry features, with the additional advantage of not requiring the scale mark in the patient photograph [28]. Therefore, 7 asymmetry features were immediately discarded.

For the remaining 23 ( $7 + 8 + 8$ ) features, the Squared Correlation Coefficient ( $R^2$ ) was computed. The  $R^2$ , also referred to as coefficient of determination, is the proportion of target variation explained by a single input variable. The coefficient ranges from 0, when there is no linear relationship between an input and the target, to 1, for an input that explains all of the target variability. Simultaneously, we also analysed the behaviour of each feature when going from the excellent to the poor class. It is expected that the average value of any asymmetry, colour or scar visibility measurements increase monotonically from the excellent to the poor class (see Table 6.2).

From the features that exhibited a monotonous behavior, we kept the 5 dimensionless asymmetry features, the 2 colour change features and the 2 scar visibility features with highest  $R^2$  coefficient:  $pLBC$ ,  $pUNR$ ,  $pBCD$ ,  $pBAD$ ,  $pBOD$ ,  $c\chi_a^2$ ,  $cEMD_a$ ,  $s\chi_a^2$ ,  $sEMD_a$ .

Finally, with the 7 remaining features, we considered all possible subsets. The performance of each subset of features was estimated using a four-fold cross-validation scheme.

Other feature selection approaches were also tried [222]. In this work is presents an adaptation of a quadratic programming feature selection technique that identifies in one-fold the redundancy and relevance on data. This approach introduces a non-probabilistic measure to capture the relevance based on Minimum Spanning Trees. However, this approach obtained worse results, and for that reason were not included in this Chapter.

### 6.1.3 Dealing with Decision Costs

In this application, the default assumption of equal misclassification costs underlying machine learning techniques is violated. Caregivers consider an error in a true excellent or true poor patient



Features	$R^2$	class avg (std. dev.)			
		1	2	3	4
$pBRA$	0.19	0.09 (0.04)	0.11 (0.06)	0.17 (0.07)	0.17 (0.05)
$pLBC$	0.52	0.04 (0.02)	0.05 (0.03)	0.11 (0.03)	0.15(0.05)
$pUNR$	0.23	0.06 (0.05)	0.09 (0.06)	0.17 (0.09)	0.17 (0.07)
$pBCE$	0.03	0.11 (0.08)	0.25 (0.25)	0.34 (0.22)	0.25 (0.25)
$pBCD$	0.40	0.04 (0.03)	0.06 (0.04)	0.11 (0.06)	0.19 (0.07)
$pBAD$	0.23	0.10 (0.06)	0.12 (0.10)	0.18 (0.12)	0.32 (0.013)
$pBOD$	0.24	0.18 (0.05)	0.21 (0.09)	0.28 (0.09)	0.37 (0.14)
$c\chi_L^2$	0.04	0.05 (0.03)	0.09 (0.06)	0.11 (0.06)	0.10 (0.04)
$c\chi_a^2$	0.07	0.03 (0.03)	0.05 (0.06)	0.08 (0.08)	0.10 (0.12)
$c\chi_b^2$	0.01	0.07 (0.06)	0.14 (0.14)	0.16 (0.16)	0.13 (0.11)
$c\chi_{Lab3D}^2$	0.01	0.04 (0.03)	0.07 (0.07)	0.09 (0.07)	0.06 (0.03)
$cEMD_L$	0.03	0.65 (0.34)	0.95 (0.53)	1.07 (0.57)	1.03 (0.45)
$cEMD_a$	0.06	0.09 (0.07)	0.14 (0.10)	0.17 (0.11)	0.19 (0.13)
$cEMD_b$	0.03	0.19 (0.11)	0.30 (0.21)	0.34 (0.23)	0.31 (0.19)
$cEMD_{Lab}$	0.02	0.10 (0.06)	0.14 (0.09)	0.17 (0.09)	0.15 (0.06)
$s\chi_L^2$	0.00	0.47 (0.32)	0.70 (0.44)	0.63 (0.40)	0.57 (0.34)
$s\chi_a^2$	0.12	0.10 (0.10)	0.18 (0.15)	0.23 (0.15)	0.35 (0.25)
$s\chi_b^2$	0.01	0.25 (0.17)	0.47 (0.33)	0.47 (0.29)	0.43 (0.29)
$s\chi_{Lab3D}^2$	0.01	0.27 (0.20)	0.43 (0.36)	0.50 (0.38)	0.34 (0.25)
$sEMD_L$	0.01	1.84 (0.89)	2.53 (1.37)	2.42 (1.17)	2.65 (1.49)
$sEMD_a$	0.11	0.18 (0.12)	0.28 (0.16)	0.32 (0.12)	0.42 (0.22)
$sEMD_b$	0.03	0.39 (0.18)	0.55 (0.27)	0.60 (0.30)	0.56 (0.26)
$sEMD_{Lab}$	0.01	0.33 (0.15)	0.42 (0.22)	0.47 (0.22)	0.41 (0.22)

Table 6.2: Correlation coefficient and average value of each class on feature selection procedure.

to be more penalising than an error in the middle classes (fair or good). Moreover, failure to a contiguous class is not as serious as failure to a non-contiguous class. From these considerations, we defined a cost matrix ( $\mathbf{C}$ ) reflecting the penalty of classifying samples from one class as another:

$$\mathbf{C} = \begin{bmatrix} 0 & 2 & 4 & 6 \\ 1 & 0 & 1 & 2 \\ 2 & 1 & 0 & 1 \\ 6 & 4 & 2 & 0 \end{bmatrix}$$

where  $C(i, j)$  is the cost of classifying a point into class  $j$  if its true class is  $i$ . The cost matrix was taken into consideration during the model building process.

#### 6.1.4 Results

The problem addressed here involves classifying examples into classes which have a natural ordering. Therefore we adopted classification methods specific for this kind of data. The SVM classifier was based on the data replication method for ordinal data [29], both with linear and RBF kernels.



We performed a "grid-search" on the parameters of the models using cross-validation in the training set (h and s parameters were left constant at 1 and 2, respectively). The test results concerning the misclassification error are summarised in Table 6.3, for the first ranked feature subsets:

**Feature Set 1:** {pLBC, pBAD,  $c\chi^2a$ ,  $s\chi^2a$ }

**Feature Set 2:** {pLBC, pBCD,  $c\chi^2a$ ,  $s\chi^2a$ }

**Feature Set 3:** {pLBC, pBCD, pBAD,  $c\chi^2a$ ,  $s\chi^2a$ }

Feature Set	kernel	C	gamma	Error	Weighted Error
1	RBF	16	0.5	0.42	0.63
2		1	0.5	0.30	0.52
3		1	0.75	0.33	0.56
1	Linear	16		0.37	0.52
2		64	-	0.37	0.58
3		4		0.38	0.60

Table 6.3: Misclassification error using the SVM classifier.

For the decision tree algorithm, the cross-validation scheme was carried out over the pruning level of the tree. The misclassification error for the test phase is summarised in Table 6.4.

feature set	# levels	Error	Weighted Error
1	7	0.47	0.63
2	7	0.43	0.61
3	7	0.46	0.64

Table 6.4: Misclassification error using the Decision Tree classifier.

In both tables, the weighted error was obtained by weighting each prediction error result with the cost of the misclassification.

A first observation is that the current RBF SVM model performs better than both the linear model and the decision tree. Nevertheless the LINEAR SVM follows closely the accuracy of the nonlinear model, particularly if we focus on the weighted error. The decision trees did not achieve the same level of performance. This is likely related with the small size of the available dataset, insufficient for reliable learning with decision trees. It is also instructive to analyse the confusion matrices of the models in Table 6.5 (note that we use data taken from 8 different users).

Table 6.6 presents confusion matrices using the equal misclassification costs for the classifier. Comparing these results with the ones presented on Table 6.5, proves that the cost matrix (see Section 6.1.3), defined taken into account caregivers consideration, helps on the correct classification and distribution of the cases.

(a) RBF SVM					
Predict \ True	Excellent	Good	Fair	Poor	
	Excellent	Good	Fair	Poor	
Excellent	30	18	0	0	
Good	10	68	2	0	
Fair	6	33	39	2	
Poor	0	0	0	32	

(b) LINEAR SVM					
Predict \ True	Excellent	Good	Fair	Poor	
	Excellent	Good	Fair	Poor	
Excellent	36	12	0	0	
Good	15	59	6	0	
Fair	8	48	24	0	
Poor	0	0	0	32	

(c) DECISION TREE					
Predict \ True	Excellent	Good	Fair	Poor	
	Excellent	Good	Fair	Poor	
Excellent	19	26	3	0	
Good	6	71	3	0	
Fair	1	56	23	0	
Poor	0	0	8	24	

Table 6.5: Confusion matrices for best classifiers, in the test set, using the cost matrix defined on Section 6.1.3.

As observed, there is no confusion between the end classes. Moreover, most of the errors are to a contiguous class and concentrated in the 'good' class, with a true 'good' being predicted either as 'excellent' or 'fair'. The elimination of these errors will likely require the integration of new features, capturing more information to distinguish these cases. We should also not forget the possible existence of inconsistencies in the reference information, due to the subjectivity inherent to the process of obtaining them.

The SVM classifier used has been generalized for multiclass problems [7, 29], in this work, and because the categories are ordered, was performed three binary classifications, (Excellent/Good, Good/Fair and Fair/Poor). This can be interpreted as a simplification of the approach suggested by Hastie and Tibshirani [100]. To design a linear discriminator between two classes we must choose a weighted sum of the individual features and the bias to differentiate classes:

$$y(x) = x^T w + b \quad (6.1)$$

After training and testing phase *Feature Set 1* presents better results with LINEAR kernel parameters presented on Table 6.3. The decision rule to classify unseen patients, using the boundary planes outputted by the ordinal classifier, can be cleanly stated as:

(a) RBF SVM					
Predict \ True	Excellent	Good	Fair	Poor	
	Excellent	Good	Fair	Poor	
Excellent	0	48	0	0	
Good	0	80	0	0	
Fair	0	54	26	0	
Poor	0	1	4	27	

(b) LINEAR SVM					
Predict \ True	Excellent	Good	Fair	Poor	
	Excellent	Good	Fair	Poor	
Excellent	0	48	0	0	
Good	0	80	0	0	
Fair	0	64	16	0	
Poor	0	0	8	24	

(c) DECISION TREE					
Predict \ True	Excellent	Good	Fair	Poor	
	Excellent	Good	Fair	Poor	
Excellent	14	34	0	0	
Good	1	62	11	6	
Fair	0	57	12	11	
Poor	0	0	8	24	

Table 6.6: Confusion matrices for best classifiers, in the test set, using the equal misclassification costs.

$$\text{if } (12.66 \cdot pLBC + 1.82 \cdot pBAD - 0.93 \cdot c\chi^2 a + 0.77 \cdot s\chi^2 a) \begin{cases} < 0.72 & \Rightarrow \text{Excellent} \\ \geq 0.72 \wedge < 1.63 & \Rightarrow \text{Good} \\ \geq 1.63 \wedge < 2.12 & \Rightarrow \text{Fair} \\ \geq 2.12 & \Rightarrow \text{Poor} \end{cases}$$

The replacement of the current model in BCCT.core by the LINEAR SVM model will increase the interpretability and acceptance of BCCT.core, enabling caregivers to validate their experimental results, and improving trust on this kind of software for aesthetic evaluation of BCCT.

## 6.2 Improving the Model with Lateral Information

Almost all methodologies that use single photographs, namely the one used on BCCT.core, only use patient frontal images. The exclusive use of face-views may discard some important information, such as volume perception and inexistence of lateral measurements. This can be enhanced with the introduction of lateral information extracted from patient side-views photographs, providing the model with more information on the breast shape. This approach was never addressed on the aesthetic outcome of BCCT and only a small number of researchers used lateral features in

their work, such as Kim *et al.* [126], who compare objective measurements based on breast ptosis with ratings on a subjective scale made by experienced clinical observers. This can be related with the difficulty of extracting robust features from the side-views.

Although the current model (see Section 6.1) presents satisfactory results, it is possible that with frontal views only we do not have sufficient information on the breast aesthetic, such as, breast volume perception. By using lateral photographs (see Figure 6.1), we increase the quantity of the data related with the breast, improving its quality. This information will be used as a complement for the currently used frontal data trying to improve the accuracy of the software.

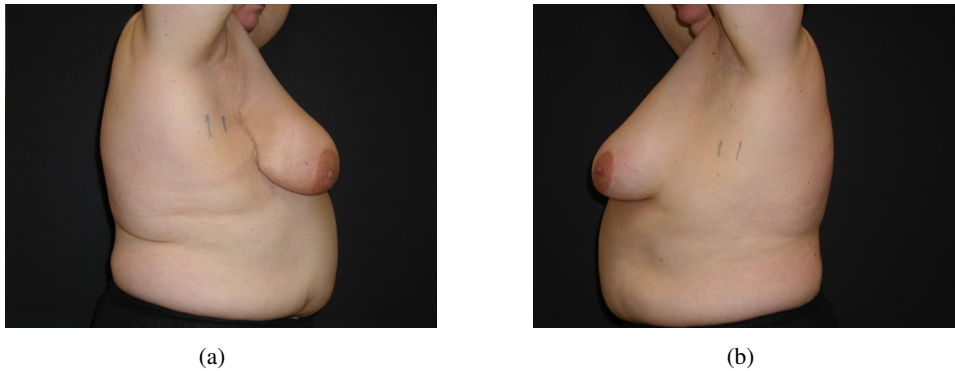


Figure 6.1: Typical lateral photographs.

Lateral features were defined from the comparison between the treated and the non-treated breast. This represents an additional problem, because we have different images for each breast, as opposed to the frontal analysis. In order to overcome this problem, we needed to find a normalisation factor between the images, making all measurements in the same scale.

### 6.2.1 Study Population and a Gold Standard

This study uses a portion of the dataset of 143 photographs with the ground truth used in Section 6.1. Only 63 photographs from the initial data were used due to the following reasons: definition of the normalization factor between images; fiducial points annotation; poor quality of some of the lateral photographs. Any little movement of the patient can lead to a not truly lateral posture, which may influence the measurements taken from the photographs. The distribution of the patients over the four different classes is summarised in Table 6.7.

Class	Excellent	Good	Fair	Poor	Total
# cases	10	35	11	7	63

Table 6.7: Distribution of the 63 patients over the four classes.

### 6.2.2 Feature Definition

Contrarily to frontal features, there are not many studies reporting the definition of features in lateral images. For that reason we had to introduce, for the first time, features that may be used to assess the aesthetic result of BCCT. As an exception, Kim *et al.* [126] refer the ptosis factor to evaluate the aesthetic results, that is, the extent to which the nipple is below the IMF, the crease beneath the breast. Unacceptable ptosis occurs when nipple and lower pole of the breast descend below the level of IMF, and is defined by four grades. A patient has no ptosis when the nipple and most of the breast gland is located above the level of submammary fold (Grade 0). In first degree or minor ptosis, the nipple is at the level of submammary fold and above the lower contour of the breast (Grade 1). Second degree or moderate ptosis is when the nipple lies below the fold but above the lower contour (Grade 2). Third degree or major ptosis is when the nipple lies at the lower breast contour and most of the breast is below the fold (Grade 3).

In this phase all fiducial points were manually marked directly on the photographs. In the future, this process could be made automatically. Figure 6.2 shows all fiducial points and measurements used for feature definition.

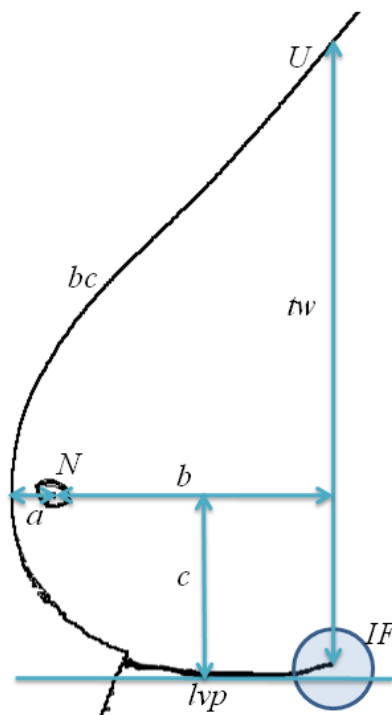


Figure 6.2: Illustration showing fiducial points and measurements.

The fiducial points marked on patient photographs were:

- **N:** Nipple;
- **IF:** Lateral terminus of IMF;

- **lvp**: Lowest visible point of the breast contour;
- **tw**: Thoracic wall;
- **U**: Projection of the IF point on the upper breast contour along the **tw**.

The IMF is a curvilinear structure that is generally hidden behind breast tissue since most women have some degree of ptosis. The Lateral Terminus of the Infra-Mammary Fold (**IF**) is the endpoint of the IMF, where it intersects the thoracic wall (**tw**). The lowest visible point (**lvp**) is located at the lower point of the breast. The centroid of nipple (**N**) is considered rather than the NAC since many women have an irregular shaped areola.

From the fiducial points, the following measurements were defined (see Figure 6.2):

- **Breast contour (bc)**: Breast exterior curve between **IF** and **U** point;
- **a**:  $\min(\text{dist}(N, bc))$  - Distance between **N** and **bc**;
- **b**:  $\min(\text{dist}(N, tw))$  - Distance between **N** and **tw**;
- **c**:  $\min(\text{dist}(N, lvp))$  - Nipple height (distance between **N** and **lvp**);
- **Breast size (bs)**:  $a + b$  - Sum of **a** and **b** distances;
- **Ptois (p)**:  $\frac{N-lvp}{IF-lvp}$  [126].

As mentioned before, a normalization factor (**nf**) is required in order to have measurements from each patient, from two photographs, on the same scale, and was manually defined on the patient's body in each picture. Features were assigned using previous measurements related with the asymmetry between treated and non-treated breast. The indices recorded to assess breast asymmetry were the following (note that all measurements are normalized by **nf**):

- **bs difference**

$$LBSD = \frac{|bs1-bs2|}{(bs1+bs2)/2}$$

- **Nipple distance to bc**

$$LNDbc = \frac{|a1-a2|}{(a1+a2)/2}$$

- **Nipple distance to tw**

$$LNDtw = \frac{|b1-b2|}{(b1+b2)/2}$$

- **Nipple distances evaluation**

$$LNDE = |a1/b1 - a2/b2|$$

- **Breast compliance evaluation**

$$LBCE = \frac{|c1-c2|}{(c1+c2)/2}$$

- **Breast ptosis difference**

$$LBPD = \frac{|p1-p2|}{(p1+p2)/2}$$

In order to minimize the influence of different scales between the images from the same patient, other features based on ratios of the aforementioned measurements were also introduced.

The ratio indices recorded to assess breast asymmetry were the following:

- **bs difference Ratio**

$$rLBSD = \min(BS1/BS2, BS2/BS1)$$

- **Nipple distance to bc Ratio**

$$rLNDbc = \min(a1/a2, a2/a1)$$

- **Nipple distance to tw Ratio**

$$rLNDtw = \min(b1/b2, b2/b1)$$

- **Nipple distances evaluation Ratio**

$$rLNDE = \min(a1 * b2 / b1 * a2, b1 * a2 / a1 * b2)$$

- **Breast compliance evaluation Ratio**

$$rLBCE = \min(c1/c2, c2/c1)$$

- **Breast ptosis difference Ratio**

$$rLBPD = \min(p1/p2, p2/p1)$$

### 6.2.3 Classifier

In this study, we applied a cascade generalization based classifier [87]. Cascade generalization uses a set of classifiers sequentially performing, at each step, an extension of the original data by the insertion of new attributes. In this particular application, the frontal features were initially applied to the model described in Section 6.1, resulting in a  $C_F$  classification. Afterwards, lateral features were introduced in the system, one by one, together with the result of the first model, resulting in a  $C_G$  classification. The scheme of the classifier is shown in Figure 6.3.

This new model is also a classifier based on machine learning techniques, namely SVMs. SVMs have proved to be being capable of representing complex classification or mapping functions. They discover the representations using powerful learning algorithms. We make use of the replication method as a tool to map this kind of data [29], because there is an inherent ordering between the classes. As already performed for the definition of the frontal model (see Section 6.1), we did not use the same misclassification costs for all classes (see Section 6.1.3).

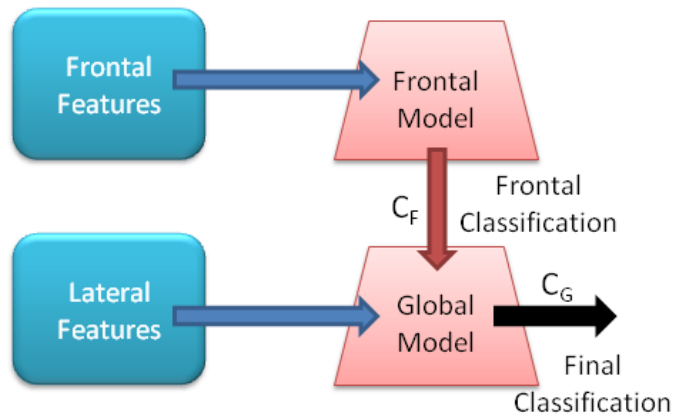


Figure 6.3: Classifier flowchart.

## 6.2.4 Results

We performed a "grid-search" optimizing each parameter using a cross-validation approach in the training set.

For all the 63 photographs the preliminary misclassification error was computed only with the frontal features using the model described in Section 6.1. The testing phase was made with the same data used in the training phase by means of a Leave One Out (LOO) scheme [65]. The test results concerning the misclassification error are summarised in Table 6.8. For the first ranked of each lateral feature:

Lateral Feature	kernel	C	gamma	Error	Weighted Error
LBSD	Linear	16	-	0.37	0.61
rLBSD	Linear	1	-	0.37	0.61
LNDbc	RBF	1	0.125	0.39	0.58
rLNDbc	RBF	1	0.125	0.39	0.60
LNDtw	Linear	64	-	0.43	0.65
rLNDtw	RBF	1	0.5	0.43	0.70
LNDE	Linear	4	-	0.39	0.65
rLNDE	Linear	1	-	0.38	0.61
LBCE	RBF	16	0.25	0.32	0.53
rLBCE	RBF	16	0.25	0.32	0.53
LBPD	Linear	64	-	0.38	0.59
rLBPD	Linear	1	-	0.40	0.60

Table 6.8: Misclassification error using the lateral features (the preliminary misclassification error for the frontal features was 0.35 and 0.57 for the weighted error).

A first observation from Table 6.8 is that with the integration of these new lateral features the classification error decreases. The accuracy of the model researched herein, compared with the model developed for the classic BCCT.core software, showed a better performance. It is also



instructive to analyse the confusion matrices of the models in Table 6.9 (note that we use data taken from 8 different users).

(a) Frontal model					
Predict \ True	Excellent	Good	Fair	Poor	
	Excellent	Good	Fair	Poor	
Excellent	62	18	0	0	
Good	74	183	17	6	
Fair	1	30	44	13	
Poor	0	6	13	37	

(b) Frontal plus lateral model					
Predict \ True	Excellent	Good	Fair	Poor	
	Excellent	Good	Fair	Poor	
Excellent	62	18	0	0	
Good	25	232	6	17	
Fair	1	30	11	46	
Poor	0	6	12	38	

Table 6.9: Confusion matrices for best classifiers, in the test set, for the different models. (frontal information only and frontal plus lateral information)

Obtained results might have been influenced by the choice of cases. Additional tests will be required to understand the relevance of the defined asymmetric lateral features on BCCT cosmetic result evaluation.

## 6.3 Conclusions

We have developed different accurate and interpretable models for the assessment of aesthetic result of BCCT. With the current nonlinear RBF SVM model deployed with BCCT.core as the foundation for our work, we compared the accuracy of a decision tree and a LINEAR SVM, which are models that facilitate the comprehension of factors with an impact on the decision. We have shown that the linear model achieves a performance very similar to the RBF SVM, with the obvious advantages of simplicity and interpretability.

After the model for the frontal-views is stabilised, we studied a new model for the assessment of the aesthetic outcomes of BCCT, incorporating lateral information. A better performance was obtained, although the improvement is not significant. The number of photographs used, 63, significantly lower than the initial dataset of 143, was due to quality and possible definition of fiducial points and features on the lateral photographs. Obtained results may have been influenced by the choice of cases. Additional tests will be needed to understand the relevance of the defined asymmetric lateral features on BCCT cosmetic results evaluation. To increase the robustness of the models presented here, we could definite new lateral features, namely, related with skin colour changes caused by radiotherapy treatment, and surgical scar appearance.

Applications using 3D imaging appear to be more efficient than methods that use single pictures, and can serve as a valuable tool for various types of breast surgery, including breast reconstruction with implant/tissue expanders, local flap reconstruction, free-flap reconstruction, breast augmentation, and breast reduction surgery. In each of these clinical scenarios, 3D analysis provides volumetric data that are of unique value for surgical planning and postoperative analysis.

## Chapter 7

# Uncalibrated-based 3D Reconstruction\*

It is recognisable that the BCCT.core can be improved by adding more measurements. Particularly, the aim is to add dimensionality to the measurements in order to reduce the limitation of only performing measurements based on what can be seen in a frontal photograph. In Chapter 6 the inclusion of side-view photographs was study; however, the improvements achieved were not meaningful, suggesting the application of 3D imaging. The ability to manipulate and measure over 3D readings from the breasts can improve the accuracy and objectivity of the tool.

3D capabilities are recognised as having high clinical potential. However, currently used techniques face two major problems: the high cost of equipment required and the need for specialised operators to work with them. Current techniques are based on especially designed cameras and hardware, mainly resorting to different lenses on the same camera or to laser scanners. Due to these special needs, 3D applications are considered costly and are not commonly implemented; thus, the benefit of 3D modelling is not availed. Therefore, it is necessary to search for new methodologies that should be affordable and easy to use. The practitioners themselves must be able to use the tool without any difficulties.

If this 3D modelling proves to be robust enough, the software can be further extended to surgical simulations. Those types of simulations are notably desired features that would provide the means for anticipating possible surgical options and outcomes, allowing a better education of patients towards a more informed choice and the improvement of surgical techniques and skills.

This chapter starts with a brief description of the basic math involved in recovering depth information from images. The reader is conveniently directed to appropriate sources for thorough information and formal proof, mostly based on [26, 228, 241]. Then, the application of the state-of-the-art algorithms is studied for feature detection, matching, rectification, stereo matching and reconstruction based on uncalibrated views from free-moving cameras when applied to female breasts. Because of the large amount of algorithms and applications available, it is necessary to select, *a priori*, the candidate applications that might prove to be successful in this particular problem. The final goal is to discover if the existing, most common and recent algorithms have

---

\*Some portions of this Chapter appeared in [182, 219]

some applicability to the type of scenes that the BCCT.core deals with, and to understand if 3D reconstruction can be performed directly using this method.

## 7.1 General Uncalibrated Reconstruction Workflow

The 3D reconstruction is one of the most important problems in computer vision. Many applications have been developed to be applied to different fields, such as: reconstructing objects from image or video sequences, image-based modelling from large photo collections, recognising 3D shapes and detecting 3D obstacles in mobile robotics. Currently there are three major approaches to multi-view stereo imaging: depth-map merging, volumetric-based, and feature point detection. The depth-map merging methods are based on the correspondences between an independent pair of depth-map images [131, 145]. In volumetric based approaches [247], the scene is firstly represented as a set of 3D voxels, which precede an energy minimisation approach to determine whether or not those voxels should be filled. Feature-point based algorithms start by extracting and matching feature points, before the surface reconstruction based on geometric, photometric or visualisation constraints [83, 144]. The approach presented here focuses on this last methodology. This kind of 3D reconstruction problem has been studied for several decades [94, 98, 117, 125, 141, 198, 202, 268]. There is a vast amount of research in the literature on this problem, with wide applications, ranging from world comprehension for robotics, aerial building reconstruction, to virtual tours for on-line tourism.

The basic workflow of an uncalibrated reconstruction from two or more photographs from a scene, using a feature-point based approach, can be described as the following sequence of steps [125]:

1. Searching for feature points;
2. Matching the highest number of those feature points in a robust way;
3. Using the matched features to find the fundamental matrix  $F$  with the lowest error possible (with the lowest inconsistency within the largest possible set of inliers);
4. Using that matrix and the inlier matches to compute rectification homographies, and then to rectify both views;
5. Performing the largest stereo matching possible, with the highest confidence available for each match;
6. If camera parameters were found somewhere in the fundamental matrix estimation or rectification algorithm, use them to return a quasi-Euclidean reconstruction; otherwise, return the projective reconstruction.

Failures in any of the previous steps (either due to lack of features, or to failure in matching enough features) will invalidate the remainder of the reconstruction. Therefore, each step is critical to the process. This workflow is depicted in Figure 7.1.

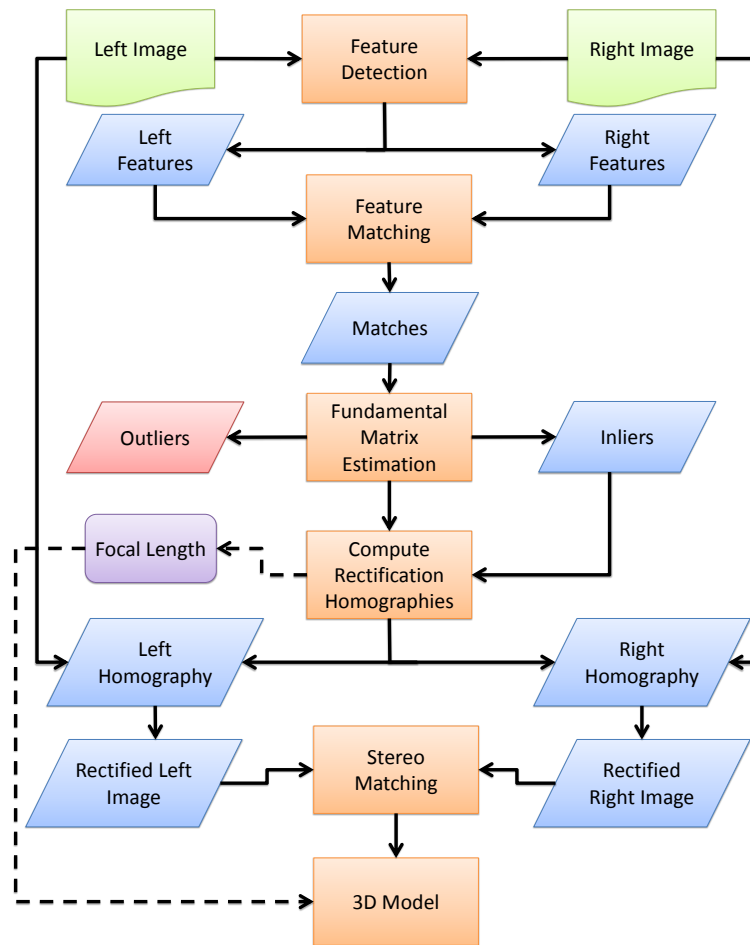


Figure 7.1: Stereo reconstruction generic workflow.

## 7.2 Camera Model

It is common in the literature to use the pinhole camera model, as it simplifies the math involved and provides an intuitive view of how the image plane is formed [26,228]. Imagine, for this model, that the projection is based on light being projected on a plane, entering from a small hole, so small only a ray can enter in each point of the scene (see Figure 7.2).

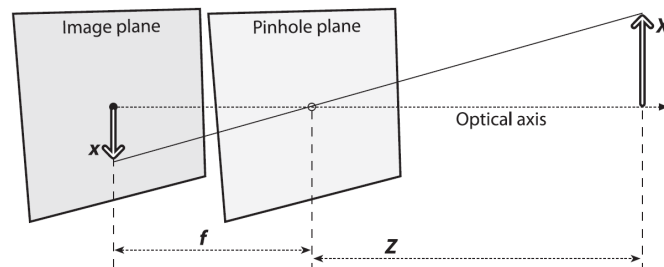


Figure 7.2: Pinhole camera model, the  $Z$  axis will be considered the optical axis,  $f$  stands for the focal length,  $X$  is a point in space and  $x$  is the projection of that point on the image plane. Please note the change in coordinates from  $X$  to  $x$  (From [26]).

This model can be rearranged so that the math comes out easily and so that the image plane is not inverted in relation to the scene (see Figure 7.3).

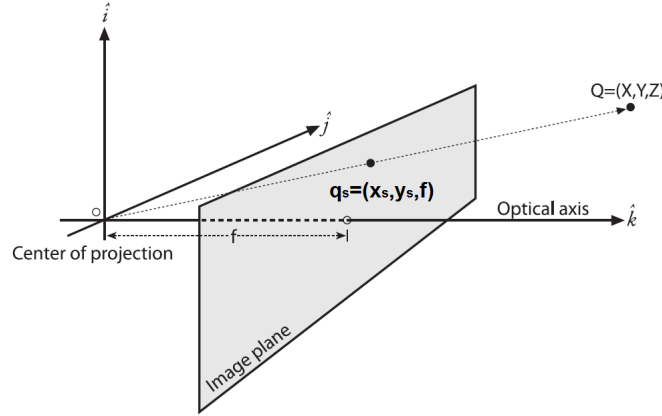


Figure 7.3: Simplified pinhole camera model (From [26]).

The projection that occurs on a camera when transforming 3D coordinates to 2D plane coordinates is a true 3D *perspective*, as confirmed by [228]. Therefore, any point  $Q$  will be transformed into an image point  $q$  (in inhomogeneous coordinates) as follows:

$$q = \mathcal{P}(Q) = \begin{bmatrix} X/Z \\ Y/Z \\ Z/Z \end{bmatrix} \quad (7.1)$$

However, this transformation is based on an ideal pinhole camera, which is extremely difficult, if not impossible, to create. Once the 3D points are transformed through the pinhole, it is still necessary to transform those coordinates into sensor space coordinates (camera coordinates). A camera can be simplified as a sensor and lens, where the lens gathers the maximum amount of light it can and redirects that light into the sensor. An example of that process can be found in Figure 7.4.

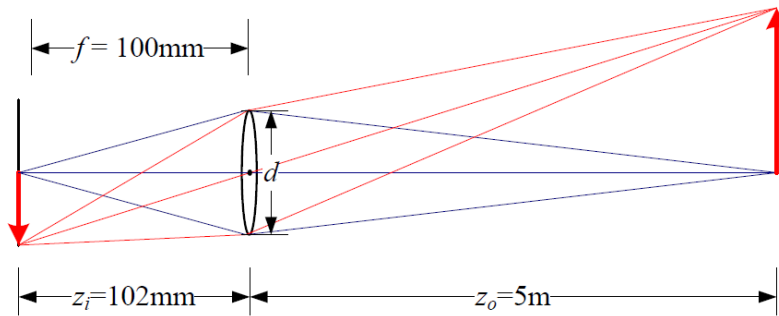


Figure 7.4: Simplified camera and lens example; note the possible similarities between this example and the pinhole model in Figure 7.2 (From [228]).

Therefore, a point  $q_s = [x_s, y_s]^T$  is formed from a scene point  $Q = [X, Y, Z]^T$  by some simple relations:

$$x_s = f \cdot \frac{X}{Z} + c_x \quad y_s = f \cdot \frac{Y}{Z} + c_y \quad (7.2)$$

Here,  $x_s$  and  $y_s$  are image coordinates,  $f$  is the focal length and  $c_x$  and  $c_y$  are the coordinates for the centre of the image plane. These two additional parameters result from the possible displacement that may exist between the sensor and the optical axis. Perfectly aligning the sensor and the optical centre of the lens would require a level of precision that is difficult to achieve, especially in low-end cameras. There is also the possibility that the sensor is not perpendicular enough to the optical axis to create a skew,  $s$ , between the sensor axis. Therefore, as for now, it is possible to build the camera calibration matrix as follows:

$$\mathbf{K} = \begin{bmatrix} f & s & c_x \\ 0 & af & c_y \\ 0 & 0 & 1 \end{bmatrix} \quad (7.3)$$

where  $a$  is the aspect ratio of the image.

In practice, for many applications, it can be assumed that  $a = 1$  and  $s = 0$ , resulting in the final calibration matrix:

$$\mathbf{K} = \begin{bmatrix} f & 0 & c_x \\ 0 & f & c_y \\ 0 & 0 & 1 \end{bmatrix} \quad (7.4)$$

This equation makes it possible to correct the ray coordinates for each point in space on the image plane, using the following relation:

$$\hat{x} = \mathbf{K}^{-1} \cdot x \quad (7.5)$$

Now that the form of the calibration matrix is agreed upon, it can be put together with the camera extrinsics, the camera orientation in space (the camera position regarding the space coordinates). That position can be parameterised by a rotation matrix and a translation vector, respectively,  $\mathbf{R}$  and  $\mathbf{t}$ . As such, it is possible to define the camera matrix:

$$\mathbf{P} = \mathbf{K} \cdot [\mathbf{R} | \mathbf{t}] \quad (7.6)$$

It is preferable to use an invertible  $4 \times 4$  matrix, defined as follows:

$$\tilde{\mathbf{P}} = \begin{bmatrix} \mathbf{K} & \mathbf{0} \\ \mathbf{0}^T & 1 \end{bmatrix} \cdot \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix} \quad (7.7)$$

With this matrix, a point in space can directly be transformed into screen coordinates using the equation:

$$\bar{q} \sim \tilde{P} \cdot \bar{Q} \quad (7.8)$$

where  $\bar{Q} = (x, y, z, 1)$  and  $\bar{q} = (x_s, y_s, 1, d)$ ,  $d$  is the point's disparity and  $\sim$  denotes equality up to a scale.

### 7.3 Binocular Disparity and Stereo Vision

Part of the problem that is addressed in this Section is that of binocular disparity, namely the different displacement objects undergo when seen through different viewpoints. The human brain uses this process to identify distances to objects. The same process is used in computer vision, where similar features are matched and their disparity is calculated. Then, with some computation and provided that there is some knowledge between the cameras (or that knowledge is inferred, as it will be seen in Section 7.6.2), it is possible to reconstruct the scene up to a metric reconstruction, where the lengths and distances of that reconstruction have a direct relation to real world measurements. If such knowledge is unknown, every reconstruction is relative to a trivial projective transformation, and thus the measurements of that projection might have no direct relation to real world measurements.

Based on [98, 197, 228], it is possible to build a mathematical model in order to relate two views (see Figure 7.5).

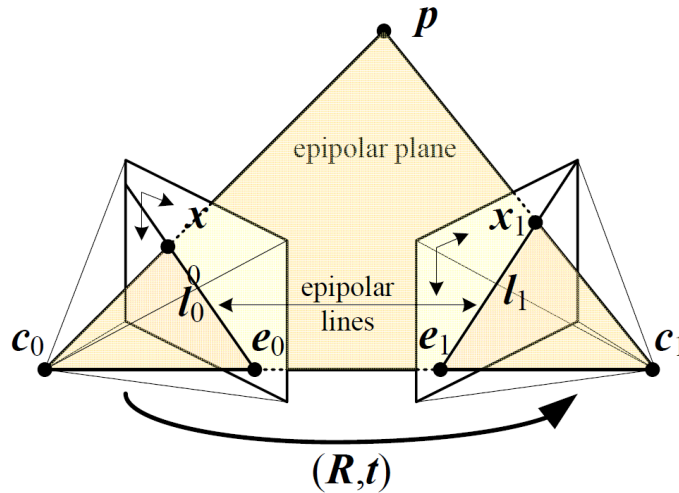


Figure 7.5: Geometric model for two-view geometry based on the pinhole camera model. The projection of point  $p$  is, respectively,  $x$  and  $x_1$  on the left and right view,  $c_0$  and  $c_1$  are the camera centres, and  $e_0$  and  $e_1$  the epipoles. Finally,  $l_0$  and  $l_1$  are the epipolar lines that pass through  $x$  and  $x_1$ , respectively, and are contained in the epipolar plane for point  $p$ . Note that all epipolar lines and planes go through the epipoles. Also note that, between the two cameras, it is assumed that there is a rotation and translation associated  $(\mathbf{R}, \mathbf{t})$  (From [228]).



Although the exact position of point  $p$  is not known, it must be located somewhere in the ray cast between the camera centres and that point, which pass through the image plane in each one of the points  $x$  and  $x_1$ . After some geometrical relations, and based on the camera matrix discussed in the previous section, it is possible to state that there is a matrix  $\mathbf{E}$ , called *essential matrix* [151],  $3 \times 3$  of rank-2, that relates every corresponding point as follows:

$$\hat{x}_1^T \cdot \mathbf{E} \cdot \hat{x} = 0 \quad (7.9)$$

where  $\mathbf{E}$  is defined as:

$$\mathbf{E} = [\mathbf{t}]_{\times} \cdot \mathbf{R} \quad (7.10)$$

where  $[\mathbf{t}]_{\times}$  is the skew-symmetric matrix [228].

This is, of course, if both  $\mathbf{R}$  and  $\mathbf{t}$  are known. In this project's case study, those matrices are not known, as it deals with free-moving hand-held cameras. Also, as it is an uncalibrated environment, there is no easy access to camera parameters and to the cameras' calibration matrix. In such situations, it is possible to define a new matrix  $\mathbf{F}$ , the fundamental matrix, which still respects the epipolar constraint, as follows:

$$\hat{x}_1^T \cdot \mathbf{E} \cdot \hat{x} = x_1^T \cdot \mathbf{K}_r^{-T} \cdot \mathbf{E} \cdot \mathbf{K}_l^{-1} \cdot x = x_1^T \cdot \mathbf{F} \cdot x = 0 \quad (7.11)$$

The matrix defined in (7.11) [77, 98, 197] has many applications, and it is of the utmost importance, as will be confirmed in this chapter.  $\mathbf{F}$  is normally found using at least 8 matches in both images, solving a system of equations using Singular Value Decomposition (SVD) [90]<sup>1</sup>. It can, anyway, be estimated using only 7 correspondences by solving a system of non-linear equations, as this matrix only has seven degrees of freedom. If more than eight points are available, then it is possible to minimise the effect the noise can introduce while estimating  $\mathbf{F}$ . If using automatically established correspondences, then there may be some outliers (wrongly-matched points) that should be addressed.

The process of stereo matching consists of finding correspondences for the maximum number of pixels in each image. Assuming a low baseline (the displacement or disparity between the two views is significantly smaller than the distance to the objects on the scene), it is possible to assume that most of the pixels from either images will match; in other words, almost all scene points are visible in both views. It is possible to facilitate the correspondence problem and the computation of correspondent distances by pre-aligning the views, a process called *rectification*, so that corresponding epipolar lines are horizontally aligned. That way, the stereo matching problem is reduced to a one-dimension search and the math is reduced to what is summarised in Figure 7.6.

From there it is possible to note that the problem, as it is posed here, assumes that the focal length  $f$  is the same for both views. That is common as even with autofocus, in low baseline problems whose value tends to be the same [26], even though the state-of-the-art methods do try

<sup>1</sup>Simple and intuitive explanation in <http://www.cs.unc.edu/~marc/tutorial/node54.html>

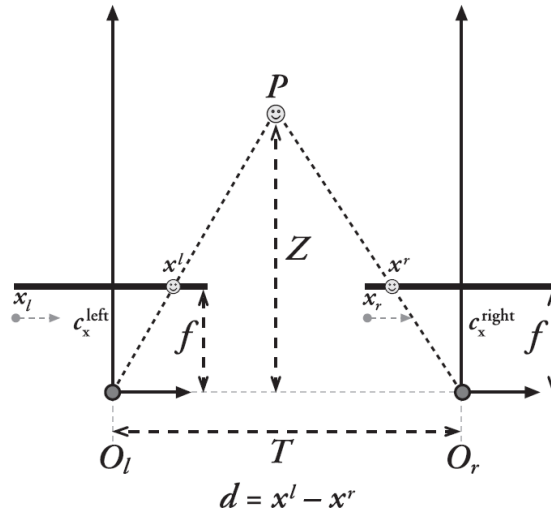


Figure 7.6: Disparity search summary assuming a low baseline ( $T \ll Z$ ). The search space for correspondences is limited to the  $x$ -dimension and the disparity value is very simple to compute (From [26]).

to transform that (and other) intrinsic parameter, assuming it is not shared by both the views, so that the model proposed holds. By simple triangle similarity, it is possible to verify that:

$$\frac{T - (x_l - x_r)}{Z - f} = \frac{T}{Z} \Rightarrow Z = \frac{f \cdot T}{x_l - x_r} \Rightarrow Z = \frac{f \cdot T}{d} \quad (7.12)$$

Since depth is inversely proportional to disparity, there is obviously a nonlinear relationship between these two terms. When disparity is near 0, small disparity differences make for large depth differences. When disparity is large, small disparity differences do not change the depth significantly. The consequence is that stereo vision systems have high depth resolution only for objects relatively near the camera (see Figure 7.7).

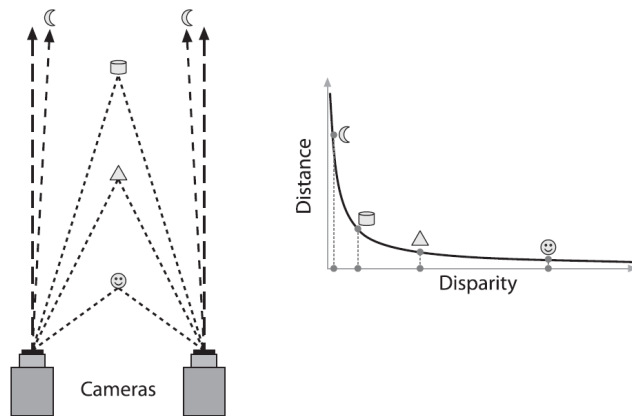


Figure 7.7: Disparity and distance relation (From [26]).

## 7.4 Feature Detection

One of the best contributions which can be found in the literature, as far as explanation and comparison of feature detection and matching algorithms, are concerned is that provided by Tuytelaars and Mikolajczyk [241]. Detailed explanations are also provided in the previously referred books [26, 228].

A feature can be defined as an image point or pattern that is relevant, different from its neighbour, and for the purpose of this project, especially interesting if visible in both views and mathematically matchable. Those features can be isolated points (pixels), but can also be edges or even other portions (such as blobs). Matching normally occurs either by performing measurements around a window or by computing especially designed descriptors. Window matching is normally not invariant to transformations, and therefore descriptors are developed in order to create that invariance.

Notice that in this part of the work we are dealing with free-moving cameras. As such, it is not known, *a priori*, what type of transformations the scene may suffer from one view to another. Thus, the ideal type of features must be invariant to the highest number of transformations possible. Good features have associated a set of commonly desired properties [241]. For this project, all properties are required, except on, which is *Efficiency*. This happens because the aim is not to create a time-critical application, and so there is more time to perform the detection. Therefore, the features to be found should have these properties:

- *Repeatability* since most of the features detected in one view should be visible and detected in the other;
- *Distinctiveness*, as features should be informative enough to allow an accurate and robust matching between all features present in an image;
- *Locality and Accuracy*. In Section 7.2 it was demonstrated that the geometrical model would require local points to be matched, and so the detected features should have an accurate location on the image plane;
- *Quantity*, even though a small minimum number of correspondences are necessary, a larger number of features would reduce noise-induced errors.

### 7.4.1 Corner Detectors

Corners, in a planar imaging sense, are points of high curvature and may not correspond to actual corners in 3D objects. The most classical algorithm for corner detection is that of Harris and Stephens [95], commonly called **Harris corner detector**. It is based on finding large changes in the derivatives signal of the image plane. A large change in both directions is a strong indication there might be a corner in that location. This algorithm was then updated to a scale-invariant version and then to an affine-invariant version [241].

The *SUSAN* [217] (*Smallest Univalve Segment Assimilating Nucleus*) is another commonly used algorithm. The principle is simple: defining a circular patch around each pixel of the image, with the centre being called the *nucleus*; each pixel in that radius is assigned with similar or different intensity values to that of the *nucleus*. Therefore, a corner is defined as a *nucleus* where the circular neighbour has less than a defined threshold of similar pixels.

Observing the images presented in the database (see Chapter 4), a good performance from the corner detectors is not expected. Visually there are not many high curvature salient points in the scenes that this project deals with.

### 7.4.2 Region Detectors

A "region" might be imagined as a region of pixels that are "darker" or "brighter" than the surrounding neighbours. Local features detected with region methods typically represent homogeneous regions. While this is acceptable for the detection step, there may be some problems for the description and matching which occur afterwards. Indeed, homogeneous regions lack distinctiveness. Fortunately, this can easily be overcome by increasing the measurement region. In other words, a larger scale region is used to compute the descriptor, such that it also contains part of the surrounding image structures, and captures the shape of the region's boundary. This usually suffices to increase the discriminative power and to match regions between images.

One possible algorithm to find "regions" that are invariant to some transformations (namely, affine) is the *Hessian-Laplace detector* [167]. This algorithm is based on a matrix of second order derivatives of the image intensities, created around points detected by the determinant of Hessian. Features are detected when the trace of such matrix attains a local maximum.

Regions may also be found using *Maximum Stable Extremal Regions* (MSER) [166]. The objective is to search for regions with distinguishable, stable and invariant shapes. To detect those regions, a threshold at all possible gray levels is applied to the image. While that threshold is changing, the area of connected components (pixels) is monitored. Regions where that area changes below a defined threshold are defined as maximally stable.

Another possibility is using the Difference of Gaussian (DoG) detector type [53]. This algorithm requires only a convolution on the images in a constant matter, depending on a Gaussian smoothing parameter. The image scale-space is developed with Gaussian filters at different scales. The features are detected searching for extremal values by comparison with the neighbours of the pixels at the various scales and octaves [156] (see Figure 7.8).

### 7.4.3 Edge Detectors

Edges are expected to vary greatly from one view to another, as a rotation and translation almost certainly would hide and/or transform the detected edges, invalidating the matching tasks. For that reason, the edge detectors were disregarded due to the time span of this project.

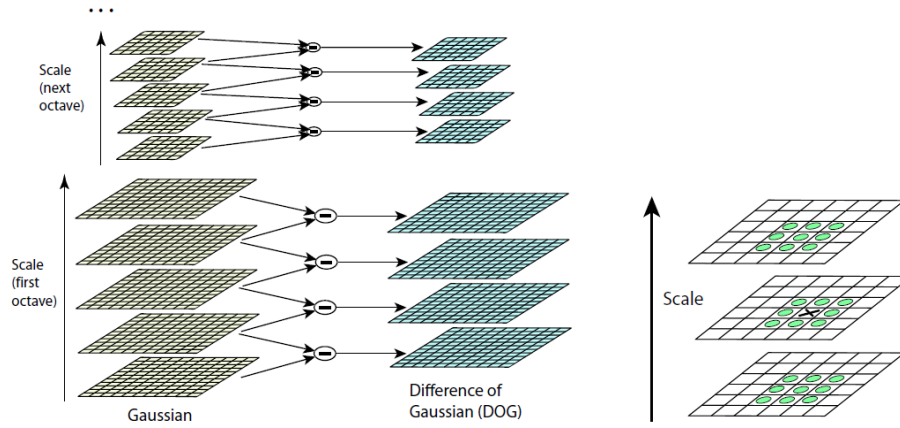


Figure 7.8: Difference of Gaussian space-scale pyramid visual description (left) with extremal pixel finding by comparison to the neighbours at each level (right) (From [156]).

#### 7.4.4 Efficient Feature Detectors

The Scale-Invariant Feature Transform (SIFT) [155] is a methodology which transforms an image into a large collection of local feature vectors, each of which is invariant to image translation, scaling, and rotation, partially invariant to illumination changes, noise, minor changes in viewpoint and robust to local geometric distortion (scale and rotation). These features are defined as maxima and minima of the result of the aforementioned DoG applied in scale space to a series of smoothed and resampled images. With Lowe's method, it is possible to extract features which are highly distinctive and relatively easy to extract, and to identify objects correctly with low probability of mismatch. They are relatively easy to match against a database of local features - however, the high dimensionality can be an issue, and generally probabilistic algorithms are used.

Bay *et al.* [16, 17] introduced the Speeded Up Robust Feature (SURF) which is partly inspired by the SIFT descriptor. The standard version of SURF is several times faster than SIFT and its authors claim that it is more robust to different image transformations comparatively to the SIFT. It is a scale-invariant feature detector based on the Hessian-matrix, as is, for instance, the Hessian-Laplace detector presented previously. However, rather than using a different measurement for selecting the location and the scale, the determinant of the Hessian is used for both. The Hessian matrix is roughly approximated, using a set of boxtype filters, and no smoothing is applied when moving from one scale to the next.

The Features from Accelerated Segment Test (FAST) [207, 208] is based on the SUSAN detector previously discussed in Section 7.4.1. The most promising advantage of the FAST detector is its computational efficiency. The SUSAN computes the fraction of pixels within a neighbourhood that have similar intensity to the center pixel, while the FAST compares pixels only on a circle of fixed radius around the centre. The test criterion operates by considering a circle of 16 pixels and radius of 3 around the corner candidate  $p$ , improved with machine learning techniques [199].

Özuyal *et al.* [185] introduced the *Ferns* method which addresses problems of illumination and affine invariant recognition. Given a query and a target image, the camera affine invariance of

the method is given by the construction of an orbit of affine and blurs deformations of the target image. Then a set of key points (using a cornerness measurement) are computed on this orbit. Only those key points that are stable under the deformations are used as effective key points. Each of them determines a class which is characterised by the comparisons (encoded in a binary vector) of the gray levels of a selected random set of pairs of surrounding points. These comparisons determine the posterior probability of a given point belonging to a class and can be used for classification using a naive Bayes approach. Thus, assuming that the priors of each class are uniform, the next step is computing of the the likelihood of each class. Since one cannot assume that each question is independent of the others, the proposal in [185] is to assume that the set of questions is organised in subsets, called Ferns, which are independent. A Fern is determined by a random selection of a subset of pairs of pixels where the values of the grey levels are compared. Several Ferns are used to characterise a key point. Each comparison produces a binary answer. This guarantees the morphological invariance of the method, which amounts to its invariance with regard to illumination changes.

## 7.5 Feature Descriptors

After finding all features, the next task is to try and match them with features in the other view. The most robust way to do that is to define and compute a descriptor for each feature found, and then search for correspondences on those descriptors. It is common for the features and the image to, at least, suffer some translation and rotation. It is also possible that there is some type of affine transformation involved. As such, it is of great importance to use robust feature descriptors in this part of the work.

Brown *et al.* [27] defined the *Multi-Scale Oriented Patches*, a descriptor based on normalised intensity patches. It defines a  $5 \times 5$  patch around the feature location, and the normalisation process transforms the intensities so that their average is zero and the variance is one. It performs well even though small image transforms occur.

Lowe [156] developed a robust descriptor using the *SIFT* algorithm. Consider a window of  $16 \times 16$  around the feature location, using the level of the Gaussian filter in which the feature was detected. The descriptor is then based on computing the gradient in each of the pixels contained in that window, downgraded by a Gaussian fall-off function (to reduce the effect of those pixels far from the feature centre). Then, in each quadrant of that window, it develops an orientation histogram by adding the value of each pixel to one of the eight orientation beans using trilinear interpolation in a  $2 \times 2 \times 2$  histogram (4 eight-bean histogram for each quadrant). That forms a 128-Dimension vector that is then normalised to unit-size, clipped to 0.2 and then re-normalised.

Inspired by the *SIFT* descriptors, Ke and Sukthankar [123] proposed a simpler approach. Around each feature, compute the  $x$  and  $y$  derivatives around a patch of  $39 \times 39$  pixels. That amounts to a 3042-dimensional vector that is then downsampled to a 36-dimensional vector using the Principal Component Analysis (PCA) [228].

Lastly, Mikolajczyk and Schmid [168], again considering the *SIFT* descriptors, proposed a small change to these descriptors. Therefore, instead of using the usual square-defined descriptor, their algorithm uses log-polar coordinates (still, centred in the feature point) to define the histogram structure. As such, the bin space are of radius 6, 11 and 15, with eight bins per coordinate, except for the centre coordinate, for a grand total of 17 bin-coordinates, each with 16 orientation bins. The 272-dimension resulting vector is then mapped to a 128-dimension vector using the PCA trained on a large database.

After detecting the features and obtaining the descriptors, it is still necessary to define a matching strategy. Normally, it is divided into two different stages: selecting a matching strategy, determining the correspondences which go to the next step taking into account their "quality"; and determining the devise efficient data structures and algorithms to perform this matching as quickly as possible. Determining the feature matches which are reasonable to process further depends on the context. The most commonly used is the simple Euclidean distance between all "target" descriptors, provided that a certain distance threshold is attained [228]. However, nearest-neighbour methods, for example, are suitable if efficiency is crucial.

## 7.6 Rectification

Section 7.3 confirms that rectification is a very important and necessary step in the reconstruction process. This step is highly based on the epipolar geometry already discussed, where the main goal is to find a way to horizontally align corresponding epipolar lines.

### 7.6.1 Based on Known Objects

If a known object is placed on the scene, visible to both views and with identifiable and matchable features, then it is possible to recover the pose of the two cameras from those measurements. In this special case, it is possible to find both the translation vector and the rotation matrix for the cameras concerning that object, and (7.10) is applicable. Basically, there are two principal ways to compute rectification terms: Hartley's algorithm [97], which can yield uncalibrated stereo using just the fundamental matrix; and Bouguet's algorithm [237, 263, 264], which uses the rotation and translation parameters from two calibrated cameras to align the epipolar lines. In practice, Hartley's algorithm produces more distorted images than Bouguet's calibrated algorithm. In situations where calibration patterns can be applied, Bouguet's algorithm is more stable [26].

### 7.6.2 Uncalibrated Rectification

There is extensive work in the literature on uncalibrated rectification. The possibility of reconstructing scenes from sets of frames or pictures is much appreciated and is extensively studied. Here only recent and recognised algorithms will be presented, among many others that can be found in the literature.



The state-of-the-art papers in uncalibrated rectification [112, 152] date back to 1999. In [152] the authors decompose each collineation into similarity, shearing and projective factors and attempt to make the projective component "as affine as possible", while in [112] the authors build upon [97] and propose a method that avoids computation of the fundamental matrix, using the same distortion criterion as in [97].

The process of planar rectification is to find a pair of homographies that will align the epipolar lines. Therefore, it is possible to formulate these assumptions:

1. All epipolar lines are parallel to the  $x$  axis;
2. All image features and points have the same corresponding  $y$  coordinate.

The algorithm from Loop and Zhang [152] is based on finding a pair of homographies that will maintain the epipolar constraint. The goal is to find those homographies by establishing a set of simpler transforms, computing each component to achieve a desired effect and satisfy some conditions. So, the homography  $\mathbf{H}$  is decomposed in two homographies, so that:

$$\mathbf{H} = \mathbf{H}_a \cdot \mathbf{H}_p \quad (7.13)$$

where  $\mathbf{H}_a$  is an affine transformation and  $\mathbf{H}_p$  is a projective transformation. Then,  $\mathbf{H}_a$  is decomposed as:

$$\mathbf{H}_a = \mathbf{H}_s \cdot \mathbf{H}_r \quad (7.14)$$

where  $\mathbf{H}_s$  is a shearing transformation and  $\mathbf{H}_r$  a similarity transformation (for more detail see the work by Loop and Zhang [152]).

All these transformations are computed using minimisation with specific criteria for each objective. This algorithm's result can be found in Figure 7.9.

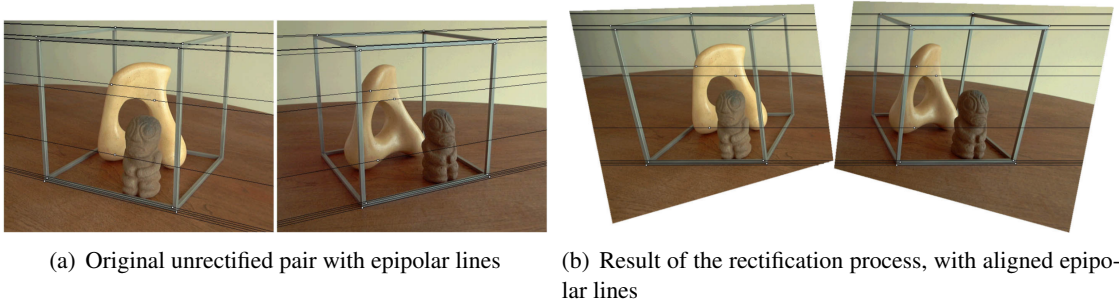


Figure 7.9: Results from the rectification algorithm developed by Loop and Zhang [152].

Hartley and Zisserman [98] created a method that is based on the process of relocating the epipoles in both images. The algorithm attempts to transform the epipoles so that their location is set at infinity (with the last coordinate as 0, in homogeneous coordinates). Start with a set of matches  $x_i \leftrightarrow x'_i$ , which can never be less than seven (though more than that is preferred). Compute



the fundamental matrix  $\mathbf{F}$  from those matches and find the epipoles  $e$  and  $e'$  so that  $e' \cdot \mathbf{F} = 0$  and  $\mathbf{F} \cdot e = 0$ . Then, compute the projective transformation  $\mathbf{H}'$  that maps the epipole  $e'$  to infinity,  $(1, 0, 0)^T$ . Using that, find the matching projective transformation  $\mathbf{H}$  that minimises the least-squares distance:

$$\sum_i d(\mathbf{H} \cdot x_i, \mathbf{H}' \cdot x'_i). \quad (7.15)$$

An example of the result from this algorithm can be found in Figure 7.10.



Figure 7.10: Uncalibrated rectification result from Hartley and Zisserman [98] for a generic real image (From [265]).

Fusiello and Irsara [85] recently developed a new rectification algorithm that tries to get close to that of the Euclidean rectification. This method uses some autocalibration notions to estimate the camera parameters (that are assumed to be equal for both views, as it is the case for the images used). From a set of correspondences  $m_l^j \leftrightarrow m_r^j$  a Sampson error function is defined for each correspondence:

$$E_j^2 = \frac{(m_r^{jT} \cdot \mathbf{F} \cdot m_l^j)^2}{(\mathbf{F} \cdot m_l^j)_1^2 + (\mathbf{F} \cdot m_l^j)_2^2 + (m_r^{jT} \cdot \mathbf{F})_1^2 + (m_r^{jT} \cdot \mathbf{F})_2^2} \quad (7.16)$$

where  $(\bullet)_i$  is the  $i^{\text{th}}$  component of the normalised vector.

The homographies are then obtained by creating a system of non-linear equations to every  $j$  correspondence ( $E_j = 0$ ). The solution is obtained through least-squares using the Levenberg-Marquardt algorithm, and through parameterisation so that:

$$\mathbf{H}_i = \mathbf{K}_{ni} \cdot \mathbf{R}_i \cdot \mathbf{K}_{oi}^{-1} \quad (7.17)$$

where  $i$  is  $l$  or  $r$ , for each one of the views.

It is shown that the rotation of one of the views around the  $x$ -axis is irrelevant, and therefore, it is discarded. Some rectification examples can be found in Figure 7.11.

This algorithm was then improved by Monasse [171] showing better results. The biggest changes are:

- Improved Jacobian matrix calculations;



Figure 7.11: Rectification examples by Fusiello and Irsara [85]. The original views are on the left, while the rectified ones are on the right.

- Some changes on the terms and geometric interpretations;
- Actual use of the Levenberg-Marquardt minimisation;
- Use of the Optimised Random Sampling Algorithm (ORSA) algorithm [170] instead of the common Random Sample Consensus (RANSAC) to find and eliminate outliers from the matched points.

Then, the relation between the fundamental matrix  $\mathbf{F}$  and the rectification process is that of finding the pair of homographies that will transform the images so that they become rectified [26, 152, 197, 228]. In the rectified case, the  $\mathbf{F}$  has the form [85, 152]:

$$\mathbf{F} = [u_1]_{\times} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \quad (7.18)$$

where  $[u_1]_{\times}$  is the skew-symmetric matrix [228].

To perform the rectification step, it is necessary to search for the aforementioned homographies ( $H_r$  and  $H_l$ ), meeting the epipolar geometry constraint (7.10) and (7.11) [85, 152]:

$$(\mathbf{H}_r \cdot \mathbf{x}_r)^T \cdot [u_1]_{\times} \cdot (\mathbf{H}_l \cdot \mathbf{x}_l) = 0 \quad (7.19)$$

where:

$$\mathbf{F} = \mathbf{H}_r^T \cdot [u_1]_{\times} \cdot \mathbf{H}_l \quad (7.20)$$

## 7.7 Stereo Matching

The stereo matching process is almost always difficult and prone to errors, mainly because it is ill-posed [211]: due to the possible colour repeatability, its solution might not be unique. There will be multiple situations where it will not be possible to define, with confidence, the correct correspondence between pixels. The colour values in some regions might be equal or very similar in some areas, thus preventing a confident matching. The classical algorithm for finding pixel correspondences is based on simple correlation, also called the "brute force" algorithm. It can be simply described as such: for each pixel in a "base" image, calculate a weight function in a window around that pixel; then, for each pixel in the same row of the "target" view, calculate the weight of the window around it. The "target" pixel that has a value closer to that of the "base" pixel is, potentially, the best match. It is not difficult to see that this algorithm can greatly suffer from problems in low or repeated textured scenes, where there would be similar valued pixels and where a unique match is not possible [228]. Common window evaluation functions are, for example:

$$\begin{aligned} \text{Normalised correlation: } & \frac{\sum_x \sum_y L(x,y) \cdot R(x,y)}{\sqrt{(\sum_x \sum_y L(x,y))^2 (\sum_x \sum_y R(x,y))^2}} \\ \text{Sum of squared differences: } & \sum_x \sum_y (L(x,y) - R(x,y))^2 \\ \text{Sum of absolute differences: } & \sum_x \sum_y |L(x,y) - R(x,y)| \end{aligned} \quad (7.21)$$

where  $L$  and  $R$  are, respectively, the images on the left and on the right (see Figure 7.12).

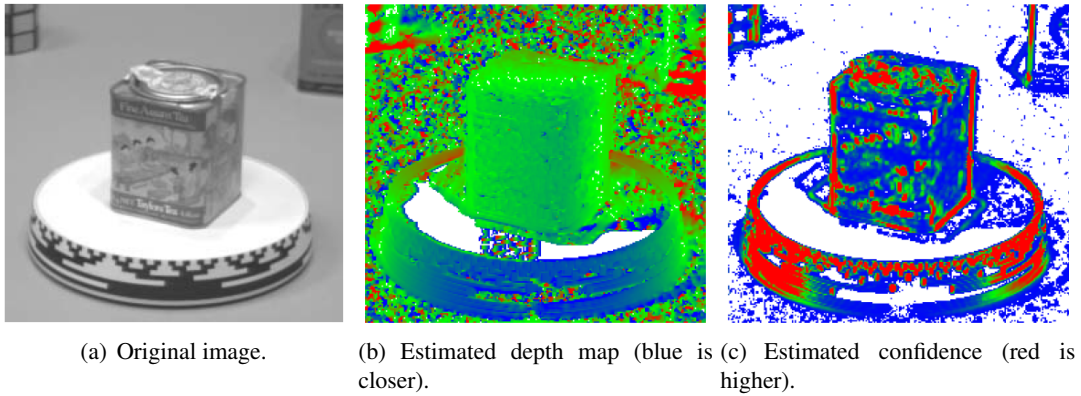


Figure 7.12: Uncertainty in stereo matching, as confirmed in the last image (From [228]).

When faced with a pair of stereo images, it could be difficult to establish a dense correspondence for each pixel in both images. This section aims to present some interesting algorithms that are considered top-notch to solve that problem in a correct and efficient way. Nevertheless, the reader should note that efficiency is not a concern at this stage, and so there will be no performance analysis in spatio-temporal or complexity terms.

Kolmogorov and Zabih [132, 133] present a well-known and recognised algorithm for stereo matching. Their approach is based on energy minimisation via graph cuts, and it is presented as

a fast and accurate matching method. They give some degree of interest to the problem of occlusion detection. Their work treats both images symmetrically, handles visibility properly, and imposes spatial smoothness while preserving discontinuities. The problem of energy minimisation is proved to be NP-hard, and therefore graph cuts are necessary so that it is possible to compute a local minimum of the energy function. Their experimental results are very promising and confirmed the performance of the algorithm. The results are presented in Figure 7.13.

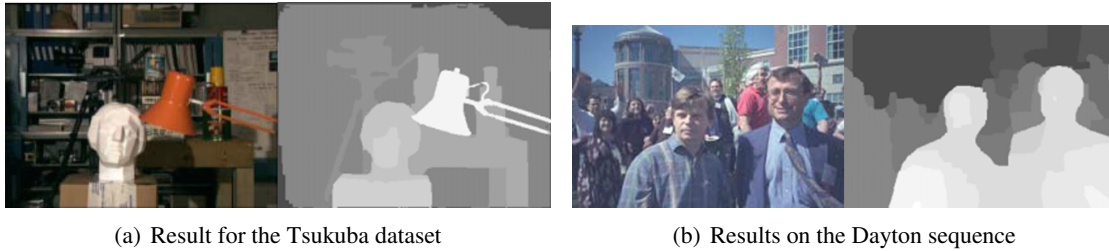


Figure 7.13: Stereo matching results from Kolmogorov and Zabih algorithm. A lighter colour means that the points are closer to the camera (From [133]).

Radim Šára contributed with various algorithms for efficient stereo matching. His work was based on trying to solve the problems of ambiguity and occlusion. Believing that studying the ambiguity related to the reconstruction would extensively help the reconstruction problem, he presented a methodology [211] describing an algorithm that is intended to find the largest unambiguous portion of matched points, identifying variables that can be manipulated to identify what is or is not ambiguous. As a result, he defined a new *stability property* that is a condition a set of matches must satisfy to be considered unambiguous at a given confidence level. The experimental results can be seen in Figure 7.14.

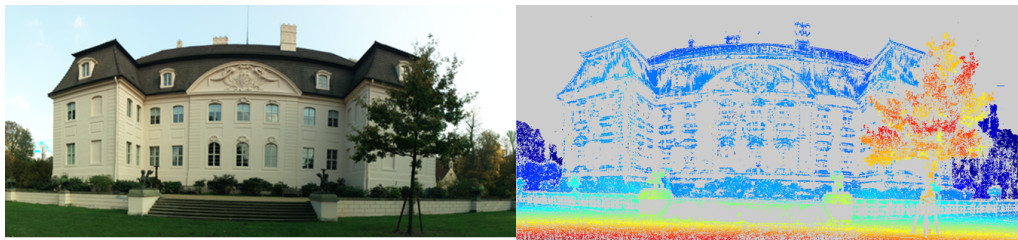


Figure 7.14: Results from the algorithms presented by Šára [211].

Later this author presented, in [212], a new set of algorithms based on his previous work. The aim was to build a robust framework for correspondence matching in computer vision. This framework is capable of adding well-structured prior knowledge. This work presents considerations and definitions on visibility, occlusions and ambiguity, and outlines the most common tasks to solve stereo problems. The paper also shows a large degree of interest in detailing that the prior knowledge or model severely affects the solution found, thus explaining that a contradicting prior model will result in artifacts and wrong disparity calculations. The framework is based on the notion of digraph kernels, which are considered excellent classification tools and have widespread applications and uses [24].



Čech and Šára [43] present a two-step stereo matching method. This algorithm visits only a small fraction of the disparity space, finding a semi-dense disparity map. From a set of seeds selected previously (using helper algorithms like [95]), the matching is performed increasing the correspondence from those seeds under a theoretically well-grounded rule, the uniqueness constraint until the result can no longer be improved. Then, a robust global optimality task is performed, selecting from the competing patterns those that are unambiguous. It is demonstrated that the algorithm is very resistant to repetitive patterns and other ambiguities. The most interesting result is that, contrarily to initial belief, there is no need for high quality seeds, as the algorithm seems to be able to perform matching in very complex scenes from a set of pseudo-random initial points. Their experimental results are very promising and can be viewed in Figure 7.15.

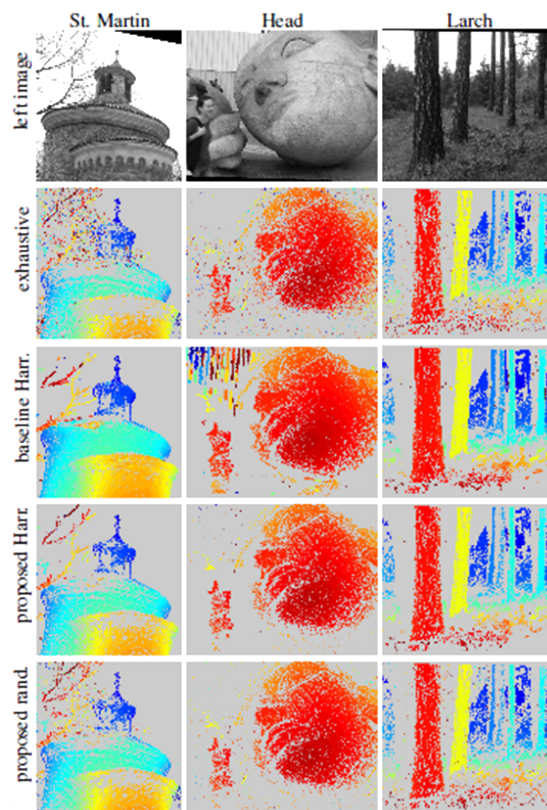


Figure 7.15: Experimental results from the algorithm presented by Čech and Šára [43] for some combinations of parameters and different images. Note that "proposed" is the proposed algorithm, "Harris" is the feature detector presented in [95] and "exhaustive" is the application of a classical exhaustive search algorithm.

Hirschmüller presented some work [109] on a stereo matching algorithm with a semi-global approach. By using an especially designed cost function, the colour variations that pixels from the same 3D point may exhibit in the two images are compensated. The matching is performed using a global cost function that propagates the matching for a determined pixel in all directions trying to reach a global optimisation value. Some post-processing steps are also considered that try to improve the solution found. Tests and comparisons confirm that this algorithm is well ranked

among the commonly existing algorithms, and the author claims that it can perform better if some subpixel accuracy is computed in the resulting disparity map. This algorithm is implemented with optimisation in the *OpenCV*<sup>1</sup> computer graphics library.

Alagoz [6] created two stereo matching algorithms based on classical global energy minimisation functions. The approach was, however, to try and add a smoothing function so as to eliminate unreliable disparity estimations. In the first algorithm, there is an attempt to smooth the energy function by applying an averaging filter a large number of times, assigning a new minimum disparity value to each pixel if this value is found in each iteration. In the second algorithm, the approach is based on region growing along image rows, defining root points where regions begin. A root point is defined as a point that does not belong to any other region, and its disparity is computed using the energy function previously defined by the author. The region is then increased from that point on while the energy function stays equal or lower than the value for the root point. Otherwise, a new root point is selected and the region starts growing from there.

Bhatti and Nahavandi [21] published another approach to the stereo matching problem. Their claim is that few attempts have been made in the area using wavelets/multiwavelets multiresolution analysis. As such, their work tries to develop a stable algorithm that uses translation invariant multiwavelet transforms to establish correspondences on the stereo pair. In the paper, they also studied some base changes and the effects they could have on the performance of the correspondence estimation. Their experimental results on some of the classical datasets are presented in Figure 7.16.

Wang and Road [252] built a stereo matching algorithm based on competitive and cooperative region optimisation. It is an energy minimisation approach. Starting by using segmentation, it defines a set of image regions, establishing initial correspondences with simple window-based methods. Those disparities are then refined through a plane fitting process that uses voting for the assignments. Finally, the disparities of all planes are refined by an inter-regional cooperative optimisation procedure based on global energy minimisation. Until this moment, according to the *Middlebury Stereo Evaluation* ranking this is the third best performing algorithm (see next paragraph). Their experimental results can be seen in the Figure 7.17.

In [129] Klaus *et al.* present another matching algorithm. Their approach is based on colour segmentation on the reference image followed by an adapting matching score, developed for maximising an unambiguous solution. After identifying areas of similar colour, the matching is performed on a local window of every point. The idea is that disparity is smooth in areas with the same colour, while large disparity variations should occur in discontinuities. After that, the disparity attribution is based on deriving a set of planes that should be adequate to represent the scene structure. Finally, each segment found in the first phase is assigned to a specific plane, using an optimisation algorithm based on energy minimisation. It is one of the top ranked algorithms found in the *Middlebury Stereo Evaluation* ranking <sup>2</sup>, right after the best algorithm, submitted anonymously, and for that reason it has been excluded from this work. The experimental results can be

<sup>1</sup>The *OpenCV* library can be found in <http://opencv.willowgarage.com/>

<sup>2</sup>The ranking can be found in <http://vision.middlebury.edu/stereo/eval/>

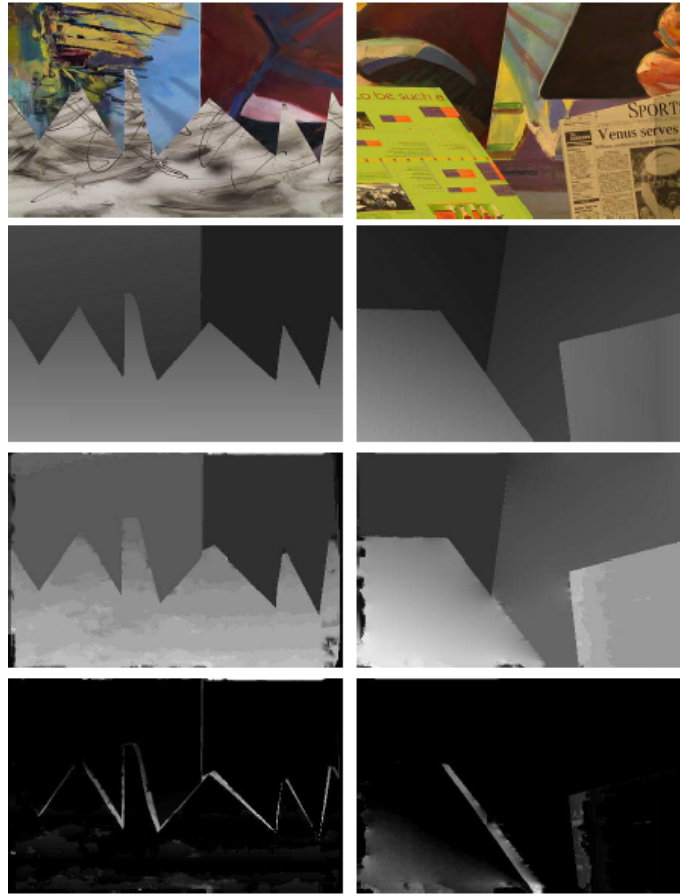


Figure 7.16: Results of the stereo matching algorithm by Bhatti and Nahavandi [21] for the classical Sawfoot and Venus stereo datasets.

found in Figure 7.18.

Geiger *et al.* [89] have very recently introduced the algorithm which they entitled *Efficient Large-scale Stereo* (ELAS) focusing on the problem of creating one algorithm that would be computationally less expensive, while achieving comparable results. Their approach is based on the Bayesian probability computation. Having in mind that most pixels in a pair of images may be ambiguous and repetitive, their method relies on a set of "support points", defined as points of strong, unambiguously matched features. Then, the assumption is that, along those ambiguous portions, the disparities should vary smoothly. After computing the disparity of those 'support points', they create a 2D mesh via Delaunay triangulation. Then, they tried to find a piecewise linear function that is able to approximate that mesh to the real disparities. Results for the classical evaluation datasets are shown in Figure 7.19.

## 7.8 Experimental Work

This section shows the results for each step of the uncalibrated 3D reconstruction based on the state-of-the-art works presented previously.

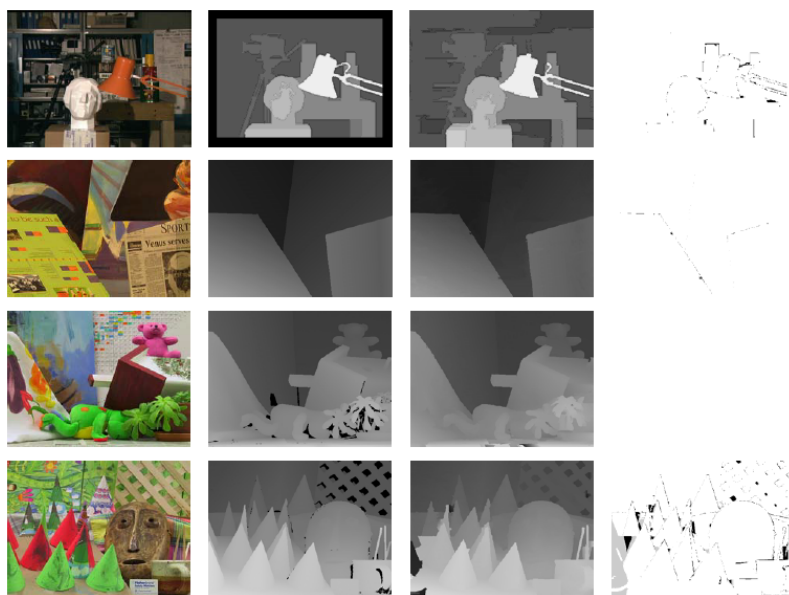


Figure 7.17: Experimental results from the stereo matching algorithm by Wang and Road [252] for the Tsukuba, Venus, Teddy and Cones databases.

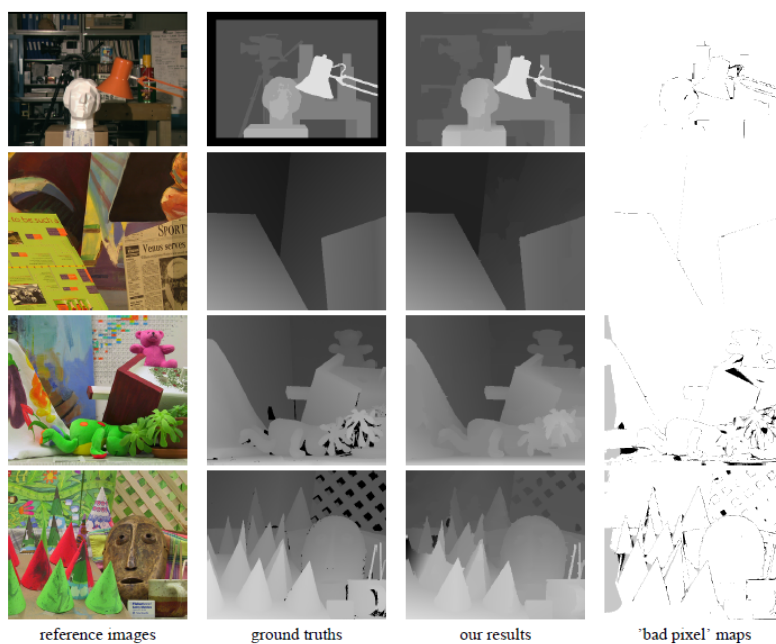


Figure 7.18: Experimental results from the stereo matching algorithm provided by Klaus *et al.* [129] for the Tsukuba, Venus, Teddy and Cones databases.

### 7.8.1 Test images

After some analyses on the state-of-the-art stereo matching descriptions, it is clear that well known and somewhat featured images have been developed and tested. The testing datasets also evolved <sup>1</sup>

<sup>1</sup>Some datasets can be found in <http://vision.middlebury.edu/stereo/data/> and the more recent have been used in the best "scoring" algorithm.



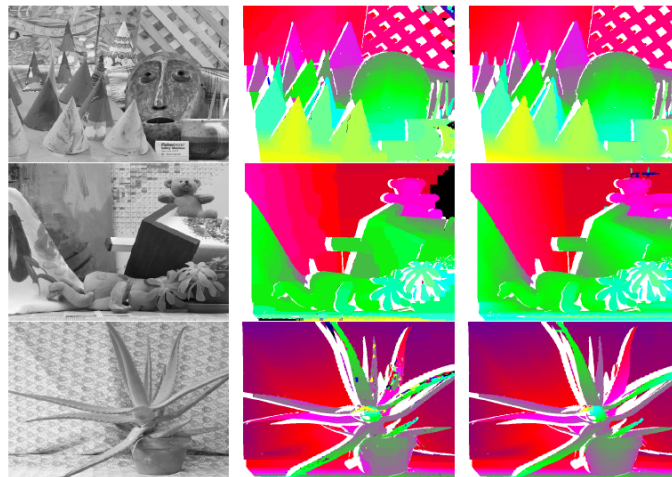


Figure 7.19: Classical result set for the stereo matching algorithm by Geiger *et al.* [89].

and it was decided that, before trying to develop any algorithm, they should be tested against a "standard" image that had some similarity with this project's images: presence of areas of high ambiguity due to low texture or repeatability. That would allow for some overview of how the algorithms would behave on those smaller and almost perfectly rectified images, before introducing the bigger, not so properly rectified views. As such, the *Aloe* and *Flowerpot* images from the 2006 *Middlebury* datasets were chosen to verify if the algorithms performed well against texture repeatability and large portions of ambiguity. The selected pairs are presented in Figure 7.20.

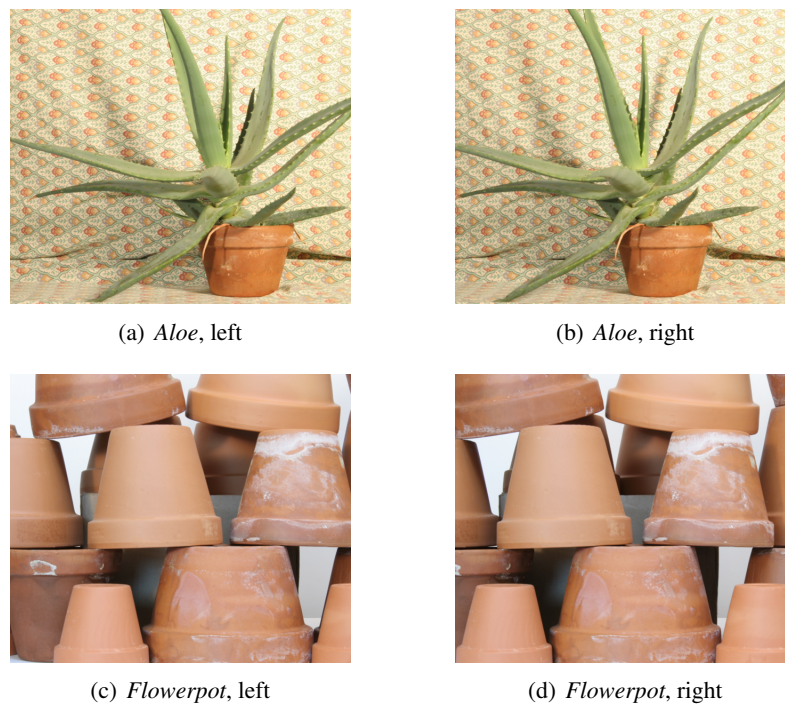


Figure 7.20: Initial test images, from the *Middlebury* dataset.

Later on, there was a brief test on two very ambiguous and virtual images. The nature of those images would pose a very hard to solve test on all known algorithms and would allow a strong starting point for this study. Those images can be seen in Figure 7.21.

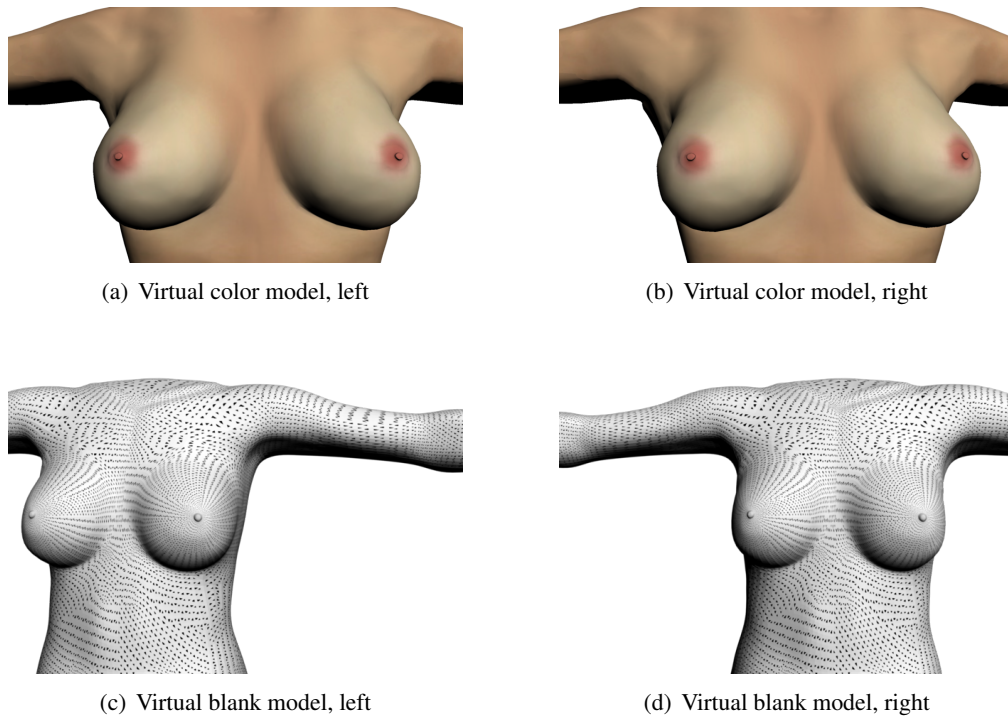


Figure 7.21: Virtual models for large ambiguity tests.

The next set of images is based on a dummy (original: 7.22(a)) that was modified to incorporate some volume change between the two breasts. Although poorly executed, the idea was to mimic the natural skin ambiguity by using almost always the same colour, although with some points of colour change that would copy the effect of skin lesions, moles, freckles or blemishes. That model and the selected views are presented in Figure 7.22.

### 7.8.2 Feature Point Detection

It is easy to understand that there are not many distinctive features in the images of this project. The most common and known algorithms were tested for the number and quality of features. The objective is to find the detector that is able to find the biggest number of quality features that can then be matched. The tested algorithms were:

- Harris corner detector [95]
- SIFT [156]
- SURF [17]
- MSER [166]



(a) Original dummy



(b) Modified dummy, left view



(c) Modified dummy, right view

Figure 7.22: Dummy model for more real tests, with induced volume difference in the breasts.

### 7.8.3 Image Rectification

The method provided by Hartley and Zisserman [98] was computed and tested using the tools present in *Matlab*<sup>®</sup> <sup>1</sup>. The Fusiello and Irsara [85] method is tested with its own *Matlab*<sup>®</sup> toolbox and, as a result, that method was improved by Monasse [171].

### 7.8.4 Disparity Map Generation

The disparity tests are performed using the author's code, either through executables or through their *Matlab*<sup>®</sup> wrappers. Not many implementations were found, while the most common code available is that of very simple matching presented in Section 7.7. Apart from those, the tested algorithms were:

- Hirschmüller [109]
- Čech and Šára [43]
- Alagoz [6]
- Klaus *et al.* [129]

<sup>1</sup><http://www.mathworks.com/products/computer-vision/demos.html?file=/products/demos/shipping/vision/videorectification.html>

- Geiger *et al.* [89]

## 7.9 Results

Here, is presented the results to the tests outlined in the previous Section. A figurative testbed was developed and there are now conditions to conduct all tests for any image that is required. The parameters used are those that were shown (by the authors) to have the best results. Note that those parameters might not be the better for these images, but parameterisation was not to be addressed here.

### 7.9.1 Feature Point Detection and Matching

The testbed was run on all images and the results are present in Table 7.1. Note that the extra column on the *SIFT* algorithm is based on matching against *SIFT* descriptors using the squared Euclidean distance. This is present because, according to Lowe [156], certain matches should be rejected when considered too ambiguous.

		Harris	SIFT		SURF	MSER	Total
		Corners	Features	Matched	Features	Regions	
Aloe	l	26	10383	5727	8606	11306	72649
	r	27	10447		8510	11024	
Flowerpot	l	8	4605	2127	1055	591	
	r	10	4611		1093	597	
Color model	l	75	3751	1748	202	9	
	r	92	3726		191	10	
Artificial model	l	500	8810	4219	9803	2364	
	r	500	8759		9409	2352	
Dummy	l	7	9832	1072	960	197	
	r	9	9663		971	205	
Total (avg. unmatched)		627	37294	—	20400	14328	
% of total		0.8	51.3		28.1	19.7	

Table 7.1: Feature detection results for the testbed and evaluation. Note that, because there are no defined matching refusal criteria for the remainder algorithms, decided decision was made to compare the global number of features detected, as a larger number raises the probability of correct matches.

The *SIFT* algorithm appears to be largely superior to the remainder, and therefore it was selected for the rest of the tests. The feature representation (as close as possible to that returned by the algorithm) is shown in Figure 7.23 and Figure 7.24. It is possible that a combination of all four algorithms would increase the accuracy; however, that depends on defining an algorithm that would:

1. Eliminate duplicate or close features;



2. Scale those features according to a certain quality level, where only the high quality features would be used.

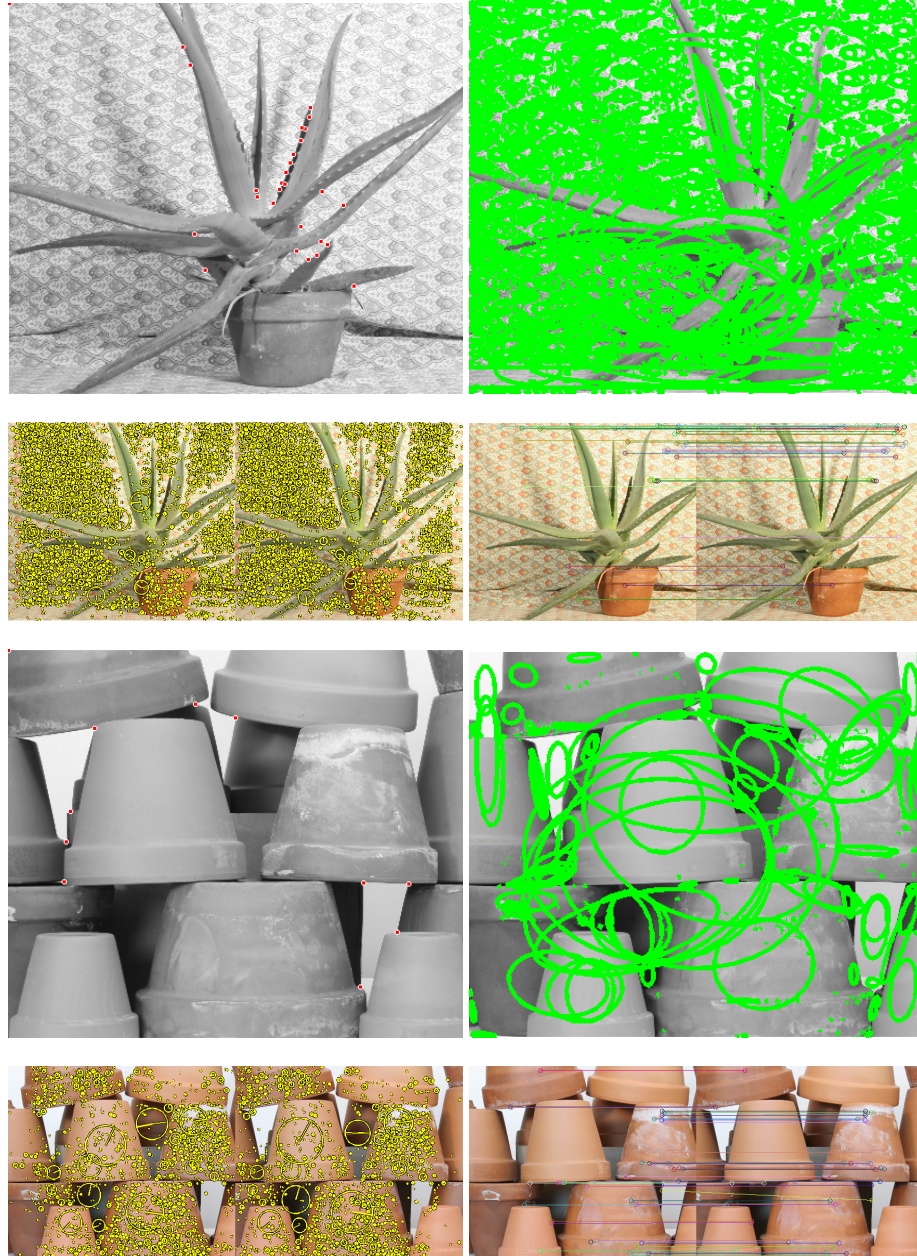


Figure 7.23: Feature detection performed in the *Aloe* and *Flowerpot*; algorithms in this order: *Harris*, *MSER*, *SIFT* and *SURF*.

### 7.9.2 Rectification

The rectification was performed on the coloured artificial 3D model and on the marked dummy as those are closer to human body representations. The other test images were discarded for

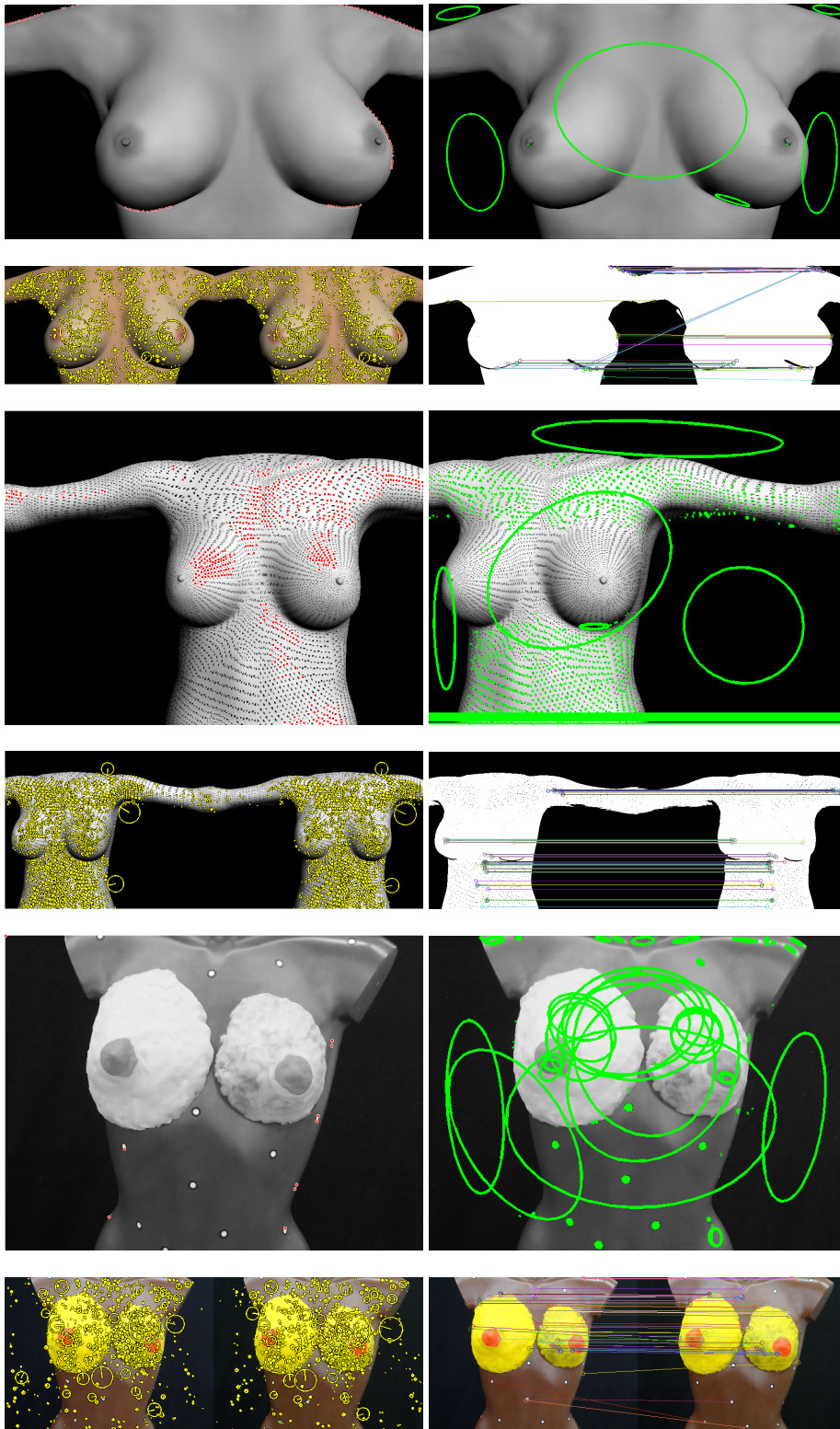


Figure 7.24: Feature detection performed in the three artificial models used; algorithms in this order: *Harris*, *MSER*, *SIFT* and *SURF*.

these tests. The pairs used for rectification were obtained using the *SIFT* both for detection and matching, in this configuration:

- Dummy: 1072 matched pairs;
- Artificial coloured model: 1748 matched pairs.

The rectification algorithm provided by Hartley and Zisserman [98] was parameterised to use different fundamental matrix generation methods. The methods that were tested were the *RANSAC*, the *Least Trimmed Squares* and the *Least Median of Squares*. All methods considered the entire set of 1072 pairs as inliers for every pair of images. The selected parameters were:

- Confidence: 99.99%
- Distance: 0.01 Sampson
- Number of trials: 10000

This method achieved positive results; however, more recent algorithms have performed better.

Fusiello and Irsara [85] achieved the best results of the test, with the lowest possible error. However, it is also the slowest to run and sometimes the result may not have been the best due to the random characteristics of the Non-linear Least Squares method. The initial algorithm does not search or eliminate outlier matches; therefore, this was done by fitting an initial fundamental matrix using the *RANSAC* at a required minimum distance to the unrectified epipolar lines for each pair set to a very restrictive 0.00001 pixel.

In this scenario, the method presented by Monasse [171] achieved the highest number of rectification errors though the author claims that the algorithm performs best. This may be due to a very restrictive set of inliers computed for the previous algorithm or it could be related to the scene or the quality of features found.

The results are summarised in Table 7.2. The rectified pairs are visually very similar, and therefore only the result with the minimum error is presented in Figure 7.25. The presented error was obtained by applying the Sampson error suggested by [84] and defined as:

$$e_{rec} = \frac{1}{N} \sqrt{\sum_j E_s^j} \quad (7.22)$$

where:

$$E_s^j = \frac{(\mathbf{m}_r^j \cdot \mathbf{F} \cdot \mathbf{m}_l^j)^2}{\|[\mathbf{u}_3]_{\times} \cdot \mathbf{F} \cdot \mathbf{m}_l^j\|^2 + \|\mathbf{m}_r^{jT} \cdot \mathbf{F} \cdot [\mathbf{u}_3]_{\times}\|^2} \quad (7.23)$$

and can be seen as the root mean squared distance of each feature point to its horizontal epipolar line.



			Color Model	Dummy
# Features			1748	1072
Hartley	RANSAC	Error	0.15	1.76
		Inliers	1748	1072
	LMedS	Error	0.18	0.689
		Inliers	1748	1072
	LTS	Error	0.17	0.78
		Inliers	1755	1069
Fusiello	Error		0.019	0.024
	Inliers		595	586
	Focal length		1006.5	2872.5
Monasse	Error		0.58	0.044
	Inliers		595	751
	Focal length		2824.61	2691.91

Table 7.2: Rectification results and comparison. **Error** values are in pixels and were established through (7.22) equation.

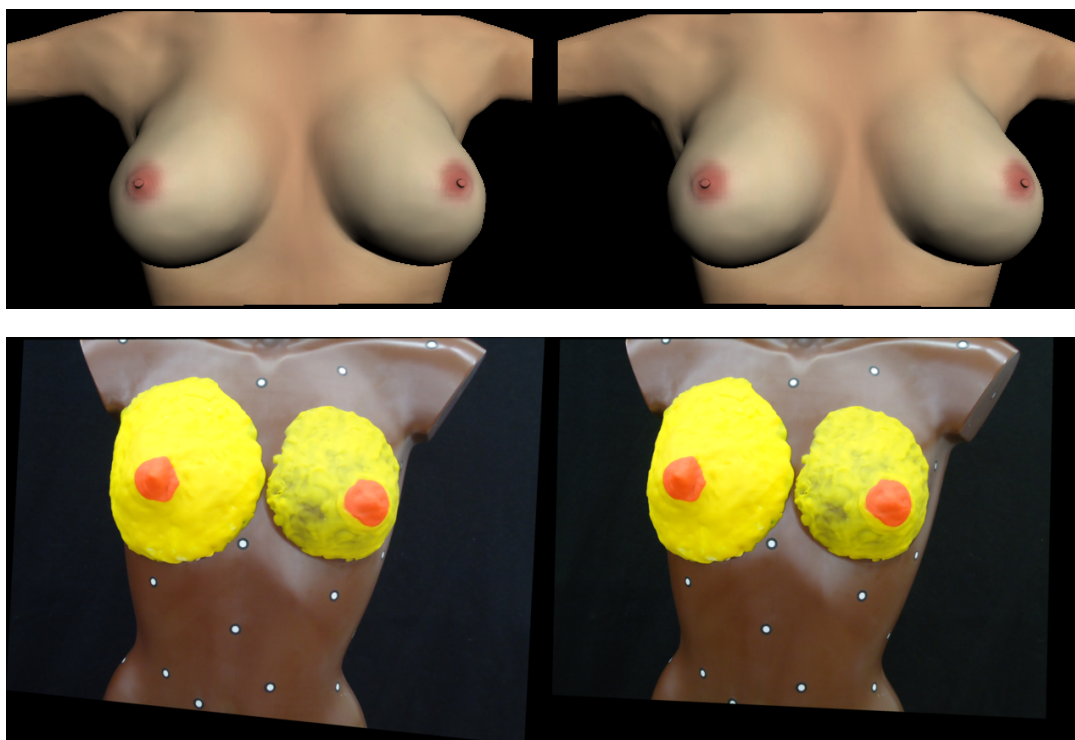


Figure 7.25: Best results from the rectification algorithms, both by Fusiello and Irsara [85].

### 7.9.3 Stereo matching

The same images used for studying the rectification algorithms will now be used to try and generate the most dense disparity maps possible. The rectification algorithm used was that with the lowest error, according to Table 7.2. There is no ground truth available and there is no clue about minimum or maximum disparity; therefore, so it is always assumed to range between 0 and a large



number, chosen to be 255.

The results for all the algorithms running over the coloured 3D model with the default parameters can be viewed in Figure 7.26. It is evident that the large ambiguity the image presents is handled poorly by all algorithms, and the resulting disparity maps are visually wrong, sparse and unintelligible.

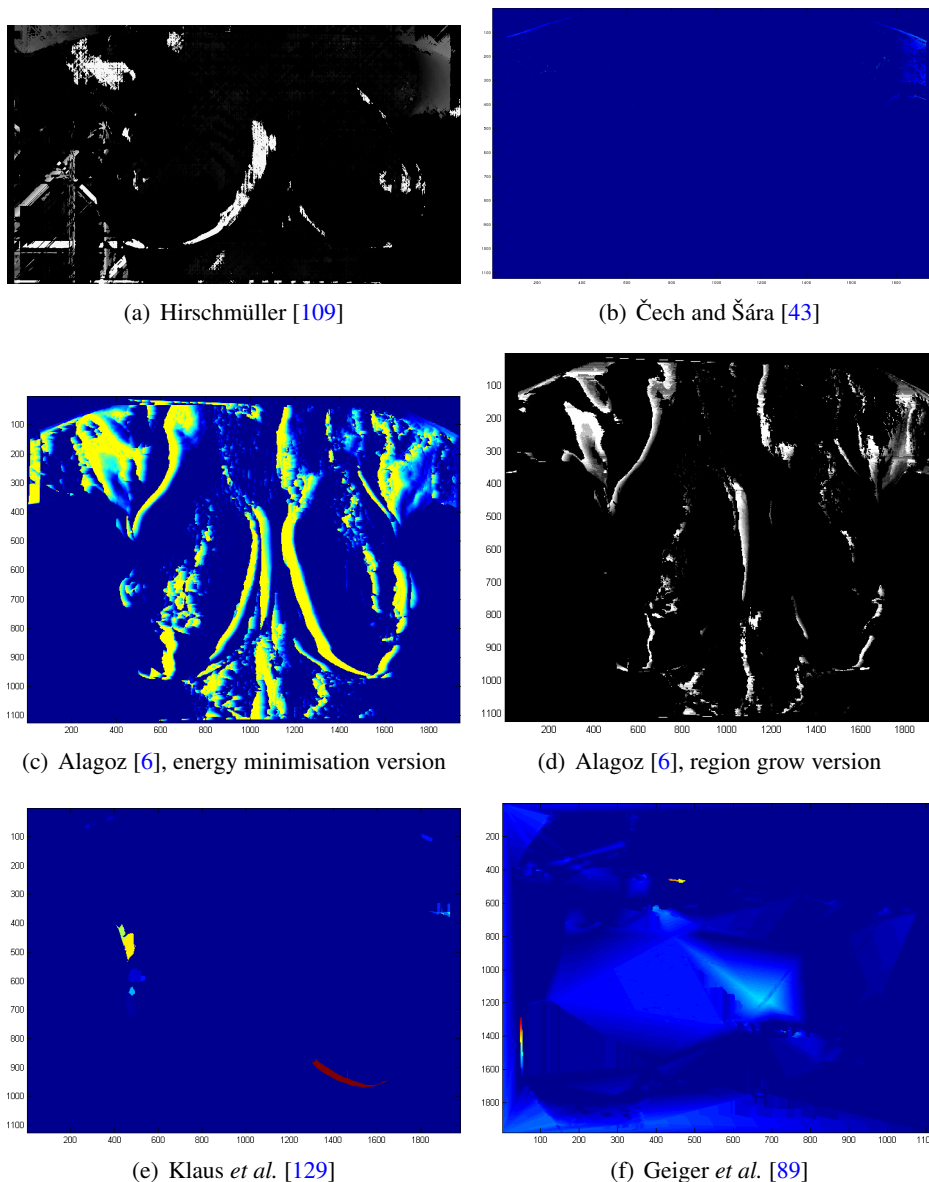


Figure 7.26: Results for stereo matching algorithms running over the 3D coloured model.

Unfortunately, for the Dummy images the maps did not improve as much. The large texture and colour ambiguity on the surface of the Dummy can be attributed to these poor results. In any case there is a disparity map that is suitable for 3D reconstruction and in most of them there are no distinguishable differences between parts of the left and right breasts, although originally they

were at different depth levels. The disparity maps can be viewed in Figure 7.27. Note that result in Figure 7.27(f) was rotated for unknown reasons by the code provided by Geiger *et al.* [89].

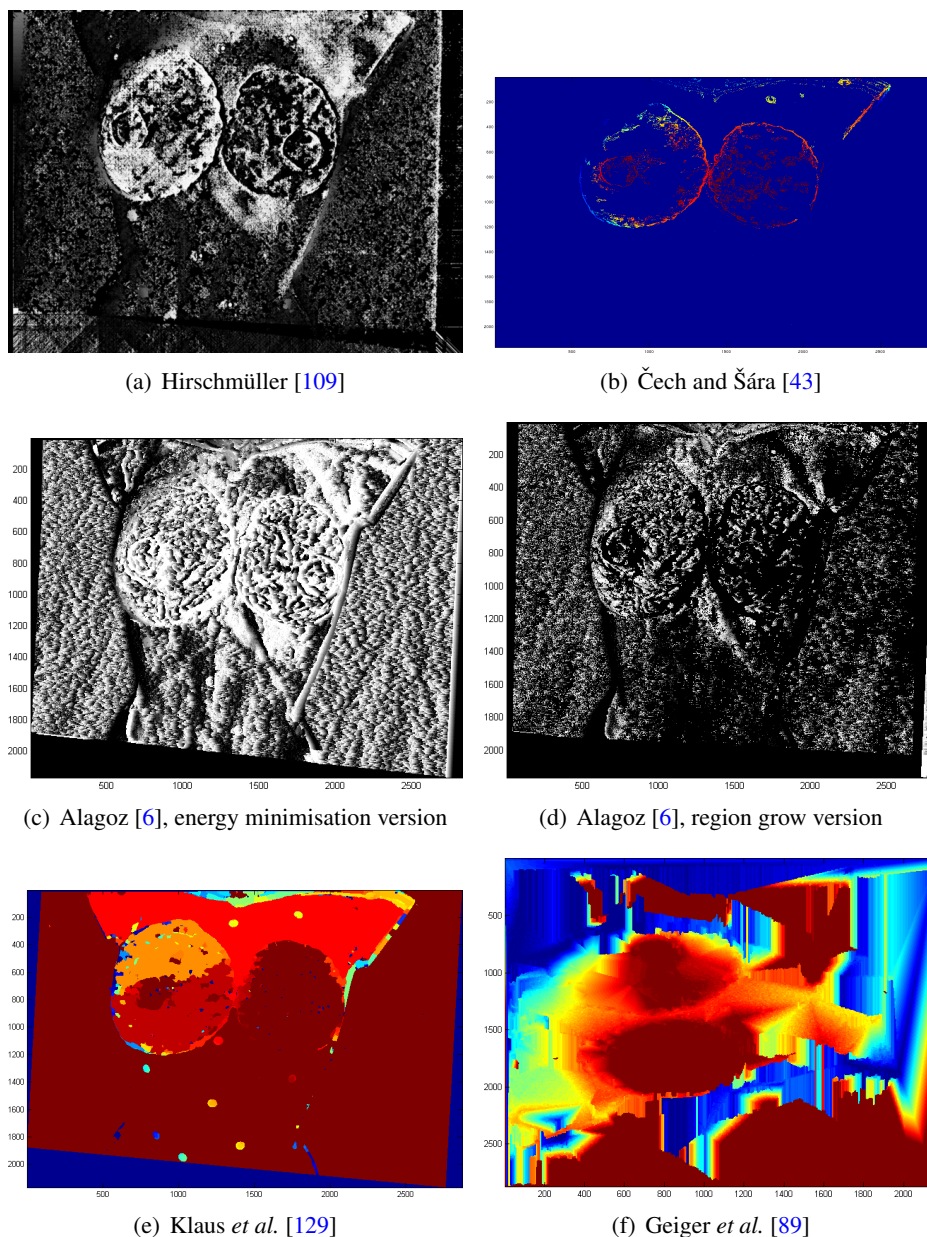


Figure 7.27: Results for stereo matching algorithms running over the Dummy model.

#### 7.9.4 Real Scenario

A small set of images was used from the database constructed as part of this project (see Chapter 4). One of the first immediate difficulties was for the physician to retain the "low baseline" constraint, which did not occur in most of the shots. Taking into account that it was specifically requested and reasoned, it is denoted that enforcing this constraint within medical staff might be a concern.

This section presents two cases, namely the best and the worst images chosen from the set of images. The images were captioned **Patient A** and **Patient B** (see Figure 7.28).

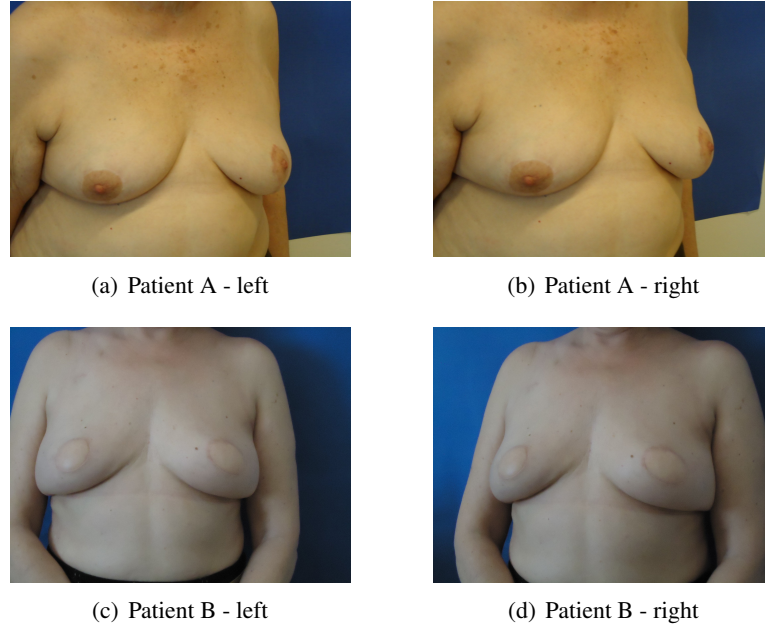


Figure 7.28: Original patient pairs, prior to processing.

The rectification results are summarised in Table 7.3 and can be seen in Figure 7.29. Both patients were tested for all the disparity algorithms that were tested in the previous section.

	Patient A	Patient B
<b>Features</b>	1377	329
<b>Inliers</b>	755	168
<b>Focal length</b>	4505.1	5185.1
<b>Error</b>	0.026	0.33

Table 7.3: Rectification values for the Patient photographs using the *SIFT* for features and Fusiello and Irsara [85] for rectification.

For **Patient A** (see Figure 7.30), in some cases the results appear to be better than the testing ones. On the other hand, for the remainder, the results are very poor. There is still no disparity map dense and correct enough for a decent reconstruction and some algorithms tend to continuously give close disparity values to both breasts and, in some cases, even to the chest.

The images from Patient B (see Figure 7.31) are more difficult to process than the rest due to the lack of features and because the images are a somewhat blurred. The results are very poor, and there are no distinguishable depth differences between the chest and the breasts. No disparity map is usable for an attempted reconstruction.

It is possible to conclude that, for this kind of images, it is not possible to use the most common state-of-the-art algorithms. It is necessary to search for a more suitable algorithm, one that does a wider, more globally oriented search that would better fit the ambiguity of these images.



(a) Patient A



(b) Patient B

Figure 7.29: Rectification results for the patient pairs.

Nonetheless, a global algorithm would increase complexity significantly, and therefore it might not be suitable for the purpose of this project.

Parameter variation was attempted for some algorithms in order to verify what changes would occur on the generated map. Those variations were done on the two most common parameters: window size and maximum disparity value. Other parameters were available the solutions provided by Klaus *et al.* [129] and Alagoz [6], but those proved to have little if no effect on the final result. With the variation of parameters, a large number of results were obtained; however, the quality of the results are similar to those previously presented, and for that reason, they were not included in this chapter. For details on the most relevant and consistent results in various runs, go to [219]

As previously suggested and somewhat predicted, the state-of-the-art algorithms could not successfully reason the correct correspondences for a too large portion of pairs, both virtual and real. According to the literature, that can be attributed to the fact that it is very difficult achieve a confident matching in many of the areas of the images used in this project. That difficulty can be verified with a simple zoom on similar areas in both images. The difficulty of reasoning is clear and even a human, with visual capabilities that overcome the actual computer possibilities in many areas, is incapable of reasoning about the correct correspondences.



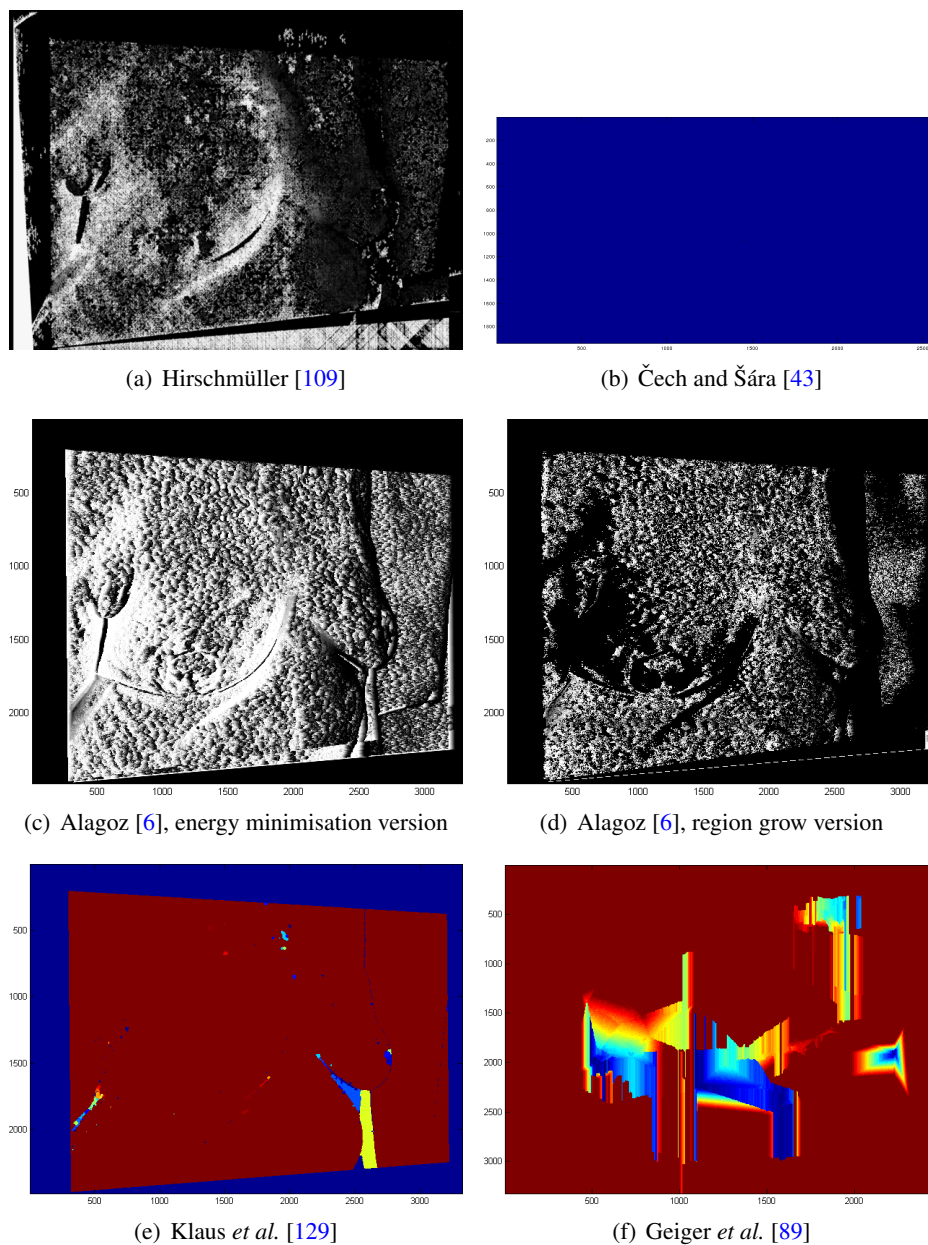


Figure 7.30: Results for stereo matching algorithms running over photographs of Patient A.

### 7.9.5 A Simplified Scenario

A simpler approach was used after the poor results obtained with the state-of-the-art stereo matching algorithms. Here, the set of test images are from a female phantom torso. In this particular test set, ground truth measurements were performed both manually and with equipment based on active stereoscopy techniques (see Figure 7.32).

This reference model was acquired by the Computer Aided Radiation Therapy group of the Laboratory of Biomedical Technologies at the Politecnico di Milano <sup>1</sup>, which developed, spe-

<sup>1</sup><http://www.tbmlab.polimi.it/index.html>

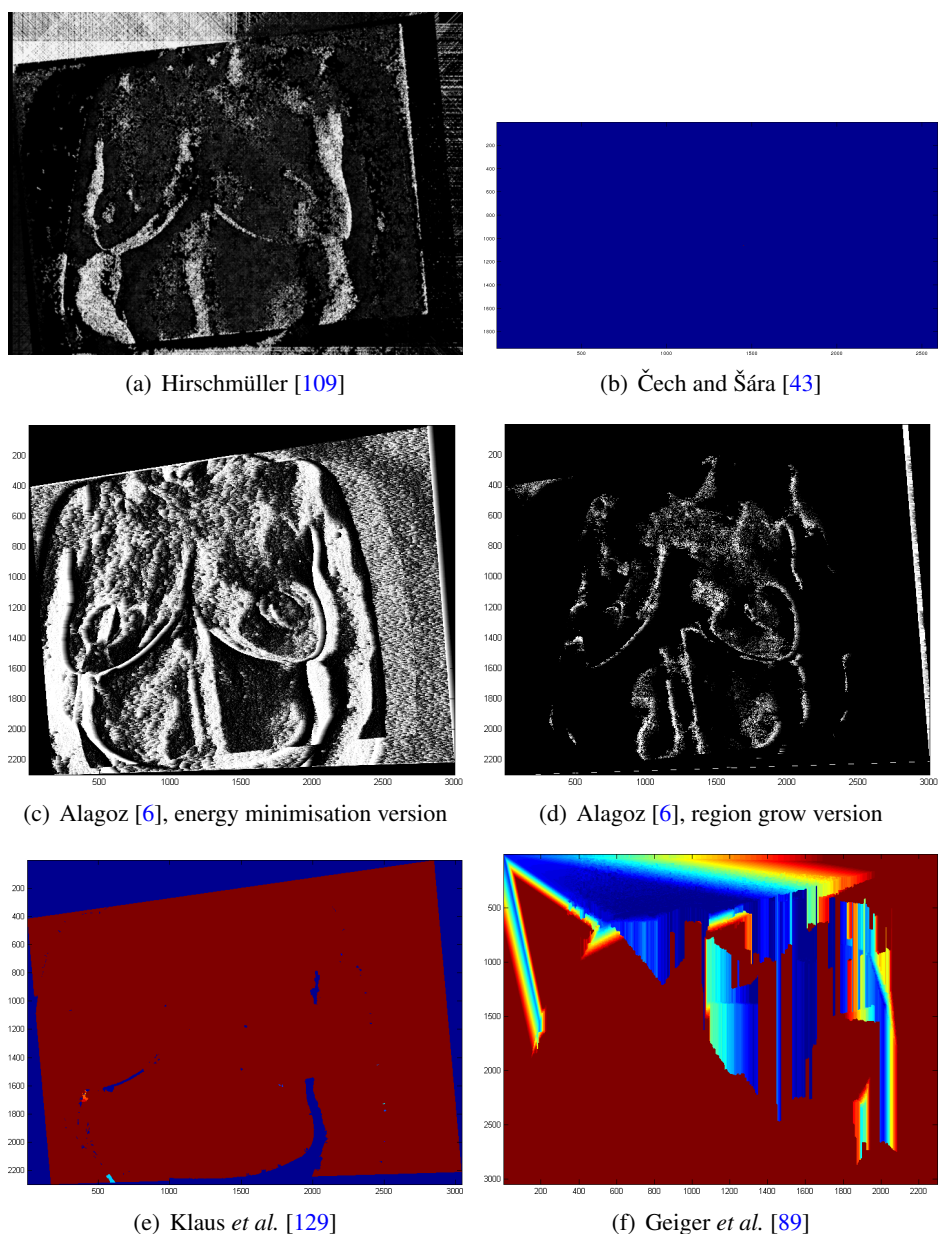


Figure 7.31: Results for stereo matching algorithms running over photographs of Patient B.

cific instrumentations and algorithms over the last years to create a tool to acquire and analyse patients' surface in order to quantitatively assess surgical outcome in breast plastic and reconstructive surgery [14]. The acquisition system is based in a programmable laser pattern static projector and an optical tracking system [188] (see Figure 7.33). The surface is obtained using a "row-wise" scanning projected onto the object that we want to acquire, and the scanning can be repeated as many times as desired to obtain a denser point cloud. The tracking system used to locate the laser spot projected on the surface is based on a stereo vision system.

In this methodology a simplified approach was implemented. Basically, in each point the depth is computed using only the difference in the **x-axis**, assuming a correct image rectification. Using



(a) Real photo.



(b) 3D surface model

Figure 7.32: Female phantom torso.



(a) Laser scanner and the female phantom torso.



(b) One of the cameras from the tracking system.

Figure 7.33: Acquisition system from Patete *et al.* [188]

a laser scanner, several light spots were projected onto the female phantom to make it easier to detect and match pixels. Several pairs of photographs were acquired, from two different points of view, using a single camera without any calibration procedure. The two views present some rotation and/or translation between them, assuming a low baseline between acquisitions.

#### 7.9.5.1 Identification and Matching of Points

Initially, the points are identified and correctly matched, using a semi-automatic approach, from the light spots projected onto the female phantom. The light spots are detected automatically, and then the users match the corresponding points manually (see Figure 7.34).

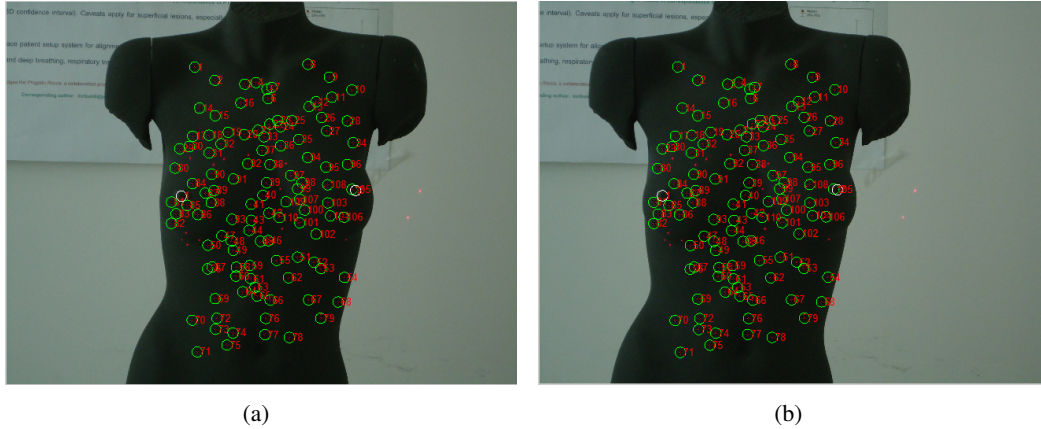


Figure 7.34: Identification of corresponding points from the female phantom torso. The nipples are also identified by the user (white circles).

### 7.9.5.2 Rectification of the Two Views

In this step, the points that have been correctly identified previously are used to rectify the images using epipolar geometry with the algorithm by Fusiello and Irsara [85] (see Figure 7.35).

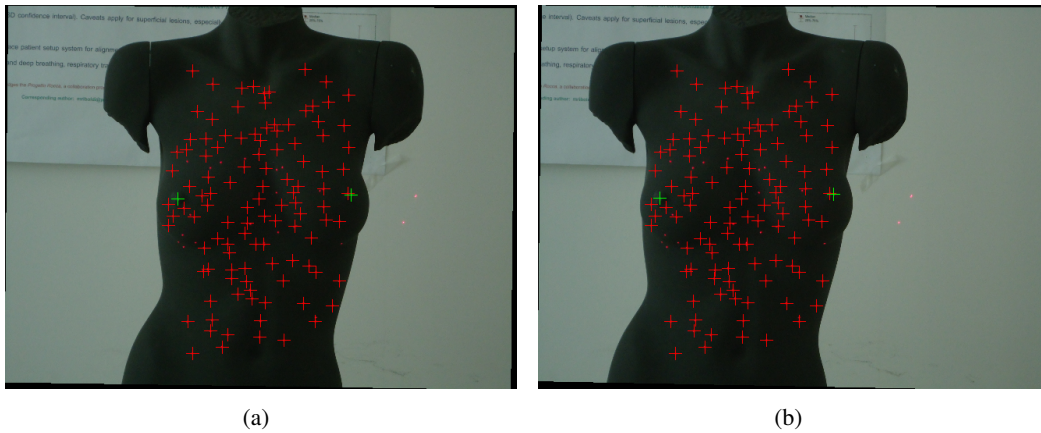


Figure 7.35: Rectified images from the female phantom torso.

The rectification algorithm used was the introduced by Fusiello and Irsara [85] because it achieved the best results, with the lowest error, as demonstrated in the previous Sections. However, it is also the slowest to run and sometimes the results may not be the best due to the random characteristics of the Non-linear Least Squares method, which initialises the focal length with a random value and then iteratively tries to improve that value. As stated before, various pairs of photographs were acquired and subject to the steps presented previously. The quality of rectification is measure by evaluating the difference, in pixels, between each pair of matched pixels and the corresponding epipolar line using (7.23). In this part, an average error of 0.24 pixels was obtained with a Standard Deviation of 0.017.



### 7.9.5.3 Depth Information

After the rectification step, all the corresponding points in the images present the same  $y$  coordinate, and now it is possible to compute the disparity between each point. From (7.12) it is possible to see that the depth information for each point is found using the focal distance, the baseline of the cameras and the disparity information. As we are working in an uncalibrated environment, the focal distance and the baseline information are unknown. For that reason, it is not possible to work with real metric values. On the other hand, it is possible to work with relative distances and with ratio values in both scenarios: real phantom and the computation uncalibrated approach. In this scenario, (7.12) is simplified to:

$$Z = \frac{1}{d} \quad (7.24)$$

### 7.9.5.4 Torso Surface Fitting

In this part, the matched points in the chest and abdomen zone are used to fit to a 3D surface. The generated models can be used to take measurements and to compare them with the reference model. Several models were tested from a simple plane to a surface with a higher level. The tested models are presented in Table 7.4, which shows the Mean Square Error (MSE) of the points used on the surface created:

Model No.	Model Equation	MSE
Model 1	$Z = a_0 \cdot x + a_1 \cdot y + a_2$	5.96E-10
Model 2	$Z = a_0 \cdot x^2 + a_1 \cdot y^2 + a_2 \cdot x + a_3 \cdot y + a_4$	2.27E-10
Model 3	$Z = a_0 \cdot x^2 + a_1 \cdot y^2 + a_2 \cdot x + a_3 \cdot y + a_4 \cdot x \cdot y + a_5$	1.72E-10
Model 4	$Z = a_0 \cdot x^3 + a_1 \cdot y^3 + a_2$	6.53E-10
Model 5	$Z = a_0 \cdot x^3 - a_1 \cdot y + a_2$	6.37E-10
Model 6	$Z = -a_0 \cdot x + a_1 \cdot y + a_2$	6.37E-10

Table 7.4: Mean square error of 3D fitting models.

These results were obtained considering all the image pairs. The MSE values obtained is an average of all the tests. As it is possible to observe from Table 7.4 *Model 3* obtained better results. The obtained surface is presented in Figure 7.36. The nipples were also modelled and are identified with red circles.

### 7.9.5.5 Results

A surface fitting model was generated for all image pairs. The nipples were also modelled using (7.24). From this information, relative distances of the nipples to the surface can be computed. The same procedure was conducted for the 3D reference model (see Figure 7.32(b)). The distances between each nipple and the chest are presented in Table 7.5.

As this work involves relative measurements, it is necessary to use ratio values to compare the model achieved in our approach with the 3D reference model. The ratio shown in Table 7.5

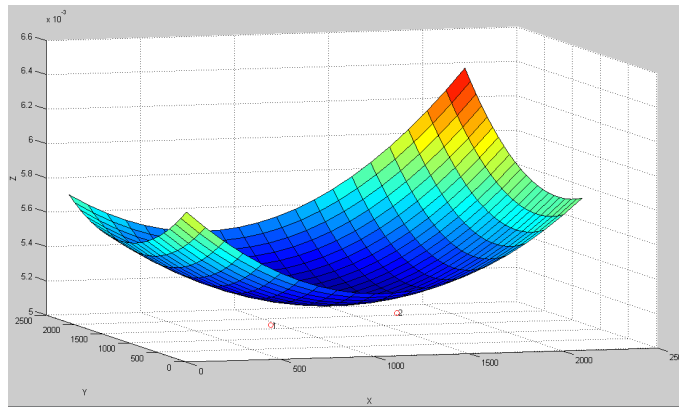


Figure 7.36: Surface fitting model.

Real Measurements	
Right Nipple	4.3cm
Left Nipple	3.7cm
Ratio	1.16

Table 7.5: Nipples measurements taken from the 3D reference model.

represents one of the features that can be extracted from the patient and used to improve the BCCT.core model. For the models generated with this approach, and making an average of the ratios obtained for image pairs, it is possible to obtain an average ratio of 1.12 and a standard deviation of 0.10.

The value obtained is not far from the real one, but there is some variation in the results, which means that this methodology is not stable. It was not possible to use this methodology with real patient; however, similar and difficult results are expected.

## 7.10 Conclusion

In this chapter, there was an attempted to create a conclusive set of tests for the 3D reconstruction based on stereo disparity search. After having defined the common mathematics associated, it was possible to establish some basis for the validity of reconstruction and the general problems that were posed.

It is therefore fair to conclude that the current state-of-the-art algorithms for stereo matching are not capable of reliably attributing a dense estimation of the disparity map from the type of images used. The estimated maps returned from these algorithms are generally very sparse. Therefore, visual inspection is enough to tell that the maps are generally wrong also, where areas of the chest sometimes have close disparity to those from the breast. This can be attributed mainly to the difficulties of reasoning to what pixel on the "target" image does one pixel on the "base" pixel belong or is correlated. Where repetitive texture is present, the difficulty is very high, and too many options are available for a correct match.

As such, it is required to search for more suitable algorithms that would make more globally oriented decisions. Those algorithms could be improved by a parametric model that would incorporate some constraints to the possible shape of the breast, reducing the complexity associated to the stereo matching process. Multi-view and/or video incorporation could allow for ambiguity reduction, using methods such as optical flow. The biggest adversary to a global solution is the increased complexity associated, and it can be a difficult matter accounting for the possible deployment on standard PCs.

For the evaluation alone it is not necessary to have a very high quality disparity map. Some measurements can be introduced into the classifier and it might be sufficient for achieving better results. Therefore, a sparse disparity map with just the right features would suffice. The right features would be those enough to add volume loss perception in some mathematical model. A semi-automated solution could improve the rectification process. By asking the user to enter some close-matched points, and then attempting to use the feature selection algorithms to search for features around those points, it could be possible to obtain better quality features in a reasonable number.

Further work is required and some upgrades are necessary, such as multi-view adaptations or plane fitting algorithms that would make it possible to adapt a sparser disparity map. The problem of maintaining the low baseline restriction should also be taken into consideration. There is a large set of methods on these subjects that were not addressed in this work due to time restrictions.

Since the goal of this thesis is to update the BCCT.core system, maintaining its simplicity of acquisition and affordable price, one solution could be using low-cost equipment which provide the required volumetric information based on active stereo vision.



## Chapter 8

# Depth-map Based Methodologies\*

As concluded in Chapter 7, the 3D model reconstruction from two uncalibrated views using the state-of-the-art algorithms failed to provide a proper reconstruction. The unsuccessful results were attributed to the high ambiguity that is found on the patient images, and to the lack of poor matching in image areas. These results require other solutions, always trying to maintain the system practical to use with a low cost.

Recently introduced at large scale in the market, essentially driven by computer gaming and home entertainment applications, RGB-D cameras are sensing systems that capture RGB images along with per-pixel depth information. RGB-D cameras are an emerging trend of technologies that provide high quality synchronised depth and colour data.

RGB-D cameras rely on either active stereo<sup>1 2</sup> or time-of-flight sensing<sup>3</sup> to generate depth information in a large number of pixels. While sensor systems with these capabilities have been custom-built for years, only now are they being packaged in form factors that make them attractive for research outside specialised computer vision groups. RGB-D cameras make it possible to capture reasonably accurate mid-resolution depth and appearance information at high (30 *fps*) data rates.

This Chapter starts by presenting the main characteristics of the RGB-D device used in this work (Kinect device). Very recent applications using Kinect are present, most of them on body scanning or health related research projects. An initial validation of Kinect depth data for the aesthetic outcome will be presented here. The great potential of this kind of devices is confronted with the relatively Low-Resolution (LR) of the generated disparity maps. For that reason, enhancing the resolution of the images could be an interesting way to improve the disparity information. This Chapter also shows a comparative study of state-of-the-art Super-Resolution (SR) algorithms applied to the Kinect depth and colour data. Since Kinect information from the two cameras (depth and colour) are not aligned, a joint calibration stage must be conducted prior to processing the information. This issue will be also discussed in this chapter.

---

\*Some portions of this Chapter appeared in [182, 183, 215]

<sup>1</sup><http://www.xbox.com/en-US/kinect>

<sup>2</sup>[http://www.asus.com/Multimedia/Motion\\_Sensor/Xtion\\_PRO/](http://www.asus.com/Multimedia/Motion_Sensor/Xtion_PRO/)

<sup>3</sup><http://www.softkinetic.com/Solutions/DepthSensecameras.aspx>

## 8.1 Depth-map Sensor

Microsoft's Kinect was adopted in this solution. It was originally designed to control games on the Xbox 360 without having to hold a device such as a controller. Released in early November 2010, the Kinect has demonstrated to effectively and accurately track human motion. Its characteristics led to its widespread popularity and adoption in areas such as robotics [105] or health-care [46].

In comparison to other devices, the Kinect has higher quality drivers, it is more stable with USB controllers and has a motor controlled remotely, which allows a more convenient positioning of the device. The main disadvantages of the Kinect are its size ( $30\text{cm} \times 7.6\text{cm} \times 6.4\text{cm}$ ) and the need for ACDC power supply.

### 8.1.1 Hardware

The Kinect is based on a sensor design developed by PrimeSense<sup>1</sup>. The Kinect consists of three optical components: a laser based near Infra-Red (IR) projector, an IR camera (PS1080 CMOS image sensor) and a colour camera. For the audio part there is a multiarray microphone, which consists of four sensors and is able to separate sound from different directions (see Figure 8.1). PrimeSense sensor is able to take some technical specifications of the Light Coding, which is the technology that allows the Kinect to construct 3D depth maps of a scene in real-time. Structured near-IR light is projected on a region of space and a standard CMOS image sensor is used to receive the reflected light.

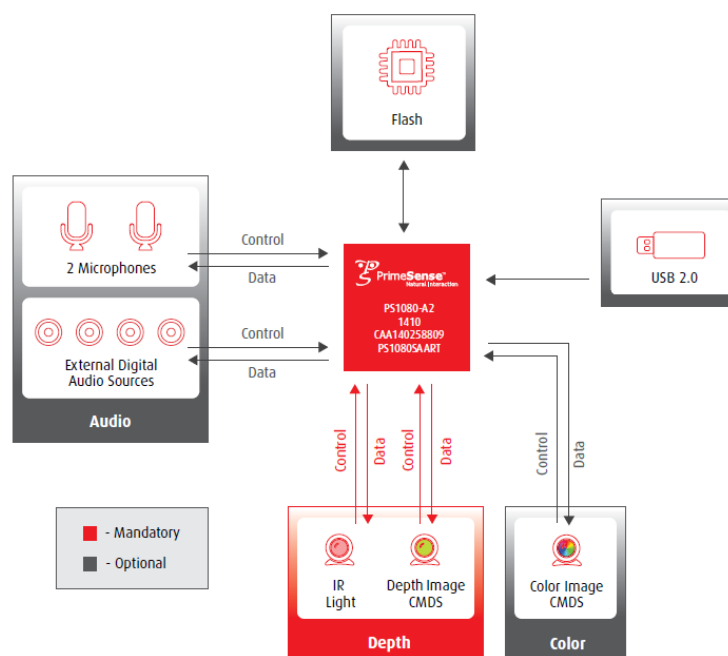


Figure 8.1: Prime sense design (From <sup>1</sup>).

<sup>1</sup><http://www.primesense.com/>

The projected speckle pattern repeats itself after 211 horizontal spots and 165 vertical spots and in every of these blocks there is a bright center point. The total pattern is composed of a  $3 \times 3$  repetition of the before mentioned spots, which results in  $633 \times 495$  spots. This pattern provides a rich source of easily extracted features. The variation of these features compared to the known pattern for a fixed distance provides a method for depth reconstruction. This represents a computationally less demanding solution compared to the more usual use of two cameras for stereo vision.

The depth image size from the IR camera has a maximum resolution of  $640 \times 480$ . At  $2m$  from the sensor it is able to resolve down to  $3mm$  for height and width and  $1cm$  for depth. It operates at a range between  $0.8m$  and  $3.5m$ . The resolution at  $0.8m$  is just over  $1.3mm/pixel$  <sup>1</sup>. Experimentation has shown that Kinect is only able to process depth data at a frame rate of  $30fps$ . The sensor also has an integrated RGB camera with a maximum resolution of  $1600 \times 1200$  to match the depth data [246].

### 8.1.2 Software

The Kinect itself actually does not calculate the depth, but returns a more abstract value for the host system to handle. There are several projects with freely available libraries and drivers that can be used to collect and process data from a Kinect sensor.

Even before the official release of Microsoft Kinect on November 4th, 2010, the Adafruit Industries <sup>2</sup> announced a contest to produce open-source drivers for the Kinect. In response to that call, Hector Martin released his code on November 10th, which marks the beginning of the OpenKinect community <sup>3</sup> that continued to evolve ever since. The OpenKinect community released the Kinect driver called libfreenect (released under Apache 2.0 or GPLv2 license) to connect and use the Kinect with a computer instead of its designated use with the Xbox.

In addition, a group of companies which included PrimeSense, responsible for developing the reference design on which the Kinect is based, launched OpenNI <sup>4</sup> as a not-for-profit organisation that aims to set an industry standard framework for the interoperability of natural interaction devices. With this framework they began to supply the OpenNI driver and the NITE (Natural Interaction Technology for End-user) Middleware software for scene perception from a compliant device to application-ready data. These libraries are available for applications written in C, C++, and C#. Because the Kinect is an OpenNI-compliant device, this software can be used to build applications for it. However, it has to be noted that the Kinect sensor is not the only device that uses the PrimeSense reference design.

Due to the success of the Kinect, Microsoft released a software development kit (SDK) beta for non-commercial purposes only to access Kinect's capabilities on a PC in June 2011. Recently,

---

<sup>1</sup><http://en.wikipedia.org/wiki/Kinect>

<sup>2</sup><http://adafruit.com/>

<sup>3</sup>[http://openkinect.org/wiki/Main\\_Page](http://openkinect.org/wiki/Main_Page)

<sup>4</sup><http://openni.org/>

as Craig Eisler announced <sup>1</sup>, Microsoft has launched not only a new Kinect hardware especially for Windows but also the Kinect for Windows commercial program <sup>2</sup>. While presenting some new improved features, such as the Kinect developer suite and near mode, the difference of this new Kinect device with the existing equipment can generally be considered negligible or nonexistent.

Both libfreenect and OpenNI projects work on Windows, Linux (Ubuntu) and Mac OS X, while Microsoft SDK is for windows only. All allow access to colour information (8-bit RGB) and depth images (11-bit) in  $640 \times 480$  with resolution at 30fps. The projects are not compatible and they cannot be used simultaneously. Differences between the libraries are motor control, access/use of the Kinect's image and depth registration, integration with middleware for higher level, and calibration requirements.

It is important to highlight that although a  $640 \times 480$  resolution gives a theoretical upper limit of  $640 \times 480 = 307200$  points (11-bit values) in a point cloud, in practice, a scene with good capturing conditions will result in a cloud of a maximum of 265000 points. The main reason for that discrepancy should be due to how the depth image will map onto the colour image, which is captured with a wider field of view. The quality of colour images captured by the Kinect is about as good as a decent webcam, and Bayer noise is noticeable.

For the purposes of this thesis, the libfreenect was selected because it can be used in multiple operating systems and is simpler than OpenNI

## 8.2 The Kinect Sensor as a Tool

Novel approaches to human or scene scanning have been addressed using the Kinect sensor device. Using the colour images combined with the depth maps rendered by the device, Zollhofer *et al.* [269] create a high-quality personalised avatar that is designed for home use. The proposed algorithm combines robust non-rigid registration and fitting of a morphable face model to generate a high-quality reconstruction of the facial geometry and texture within seconds (see Figure 8.2).

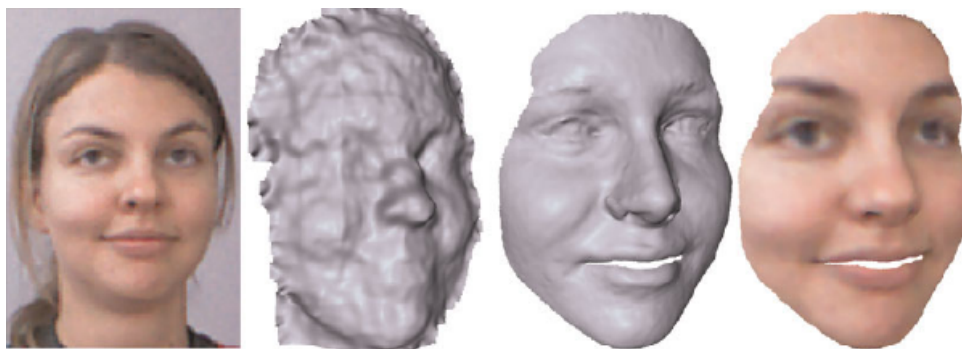


Figure 8.2: Avatar reconstruction using the algorithm by Zollhofer *et al.* [269].

<sup>1</sup><http://blogs.msdn.com/b/kinectforwindows/archive/2012/01/09/kinect-for-windows-commercial-program-announced.aspx>

<sup>2</sup><http://www.microsoft.com/en-us/kinectforwindows/>



In [55], the authors describe a method to scan a human body using a single Kinect by aligning depth and colour scans provided by the device. They achieve a full 360 scan by moving the device freely around the object doing a full rotation in about 30 seconds. To process the data, they apply a 3D SR algorithm followed by a loop-closing method specifically tailored for the Kinect. Finally, they execute non-rigid registration to correct residual errors of movements. The result is a complete model with smooth surfaces and very good detailed structures (see Figure 8.3). By using the chosen motion device, the authors add that the main difficulties encountered are the LR and low accuracy of the depth sensor, plus its random noise.

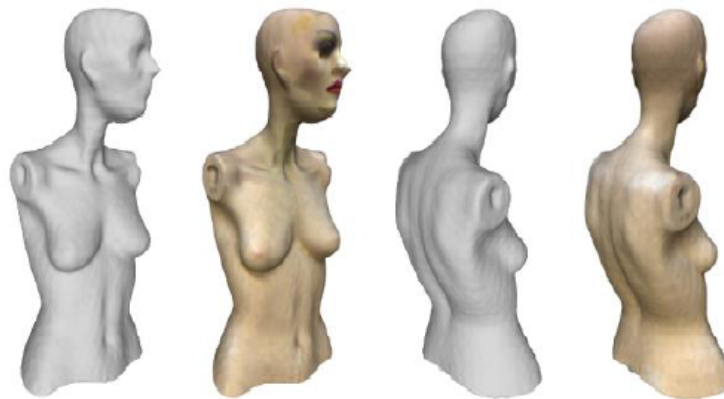


Figure 8.3: Results example of mannequin using the algorithm by Cui and Stricker [55].

Another interesting application was conducted by Microsoft's research group Kinect Fusion [114]. By using one Kinect and with hand-held movement, it is possible to reconstruct a high-quality 3D model of the scene. The system captures the scene and for every viewpoint available it renders enhancements onto the model.

Weiss *et al.* [254] present a method for human shape reconstruction combining LR image silhouettes with coarse range data to estimate a parametric model of a human body. 3D shape estimation is obtained by combining multiple views of a person moving in front of the sensor. To deal with varying body pose, authors use a SCAPE body model [9] which factors 3D body shape and pose variations, enabling the estimation of a single consistent shape while allowing pose to vary (see Figure 8.4).

In the health-care area, Chang *et al.* [46] developed a rehabilitation system to assist therapists. The proposed system for physical rehabilitation, called the Kinerehab (a portmanteau of the words "Kinetic" and "rehabilitation"), enabling users to control and interact with a specific application. This application automatically recognises human gestures to determine whether a user performed the exercises correctly in physical rehabilitation. Using Kinect means that the users do not need be bothered with body sensors and the system can save the users from wearing sensors that can be intrusive.

In scene scanning and modelling, Henry *et al.* [105] built 3D dense maps of indoor environments to be used in robotics (see Figure 8.5). They use PrimeSense's camera that is much equiva-

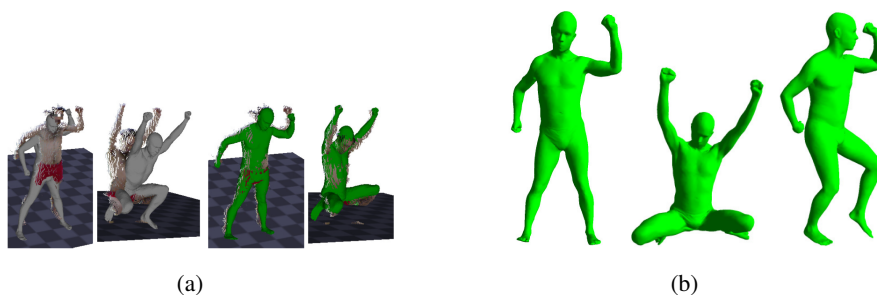


Figure 8.4: Widely varying poses from Weiss *et al.* [254] work. a) Initialisation; b) Final resulting model.

lent to the one installed in Microsoft Kinect, and employ a novel joint optimisation algorithm and combine visual features plus shape-based alignment.



Figure 8.5: Surfel update demonstration by Henry *et al.* [105]. (a) Initial frame; b) Complete updated surfel with 95 RGB-D images.

Tong *et al.* [234] present a system to scan 3D full human body shapes using multiple Kinects. Each Kinect scans a different part of the body and therefore they can be put close to the body to obtain higher quality data. The authors faced two main problems with this application: the quality of the data overlapped between two regions and body movements during the capturing process, which required non-rigid registration of the captured data (see Figure 8.6).

Wang *et al.* [250] propose a body modelling system with a single fixed Kinect. An initial model is obtained with a registration process from four key poses out of the entire depth video sequence, which are front (reference pose), back, and two profiles. After registering four key frames, an articulated part-based cylindrical body model is used to process the rough and noisy registered points in the body to obtain the final 3D model (see Figure 8.7).

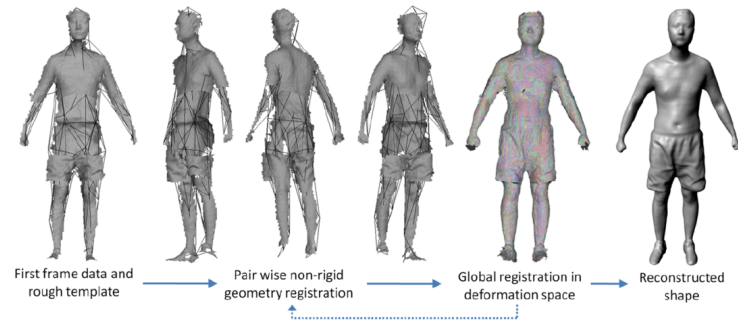


Figure 8.6: Overview of the reconstruction algorithm using the application by Tong *et al.* [234].

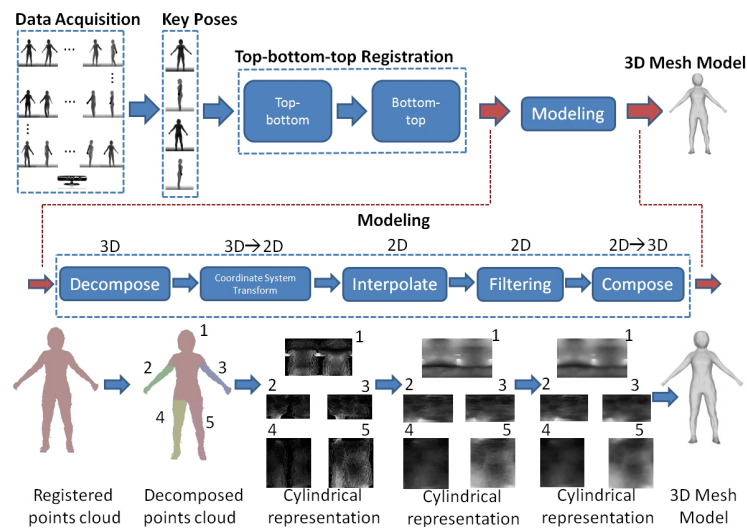


Figure 8.7: General pipeline using the system by Wang *et al.* [250].

## 8.3 Validation of the Kinect as a Tool to Assess Aesthetic Outcomes

By defining the use of the Kinect device for 3D data acquisition, the knowledge of its operations conditions and technical restrictions is peremptory. The device is capable of operating under any interior lighting conditions, and is a compact and light equipment. The device renders depth maps with  $640 \times 480$  pixels, but it is also necessary to know its range accuracy and depth precision.

### 8.3.1 Transformation of Raw to Metric Data

The information of each pixel from the depth image, obtained with the Kinect, represents specific disparity information, but the conversion to metric distances is not proportional. Transformation equations must be used to linearise the raw data generated by the Kinect, which is represented on  $2^{11}$  levels. For this, there are few different equations to transform Kinect data. From the literature

three different equations were found to transform the Kinect disparity information:

$$d_m = \left( \frac{d_r}{2048} \right)^3 \times 9216 \quad ^1 \quad (8.1)$$

$$d_m = K_a + \tan \left( \frac{d_r}{K_2} + K_1 \right) \quad ^2 \quad (8.2)$$

with  $K_1 = 1.1863$ ,  $K_2 = 2842.5$  and  $K_a = 0.1236$

$$d_m = \frac{1}{d_r \times K_1 + K_2} \quad ^3 \quad (8.3)$$

with  $K_1 = -0.0030711016$  and  $K_2 = 3.3309495161$

where  $d_r$  is the raw data from the Kinect and  $d_m$  represents the transformed data.

To compare the equations, four objects were placed at a different distance from the Kinect device, as shown in Figure 8.8.



Figure 8.8: Kinect disparity map from the test scene. The object on the right (object#4) is the closest to the camera, and the object on the left (object#1) the farthest.

The results achieved in this study are shown in Table 8.1.

Then, ratios taking the first object as a reference were computed to validate the results (see Table 8.1). The MSE between the obtained ratios and the real ratio were computed to compare the transformation functions.

From Table 8.2 it can be seen that all of the different transformation functions present good results, in contrast to the values obtained using raw data. However, when considering the computed MSE, the values obtained with the first transformation function (8.1) present better results.

<sup>1</sup>[http://groups.google.com/group/openkinect/browse\\_thread/thread/31351846fd33c78/e98a94ac605b9f21?lnk=gst&q=stephane&pli=1](http://groups.google.com/group/openkinect/browse_thread/thread/31351846fd33c78/e98a94ac605b9f21?lnk=gst&q=stephane&pli=1)

<sup>2</sup>[http://groups.google.com/group/openkinect/browse\\_thread/thread/31351846fd33c78/e98a94ac605b9f21?lnk=gst&q=stephane&pli=1](http://groups.google.com/group/openkinect/browse_thread/thread/31351846fd33c78/e98a94ac605b9f21?lnk=gst&q=stephane&pli=1)

<sup>3</sup><http://nicolas.burrus.name/index.php/Research/KinectCalibration>

Object#	Distance (cm)	Pixels values			
		Raw data	Function#1	Function#2	Function#3
1	0	714	390.52	0.92	0.88
2	5	694	358.62	0.88	0.83
3	10	673	327.04	0.83	0.79
4	15	650	294.64	0.79	0.75

Table 8.1: Comparison of different disparity transformation functions. The object on the left was used as reference.

Objects Ratio#	Distance Ratio	Raw Data	Function ratios		
			Function#1	Function#2	Function#3
#4/ #3	1.5	1.56	1.51	1.48	1.48
#4/ #2	3	3.20	3.01	2.88	2.88
#3/ #2	2	2.05	1.99	1.95	1.95
MSE		0.054	0.004	0.033	0.033

Table 8.2: Validation of transformation functions.

### 8.3.2 Kinect Sensitivity Assessment

An experimental procedure was conducted in order to retrieve the closest capture point and the depth curve according to the distance to the device. *Lego* blocks were used and set at known positions as object surfaces. Through various measurements, it was verified that the closest point of capture was around 50cm from the sensor. Blocks were put 5cm apart, and from the 50cm to the 145cm mark (see Figure 8.9(a)).

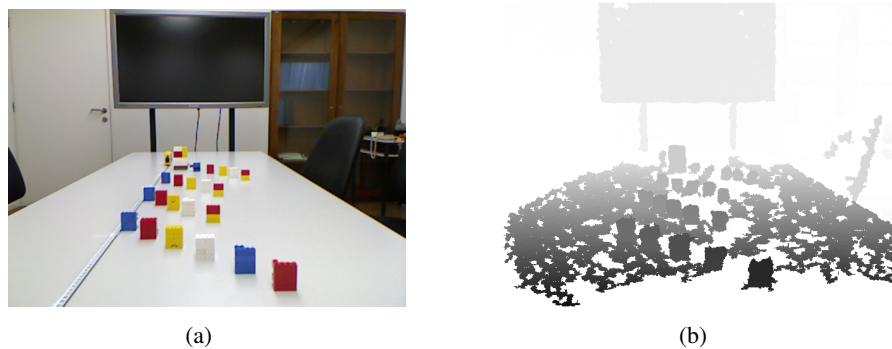


Figure 8.9: Kinect capture setup. (a) RGB image; b) Depth image.

The depth of each block was calculated as the mean value gathered from a window, defined by user input, for every surface block. The depth values are obtained using the transformation function (8.1) presented in Section 8.3.1. From this point it was possible to construct an estimative depth curve for the distances tested, as well as to calculate the mean resolution sensitivity step of distance per depth value. Figure 8.10 provides the depth information between 50cm and 145cm,

for raw data and the depth information after applied the transformation function.

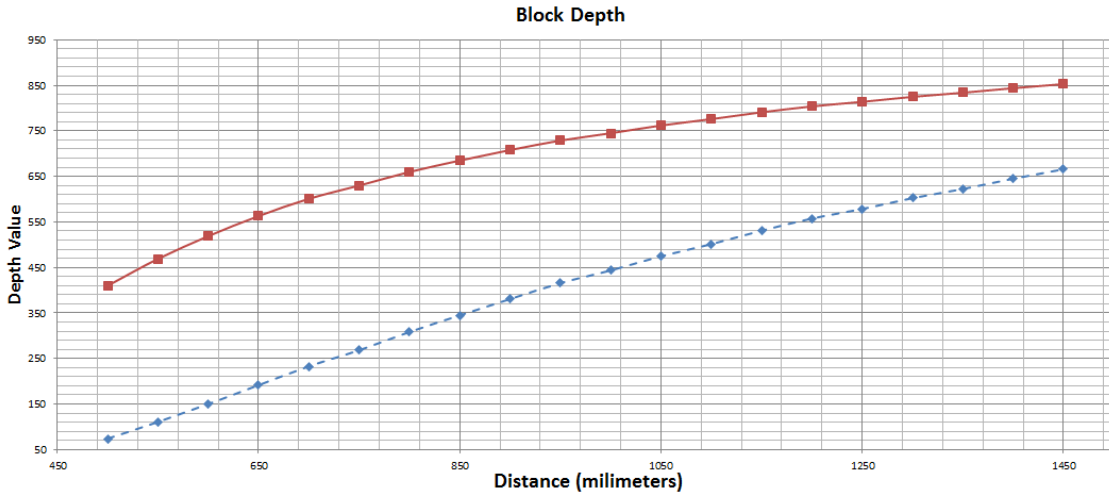


Figure 8.10: Kinect depth curve vs distance. (Raw data - solid red curve; Transformed data - dashed blue curve.)

With the extreme values registered, it is possible to achieve a mean step of  $1.60mm$  per depth value.

$$sensitivity_t = \frac{1450 - 500}{665.89 - 73.94} \approx 1.60mm \quad (8.4)$$

As verified in Figure 8.10, the Kinect renders higher depth resolution the closer the object is to the device, and also after the transformation function is applied. In this scenario the curve is almost linear; however, for higher distance values the sensitivity is lower. Therefore, the ideal set presented in ( 8.4) is not constant, where in a close region a given depth range corresponds to a small range of distances, while the same range at a farther distance corresponds to a smaller range of distance values. Likewise, the closer the object is the better will be for the 3D model reconstruction. Anticipating the scenario for our problem, where acquisitions are made at  $\approx 80cm$  to the patient, the sensitivity is:

$$sensitivity_{80cm} = \frac{850 - 750}{344.85 - 268.27} \approx 1.31mm \quad (8.5)$$

The value obtained with (8.5) is extremely useful because it allows us to understand the minimum differences detected by the Kinect device.

In order to study this phenomenon and to understand the possible implications, a few more experimental setups were explored. In the first, 10 blocks were positioned in the front (starting at  $50cm$ ) and back (starting at  $150.9cm$ )  $0.3cm$  apart. As observed in Figure 8.11, the closer blocks follow a linear tendency with short deviation (see Figure 8.11(a)) comparatively to the more distant

blocks (see Figure 8.11(b)). This confirms that in the back region the linear ideal step is not the actual distance per depth, and the same depth value can correspond to slightly different distance values. Even more, the accuracy for acquisitions at higher distances is very low as shown in Figure 8.11(b). Therefore, as previously stated, the device did not render each depth value per  $1.60mm$  step, while this value may be slightly lower at short distances, but equal or higher at higher distances. In this particular case, a  $1.18mm$  slope was obtained in the first experiment and  $1.47mm$  in the second.

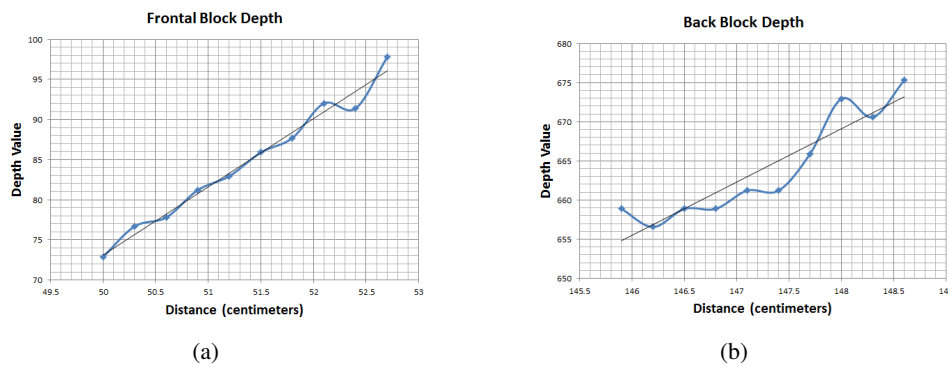


Figure 8.11: Kinect sensitivity analysis at extreme distances. a) Frontal blocks; b) Rear blocks.

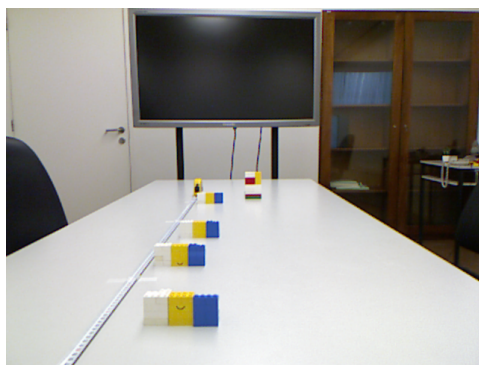
Another experiment was conducted, but with the blocks closer to each other (no parallel distance) as depicted in Figure 8.12(a).

The experiment was re-conducted with the blocks  $0.2cm$  and  $0.6cm$  apart from each other to verify variations to these data. The corresponding depth curves obtained are represented in Figure 8.13.

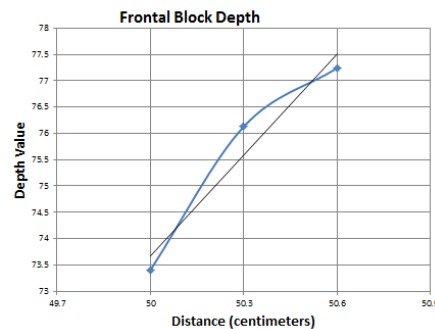
From these results, it is possible to conclude that the device is more capable of detecting depths with adjacent objects or within a larger one, than with isolated object. This is a highly significant aspect to this work since the breast shape and the women themselves are one large continuous surface. Also, it is possible to confirm that the device fails to provide the same depth at the same distance in different captures. This may be due to imprecisions of the sensors, and to the high ability to detect small variations in close distances, or on the contrary, due to the low performance with higher distances. The presence of a pair of objects of the same value at a higher distance is confirmed in Figure 8.13(e), with only a  $0.2cm$  distance separating the blocks. However, there is a high (+3) depth value in the next step distance. Whether this happens due to sensor limitations or acquisition conditions is not clear.

In conclusion, using a non-linear model to estimate distance should improve the consistency, but that would be too complex. Besides, the use of such model could prove inaccurate due to slight differences in devices.

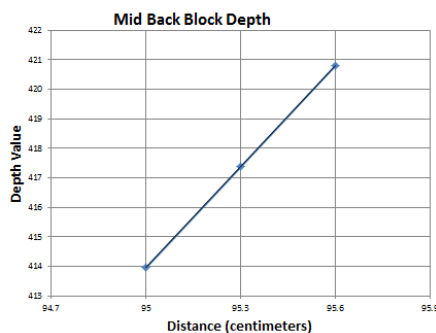




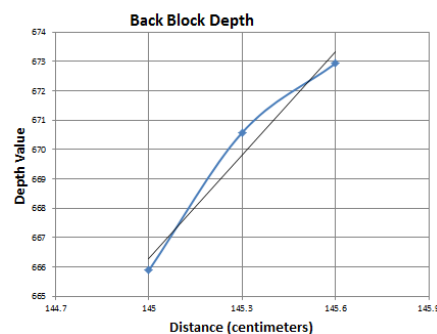
(a)



(b)



(c)



(d)

Figure 8.12: Detailed Kinect sensitivity analysis. a) RGB Image; b) Frontal blocks; c) Middle blocks; d) Rear blocks.

### 8.3.3 Kinect Measurements on a Female Phantom Torso

As in Section 7.9.5, the Kinect was applied to the female phantom torso. The transformation equation gives us metric depth values, as previously presented; however,  $x$  and  $y$  information cannot be used because it does not correspond to real values. For that reason, it is not possible to compare the results of the Kinect with other reference models, such as, the ones obtained with the laser scanner. However, some results were obtained from the data acquired. The phantom was acquired using the Kinect device in three different occasions. In these three acquisitions, there are some differences, such as: some rotation of the phantom, translation of the Kinect device and also different distance between Kinect and the object (see Figure 8.14).

To evaluate the quality of the depth map, this depth information was compared to measurements taken from 3D model generated with the laser scanner, as presented in Section 7.9.5.5 (see Table 7.5). Those measurements represent the height of each breast by measuring the distance between the nipple and the chest. In the depth map, the position of the nipples was defined semi-automatically based on depth-intensity information in an area delimited manually. As in Section 7.9.5.5, a ratio of measurements was used to be compared with the reference. The ratio values obtained from the depth map are presented in Table 8.3 – without compensation of rotation



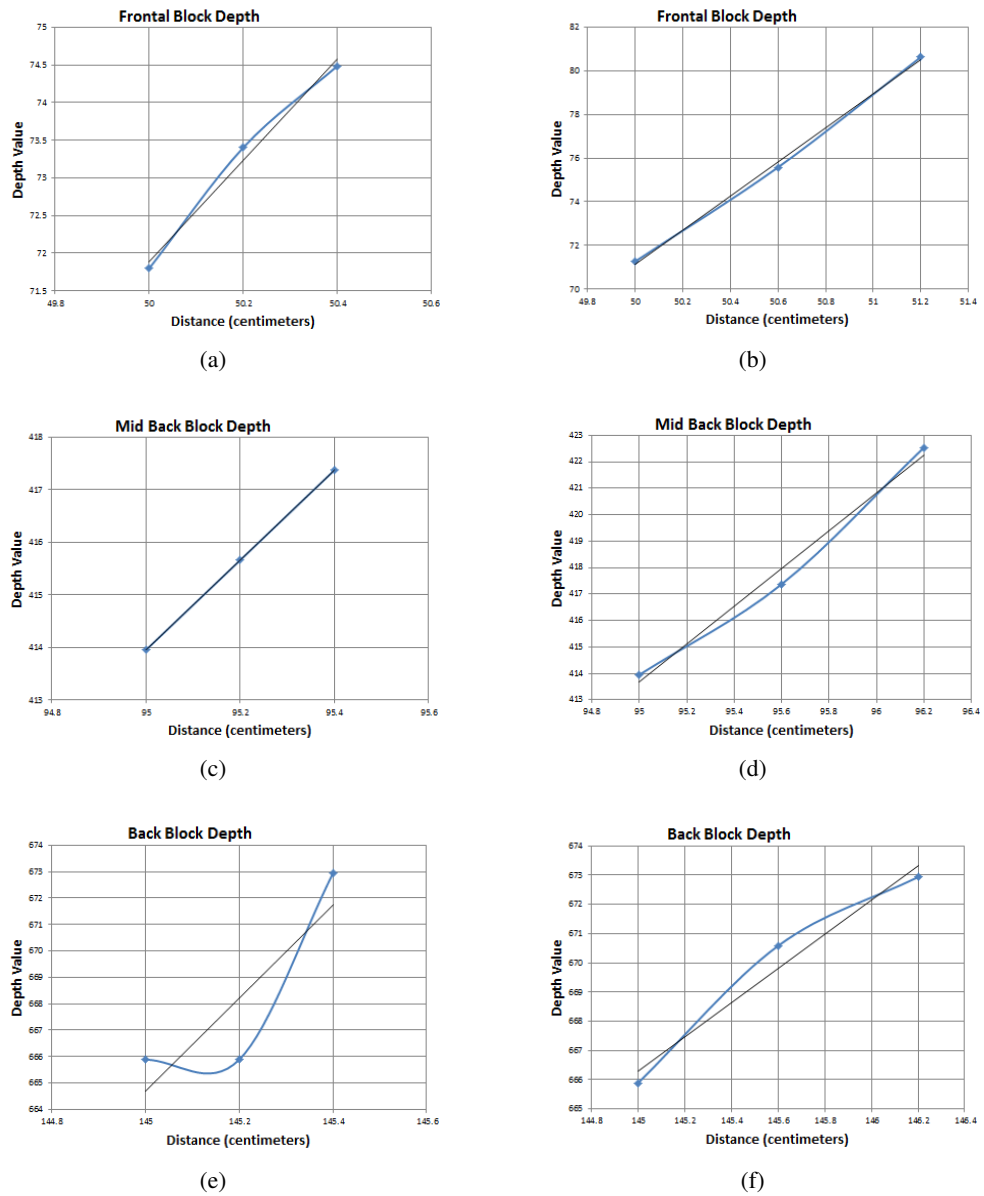


Figure 8.13: Detailed Kinect sensitivity analysis of blocks spaced 0.2cm and 0.6cm. a) Frontal blocks (0.2cm); b) Frontal blocks (0.6cm); c) Middle blocks (0.2cm); d) Middle blocks (0.6cm); e) Rear blocks (0.2cm); f) Real blocks (0.6cm).

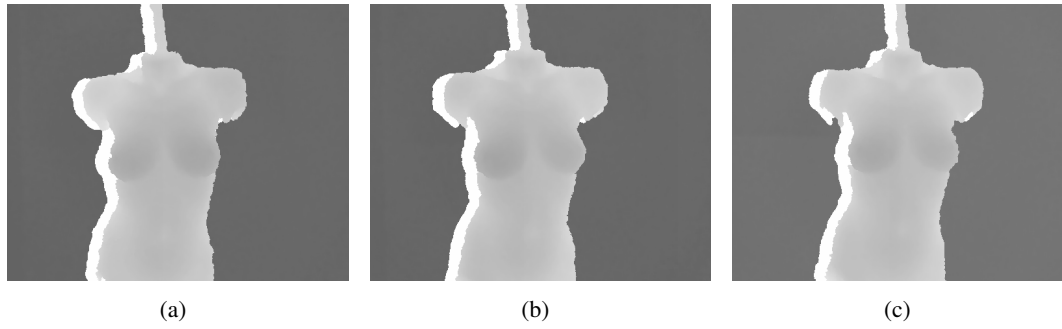


Figure 8.14: Different acquisition of phantom with Kinect.

(there are several acquisitions for each period):

Acquisition Period	No Compensation Rotation		Compensation Rotation	
	$\mu$	$\sigma$	$\mu$	$\sigma$
1	1.07	0.01	1.15	0.06
2	2.04	0.06	1.11	0.05
3	1.64	0.02	1.15	0.10

Table 8.3: Ratio values obtained with Kinect using the female phantom torso.

By comparing the ratio values obtained with Kinect device and the value obtained with the reference model (see Table 7.5), it is possible to see that there are great discrepancies between values. In the first acquisition, the phantom does not present too much rotation and, therefore, the ratio is similar to the measurement performed using the reference model. The other results are far from the reference because there is some rotation and this fact has to be compensated. For that reason, rotation compensation was implemented based on value points extracted from the stomach, assuming that that region presents a similar shape on both sides separated by a vertical line that crosses the middle of the breasts. By looking again at Table 8.3, it is possible to compare the results with and without compensation and comparing with the values obtained with the reference 3D model. As it is easy to observe, the approach using compensation of rotation presents very satisfactory results.

## 8.4 Super-Resolution Application

As shown in Section 8.2 and Section 8.3, the use of disparity maps rendered by the Kinect may be very promising for the BCCT.core model, with the inclusion of 3D information. Furthermore, the possibility of using such apparatus is well received by clinicians in favour of too expensive solutions for 3D imaging. Still, this acquisition system may pose difficulties due to its own limitations. Since the system would operate on disparity map images the measurements, and consequently the obtained ratios, are directly related to the pixels on which they were performed. The disparity maps obtained have relatively LR, which may obstruct the construction of reliable models.

Therefore the autonomous retrieval of independent breast features suffers from that lack of HR images. For that reason, the enhancement of the captured images may be essential before their use in subsequent steps in the processing chain.

SR is the process which employs techniques to enhance the resolution of a single or a set of images [64, 76, 111, 186, 187, 191]. The goal of SR is to apply signal processing techniques to increase the spatial resolution and to recover details that are present in LR images in an aliased form (see Figure 8.15). Due to physical constraints, imaging systems hardware have limitations resulting in a spatial resolution that would not be ideal. Moreover, the costs of HR imaging system are too high, which makes it impractical for common use. SR is often seen as an alternative to the usage of such systems, representing a processing stage after image [111].

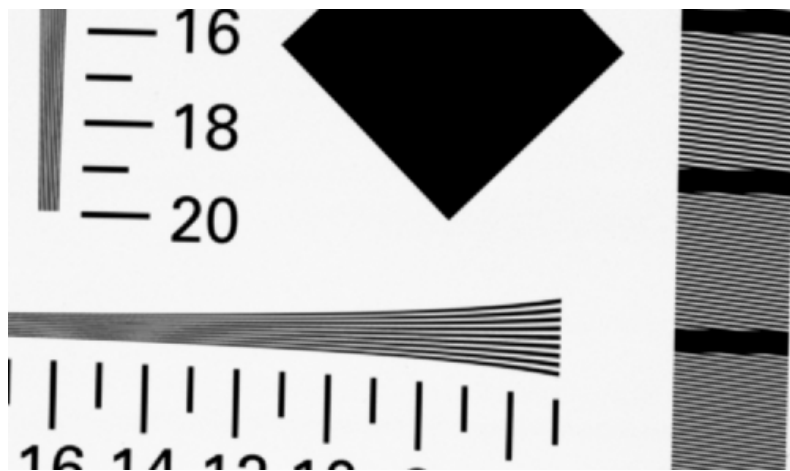


Figure 8.15: Aliasing in high frequency zones in the ISO 12233 chart [224].

This segment of signal processing has been studied for several decades, gathering more and more attention with the increase in computational capability verified by today's standards. Authors usually refer to SR as a 3-step process: (1) image registration, (2) image interpolation and (3) image regularisation. Some authors call it a 2-step process wherein they consider that image interpolation and image regularisation to be a single normalisation step.

Image restoration is oftentimes associated with SR and to recover a degraded image. Most imaging capture devices render LR images, and due to technology limits, these images present blurring, noise and aliasing effects. Therefore, part of the enhancement of the captured images is rendered by eliminating artifacts. For that matter, most SR techniques use image restoration for deblurring and denoising as part of their processing chain.

The registration step corresponds to the motion estimation that is verified in a set of images. This assumes that there are sub-pixel misalignments, where the aliased information resides. If the shifts are at integer levels, the images would contain the same information, not presenting any gain in detail to the HR image. Also, this step is crucial for the entire process chain. If this step fails to accomplish a good registration, the entire process is rendered void and a poor, and often very bad, output image is produced. Due to image shifts and to the fact that their occurrence is not

predictable, the interpolation step is responsible for creating a uniformly spaced HR grid through non-uniform interpolation. Image restoration is then performed on this resulting HR image to remove blur and noise artifacts, resulting in a final HR output image.

SR image reconstruction is typically an ill-posed inverse problem and is generally modelled as [64]:

$$y = D \times H \times x + v \quad (8.6)$$

where  $y$  is the low resolution image and  $x$  is the unknown high resolution image.  $D$  and  $H$  are degrading operators regarding the downgrading blur and noise present in the images, and  $v$  is additive noise.

The SR methods can be roughly classified into multi-frame, single-frame or example-based variants, either applying algorithms in the frequency or spatial domain. Multi-image methods try to take advantage from the aliased high frequency information that is split across multiple images, while example-based methods consider that it is possible to reconstruct HR information from existing database patches. Single-image methods employ a combination of the other two approaches. They assume that patches can recur inside an image and can be found in the same scale (multi-image) or across different scales (example-based).

Frequency domain algorithms take advantage of redundant information present in the LR images by the shifting and aliasing properties of the Fourier transform, assuming that aliasing has occurred or attempted to extrapolate information to retrieve values within intervals. Spatial domain algorithms usually employ regularisation based approaches which enable a priori knowledge inclusion, non-uniform interpolation to have the advantage of low computational cost, Projection Onto Convex Sets (POCS) [187] usually designed for simplicity, Iterative Back Projection (IBP) [111, 191] for straightforward implementations, and Maximum A Posteriori (MAP) estimation. The stochastic and example-based approaches also include a priori knowledge or may use the recurrence of patches within images or even the same image.

#### 8.4.1 Super-Resolution Applied to Depth Maps

Usually, SR methods were not designed or tested on depth images. But more recently, presumably due to the wider spread of sensors capable of providing depth images, there have been projects using SR and more research on techniques for the specific use of SR in this type of images. Li *et al.* [146] proposed a multi-frame regularisation-based method with new regularisation items to yield better results in images with sharp edges and smooth flat regions. Although the performance of the proposed method led to "better results with noise suppression and detail preservation", experiments were not conducted on depth maps. Moreover, new techniques have been proposed for range-enhanced oriented methods, such as [257], reportedly registering enhances of spatial resolution up to 100 times, and in [88] using both depth maps and colour images to increase resolution. This proved to be computationally efficient when compared to probabilistic approaches. In [257]

(see Figure 8.16), the authors propose a method to enhance the resolution by iteratively refining the LR image using HR images as reference. They up-sample the LR input image to the high resolution camera image and process an iteratively refinement over the first image. The refinement module iteratively applies bilateral filtering to each slice of 3D volume of depth probability referred as cost volume. The HR image is obtained by selecting the depth hypothesis with the minimal cost followed by sub-pixel estimation.

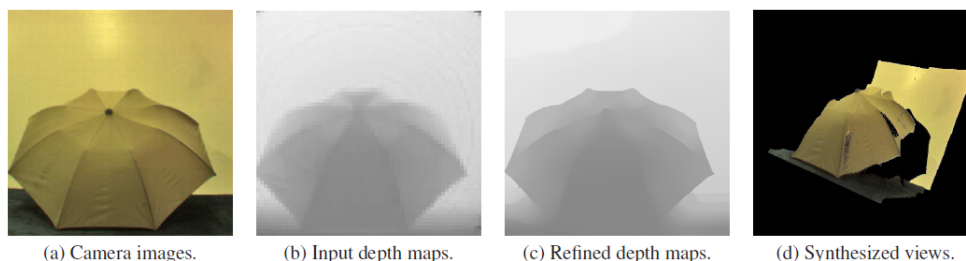


Figure 8.16: Experimental results using the methodology by Yang *et al.* [257].

The approach presented in [88] uses an original interpolation technique exploiting side information from a single colour camera image to increase resolution. The proposed method is performed by first pre-processing the colour image through a graph-based segmentation algorithm to identify main surfaces. Then, a projection of points of the LR depth map is made on the segmented colour image to create a depth grid. The points within each grid are then the initial values of the LR depth map. The value of each pixel in the HR image is then estimated by interpolation of the grid samples belonging to the same regions (see Figure 8.17).

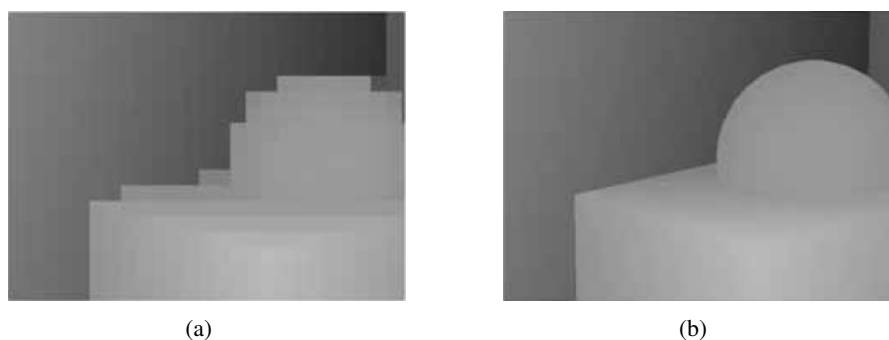


Figure 8.17: Experimental results using the methodology by Garro *et al.* [88]. (a) Input image; b) Output image.

SR presents many advantages with several use cases presenting successful implementations. In theory, multi-image (MI) based methods are able to achieve a substantial increase in detail, but require strict conditions to perform adequately. Most single-image (SI) based methods employ stochastic approaches and the inclusion of a priori knowledge. While achieving very good reconstructions, they also take high computational load and may see its performance restrained

to the quality of training dictionaries used. The approaches that use redundant information also show good reconstructions without needing a priori knowledge, but have serious limitations if no redundant information exists. Moreover, the cases of SR application on depth maps are very encouraging, and it is hoped that good outcomes are achieved with the type of images in this project.

## 8.4.2 Performance Tests

Several state-of-the-art SR algorithms were tested in order to understand the impact of SR application in patient images acquired using a Kinect. Eventually, tests are restricted to more recent or updated versions of the accessible implementations. Although this confined the general set of results, it would be impractical to test all implementations that could be found. Not relying on eventual published results for any given algorithm, we have decided to conduct tests with our own images. Then, we are in the position to provide an overview of performance and adaptability over the results, and decide which suits best the project's needs. A total of ten programs were tested, nine of which were Matlab<sup>®</sup> implementations and the other was a binary executable written in C/C++. Out of the ten, only four were deemed to execute or produce viable results with depth images. Programs were named accordingly to the order they were tested, assuming a nomenclature of "program#<number>".

The experiments were conducted with a total of 131 patients (see Section 4). A synthetic, gradient-like image was also put under scrutiny of the tested methods. Due to its linearity, any method should cope far better with it than any other type of image. This was used to confirm the applicability and results. Tests with RGB images were also subjected to SR algorithms.

The respective identifiers for each program, author, year, paper, number of images used and the web address where they have been retrieved are shown in Table 8.4.

Program	Info	#imgs	Web Address
#1	Elad'97 [69]	MI	<a href="http://www.mathworks.com/matlabcentral/fileexchange/33839">http://www.mathworks.com/matlabcentral/fileexchange/33839</a>
#2	Barzigar'12 [15]	SI	<a href="http://www.mathworks.com/matlabcentral/fileexchange/30488">http://www.mathworks.com/matlabcentral/fileexchange/30488</a>
#3	Farsiu'04 [76]	MI	<a href="http://www1.idc.ac.il/toky/videoproc-07/projects/superres/srproject.html">http://www1.idc.ac.il/toky/videoproc-07/projects/superres/srproject.html</a>
#4	Yang'10 [256]	SI	<a href="http://www.ifp.illinois.edu/~jyang29/codes/ScSR.rar">http://www.ifp.illinois.edu/~jyang29/codes/ScSR.rar</a>
#5	Pickup'08 [195]	MI	<a href="http://www.robots.ox.ac.uk/~elle/SRcode/index.html">http://www.robots.ox.ac.uk/~elle/SRcode/index.html</a>
#6	Cheung'05 [47]	SI	<a href="http://www.psi.toronto.edu/~vincent/sourcecode.html">http://www.psi.toronto.edu/~vincent/sourcecode.html</a>
#7	Vandewalle'06 [243]	MI	<a href="http://rr.epfl.ch/3/">http://rr.epfl.ch/3/</a>
#8	Vandewalle'06 [242]	MI	<a href="http://rr.epfl.ch/6/">http://rr.epfl.ch/6/</a>
#9	Dong'11 [64]	SI	<a href="http://www4.comp.polyu.edu.hk/~cslzhang/ASDS_AReg.htm">http://www4.comp.polyu.edu.hk/~cslzhang/ASDS_AReg.htm</a>
#10	Dong'09 [63]	SI	<a href="http://www4.comp.polyu.edu.hk/~cslzhang/">http://www4.comp.polyu.edu.hk/~cslzhang/</a>

Table 8.4: Super-Resolution tested programs.

## 8.4.3 Performance Evaluation

The performance of each SR method was evaluated for a scale factor of 2 and, accordingly to their nature, using single or multi-image based approaches. Some multi-image based methods were also subjected to sets of images comprising the same image to verify the performance impact of such use case.

Since we do not hold HR depth images from the scene, the acquired images using the Kinect were the ground truth (see Chapter 4). Images were then downsized with the same scale factor of 2, and then subjected to SR methods so that the final spatial resolution would match the original. Images were downsized assuming a loss of integer pixel information for every 2 pixels, thus not presenting any interpolation that could have an impact on the outcome.

All output images were subjected to visual inspection and the Peak Signal to Noise Ratio (PSNR) value was registered in comparison with the original image. The PSNR calculation over multi-image based methods output was calculated with the image, of that set, which presented the best linearity in breast contours.

All experiments were conducted using Matlab<sup>®</sup> R2011a, running on an Intel<sup>®</sup> i5 M450 (@ 2.4 GHz) with 4 GB of RAM.

## 8.4.4 Results

### 8.4.4.1 Depth Images

Most methods showed the ability to choose from several registration, motion estimation or reconstruction algorithms, as well as changing some parameters for these steps. Throughout the experiments, some programs were not deemed to run or produce viable results. That was the case of programs #1, #2, #3, #5, #6 and #8, whereas programs #1, #2 and #3 produced undesirable output resolutions and program #6 produced an output image that did not truly correlate to the original. Besides all efforts, programs #5 and #8 were not able to run effectively, and similarly to the other programs, they were excluded from further tests.

The remaining programs, #4, #7, #9 and #10, were able to run and produce the expected results. Program #4, a single-image based method, used of dictionaries to improve image reconstruction [256]. This method relies on patches from the input image; however, instead of working directly with the image patch pairs sampled from the HR and LR images, authors used a compact representation for these patch pairs to capture the co-occurrence beforehand, significantly improving the speed of the algorithm. Program #7 represented an updated version of program #8 [242] and used three other registration methods [124, 157, 162] for comparison with the original method [243]. Unlike other motion estimation algorithms whose performance is often very low for noisy or highly aliased images, their algorithm only uses low-frequency information. They developed a computationally method to estimate planar rotations. To reconstruct the HR image, they apply bicubic interpolation on a high-resolution grid.

Finally, programs #9 and #10 were tested last. Program #9 is also a Matlab<sup>®</sup> implementation while program #10 is a binary executable program. Program #9 used internal procedures to simulate a LR image presenting blur and noise [64]. Considering that the contents can vary significantly across different images or different patches in a single image, they proposed to learn various sets of bases from a precollected dataset of example image patches, and then, for a given patch to be processed, one set of bases is adaptively selected to characterise the local sparse domain. Program #10 was the only binary executable to be tested that was not a Matlab<sup>®</sup> Graphical

User Interface (GUI) or code-only program. The program's files were in *C/C++* language and all project material were provided for the program to effectively run. It accepted single-image inputs of type Portable GrayMap Format (PGM) which was not a limitation for now, since the actual output from the Kinect is given in that file format. This algorithm is based on the IBP technique, which is argued by authors to often produce "jaggy" and "ringing" artifacts because the reconstruction errors are back projected to the reconstructed image isotropically and locally. In their technique they propose to adaptively incorporate the non-local information into the IBP process so that the reconstruction errors can be reduced.

Within the set of programs that did provide usable results, Table 8.5 shows the performance obtained by applying SR algorithms.

Table 8.5: Performance Overview

Program	Exec. Time (s)	PSNR (dB) - $\mu$ ( $\sigma$ )
#4	444.42	40.1(1.3)
#7	3.00	46.0(0.8)
#9	380.99	51.4(1.2)
#10	7.35	40.1(1.0)

The average PNSR value obtained for each program is shown, as well as the respective standard deviation. The medium execution time to produce a single image is also provided. From these results it is clear that program#9 shows the best results. Even though the processing time was not the expected, this program presents clear reconstructions above the remaining contenders.

The results are shown in Figure 8.18.

From this evaluation, both from Table 8.5 and Figure 8.18, it is possible to clearly state that program#9 produces the best results. By looking at Figure 8.18(d) it is possible to confirm that edges are preserved with program#9 and are similar to the original image. Program#10 creates images that are also similar to the original, but the colour of the pixels is modified, mainly near to the edges. Program#7 presents worse results than program#9; however, the results are better than program#10, producing an image similar to the original one, but slightly blurred. Program#4 provided the worst results, which is confirmed in Figure 9.19(b). With this method, artifacts are introduced in the images in the form of pixellation.

#### 8.4.4.2 Synthetic Images

To acknowledge the performance of the tested methods, a synthetic image was used (see Figure 8.19(a)). The relationship between the outputs provided by each method should confirm the data gathered from the set of tests, also expecting an increase in performance with this simpler and linear image (see Figure 8.19).

Tests showed the best expected performance achievable in comparison with results presented in Table 8.5. The best output was still from program#9, with a 70.40 dB PSNR value. Again, it outperformed the remaining methods in PSNR value; nevertheless, with the visual inspection this comparison is not simple. Table 8.6 shows the results obtained.



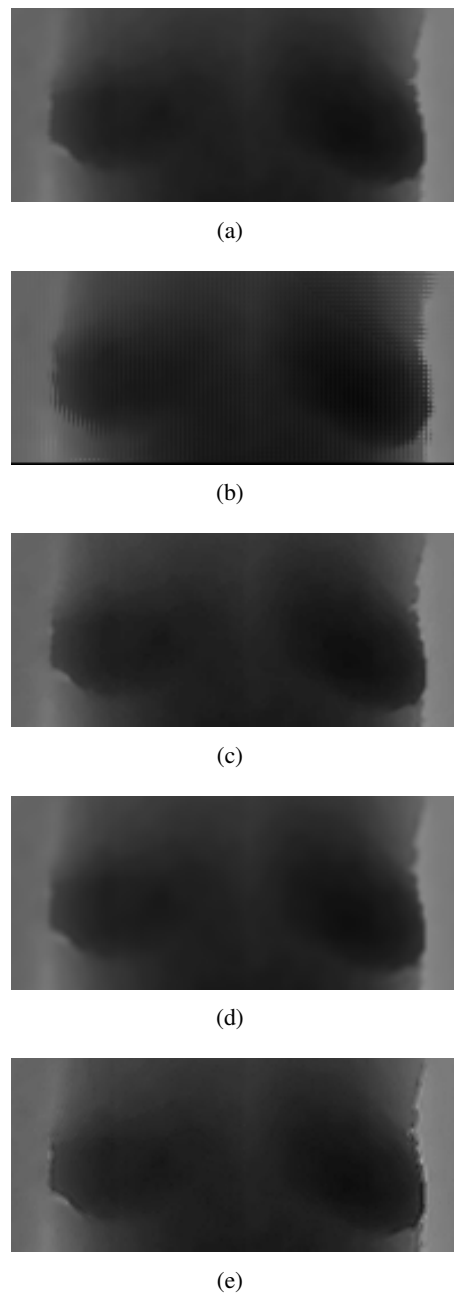


Figure 8.18: Results from the application of Super-Resolution algorithms to an example depth image. a) Original; b) Output of program#4; b) Output of program#7; c) Output of program#9; d) Output of program#10.

#### 8.4.4.3 RGB Images

For the automation of BCCT.core and in the interest of using the images captured from the Kinect, using RGB images captured by the device may prove to be extremely beneficial. This would render void the use of a separate HR digital camera. At the beginning, the LR of the RGB images from the Kinect is a setback for that scenario. However, if the images are successfully enhanced,

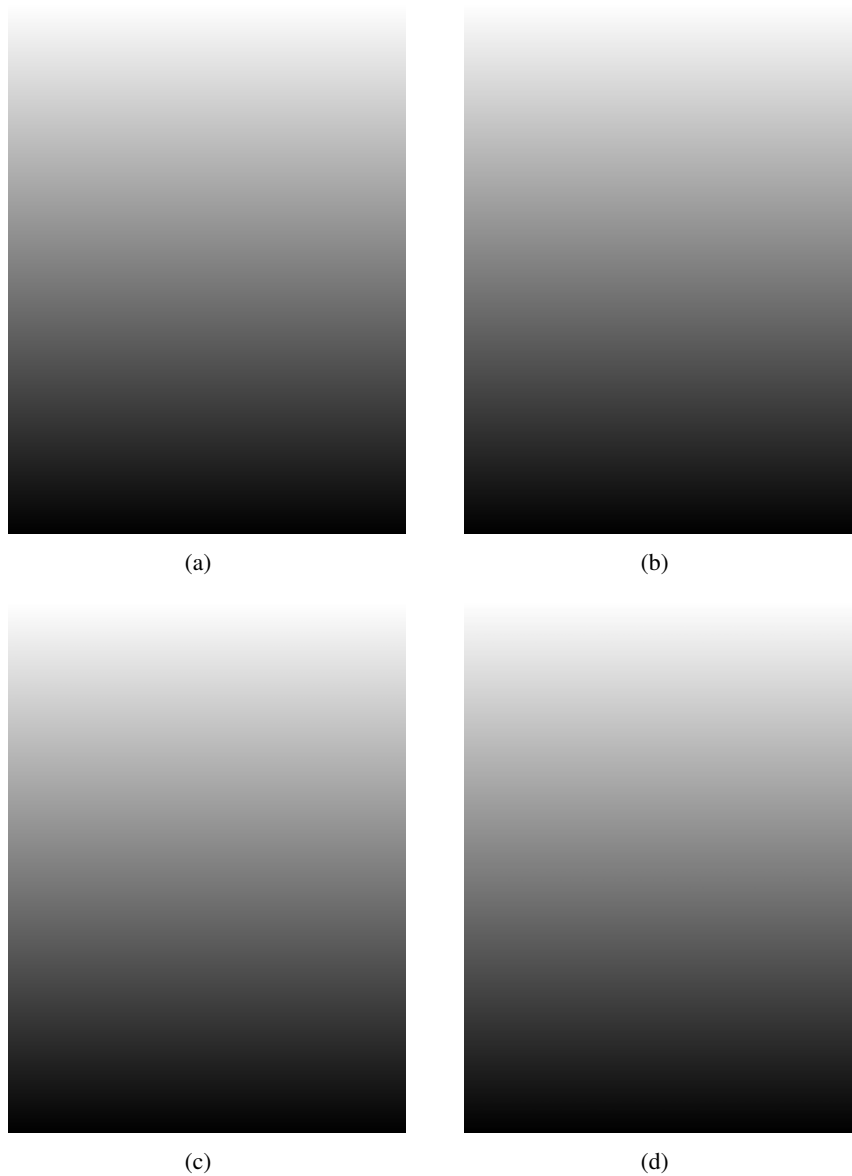


Figure 8.19: Example of application of Super-Resolution in synthetic images. a) Original; b) output of program#7; c) Output of program#9; d) Output of program#10. (The output obtained by the program#4 was not included due to its poor results both for the execution time and PSNR.))

Program	Exec. Time (s)	PSNR (dB)
#4	358.10	50.94
#7	4.05	55.05
#9	353.75	70.40
#10	5.01	53.2

Table 8.6: Synthetic Image Performance Overview

it would be a strong breakthrough. For that reason the study of SR methods on RGB images is welcome. The best performing methods were tested using data from a total of 79 patients (see Chapter 4)

One last obstacle was imposed by the methods themselves. Program#10 was not able to use colour images because it assumes inputs of 8 – *bit* gray scale images only.

After several efforts, the only output sets came from programs #4, #7 and #9. Results for this experimental procedure are shown in Table 8.7, and their respective output can be visualised in Figure 8.20, where program#7 used single-image based input similarly to programs #4 and #9.

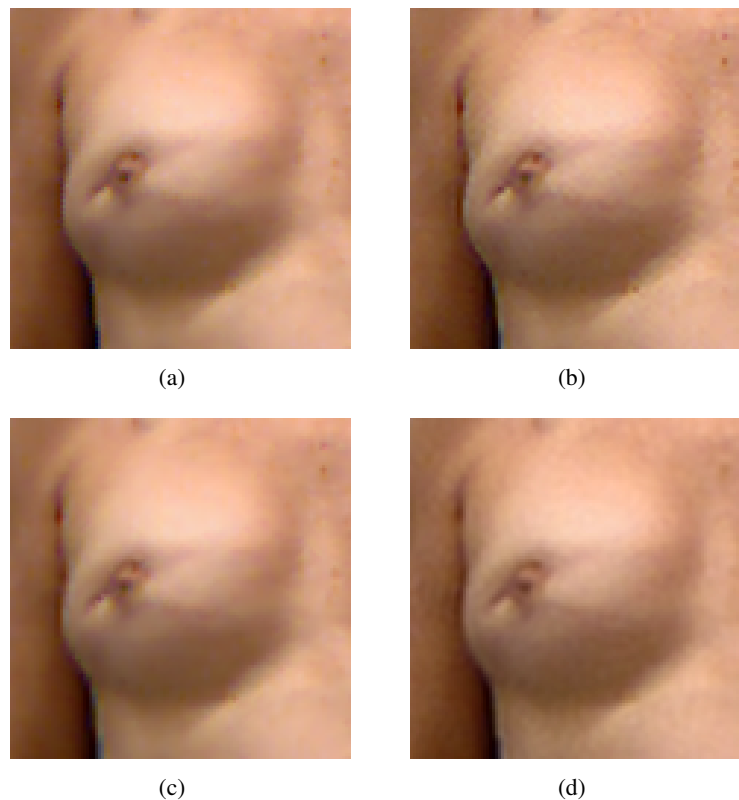


Figure 8.20: Example of application of Super-Resolution in RGB images. a) Original; b) output of program#4; c) Output of program#7; d) Output of program#9.

Program	#4	#7	#9
Average PSNR (db)	35.57	38.57	38.18

Table 8.7: RGB Performance Overview

The PSNR value presented was obtained using the average of the individual PSNR channel (Red,Green,Blue). Although they are not poor results, they are clearly inferior to those with depth images, most likely due to the higher complexity needed to enhance these images.

### 8.4.5 Discussion

Through the experimentation sets with SR methods, we faced different results that may vary substantially; however, there were some similar results. In some cases, when only the methodology of one processing step is applied, the overall result can be similar to the combination of others. In any case, the general standard deviation throughout the experimentations on the set of patients is very satisfactory with values to be around plus or minus 1 dB, within methods applied to depth images. From the general results it is possible to detect the potential of using SR methods to enhance image resolution, before proceeding to further processing. Two programs recorded an average PSNR value above 45 dB, one of which even recorded an average value above 50. This is a very satisfactory reconstruction as confirmed by visual inspection. However, the execution time, mainly for the algorithm which performed better, is not the desired, obviously for on-line utilisation.

It is also important to highlight the fact that multi-image based approaches did not produce the expected results, in most cases, representing a poorer reconstruction in comparison to single-image application, even within the same program. This was attributed to the fact that some images presented a substantial motion, compared to the remaining images of the set. Also, that motion may be significant enough to not be at sub-pixel misalignments; therefore, the registration step may compromise the entire process. In order to reach a conclusion on this topic, a clear study should be performed to verify if the multi-image based approach may be beneficial to this project application. This would require the source images to be captured within fractions of a second, or noticeably with minor shifts.

With regard to RGB images, the study was not conclusive. In fact, the results were poorer than for depth-map images, probably due to more complex processing. This may compromise the immediate use of these algorithms for RGB images.

## 8.5 Joint RGB-D Calibration

As previously presented, the Kinect consists of a colour camera rigidly attached to a depth sensor which comprises a projector-camera pair. Using the capabilities of the Kinect device to the solution of the problem proposed in this thesis, the hypothesis is that the simultaneous detection of characteristics on RGB and depth images. However, it has been observed that such systems suffer from geometric distortions due to the processing performed and the inevitable tolerances in their manufacturing. While a radial and tangential distortion model is usually sufficient to correct the 2D pixel position in colour cameras, depth cameras require a more complicated model to correct the 3D measurement volume [108]. In order to reconstruct a scene from the camera pair measurements the system must be calibrated. This includes internal calibration of each camera, as well as relative pose calibration between the cameras. The calibration of a colour camera is a well-studied problem and it is the most used approach based on Zhang's method [264]. However, colour discontinuities typically used in calibration approaches are not visible on the depth image.

It is then essential to get the best compromise possible in order to find correspondences between the information typically extracted from checkerboard corners present in a colour image and the depth image of the same scene in which the corners are not visible.

In this case, the standard approach is to calibrate the cameras independently and then calibrate only the relative pose between them [214, 262]. However, with this solution the calibration of each camera cannot be improved with the calibration of the other. Furthermore, the calibration of a depth camera may require a high precision 3D calibration object that can be avoided using joint calibration. Fuchs and Hirzinger [82] propose a multispline model for Time-of-Flight (ToF) cameras. This model requires a large number of parameters and has to know the exact pose of the camera. A method for fusing depth from stereo and ToF cameras was described by Zhu *et al.* [266]. Their calibration is based on triangulation from the stereo cameras as ground truth, which cannot be an optimal solution as it ignores possible errors from the stereo triangulation and measurement uncertainties.

ToF cameras are characterised by distortions both on the optical ray direction and on the measured depth. The work by Kim *et al.* [128] presented a principled approach calibration which enables simultaneous recording of dynamic scenes with multiple HR video and LR ToF depth cameras. The authors created this methodology to serve as a testbed for the development of new algorithms for high-quality multi-view dynamic scene reconstruction and 3D video. Cui *et al.* [54] show in their work that the depth distortion of ToF cameras is radially symmetric and scene dependent. Hence, they estimate new distortion correction parameters for each image.

The Kinect device has also shown radially symmetric distortions [216]. However, being a structured light sensor, the nature of the distortions is different. Kinect devices are calibrated during manufacturing with a proprietary algorithm, and calibrated parameters are stored in the device's internal memory to be used by the official drivers to perform the reconstruction. However, manufacturer calibration does not correct the depth distortion, and each Kinect has its own calibration parameters. For this reason, many authors developed their own Kinect calibration. In of the first algorithms <sup>1</sup> only the intrinsics (focal length and principal point) of the infrared camera were calibrated, and not the parameters to convert Kinect disparity units to meters. Herrera *et al.* [107, 108] performed a comprehensive calibration of all parameters of the camera pair. A similar formulation improved previous work with correspondences between the colour and depth images [261].

The approach introduced by Herrera *et al.* [107, 108] was used here. Although this toolbox is not the only possible solution <sup>2 3</sup>, and still presents obstacles, it is one of the most flexible and complete. The toolbox authors proposed a technique to calibrate Kinect cameras by observing a planar pattern shown at a few different orientations. It requires the pattern to be printed and attached to a planar surface (see Figure 8.21). In such approaches the checkerboard corners provide suitable constraints for the colour images, while the planarity of the points provides constraints to

<sup>1</sup><http://nicolas.burrus.name/index.php/>

<sup>2</sup><http://nicolas.burrus.name/index.php/>

<sup>3</sup><http://www.rgbdtoolkit.com/>

the depth images. The pixels at the borders of the calibration object are ignored and thus depth discontinuities are not required.

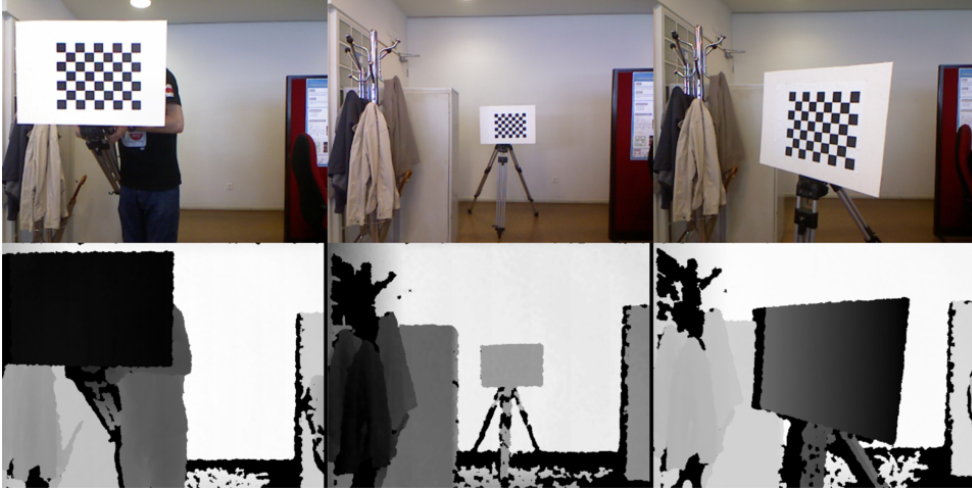


Figure 8.21: Samples from the calibration images. **Top:** Kinect colour images. **Bottom:** Kinect depth images.

Figure 8.22 shows the different reference frames present in the scene. Points from one reference frame can be transformed into another using a rigid transformation denoted by  $T = f(R, t)$ , where  $R$  is a rotation and  $t$  is a translation. For example, the transformation of a point  $x_w$  from world coordinates  $W$  to colour camera coordinates  $C$ , follows  $x_c = R_c \times x_w + t_c$ . Reference  $V$  is anchored to the corner of the calibration plane and is only used for initialisation. The relative pose  $T_r$  is constant, since each image has its own world pose  $T_c$ .

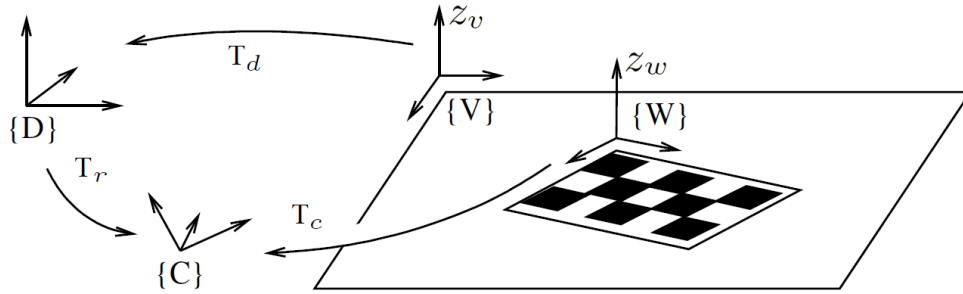


Figure 8.22: Reference frames and transformations presented on the scene.  $C$  and  $D$  are the colour and depth camera reference frames respectively.  $V$  is the reference frame anchored to the calibration plane and  $W$  is the world reference frame anchored to the calibration pattern. Adapted from [107]

In a concise manner the method proposes an initial estimation for the calibration parameters obtained by independently calibrating each camera. The independent calibrations give an estimation of the transformations  $T_c$  and  $T_d$ . The depth intrinsic parameters of the depth camera and the

relative pose  $T_r$  are then refined using a non-linear optimisation. Finally, all parameters are refined simultaneously.

### 8.5.1 Calibration Method

Based on a toolbox released by Daniel Herrera<sup>1</sup> accompanying the work published in [107, 108], we conducted experiments that show an improved accuracy with regard to colour/depth alignment, seeking to pave the way for fusion and interoperability of the two types of information that the Kinect provides.

Nevertheless, some repairs may have to be performed, mostly related to its lack of automatic segmentation methods in the depth images used to extract information in order to calibrate the depth camera. Furthermore, the approach largely depends on the image set provided for training and lacks flexibility in the reciprocity of possible conversions, from RGB to depth and vice-versa.

Figure 8.23 depicts some of the efforts made in order to achieve the desired depth colour camera calibration. The data acquired (see Figure 8.21) cover the entire viewing area, providing an overall operating range of the Kinect device and introducing variability to the relative orientation of the reference surface and the camera.

The calibration procedure has five steps, after acquisition:

1. Select image: a set of images are chosen from the complete set of images acquired;
2. Select corners in RGB images: the corner extraction can be performed manually or automatically;
3. Select planes in disparity images: a polygon is drawn to select the area of the image where the plane is visible;
4. Perform initial estimation: the initial estimation is performed by calibrating the RGB cameras individually and taking well-known calibration values for the depth camera. This provides an approximate calibration that serves as an initial guess for the non-linear minimisation;
5. Calibration: the final calibration is performed here.

Two Kinect devices were used to create the database (see Section 4). As stated before, each device has its own calibration. With the methodology proposed by Herrera *et al.* [107, 108], it is possible to compute the disparity uncertainty, which was modelled by a Gaussian distribution. Using Gaussian distribution and the estimated disparity variance, it is possible to numerically obtain the expected variance in depth for each disparity value (see Figure 8.24).

From Figure 8.24, it is possible to see, as expected, that uncertainty increases as the depth increases. By looking the graphs, it is possible to state that the two devices present more or less the same uncertainty curve. Each point in the graph represents an image used during the calibration

<sup>1</sup><http://www.ee.oulu.fi/~dherrera/kinect/>

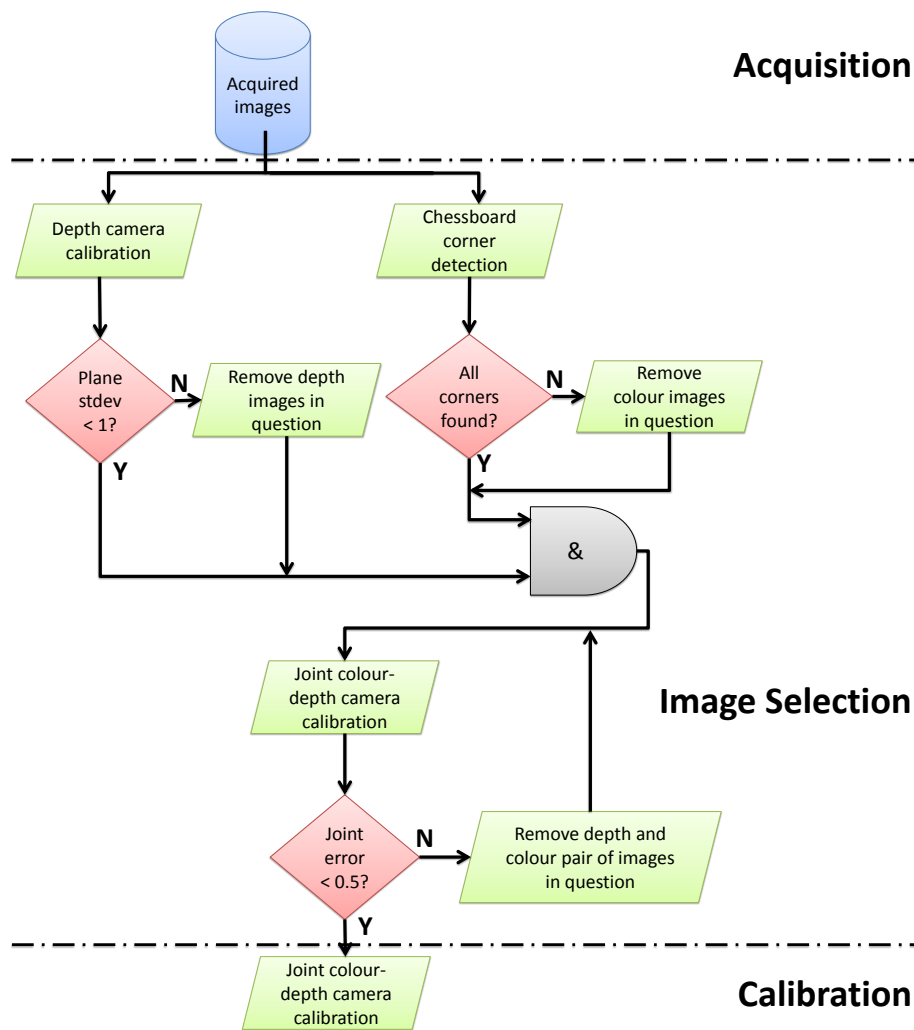


Figure 8.23: Flowchart from the calibration methodology used.

process. During the acquisition procedure the quantity of images acquired in lower and higher distances are the same; however, the graph shows more values at lower depths. This is due to the conditions presented in the calibration workflow (see Figure 8.23).

Figure 8.25) presents histograms of the reprojection errors for the set of images used during the calibration.

As in the uncertainty, the reprojection error is very similar in both devices. The obtained values are very good, since standard deviation for RGB images is less than one pixel and for the depth images it represents less than one Kinect units (see Table 8.8).

Figure 8.26 shows the alignment of depth and colour information performed after the Kinect calibration.

Figure 8.27 provides a 3D perspective of two patients.



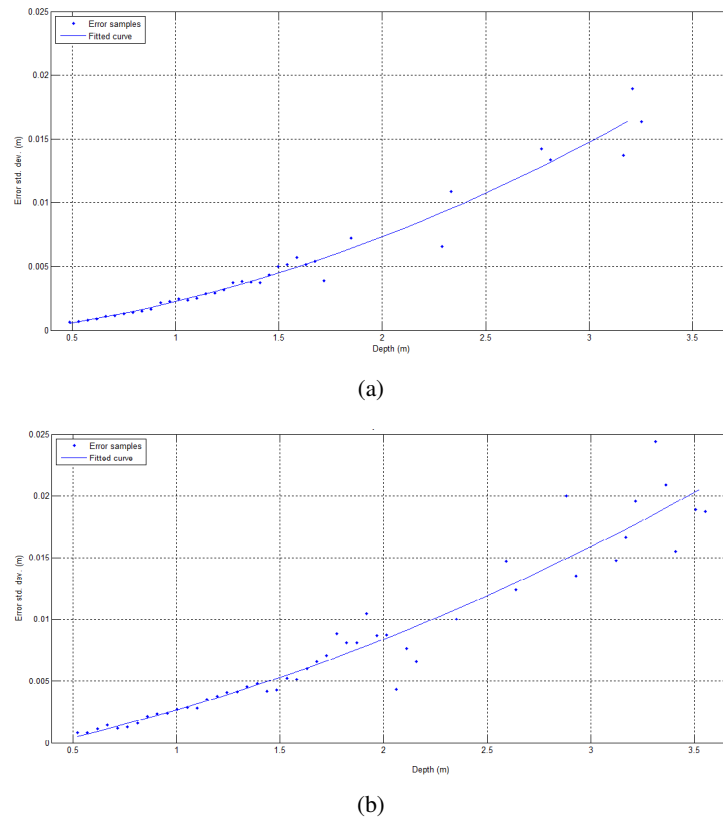


Figure 8.24: Measurement uncertainty for varying depths for the two Kinect devices.

Kinect device	Image type	$\mu$ ( $\sigma$ )
#1	RGB	0.00 (0.47)
	Depth	0.00 (0.72)
#2	RGB	0.00 (0.42)
	Depth	0.01 (0.85)

Table 8.8: Reprojection error after Kinect device calibration.

## 8.6 Conclusion

It is clear that the BCCT.core can be improved by adding more features. Particularly, the aim is to add dimensionality to the measurements in order to decrease the limitation of only measuring based on what can be seen in a frontal photograph. The ability of manipulate and measure over 3D readings from the breasts can improve the accuracy and objectivity of the tool.

3D capabilities are recognised as having high clinical potential. However, the current techniques, applied to the aesthetic evaluation of breast surgical interventions, face two major problems: the high cost of required equipment and the need for specialised operators to operate them. Current techniques are based on especially designed cameras and hardware, mainly resorting to different lenses on the same camera or to laser scanners. Due to these special needs, 3D applica-

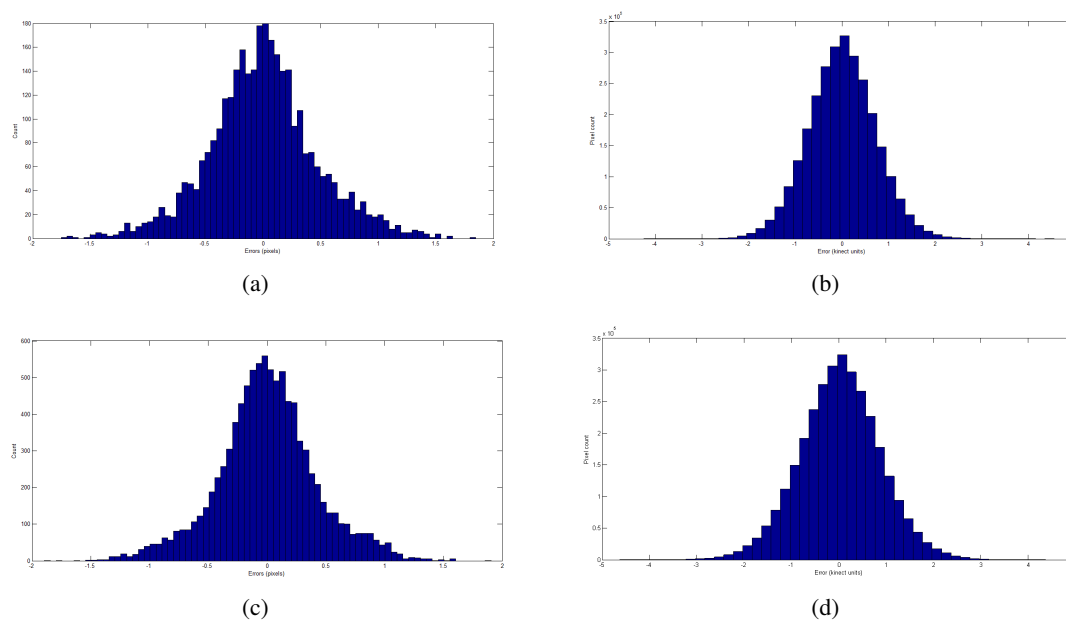


Figure 8.25: Reprojection error after the Kinect device is calibrated. a) and b) are from the first Kinect, while c) and d) are from the second.

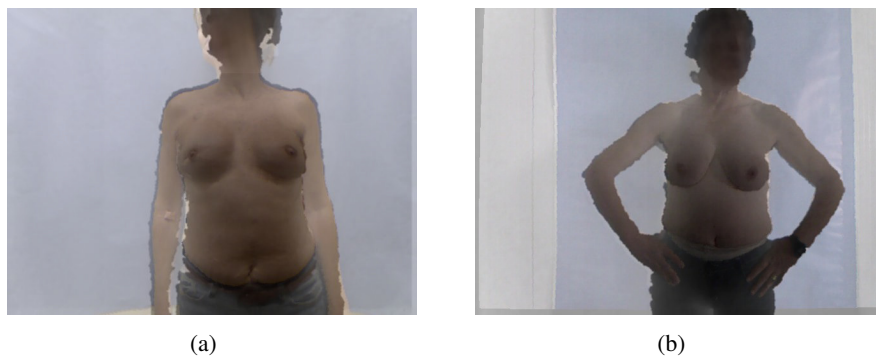


Figure 8.26: Depth and RGB data overlap.

tions are considered costly and are not commonly implemented and, therefore, the benefit of 3D modelling is not availed.

This chapter presents a study on RGB-D low cost devices, their main characteristics, problems and software to operate these devices. Many research applications have been presented, most of them in the area of body scanning, in order to demonstrate the applicability to the problem presented here. Additionally, a preliminary study is demonstrated in which the image of a female phantom torso is acquired using a Kinect device. The device has shown very satisfactory results, making possible to detect volumetric differences of the breasts using the disparity map generated. The results were very similar to the reference performed manually, and for that reason we believe that this approach presents a huge potential as it is a low-cost, user friendly, infrared laser projector

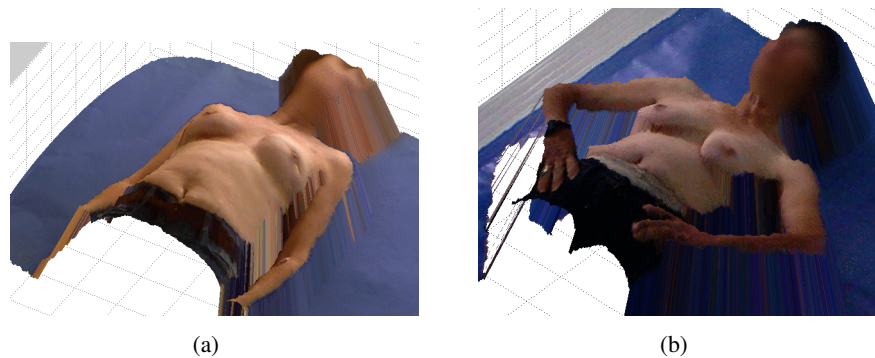


Figure 8.27: A 3D perspective of the scene after correct calibration.

that can be used for aesthetic evaluations after breast surgery.

Therefore, the importance of improving the quality of images obtained with Kinect is emphasized with the application of SR methods. For that purpose, an initial study was conducted on the applicability of SR. Even though the study was limited to the publicly available programs, the general output images were considered to be fairly good. Nonetheless, due to the unavailability of open SR solutions, it was not possible to study state-of-the-art methods specifically tailored for depth images. Therefore, there is no true knowledge on the performance they could achieve on these projects' images. Furthermore, applying SR onto this project's RGB images may be considered standard. However, their use in the BCCT.core software needs to be studied in order to render unnecessary the use of a HR digital camera. Moreover, single-image based methods provided the best results.

The database of input images may show inappropriate for multi-image methods, and therefore it is necessary to study the required characteristics and if they can be acquired using the device, or even if the process of acquisition can be changed to render more consecutive images. The use of SR methods onto the raw depth images did not provide the expected improvement. Thus, no beneficial gain was verified. Nevertheless, the data provided by enhanced images are identical. The reconstructed 3D models were sufficient, but other techniques have been applied in order to render better approximation models [55, 105, 114]. The introduction of these enhanced models could provide better results than those achieved using the statically acquired model. Furthermore, applying the SR onto other processing chain may render a higher and valuable significance.

Lastly, a discussion was provided on the Kinect calibration method, which is necessary to align Kinect data from depth and colour images. Furthermore, the existence of differences between devices were illustrated, an aspect that causes calibration parameters to change (which includes lens correction model, and parameters intrinsic to each of the cameras), and consequently a poor alignment result when transforming the images with resulting calibration parameters of a device different from the images under analysis.



## Chapter 9

# A 3D Low-cost Solution for the Aesthetic Result Assessment\*

As stated in the previous Chapters, the BCCT.core software was developed due to the limited reproducibility of standard aesthetic evaluation methods. Although results are satisfying, there are still limitations concerning complete automation and the inability to measure volumetric information. With the fundamental premise of maintaining the system at a low-cost, this thesis studied the incorporation of the Microsoft Kinect sensor (see Chapter 8).

The current methodologies for assessing aesthetic results after breast cancer surgery present two major challenges. Methods based on photographs do not allow a complete characterisation of the aesthetic result, while techniques based on 3D information are not practical on a daily basis as they are very expensive and require specialized staff. In both cases, automation is a complicated process because breasts are characterised as feature less surfaces.

### 9.1 A Kinect Based System for the Aesthetic Evaluation of BCCT

Over the last years, there have been dramatic improvements in the capabilities of RGB-D cameras, which provide both RGB and Depth information in each image pixel (as in Microsoft Kinect). Combining depth and visual information brings new challenges and opens new possibilities in different fields of application, including medical applications. Several research communities have been making considerable progress in dealing either with 3D depth scans or camera images; now it is time to take advantage of these advances and study the combination of them both.

This Chapter addresses the development of a Kinect based system (BCCT.kore) to predict the overall outcome from data acquired with the Kinect (see Figure 9.1). The proposed methodology requires: a) the automatic identification of fiducial points on the image; b) suitable features must be chosen capturing the main factors (colour, shape, geometry, irregularity of the visual appearance of the treated breast) with an impact on the overall aesthetic result; finally, c) it is necessary to find the optimum rule to map the individual features which will help make the overall decision.

---

\*Some portions of this Chapter appeared in [177, 178, 183]

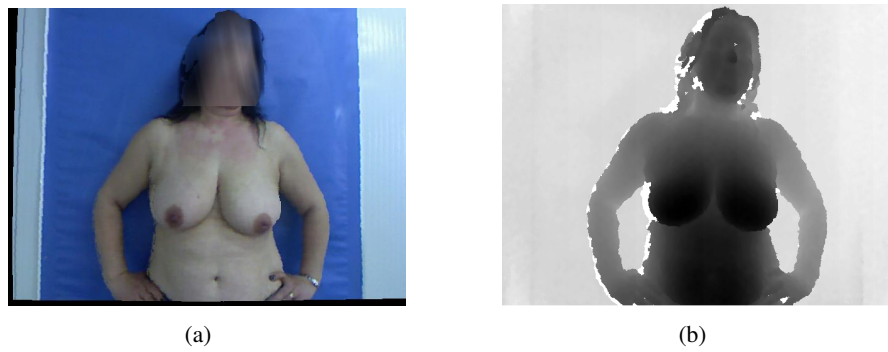


Figure 9.1: Microsoft Kinect data for patient#62. a) Patient photograph; b) Generated disparity map (normalised values).

### 9.1.1 Overview

The proposed system can be implemented as a sequence of a few high-level operations, as represented in Fig. 9.2.

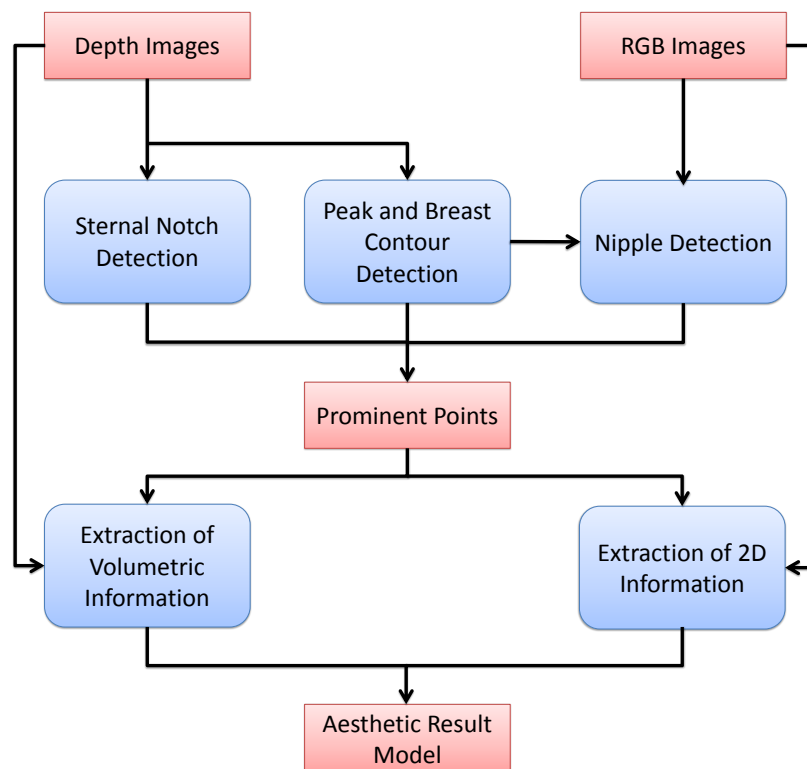


Figure 9.2: System flowchart for the joint RGB-D based system.

In the first step, the *breast peak* and *breast contour* are detected simultaneously based on the gradient vector flow theory and on a shortest path approach, respectively, in depth Kinect

data. *Nipples* are detected based on the local invariant feature detectors theory, namely corner and regional detectors. Nipple detection is performed using RGB Kinect data, assisted by the breast contour previously obtained. Next, *volumetric features* are defined over Depth data, using all prominent points previously detected. Finally, using previous 2D measurements (see Section 3.3) and the new 3D features introduced in this work, a predictive model is built using a machine learning methodology to evaluate the overall cosmetic result. With the inclusion of measurements extracted from the 3D model, the global assessment result was improved over the 2D version in the BCCT.core without increasing complexity significantly.

## 9.2 A Joint RGB-D Based Detection of Prominent Points

### 9.2.1 Simultaneous Detection of Breast Contour and Breast Peak Points

Researchers are now paying more attention to the context to aid visual recognition processes. Context plays an important role in the human visual system recognition processes and many important visual recognition tasks critically rely on context. Here, the mutual contexts of *breast contour* and *breast peak* (the area in the breast closer to the camera or further away from the chest wall, not necessarily the nipple) are modelled so that each one can facilitate the recognition of the other. When performed independently, both tasks are non-trivial since many other parts of the image may be falsely detected. However, the two tasks can benefit greatly from serving as context to each other. To model the breast peak point, a filter is used which evaluates the degree of divergence of the gradient vectors within its region of support from a pixel of interest. The breast contour is modelled as a short path in a graph whose nodes correspond to the pixels on the image and the edges connect neighbouring pixels. The weight function on the edges is defined so that short paths correspond to paths that maximise the amount of gradient strength on the image along the path. The quality or probability of the joint model for the co-occurrence of breast peak and breast contour will be proportional to the individual qualities of the two parts. Next, the proposed algorithm is described in detail.

In this part of the algorithm depth images are used. Contour detection is a very difficult task to perform on RGB images to the fact that its boundary is not clearly visible. Using depth images the boundary contrast increases. This process was implemented as a sequence of a few high-level operations over the depth information: 1) background segmentation, 2) breast peak candidate detection, 3) contour detection and 4) peak and breast contour decision (see Figure 9.3).

Two simplifications were introduced with regard to the main steps presented above. Firstly, the detection is performed only over the patient's body, just to speed up the computation process. Secondly, the mutual detection of prominent points is addressed first by over-detecting peak candidates, and then by detecting contours using all peaks found.

As in other depth technologies, such as laser scanners or range finding cameras, the Kinect's output contains a large amount of missing data and noise as a result of occlusions. These undesirable conditions are typically represented as white pixels (see Figure 9.1(b)). These missing values

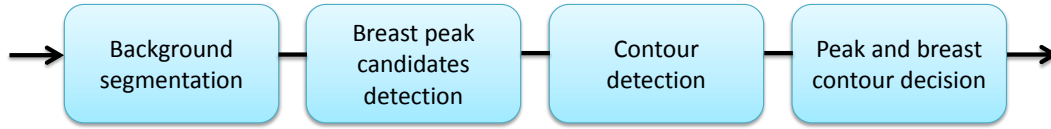


Figure 9.3: Algorithm flowchart for the simultaneous detection of breast contour and breast peak points.

were excluded from all of the following operations. The remaining pixels were scaled between 0 and 1. Finally, a  $3 \times 3$  smooth filter was applied to eliminate noise resulting from the acquisition process. Figure 9.4(c) and Figure 9.4(d) already shows the output of this pre-processing task.

### 9.2.1.1 Adaptive Background Segmentation

Acquisitions are mainly performed in a uniform background. However, occasionally the background is cluttered with the presence of different objects at different depths. Objects can appear at depths similar to the patients', or just the different body parts at different depths, dismissing the application of traditional thresholding methods such as Otsu's (see Figure 9.4(c)). In this work, it was considered that the patient is at a somewhat central position on the image and is likely the 'object' closer to the camera.

This property was exploited using a density image defined by transforming the depth information on the  $XY$  plane ( $D$ ) to the  $XZ$  plane ( $\Phi$ ) [127]. The value at position  $(x, z)$  of the density image denotes the number of points at depth  $z$  in column  $x$  of the depth image (by counting along the  $Y$  direction), (see Figure 9.4(a)).

$$\Phi(x, z) = \sum_{y=1}^n (D(y, x) == z) \quad (9.1)$$

Intuitively, each column of the density image is the histogram of the corresponding column in the depth image. A global thresholding method of the original  $XY$  image corresponds to a horizontal line in the  $XZ$  image, differentiating background from foreground (see Figure 9.4(a) and Figure 9.4(c)). A more adaptive method is allowing the threshold to vary from column to column. Since it is not expected that the threshold will change significant between consecutive columns, the method can be constrained to enforce this knowledge. The constrained adaptive thresholding leads to a continuous curve in the  $XZ$  image from left to right.

The constrained thresholding can be easily tackled using graph based concepts. For that, the quality (cost) of each candidate threshold needs to be estimated; in the resulting cost image (interpreted as a graph), the computation of the shortest path between left and right margins yields the individual thresholds (see Figure 9.4(d)).

In our application scenario, the information along each column can present one of three different patterns: (1) background only, (2) 'object' and background and (3) 'object' only. The cost will



be set according to an initial estimate of the column pattern. The specific patterns are predicted by analysing the mean and the variance of a Gaussian Distribution  $\mathcal{N}$  fitted to XZ data ( $\Phi$ ). Table 9.1 shows the qualitative indication of the values expected for each pattern type.

Pattern	$\mathcal{N}(\sigma^2)$	$\mathcal{N}(\mu)$
(1)	low	high
(2)	high	high and low
(3)	low	low

Table 9.1: Fitting parameters.

Furthermore, for each depth  $z$  and position of column  $x$ , the variance above:  $Var_a$  (the variance in the column  $x$  of the depth values greater than  $z$ ); and the variance below:  $Var_b$  (the variance in the column  $x$  of the depth values less than  $z$ ) are computed. The cost for each possible threshold is then set according to the following rule, according to the predicted pattern: (pattern 1) cumulative value of the minimum of the variance ( $Var_{min}$ ) from 1 to  $Nbins$ ; (pattern 2) average of the two variances ( $Var_{mean}$ ); (pattern 3) cumulative value of the minimum of the variance ( $Var_{min}$ ) from  $Nbins$  to 1 (see Figure 9.4(b)).

$$Var_{min}(x, z) = \min(Var_a(x, z), Var_b(x, z)) \quad (9.2)$$

$$Var_{mean}(x, z) = \frac{Var_a(x, z) + Var_b(x, z)}{2} \quad (9.3)$$

### 9.2.1.2 Breast Peak Candidate Detection

The breast peak is modelled using a filter that evaluates the degree of divergence of the gradient vectors from a pixel of interest. A major benefit of using depth-map images is the high contrast observed in the breast area. The peak point of the breast corresponds to the point in the breast where disparity attains the lowest value. The typically round or tear drop shape of a breast leads to a distinctive pattern in the gradient vector field, where the gradient diverges in all directions (see a toy example in Figure 9.5).

Breast peaks (see Figure 9.6(a)) are detected using an approach similar to the Convergence Index Filter [130], in which the image is filtered with a radial vector field pattern, as presented in Figure 9.6(b).

Gradient  $g(x, y)$  is computed firstly calculating the horizontal  $G_x$  and vertical  $G_y$  components of the image gradient. In this work, a Gaussian filter was applied, and only the gradient orientation angle  $\theta(x, z)$  was used by previously normalising the norm of the gradient vectors to unity.

$$\theta(x, z) = \tan^{-1} \frac{G_y}{G_x}, \text{ with } |g(x, y)| = 1 \quad (9.4)$$

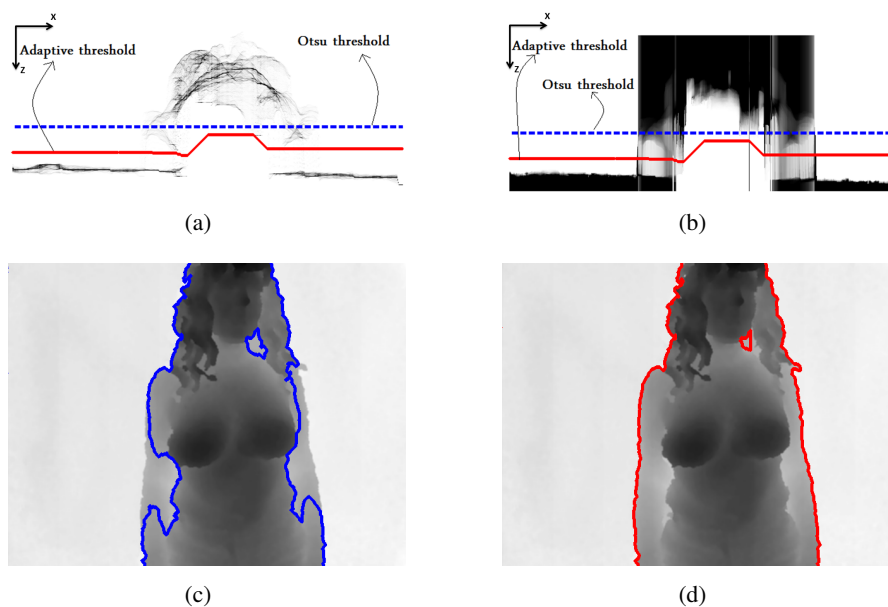


Figure 9.4: Adaptive Background Segmentation. a) XZ plane depth information; b) XY variance plane; c) Otsu's segmentation; d) Adaptive segmentation.

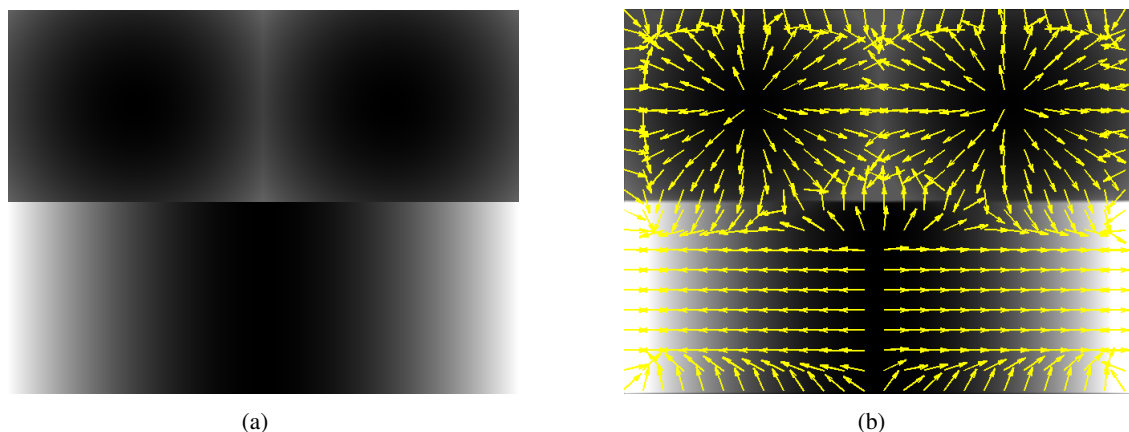


Figure 9.5: Toy example of different gradient shapes. a) Gradient example; b) Gradient vector field from toy example in Figure 9.5(a).

The similarity between the template and the image was assessed using two different measurements:

- cross-correlation

$$(f * g)[\mathbf{n}] \stackrel{\text{def}}{=} \sum_{\mathbf{m}} f^*[\mathbf{m}]g[\mathbf{n} + \mathbf{m}] \quad (9.5)$$

where  $f$  and  $g$  represent the gradient orientation vector field and the template vector field respectively.  $f^*$  denotes the complex conjugate of  $f$ .

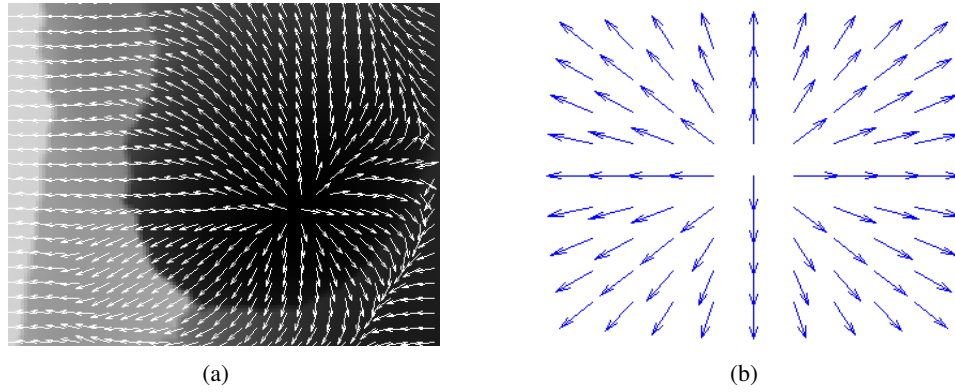


Figure 9.6: Breast peak candidate detection. (a) Breast gradient vector field (5-pixel spacing); (b) Template vector field

- circular correlation, introduced by Nicholas Fisher [79].

The result of cross-correlation and circular correlation, with regard to Figure 9.5(a), can be graphically represented as shown in Figure 9.7, where the values range from  $-1$  to  $1$ , with  $-1$  represented in blue and  $1$  in red.

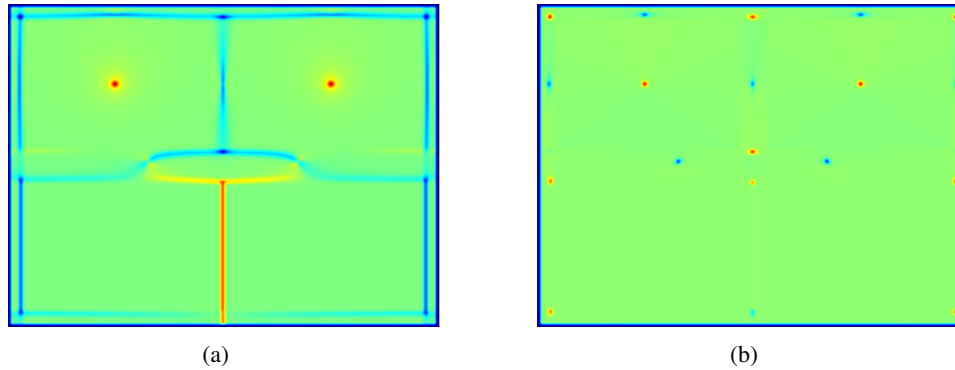


Figure 9.7: Graphical demonstration of correlation approach. (a) Cross-correlation result; (b) Circular correlation result

Breast peak candidates correspond to all local maximum positions.

### 9.2.1.3 Contour Detection

Breast contour detection was modelled using a shortest-path in a graph whose nodes correspond to the pixels on the image and the edges connect neighbouring pixels, an approach that is similar to [34, 176]. Breast boundary can be seen as a change in the grey-level values of the pixels, originating an edge in the resulting image. Interpreting the image as a graph with pixels as nodes

and edges connecting adjacent pixels, the low-cost path through edge pixels corresponds to the breast contour, using an appropriate weight function.

In this work, the image grid is considered as a *graph*  $G = (V, A)$  with pixels  $V$  as nodes and arcs  $A$  connecting neighbouring pixels  $(p, q)$ ,  $p, q \in V$ . Each arc is associated with a *weight*  $w(p, q)$ , which is a function of pixel values and pixel relative positions. A path from vertex (pixel)  $v_1$  to vertex (pixel)  $v_n$  is a list of unique vertices  $v_1, v_2, \dots, v_n$ , with  $v_i$  and  $v_{i+1}$  corresponding to neighbour pixels. The total cost of a path is the sum of each arc weight in the path  $\sum_{i=2}^n w(v_{i-1}, v_i)$ .

A path from a source vertex  $v$  to a target vertex  $u$  is said to be a *shortest path* if its total cost is minimum among all  $v$ -to- $u$  paths. The distance between a source vertex  $v$  and a target vertex  $u$  on a graph,  $d(v, u)$  is the total cost of a shortest path between  $v$  and  $u$ .

A path from a source vertex  $v$  to a sub-graph  $\Omega$  is said to be a shortest path between  $v$  and  $\Omega$  if its total cost is minimum among all  $v$ -to- $u \in \Omega$  paths. The distance from a node  $v$  to a sub-graph  $\Omega$ ,  $d(v, \Omega)$  is the total cost of a shortest path between  $v$  and  $\Omega$ :

$$d(v, \Omega) = \min_{u \in \Omega} d(v, u). \quad (9.6)$$

A path from a sub-graph  $\Omega_1$  to a sub-graph  $\Omega_2$  is said to be a shortest path between  $\Omega_1$  and  $\Omega_2$  if its total cost is minimum among all  $v \in \Omega_1$ -to- $u \in \Omega_2$  paths. The distance from a sub-graph  $\Omega_1$  to a sub-graph  $\Omega_2$ ,  $d(\Omega_1, \Omega_2)$  is the total cost of a shortest path between  $\Omega_1$  and  $\Omega_2$ :

$$d(\Omega_1, \Omega_2) = \min_{v \in \Omega_1, u \in \Omega_2} d(v, u). \quad (9.7)$$

Since the breast contour is approximately circular and centered on the breast peak candidates, the computation is more naturally performed by adopting polar coordinates, with the origin of the coordinates in the peak candidate. The computation of the shortest path is performed using dynamic programming to increase efficiency. Assuming the simplifying assumption that the horizontal paths do not zigzag back and forth, up and down, in the transformed image, the search may be restricted among connected paths containing one, and only one pixel in each column of the path.

Formally, let  $I$  be an  $N_1 \times N_2$  image and define an admissible path as:

$$\mathbf{s} = \{(x, y(x))\}_{x=1}^{N_1}, \text{ s.t. } \forall x |y(x) - y(x-1)| \leq 1,$$

where  $y$  is a mapping  $y: [1, \dots, N_1] \rightarrow [1, \dots, N_2]$ .

This means that an admissible path is an 8-connected path of pixels on the image from left to right, containing only one pixel in each column of the image. The first step is to traverse the image from the second column to the last and to compute the cumulative minimum cost  $C$  for all possible connected staff lines at each entry  $(i, j)$ :

$$C(i, j) = \min \begin{cases} C(i-1, j-1) + w(p_{i-1, j-1}; p_{i, j}) \\ C(i-1, j) + w(p_{i-1, j}; p_{i, j}) \\ C(i-1, j+1) + w(p_{i-1, j+1}; p_{i, j}) \end{cases},$$

where  $w(p_{i, j}; p_{l, m})$  represents the weight of the edge incident with pixels at positions  $(i, j)$  and  $(l, m)$ .

At the end of this process,

$$\min_{j \in \{1, \dots, N_2\}} C(N_1, j)$$

indicates the end of the minimal connected path. Hence, in the second step, one backtrack from this minimum entry on  $C$  is applied to find the optimal path.

Each column in the polar image corresponds to the derivative along each radial line in the original space:

$$G_\theta(r) = \frac{f(r+h) - f(r-h)}{2h} \quad (9.8)$$

where  $h = 1$  and  $r$  is the radius.

Each 4-neighbour pixel arc corresponds to a weight determined by the derivative value of the two incident pixels, expressed as an exponential law:

$$f(g) = f_l + (f_h - f_l) \frac{\exp(\beta(255 - g)) - 1}{\exp(\beta 255) - 1} \quad (9.9)$$

with  $f_l, f_h, \beta \in \mathfrak{R}$  and  $g$  is the minimum of the gradient computed for the two incident pixels. For 8-neighbour pixels the weight was set to  $\sqrt{2}$  times that value. The parameters  $f_l$ ,  $f_h$ , and  $\beta$  were fixed at  $f_l = 2$ ,  $f_h = 128$ ,  $\beta = 0.0208$ . This cost function and parameters were optimised in a previous work [176].

Since the start and ending of the breast contour are unknown, at this phase a conservative approach is used, searching only for angles between  $\pi$  and  $2\pi$  (see Figure 9.8(a)). Later the contour will be completed. The candidate contour is then the output of the shortest path algorithm in the polar image (see Figure 9.8(b)). The shortest path was computed between the whole external margin and a single point (point of highest gradient) in the internal margin.

#### 9.2.1.4 Peak and Breast Contour Decision

The quality or probability of the joint model for the co-occurrence of breast peak and breast contour is computed maximising the joint probability of the individual parts. It is assumed that the joint probability is a monotonous function of the product of the correlation outputted by the divergence filter  $\rho_p$  and the quality of the detected contour  $\mu(\Delta C)$ . The quality of the contour  $\mu(\Delta C)$

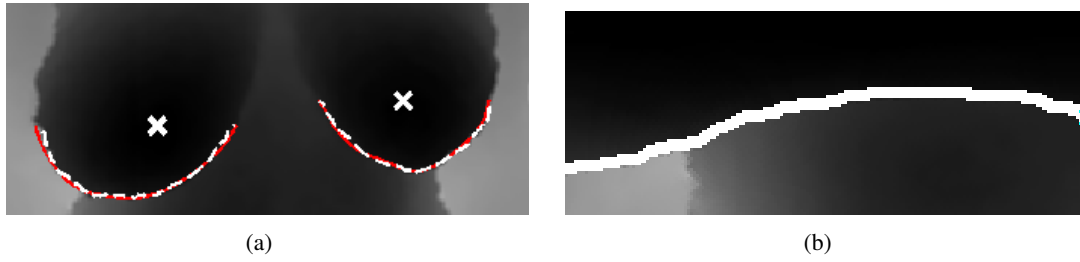


Figure 9.8: Breast contour detection. (a) Breast contour - ground truth (solid red line), detected (dashed white line); (b) Polar image and detected contour (white line).

is evaluated by the average magnitude of the directional derivative along it and  $\rho_p$  is the cross-correlation or circular correlation of the centre candidate. Therefore, the final decision consists of selecting the pair (peak, contour) that maximises the quality measurement:

$$Q(Cp) = \mu(\Delta C) \cdot \rho_p \quad (9.10)$$

Taking advantage of body symmetry with regard to the sagittal plane, a pair (peak, contour) was selected for the right and another for the left of the body centre of mass. The centre of mass is computed using the patient body mask obtained in Section 9.2.1.1.

#### 9.2.1.5 Complete Breast Contour Detection

The complete detection of the breast contour is an extension of the methodologies presented in Section 9.2.1.3. The algorithm is divided into the following operations: a) defining contour endpoints and b) defining contour between detected endpoints.

Although the definition of the external endpoints is not consensual, even among the medical community, the position of the external endpoints can be assumed to be at the point of the body where the arm contour intersects the trunk contour. However, when the patient's arms are alongside the body, the arm's contour is almost indistinguishable from the trunk's contour. Therefore, the aim was to find the external endpoint by simply measuring the 'strength' of vertical edges in the external boundary of the breast. External endpoints are detected with the following operations:

1. Computing horizontal derivative;
2. Extracting vertical edges with an appropriate threshold (experimentally tuned);
3. Detecting connected components;
4. Extracting extremities from connected components;
5. Obtaining contour endpoints from detected extremities.

These operations are illustrated in Figure 9.9(a).

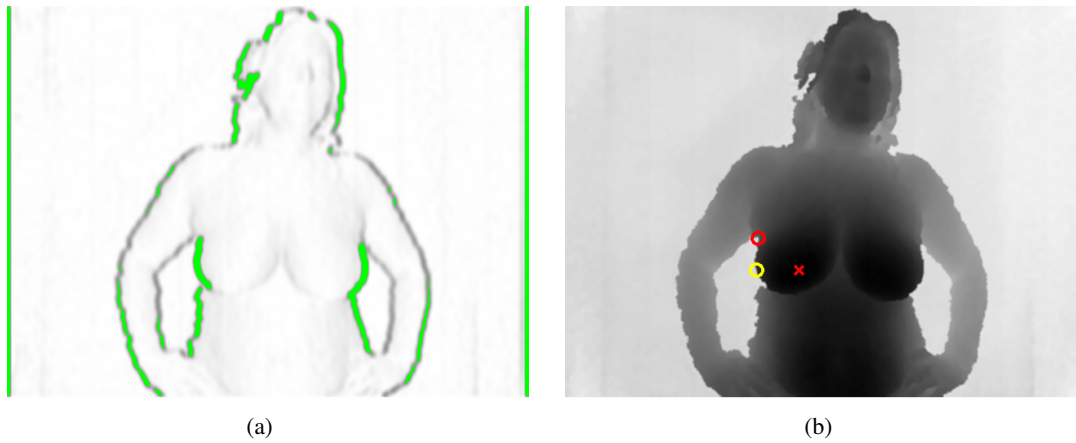


Figure 9.9: External endpoint detection. (a) Horizontal derivative and vertical 'strong' edges detection (green lines); (b) Detected external endpoint (red circle) - previous outer endpoint (yellow circle) - breast peak point (red cross).

After detecting all extremities from vertical edges, the outer contour endpoint is defined as the closest extremity to the endpoint obtained during the partial contour detection phase, presented in Section 9.2.1.3 (see Figure 9.9(b)). These operations are performed separately for each breast according to the sagittal plane defined previously.

The internal endpoint of each breast contour is defined during the complete contour detection. This task is performed following the theory presented previously in Section 9.2.1.3 with the following constraints: the starting point contour is defined as the external endpoint detected previously; and the search for the internal endpoint is limited to the sagittal plane or to the horizontal plane defined as the plane where the distance to the two external endpoint is the same (see Figure 9.10).



Figure 9.10: Complete contour detection and internal endpoint definition - external endpoint detected (red circle) - previous outer endpoint (yellow circle) - breast peak point (red cross).

In conclusion, the internal endpoints are obtained by intersecting each breast contour and the sagittal plane or the average horizontal plane defined by the external endpoints.

### 9.2.2 Sternal Notch Detection

The jugular notch is a visible dip found at the superior border of the manubrium of the sternum, between the clavicular notches. To automatically detect the sternal notch, it is necessary to locate the midpoint of the upper portion of the region at the intersection of the neck and the torso. To determine the point where the torso intersects with the neck, the depth values are projected (summed) along the rows image (see Figure 9.11).

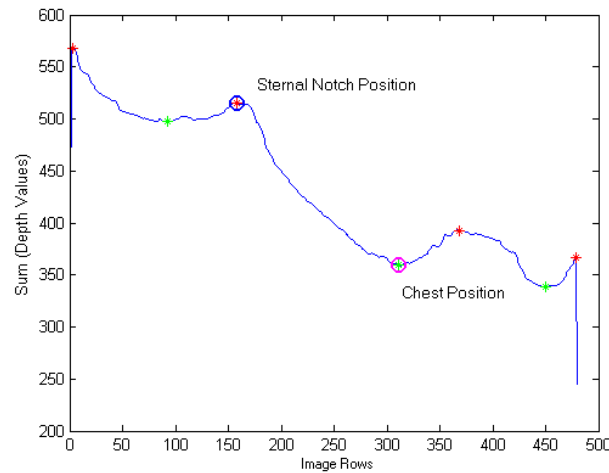


Figure 9.11: Neck position detection. (red \* - local maximums; green \* - local minimums; blue circle - neck position; magenta circle - chest position)

The graph from horizontal projection presents a characteristic shape. Specific parts of women's bodies can be found by detecting local minima (green \*) and maxima (red \*). The pair sternal notch (*local maximum*) - chest (*local minimum*) is the pair of consecutive maximum and minimum at the highest distance, yielding the location of the sternal notch (blue circle).

### 9.2.3 Nipple Detection

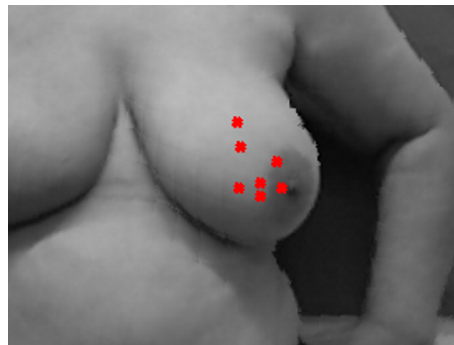
Nipple detection was performed using RGB images, converting them to gray scale, after a correct alignment of depth and colour information. As seen in Section 8.5.1, the Kinect colour sensor is separated and not synchronised with the depth sensor. Furthermore, it has a slightly wider field of view and is offset (by about 2 cm) from the depth sensor. Nipples detection was help by the delimitation of breast contours previously performed in the depth images, which obviously simplifies the process.

The breast surface is generally characterised as a featureless shape, and typically the nipple is the most prominent feature. However, many other characteristics can be mistaken by the nipple when using traditional feature detectors based on colour or more complex features such as texture. Since the nipple is part of the areola complex, a closed region enclosing the nipple, regional information can make it easier to detect the nipple.



We start by over-detecting nipple candidates; then, *one or more* closed contour(s) were found enclosing each candidate. Finally, the best pair candidate/closed contour was selected combining information from the contour and the nipple candidate.

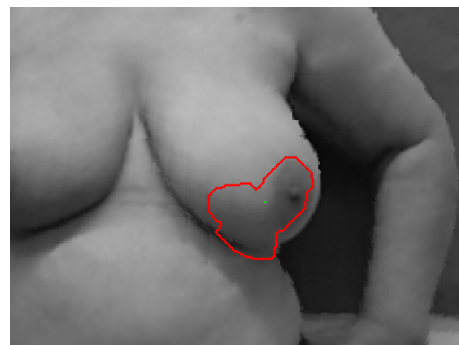
The areola complex is typically more textured than the rest of the breast surface. Therefore, a corner-based method was applied to detect the nipple. However, it was quite difficult to achieve a good detection rate without detecting multiple spurious points; controlling the spurious points leads to a high miss detection rate. Therefore, a corner descriptor (Harris detector [95]) is used as an initial step by over-detecting possible nipple locations (see Figure 9.12(a)). Afterwards, the proposed closed contour method (see Chapter 5) is applied using nipple candidate as centre of coordinates, with the aim of finding the areola contour (see Figure 9.12(b) and Figure 9.12(c)). Since the areola complex is often rich in edges and false contours, the algorithm was implemented to produce all closed contours found in the process. Instead of just keeping the closed contour with the smallest cost, all of them are kept to perform a joint analysis of the contour features (such as shape) and centre candidate. This approach has two main assumptions: firstly, at least a nipple candidate lies within the true areola contour; secondly, the areola contour represents a closed path over strong gradient pixels. The areola contour is not necessarily circular and the candidate nipple does not need to match the true position of the nipple to correctly detect contour.



(a)



(b)



(c)

Figure 9.12: Nipple detection. a) Harris corner detection inside the breast; b) Areola detection (true corner); c) Areola detection (false corner).

After completing these two steps, several pairs (nipple candidate/areola contour) are detected. The challenge now is to discriminate the correct candidate/contour pair. A combination of several features and a classifier-based approach was implemented. The extracted features include:

- The Harris corner quality factor,  $\rho_h$ ;
- The average magnitude of the directional derivative of the areola contour,  $\mu(\Delta C_a)$ ;
- The shape factor,  $S_a$ , of the contour;
- The equivalent diameter,  $d_a$ , of the contour.

The Harris corner quality factor  $\rho_h$  measures the confidence of the detector in the corner or junction found. The average derivative  $\mu(\Delta C_a)$  measures the magnitude (strength) of the directional derivative over the areola contour; the circularity of the contour is measured by the shape factor  $S_a$ :

$$S_a = \frac{4 \cdot \pi \cdot A}{P^2} \quad (9.11)$$

where  $P$  is the perimeter and  $A$  is the area enclosed by the contour. The shape factor takes values in the interval  $[0, 1]$ , where 1 represents a perfect circle. The equivalent diameter  $d_a$  represents the diameter of a circle with the same area as the inner contour region, computed as:

$$d_a = \sqrt{\frac{4 \cdot A}{\pi}} \quad (9.12)$$

Classification methods based on SVM classifiers were adopted, using the popular LIBSVM [45] implementation and a probability estimation approach.

### 9.3 Volumetric Feature Definition

Previous techniques to evaluate the cosmetic result are based on measurements taken from representative frontal photographs of the patients, being therefore limited by the information present in the frontal view. Besides the appearance of the surgical scar and the skin colour change induced by radiotherapy, the measurements are mostly based on asymmetries between the breasts.

The use of volumetric data enables a more complete understanding and reliable aesthetic assessment. Nevertheless, the volumetric features must be objectively and consistently processed, a requirement for a reproducible assessment. Furthermore, the use of enhanced images provided by SR on depth images may improve the definition of features (see Section 8.4). In this work, the thirty features extracted from the frontal view in the BCCT.core (see Section 3.3) are complemented with volume, surface area, nipple height and 3D Nipple Retraction Assessment measured in the 3D information taken from the Kinect.

Instead of limiting any subsequent analysis to an initial choice of a volumetric feature, it was decided to record multiple indices (some of which already introduced in the literature [67]) with

the purpose of proceeding later to a feature selection process). The measurement of the volumetric features relied on the prominent points detected as reported previously (see Section 9.2). The indices recorded to assess volumetric information were the following:

- Volume Difference Assessment (VDA): quantifies the ratio of the total volume between the two breasts. The volume is defined as the sum of unit values between the estimated body plane and the breast surface (see Figure 9.13(a) - red blocks);
- Surface Area Assessment (SAA): quantifies the ratio of the surface area between the two breasts. To compute the area of a surface, it is necessary to compute the area of each rectangular face of the surface and then put them together (see Figure 9.13(a) - green rectangular faces). To compute the area of each rectangular face, it is convenient to imagine the face as being split into two triangular pieces. The cross product can be used to compute the area of a triangle in a 3D space. If the coordinates of the vertices are given  $v_i = (X_i, Y_i, Z_i)$  then the surface area  $A_s$  is computed as:

$$A_s = 0.5 \cdot |(v_2 - v_1) \times (v_3 - v_1)| \quad (9.13)$$

- Nipple Height Assessment (NHA): quantifies the ratio of the nipple height between the two breasts. The nipple height is defined by measuring the distance (in depth) between nipple and chest position;
- 3D Nipple Retraction Assessment (3DNRA): quantifies the ratio between the retraction of the two nipples. The nipple retraction index is computed based on the 3D information of three fiducial points: right nipple (RN), left nipple (LN) and the sternal notch (ST) (see Figure 9.13(b)). The 3DNRA is the ratio of the 3D Euclidean distance between the nipples and the sternal notch.

Three of the features considered (volume, surface area and nipple height) have to take the chest wall plane (inter mammary fold) as reference. The chest wall plane is considered to be parallel to the camera plane at the minimum depth that is found in the area between the inner breast contour extreme points and the lowest bottom point of the breasts.

## 9.4 Aesthetic Model

The problem addressed here, which deals with predicting the overall Harris score, involves classifying examples into classes that have a natural order. Therefore, techniques were used which are specifically designed for that purpose [29] instantiated in SVMs. Two schemes were investigated to integrate 3D and 2D features. The first approach is based on a typical scheme where 2D and 3D features are combined in a simple vector without any restriction. A classifier is then trained on this enlarged vector. A second approach is based on a cascade scheme [87]; here, the output of

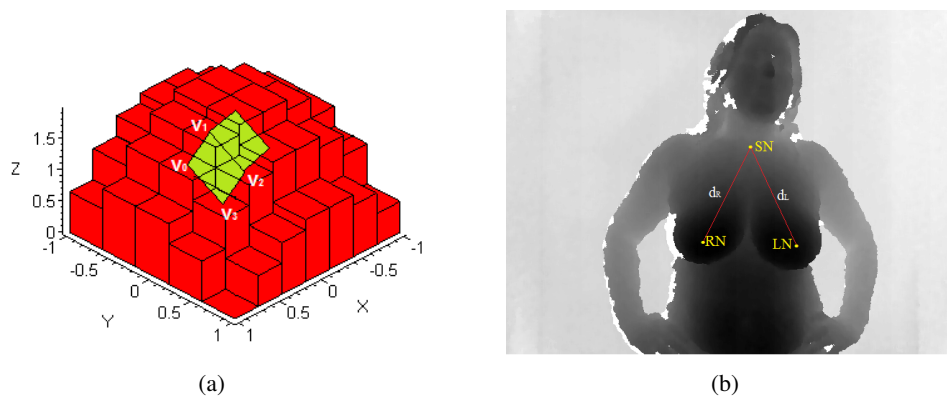


Figure 9.13: Volumetric measurements. (a) Illustration of VDA and SAA measurement based on voxels; (b) Illustration of 3DNRA measurement.

the model based on 2D features is used only as input together with the 3D features for a second model (see Figure 9.14).

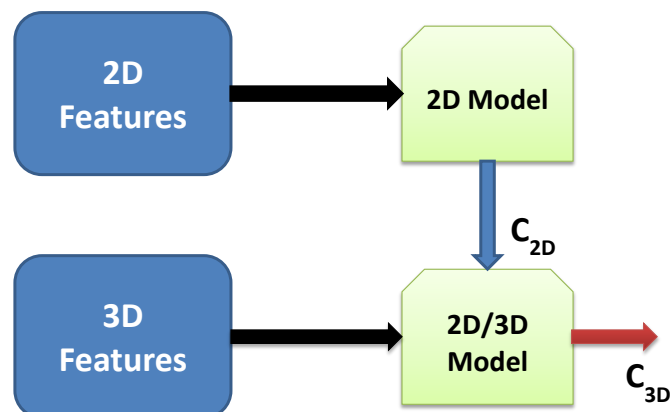


Figure 9.14: Cascade classifier flowchart.

## 9.5 Results

As showed in Chapter 4 the database includes data from 135 patients acquired during several sessions by an expert on breast cancer. For each patient the data include several depth and colour frontal images for each patient acquired with Kinect and images acquired with a standard 5MP portable camera. However, the database does not include Kinect colour and/or depth-map images for all patients. The height of both nipples (distance between the medial projection of the nipple and the sternum measured with 2 rulers) was manually obtained by the physician during acquisition period.

Manual ground truth annotation was performed both for the Kinect data and for the data acquired with 5MP portable camera (see Chapter 4). The annotation of prominent points in the Kinect data, was performed in the type of data where their detection is facilitated. The positioning of the breast peaks, defining breast contour and locating endpoints were annotated in depth-map images, as well as nipple position and suprasternal notch position in the Kinect colour image. For data acquired with portable camera ground truth annotation was performed with BCCT.core software, only for frontal patient data, defining the breast contour, the nipple position and suprasternal notch position. Subjective evaluation for the aesthetic result is also provided, based on the Harvard scale (see Chapter 4).

### 9.5.1 Breast Peak Detection

Breast peak detection was evaluated in 487 depth-map images from 131 patients. In this study only 131 patients were used, since 4 of them were not acquired using the Kinect device. The accuracy of breast peak point detection was measured using the Euclidean metric distance (see Table 9.2).

Metric	Breast	Standalone		Simultaneous detection	
		$\mu$ ( $\sigma$ )	# Miss (%)	$\mu$ ( $\sigma$ )	# Miss (%)
Circ. corr.	Right	18.05 (51.56)	73 (7.50)	10.05 (34.43)	27 (2.77)
	Left	11.39 (26.80)		6.95 (15.49)	
Cross-corr.	Right	12.08 (37.80)	43 (4.41)	4.94 (3.38)	0 (0.00)
	Left	8.71 (19.63)		4.87 (2.88)	

Table 9.2: Breast peak point detection error (in pixels).

The first column (standalone) shows the detection error that would be obtained by making the decision based on maximising the output of the convergence filter. The second column (simultaneous detection) depicts the performance for the proposed scheme. Miss detection means that breast peak point was detected outside the breast area. Therefore, using context in this task is clearly advantageous. Moreover, cross-correlation attains better results both in the mean error and in the miss detection. Since the Kinect resolution is  $1.3\text{mm}/\text{px}$ , the average error corresponds to  $6.38\text{mm}$ . Some results using simultaneous detection and cross-correlation, are shown in Figure 9.15.

### 9.5.2 Breast Contour Endpoints Detection

Breast contour endpoints detection was also evaluated in 487 depth-map images. The endpoint detection accuracy was measured using the Euclidean metric distance (see Table 9.3).

The average error of 12.6 mm is a consequence of the ambiguity of the endpoint concept. Nevertheless, as will be presented next, the impact on the remaining tasks is low. Some results for the contour endpoints detection are shown in Figure 9.16.

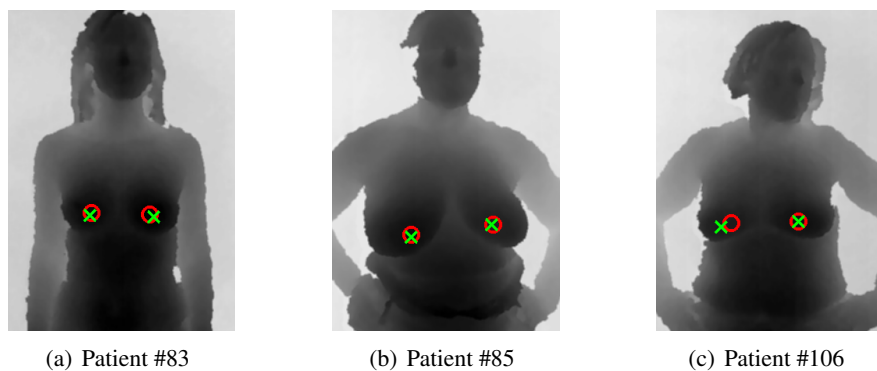


Figure 9.15: Breast peak detection examples. Detected breast peak (red circle) - ground truth breast peak (green cross).

Breast	End Point	$\mu$ ( $\sigma$ )
Right	Outer	9.15 (9.39)
	Inner	10.16 (8.86)
Left	Outer	9.18 (9.03)
	Inner	10.26 (7.89)

Table 9.3: Endpoint detection error (in pixels).

### 9.5.3 Breast Contour Detection

The breast contour detection error using the joint model for the co-occurrence of breast peak and breast contour was evaluated based on the Hausdorff and average distances (Table 9.4). The Hausdorff distance is defined as:

$$h(A, B) = \max_{a \in A} \min_{b \in B} \|a - b\| \quad (9.14)$$

where  $B$  represents the set of pixels of the ground truth and  $A$  is the segmented breast contour;  $\|\cdot\|$  is the Euclidean distance. The motivation for using this metric is that it represents the worst case scenario. As in the previous steps, breast contour detection was evaluated in 487 depth-map images.

	Detected $\Rightarrow$ Ground truth				Ground truth $\Rightarrow$ Detected			
	Right breast		Left breast		Right breast		Left breast	
	Mean	Max	Mean	Max	Mean	Max	Mean	Max
<b>Average</b>	2.51	10.86	2.45	9.64	2.25	10.83	2.17	9.41
<b>Stdev</b>	1.39	8.09	1.65	7.84	1.30	7.10	1.60	7.54
<b>Max</b>	18.37	48.62	17.02	62.85	18.27	46.87	14.22	77.73
<b>Min</b>	0.98	3.03	0.88	2.79	0.76	2.40	0.77	2.18

Table 9.4: Breast contour detection error (in pixels).

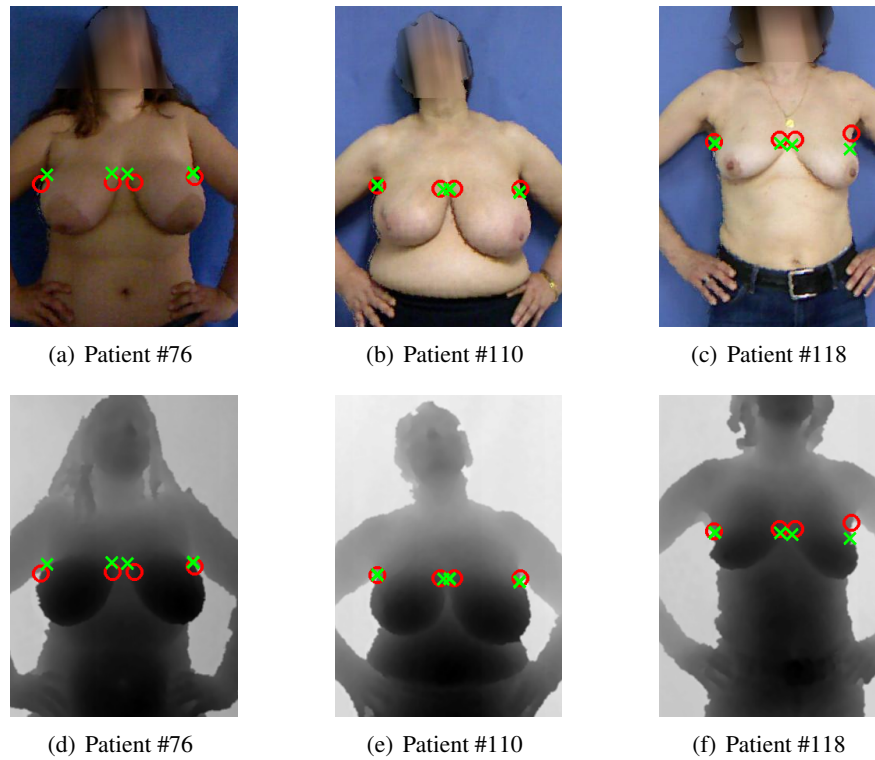


Figure 9.16: Contour endpoints detection examples. Detected endpoints (red circle) - ground truth endpoints (green cross).

From Table 9.4, it is possible to observe that the proposed algorithm presents a very interesting performance. The average error is quite low, around 3mm. It represents the result after the two phases of the algorithm: contour endpoint detection and complete contour definition. Although a relatively high error is found on the location of the endpoints, the breast contour algorithm recovered the correct contour rapidly, thus obtaining a lower average error and a high Hausdorff distance. Some results for the complete contour detection are shown in Figure 9.17.

#### 9.5.4 Nipple Detection

A training set to design a binary classifier was built by manually labelling all pairs of candidates/contours detected in a set of 76 images (yielding 1740 candidates/contour pairs). An SVM classifier was optimised by performing a grid-search on the parameters of the model (C parameter, degree of the polynomial kernel and bandwidth of the RBF kernel) using cross-validation with 4 folds. Finally, the performance of the model was estimated in a test set composed of 248 colour images. When more than one pair nipple/contour is predicted as positive in an image, the prediction with the highest probability is selected (SVMs were implemented with LIBSVM [45], enabling the estimation of the probability).

The cross-validation results in the training set are presented in Table 9.5), for the first ranked parameterisation:.

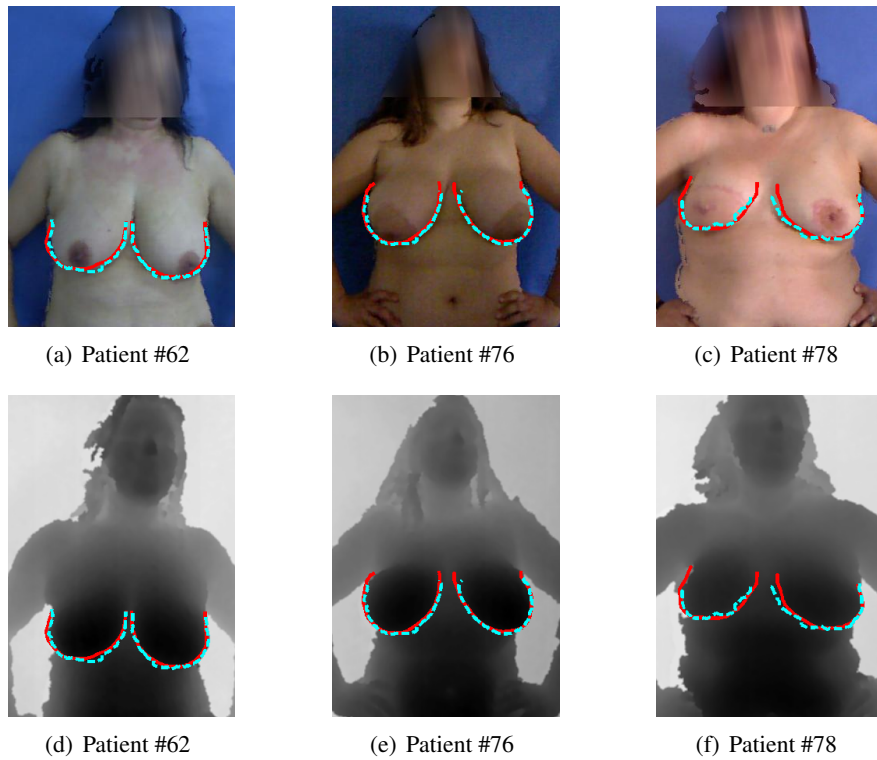


Figure 9.17: Complete contour detection examples. Detected contour (dashed cyan line) - ground truth contour (solid red line).

Features/ Kernel (Parameters)	Linear	Polynomial (degree)		RBF (gamma)				
	-	2	3	0.2	0.4	0.6	0.8	1
$\{S_a\}$	0.14	0.14	0.13	0.14	0.14	0.14	0.14	0.14
$\{\mu(\Delta C_a)\}$	0.19	0.19	0.21	0.19	0.19	0.19	0.19	0.19
$\{\rho_h\}$	0.18	0.19	0.21	0.15	0.16	0.16	0.16	0.15
$\{d_a\}$	0.21	0.18	0.19	0.20	0.18	0.19	0.17	0.18
$\{S_a \mu(\Delta C_a)\}$	0.10	0.09	0.10	0.12	0.11	0.11	0.10	0.11
$\{S_a \rho_h\}$	0.11	0.12	0.36	0.13	0.11	0.11	0.11	0.11
$\{\mu(\Delta C_a) \rho_h\}$	0.17	0.16	0.21	0.12	0.11	0.11	0.11	0.11
$\{\rho_h d_a\}$	0.13	0.17	0.24	0.11	0.11	0.11	0.11	0.11
$\{S_a d_a\}$	0.08	0.08	0.08	0.09	0.08	0.08	0.08	0.08
$\{\mu(\Delta C_a) d_a\}$	0.13	0.15	0.11	0.10	0.10	0.10	0.10	0.09
$\{S_a \mu(\Delta C_a) \rho_h\}$	0.08	0.07	0.14	0.10	0.09	0.08	0.08	0.08
$\{S_a \mu(\Delta C_a) d_a\}$	0.05	0.04	0.05	0.05	0.05	0.05	0.05	0.05
$\{\mu(\Delta C_a) \rho_h d_a\}$	0.13	0.13	0.14	0.11	0.10	0.10	0.09	0.09
$\{S_a \rho_h d_a\}$	0.06	0.08	0.10	0.08	0.07	0.06	0.06	0.06
$\{S_a \mu(\Delta C_a) \rho_h d_a\}$	0.05	0.04	0.16	0.07	0.06	0.06	0.05	0.05

Table 9.5: Miss classification error (cross-validation).



A first observation from Table 9.5 is that more complex models obtained better results. These results are important only to reduce the modes used during testing phase.

The performance in the test set, presented as the average distance between the manually marked nipple and the predicted nipple, using the best sub-set of features, is summarised in Table 9.6. Miss detection means that the nipple was detected outside the areola. The table includes the results for the best parameterisation of each kernel, and for different subsets of features. The best results show an average error around 6 mm and a single detection outside the areola complex.

Features/ Kernel (Parameters)	Linear	Polynomial	RBF
	-	degree (2)	gamma(1)
	$\mu$ ( $\sigma$ ) # Miss (%)	$\mu$ ( $\sigma$ ) # Miss (%)	$\mu$ ( $\sigma$ ) # Miss (%)
$\{S_a \mu(\Delta C_a)\}$	0.80 (1.00) 26 (5.66)	0.78 (0.95) 24 (5.23)	0.79 (0.92) 27 (5.88)
$\{S_a \rho_h\}$	0.84 (1.08) 41 (8.93)	0.73 (0.85) 25 (5.45)	0.84 (1.00) 35 (7.63)
$\{S_a d_a\}$	0.92 (1.40) 40 (8.72)	0.92 (1.41) 39 (8.50)	0.89 (1.35) 37 (8.06)
$\{S_a \mu(\Delta C_a) \rho_h\}$	0.77 (0.96) 35 (7.63)	0.65 (0.62) 11 (2.40)	0.80 (0.92) 31 (6.75)
$\{S_a \mu(\Delta C_a) d_a\}$	0.74 (1.00) 19 (4.14)	0.65 (0.66) 10 (2.18)	0.65 (0.65) 12 (2.61)
$\{S_a \rho_h d_a\}$	0.79 (1.15) 25 (5.45)	0.65 (0.69) 11 (2.40)	0.74 (0.84) 23 (5.01)
$\{S_a \mu(\Delta C_a) \rho_h d_a\}$	0.74 (1.00) 24 (5.23)	0.64 (0.62) 16 (3.49)	0.61 (0.47) 1 (0.22)

Table 9.6: Average distance between the manually marked nipple and the predicted nipple in cm (Results are presented in the format "mean (standard deviation)", and miss classification error.

Some results for the nipple detection are shown in Figure 9.18.

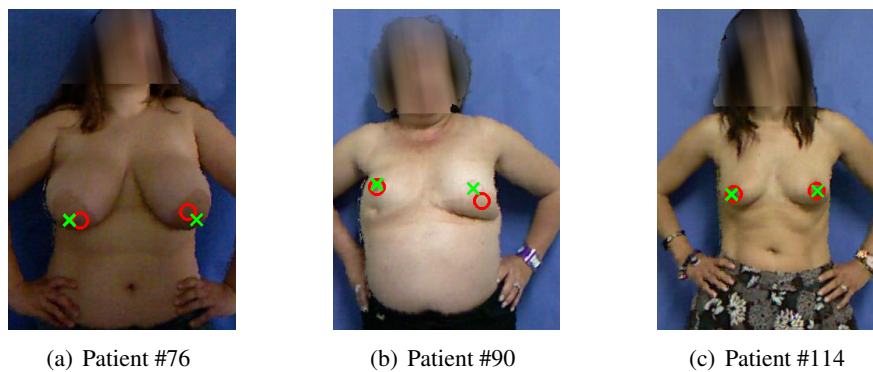


Figure 9.18: Nipple detection examples. Detected nipple (red circle) - ground truth nipple (green cross).

### 9.5.5 Volumetric Feature Definition

This section presents the results for the computation of volumetric features. Patients without NAC were removed from the study, totalling 70 patients, however in 5 of them the nipple height was not measured by the physician. Table 9.7 shows the ratio values obtained for the nipple height, volume, surface area and 3D nipple retraction for each patient, as well the ground truth ratio for the nipple height.

Here, it is presented a direct evaluation of the nipple height feature (based on the ratio between the treated and the untreated breast) by comparison with the measurement taken by the physician. Due to the absence of a ground truth, the quality and effectiveness of the other features (volume, surface and 3D Nipple Retraction) will be evaluated indirectly in Section 9.5.6 through the improvements to the final overall aesthetic evaluation. Additionally, and based on the results obtained in Section 8.4, nipple height was also computed in SR resolution images ( $\times 2$  the initial resolution) obtained with the SR algorithm that obtained best performance.

Figure 9.19 presents a comparison between the ratios detected automatically and the ground truth ratios for the nipple height feature, for the 65 patients where the physician performed the ground truth measurement.

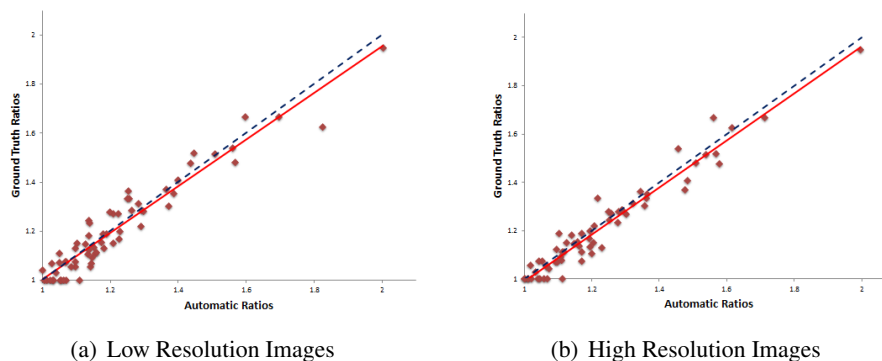


Figure 9.19: Scatter plot of depth height ratios. Red line is the trendline for graph data. Blue dashed line represents the ideal trendline.

The ratios automatically detected are very similar to those obtained by the physician. This can make clinical experts approve the use data acquired with Kinect device to evaluate the aesthetic outcome. Aspects that can interfere with the data include smaller patient breasts or similar breast shapes. It is also important to highlight the errors introduced by the physician during manual ground truth measurement. Comparing the results obtained with the LR and SR images, the second presented better results, although they are not significant. With LR images a correlation of 0.92 was obtained, whilst with SR images obtained a value of 0.95.

Patient #	Nipple GT	Nipple ratio	Nipple ratio (SR)	Volume ratio	Surface area ratio	3D Nipple Retraction
52	1.08	1.07	1.11	1.07	1.66	1.09
53	1.19	1.19	1.10	1.28	1.81	1.05
54	1.35	1.39	1.36	1.33	2.40	1.03
55	1.95	2.00	2.00	1.33	1.28	1.07
61	-	1.18	1.07	1.25	2.27	1.06
62	-	1.25	1.37	1.11	1.99	1.03
63	-	1.10	1.12	1.73	1.54	1.17
64	-	1.14	1.23	1.15	1.93	1.01
66	1.36	1.25	1.34	1.52	1.56	1.08
67	1.00	1.11	1.04	1.40	2.08	1.03
69	-	1.06	1.05	1.14	2.29	1.04
70	1.11	1.05	1.11	2.06	1.82	1.05
71	1.63	1.83	1.62	1.15	1.94	1.02
73	1.09	1.15	1.11	1.44	2.28	1.02
74	1.24	1.14	1.25	1.41	1.98	1.03
75	1.15	1.10	1.21	1.43	2.09	1.08
76	1.00	1.01	1.01	1.11	1.97	1.00
77	1.23	1.14	1.28	1.00	1.25	1.03
78	1.48	1.44	1.58	1.98	1.38	1.04
79	1.06	1.10	1.02	1.58	2.09	1.01
80	1.07	1.03	1.09	1.00	1.62	1.02
81	1.05	1.14	1.06	1.51	1.79	1.02
83	1.04	1.00	1.07	1.23	1.50	1.02
84	1.15	1.13	1.15	1.13	1.64	1.03
85	1.19	1.18	1.17	1.48	1.59	1.02
86	1.37	1.36	1.48	1.22	1.57	1.01
87	1.31	1.28	1.32	1.06	2.29	1.11
88	1.27	1.21	1.30	1.10	1.62	1.04
89	1.33	1.25	1.22	1.01	1.34	1.00
90	1.67	1.60	1.56	1.02	1.25	1.09
91	1.52	1.45	1.57	1.05	1.20	1.00
92	1.13	1.10	1.20	1.36	1.26	1.20
93	1.67	1.70	1.71	1.19	1.18	1.20
94	1.00	1.01	1.05	1.08	1.58	1.01
95	1.00	1.05	1.01	1.06	1.24	1.01
96	1.08	1.07	1.17	1.06	1.20	1.02
98	1.00	1.03	1.06	1.13	1.62	1.03
99	1.29	1.29	1.29	1.07	1.67	1.02
100	1.07	1.05	1.05	1.75	1.94	1.04
101	1.07	1.14	1.10	1.24	1.71	1.00
103	1.11	1.16	1.17	1.29	2.14	1.02
104	1.13	1.18	1.23	1.09	1.41	1.01
106	1.41	1.40	1.48	1.32	1.51	1.11
107	1.06	1.09	1.07	1.32	2.00	1.03
108	1.17	1.22	1.19	1.17	2.13	1.04
109	1.00	1.07	1.07	1.54	1.24	1.01
110	1.00	1.06	1.11	1.19	1.61	1.09
113	1.28	1.20	1.28	3.02	2.31	1.01
114	1.00	1.03	1.00	1.14	1.92	1.01
115	1.48	1.57	1.51	1.84	1.97	1.06
116	1.33	1.25	1.36	1.08	1.56	1.03
117	1.14	1.15	1.16	1.67	1.49	1.06
118	1.18	1.14	1.14	1.81	1.60	1.08
119	1.11	1.13	1.20	1.38	1.74	1.04
120	1.52	1.51	1.54	1.73	1.53	1.06
121	1.00	1.02	1.02	2.07	1.78	1.03
122	1.08	1.10	1.04	1.05	1.24	1.17
123	1.54	1.56	1.46	1.37	1.03	1.07
124	1.29	1.26	1.29	1.04	1.64	1.04
125	1.27	1.22	1.26	1.36	2.65	1.06
126	1.30	1.37	1.36	1.62	1.83	1.06
127	1.15	1.17	1.16	1.02	2.27	1.01
128	1.03	1.04	1.03	2.03	1.38	1.05
129	1.00	1.06	1.00	1.54	1.75	1.04
130	1.15	1.21	1.13	1.73	1.63	1.03
131	1.22	1.29	1.21	1.16	1.53	1.01
132	1.11	1.16	1.12	1.12	1.93	1.04
133	1.20	1.23	1.20	1.32	1.20	1.08
134	1.28	1.30	1.25	1.58	2.05	1.02
135	1.12	1.14	1.10	1.14	1.56	1.01

Table 9.7: Volumetric feature values.

### 9.5.6 Aesthetic Evaluation

The SR methods presented in this thesis did not seem to improve the result and therefore they were not included in this section.

As in previous work (see Section 6.1), the problem involves classifying examples into classes

which have a natural order. Thus, ordinal SVMs classifiers were adopted based on the data replication method for ordinal data [29].

### 9.5.6.1 Database Analysis

First of all, subjective evaluation is necessary to build and compare the models. As presented in Chapter 4, the subjective evaluation was provided by two experts in the field of breast cancer. Table 9.8 shows the distribution of the 70 patients over the four classes for the two evaluators.

Class		Excellent	Good	Fair	Poor	Total
# cases	Phy. #1	24	26	15	5	70
	Phy. #2	47	18	4	1	

Table 9.8: Distribution of the 70 patients over the four classes.

By looking at Table 9.8, it is easy to understand that agreement between the experts is low. Moreover, class #4 is not significantly represented. For that reason, there was also an attempt to create models using only three classes, merging class #3 and #4. Table 9.9 shows the new distribution of the 70 patients over the classes.

Class		Excellent	Good	Fair	Total
# cases	Phy. #1	24	26	20	70
	Phy. #2	47	18	5	

Table 9.9: Distribution of the 70 patients over the three classes.

To better understand the relation between the assessment between the experts, the disagreement, the correlation coefficient and Cohen's kappa coefficient are shown in Table 9.10.

# of classes	Disagreement	Correlation Coefficient	Kappa	Weighted Kappa
4	0.59	0.62	0.11	0.28
3	0.56	0.54	0.15	0.28

Table 9.10: Disagreement between evaluators for the subjective evaluation of the 70 patients.

From Table 9.10 it easy to observe that agreement is not high, which highlights the need for a panel of experts to perform the evaluation. As in Section 4, the Weighted Kappa coefficient  $wk$  was used. For the analysis with three classes, the weight matrix was set as follows:

$$wk = \begin{bmatrix} 1 & 2/3 & 1/3 \\ 2/3 & 1 & 2/3 \\ 1/3 & 2/3 & 1 \end{bmatrix}$$

### 9.5.6.2 Previous Models Analysis

The BCCT.core model and the new Linear model presented in Section 6.1 were not designed for the data presented in the database. Those models were designed only for patients who underwent a BCCT surgical intervention, and our database includes patients who have been subjected to other type of interventions. In any way, using the data acquired from the portable camera, it was decided to evaluate these models (see Table 9.11), as well as compare the output between each model (see Table 9.12) and compare them with the subjective evaluation performed by the specialists (see Table 9.13 and Table 9.14).

Class		Excellent	Good	Fair	Poor	Total
# cases	4 classes	BCCT.core	26	31	13	0
		Linear	33	24	9	4
	3 classes	BCCT.core	26	31	13	-
		Linear	33	24	13	-

Table 9.11: Distribution of the 70 patients for BCCT.core and linear model.

# of classes	Miss Classification	Correlation Coefficient	Kappa	Weighted Kappa
4	0.30	0.72	0.54	0.61
3	0.26	0.74	0.60	0.66

Table 9.12: Disagreement between output from BCCT.core and linear model.

Model	Phy. #	Miss Classification	Correlation Coefficient	Kappa	Weighted Kappa
BCCT.core	1	0.51	0.48	0.23	0.33
	2	0.47	0.64	0.25	0.38
Linear	1	0.46	0.62	0.33	0.46
	2	0.41	0.65	0.29	0.42

Table 9.13: Disagreement between output from previous model and the subjective evaluation performed by the physicians with 4 classes.

Model	Phy. #	Miss Classification	Correlation Coefficient	Kappa	Weighted Kappa
BCCT.core	1	0.47	0.46	0.28	0.36
	2	0.46	0.65	0.27	0.39
Linear	1	0.43	0.54	0.35	0.43
	2	0.40	0.61	0.31	0.42

Table 9.14: Disagreement between output from previous model and the subjective evaluation performed by the physicians with 3 classes.

One first comment from previous Tables is that the BCCT.model did not classify any patient with "Poor" class, which was very similar to the subjective evaluation from Physician #2. From the comparison between these two models (BCCT.core model and the linear model designed in Section 6.1), it is possible to state that they are very similar, presenting a Miss Classification around or below than 30%, which is similar to that obtained in previous studies [28, 180]. The high agreement between these models is validated with the weighted Kappa value, which is above 60%. Comparing the output of these models with the subjective evaluation from physicians, in terms of Miss Classification the best agreement is between Physician #2 and the Linear model with a value around 40%. By looking at weighted Kappa values, the best agreement is between Physician #1 and the Linear model with a value around 45%. This analysis is very important to understand how much the output classifications are related: however, new models have to be created, not only to incorporate 3D information but also for the 2D analysis - which happens because Kinect RGB data differ substantially from previous data used in the BCCT.core -, but also to run the automatic detection of fiducial points developed in this thesis.

### 9.5.6.3 Feature Selection and Models Creation

Before designing a new model, it was necessary to select the set of features to be used in the study. The selection of 2D features was based on previous studies [28, 180], whereas the 3D features were those presented in this work.

The model was designed using the selected features, considering all the possible subsets, using a LOO scheme [65]. The subjective evaluation performed by the physicians was used as reference to train the models. Since the two experts did not reach consensus, the evaluation was performed separately for each of the physicians. Models considering three and four classes were developed for the first physician evaluation, while for the second physician the evaluation was only performed with three classes due to the lack of patients evaluated as "Poor". Taking into account previous work on the same problem, the SVMs proved to be a suitable solution, mostly due to the limited amount of data available [30, 31]. Linear and RBF kernel were tested, performing a "grid-search" on the parameters of the models (h and s parameters of the data replication method were left constant at 1 and 2, respectively). Exponentially growing sequences of  $C$  were tested:  $C = 2^0, 2^2, 2^4, 2^6$ , while  $\gamma$ , for the RBF kernel, was tested with:  $\gamma = 0.125, 0.25, 0.5$ .

The test results concerning the misclassification error (MER), the correlation (Corr) and Kappa statistics for the first ranked feature subset are summarised in Table 9.15 for 2D features, Table 9.16 for the cascade scheme and Table 9.17 for the cascade scheme and Table 9.17 for the non-restricted scheme, which combines 2D and 3D features.

The discrimination of the features are presented in TABLE 9.18.

In some of the tests, more than one subset of features obtained the same miss-classification error. The Tables present the subset with less complexity. A first observation is that the introduction of 3D features improved the aesthetic evaluation. Generally, as expected, 3D features contribute to a correct evaluation, both for Physician #1 and #2. Particularly, the result obtained for Physician #2 is remarkable: a miss detection of 0.14, which is very low in comparison with the error

GT	# classes	Kernel	C	$\gamma$	MER	Corr	Kappa	wKappa	Feat. Set
Phy.#1	4	RBF	64	0.5	0.44	0.58	0.33	0.43	[10,12,13,14,16]
	3	RBF	64	0.5	0.36	0.51	0.45	0.47	[10,12,13,14,16]
Phy.#2	3	RBF	64	0.25	0.20	0.66	0.54	0.58	[8,10,12,16,24]

Table 9.15: Results obtained considering 2D features only.

GT	# classes	Kernel	C	$\gamma$	MER	Corr	Kappa	wKappa	Feat. Set
Phy.#1	4	RBF	64	0.5	0.40	0.63	0.40	0.50	[34]
	3	RBF	64	0.5	0.36	0.51	0.45	0.47	[34]
Phy.#2	3	RBF	64	0.25	0.17	0.71	0.59	0.63	[33]

Table 9.16: Results obtained for the cascade scheme.

GT	# classes	Kernel	C	$\gamma$	MER	Corr	Kappa	wKappa	Feat. Set
Phy.#1	4	RBF	64	0.5	0.43	0.63	0.41	0.51	[9,10,13,14,16,34]
	3	RBF	64	0.5	0.36	0.48	0.45	0.46	[9,10,13,14,16,34]
Phy.#2	3	linear	64	-	0.14	0.81	0.68	0.73	[9,24,32,33,34]

Table 9.17: Results obtained for the non-restricted scheme of 2D and 3D features.

Table 9.18: Selected features.

#	Acronym	Description
8	pBRA	Dimensionless Breast Retraction Assessment
9	pLBC	Dimensionless Lower Breast Contour
10	pUNR	Dimensionless Upward Nipple Retraction
12	pBCD	Dimensionless Breast Contour Difference
13	pBAD	Dimensionless Breast Area Difference
14	pBOD	Dimensionless Breast Overlap Difference
16	$c\chi_a^2$	$\chi^2$ distance between the histograms in the $a$ channel of the CIE $L^*a^*b^*$
24	$s\chi_a^2$	surgical scar visibility measured with $\chi^2$ in the $a$ channel of the CIE $L^*a^*b^*$
32	VDA	Volume Difference Assessment
33	SAA	Surface Area Assessment
34	3DNRA	3D Nipple Retraction Assessment

around 30% of the BCCT.core [28]. However, this result may be biased by the evaluation of this specialist, corroborating the need for a panel of experts to reach consensus. The cascade scheme, using the aesthetic evaluation from Physician #1 as reference, and using three classes only, did not improve the initial 2D result. From the results, it is possible to observe that 3DNRA and SAA features appear in almost all selected models, meaning that they have a greater influence on the aesthetic evaluation. However, the results need to be read with caution. More data, both in number and in diversity, with a more robust reference evaluation by specialists, are required to validate the

results.

## 9.6 Conclusion

The work presented in this Chapter demonstrated the use of low-cost equipment, the Kinect device, to build a system for the aesthetic evaluation of BCCT with improved automation, robustness, reproducibility and overall cosmetic accuracy. This work takes advantage of both the depth and RGB information acquired with the Kinect. The two are combined in order to take advantage of the best of each data modality.

Firstly, depth information is used to extract the complete breast contour and the breast peak point. With this information, after an appropriate RGB-D joint calibration, the breast contour is superimposed on the RGB image to help detect the nipples. All these prominent points are detected automatically, as well as the volumetric information obtained afterwards on the depth map images. It was demonstrated that depth-map images facilitate the automation of the overall system, maintaining BCCT.kore an affordable and easy-to-use tool. Results also confirm an excellent performance and robustness for a wide variety of patients.

Afterwards, it was confirmed that the detected volumetric features on Kinect data (nipple height, volume, surface area and 3D nipple retraction) have enough quality for aesthetic quantification. For the particular case of the breast height, the comparison with measurements performed by physicians provided evidence of the process reliability. The results prove that it is possible to extract reliable 3D information from the Kinect, improving the aesthetic evaluation comparatively to the subjective evaluation performed by specialists.

Regarding the two schemes tested for the inclusion of 3D information, it seems natural that the model integrating 2D and 3D features should present better results. Looking at the results, it is possible to confirm that this happens when the reference classification used is from the Physician #2, but not from the first physician. To understand the impact of 3D features, some examples will be presented next, by analysing the classification output performed by the different models. The examples were chosen based on visual inspection. The chosen patients were the ones whose breast deformation is easier to detect using volumetric information. Figure 9.20 presents the frontal photographs of three patients, and Table 9.19 shows the different model output, as well the subjective evaluation performed by the physicians.

The subjective classification performed by the Physician #1 classified all the three examples as "Poor", while Physician #2 evaluated two of them as "Fair" and the other as "Good". A first observation, is that there is no consensus in some patients. From the results presented in Table 9.19, considering the reference classification by Physician #1, it is possible to confirm that models that used only 2D information evaluated all the patients as "Fair", which is different from the reference classification. With the inclusion of 3D information, both with Cascade and the normal scheme, the classification now matches the reference. With Physician #2 as reference, the behaviour is not linear. For example, the initial classification of patient #122 only with 2D features is "Good" and the reference is "Fair". While the cascade model did not alter the classification value, the normal



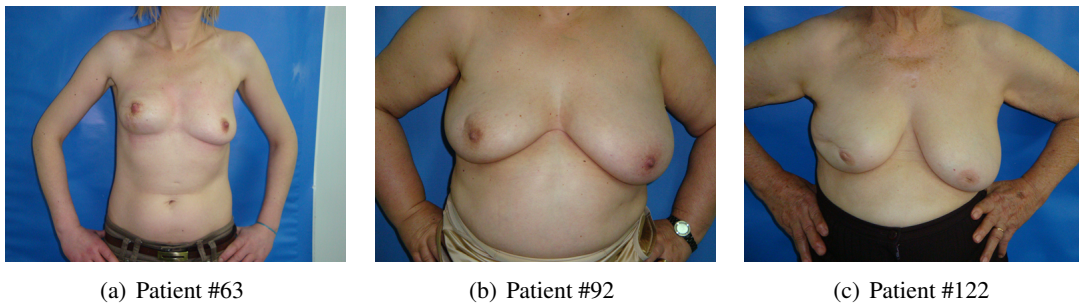


Figure 9.20: Patient examples.

	Automatic Classification							
Model	2D		Cascade		2D+3D		GT	
Phy. #	Phy. #1	Phy. #2	Phy. #1	Phy. #2	Phy. #1	Phy. #2	Phy. #1	Phy. #2
# classes	4	3	4	3	4	3		
Pat.#63	3	3	4	3	4	3	4	3
Pat.#92	3	2	4	2	4	3	4	2
Pat.#122	3	2	4	2	4	3	4	3

Table 9.19: Impact of 3D features. The output of the models trained with three classes and with reference assessment performed by Physician #1, were removed from this analysis as they were not relevant.

scheme corrects it to the "true" value. On the other hand, for patient #92 the classification with 2D features matches the reference, and when 3D features are added the classification is changed. From the global results, it is possible to state that 3D features help the overall classification of the aesthetic result; however, more tests, with a broader database with different cases, must be performed to validate the obtained results.



## Chapter 10

# Conclusions

Breast size and shape are a significant part of female body image and sense of femininity, and they are determined by physical characteristics such as tissue volume, skin dimensions and chest wall circumference. This is a dynamic structure that is easily deformed by position, gravity and external pressure. The configuration of a mature breast in an individual, changes in response to physiologic alterations (such as pregnancy, menstrual cycle) and age.

Breast Cancer Treatments alter the shape of the tissues. While some treatments change considerably the size of the breast, some alter not so much the size, but the shape and location of the Nipple-Areola Complex. Some procedures result in a combination of these effects; however, the most recent techniques, in particular the oncoplastic ones, try to recreate the entire breast (for instance, after mastectomy) using breast implants or tissue transferred from other parts of the body. In the conservative approach, however, approximately one third of the patients will have a fair or poor aesthetic outcome, affecting directly the psychosocial recover and, consequently, patients' QOL.

The lack of a standard method for measuring the aesthetic outcome has been considered an obstacle in the assessment and evaluation of the techniques applied. Until recently most used methods were based on a subjective evaluation, made by one or more observers by visual inspection; however, these methods' lack of reproducibility led to the introduction of objective methods. They are based on measurements taken directly from the patient or from patient photographs which that are essentially based on asymmetries between treated and non-treated breasts. In order to overcome the idea that objective asymmetry measurements were insufficient, other groups introduced a method based on the sum of the individual scores of subjective and objective indices. All these additions improved the aesthetic evaluation methods but were considered insufficient due to a great intra- and inter-observer variability. This fact corroborates the idea that it is necessary to replace or enhance human expert evaluation of the aesthetic result of BCCT. A possible solution should be supported by an objective tool, which has to be easy to use, highly reproducible and acceptable by those who would be evaluated. To overcome the lack of objectivity of the evaluation of the aesthetic outcome of BCCT, a semi-automatic tool named BCCT.core was developed, which is based on the comparison between breasts on frontal patient photographs. Although the BCCT.core sys-

tem obtained satisfactory results, it presents some limitations related with interpretability of the models and the fact that only patient frontal photographs are used.

The female breast is a complex 3D object and its boundaries are rather fuzzily defined in 2D pictures, thus making it difficult to identify body landmarks. On the contrary, the use of a 3D model would make it possible to compare real geometrical characteristics of the breasts including the possibility of estimating volume and 3D surface differences. Some authors have already tried using 3D methodologies in this field and, although the benefits of these methods are known, some drawback has been identified. The main drawback of 3D techniques is the need for specific hardware, software and personnel. The high cost and the difficulty of operating the equipment on a daily basis prevents its widespread use. Furthermore, these kinds of techniques do not try to predict the aesthetic result for a more informed choice of treatment, neither are they able to provide an automatic evaluation of the aesthetic result.

There was two principal focus on this thesis, regarding the aesthetic evaluation of the cosmetic result. Firstly, compare the BCCT.core model with new interpretable models and models that incorporate side-views from the patient. Secondly, develop a simple 3D model of a female torso, using low-cost solutions, namely: a reconstruction algorithm from two uncalibrated views, using the epipolar geometry approach and a Kinect sensor device. The idea is that the created model will be used in an updated version of a BCCT.core to obtain a full 3D aesthetic assessment of the surgical outcome. With the inclusion of measurements extracted from the 3D model, the aim is to improve the global assessment result, without increasing its complexity, as the pictures are acquired with a single camera or with a portable device.

This work was supported by the creation of an original multi-modal database. This database could be used in the future by breast cancer researchers in order to compare different methodologies. The database created is a unique tool and can lead to new developments in the area.

In the initial study, interpretable models were proved to be an alternative to the current BCCT.core model, since similar results were obtained with simpler models. The inclusion of lateral-views to perform an aesthetic evaluation of the cosmetic result, was introduced for the first time during this thesis. Results have demonstrated that this type of images can be useful to perform a correct evaluation; however, the improvements have not been significant. More information on the shape of the breasts is required, adding dimensionality with the acquisition of volumetric information.

An uncalibrated methodology based on epipolar geometry was used in order to maintain the acquisition of patient data with a normal camera. The idea was to acquire several views from the patient using the same camera, and reconstruct the scene only using the information on the photographs, without any calibration procedure. After applying uncalibrated state-of-the-art algorithms to our database, it was possible to conclude that current methodologies are not able to correctly reconstruct the breast. The main reason to this drawback is related to the lack of features in the breast and the repetitive texture which is typical of the skin. These methodologies do not work also when features are created artificially on a female phantom torso. This indicates that more investigation is needed, with specific algorithms designed to this kind of data, incorporating parametric models introducing some restrictions on the shape of the breast, or using specific pat-

terns to generate artificial features, trying to simplify the reconstruction process. The volumetric methodology presented here was based on this second approach.

Aiming to provide an affordable and practical tool, and taking advantage of the available low-cost systems on the market, a decision was made to use the Microsoft Kinect device. Note that the Kinect was introduced after an initial validation of Kinect depth data, which proved to be accurate to measure volumetric differences in this specific data, the limitations and capabilities of Kinect device were studied.

One of the great assets of this device is that it provides synchronised depth and colour data, which is very useful to detect specific characteristics simultaneously. However, depth and colour information are not aligned, requiring a joint calibration stage, specific for each device. This step was successfully addressed, using one of the latest state-of-the-art algorithms. It is well acknowledged that these devices are of great value; however, these advantages are confronted with the relatively LR of generated disparity, which can be a limitation to produce a clear model. This disadvantage was tackled with a comparative study of state-of-the-art SR algorithms applied to Kinect data. The study conducted proved that SR algorithms produced output images with good quality; however, the algorithms that achieved the best results are computationally expensive, which prevents them from being used in real-time applications.

Lastly, and looking to the BCCT.core limitations concerning the complete automation and inability to measure volumetric information, new algorithms were developed to detect fiducial points and breast contour without user intervention and new models were created for the aesthetic evaluation of the cosmetic result incorporating volumetric information. The prominent points were detected using synchronised depth and colour information. Each characteristic point was detected on the type of data where their detection is facilitated.

This innovation makes it easier to automate the software without increasing the complexity of the system, enhancing the reproducibility of the software since the user does not have any intervention during the process. The process of locating the nipple is based on detecting the aureola, using a closed contour methodology. This thesis introduced an algorithm which computes the closed contour in the original coordinates, in contrast to the typically approach which transforms the image into polar coordinates and the closed contour is transformed into an open contour between two opposite margins. The volumetric information is obtained using fiducial point localisation and breast contour delimitation in the depth image. Finally, using 2D features and the new volumetric information, new models were developed to evaluate the aesthetic result. With the inclusion of measurements extracted from the 3D model, it was possible to improved the overall assessment result, without increasing complexity significantly. The results proved that it is possible to extract reliable 3D information, comparatively to the subjective evaluation performed by physicians.

During this thesis, the work has been published in several technical conferences and recently in specialised journals. The work was also validated by the medical community, with the publication of several papers in top medical conferences [159–161].

## 10.1 Future Work

Regarding the contributions presented in this thesis, we identified the necessity of developing specific 3D reconstruction algorithms for uncalibrated environments, to be used in our data. Furthermore, we propose the acquisition of video frames using a Kinect device, in order to create a real 3D model, obtained from the registration of the depth images.

It is difficult to predict exactly how the breast will be changed by a specific procedure in any particular patient. How long should incisions be and where should they be placed? What size of breast implant will yield the best results? Is there enough tissue to recreate a breast that would meet the expectations of the patient? What does the patient want her breast to look like? Answering these questions is essential in pre-operative planning.

The uncertainties associated with breast surgery may be reduced by applying modelling and simulation process. A computer-aided breast simulator may enhance the practice of breast surgery at multiple points, allowing the patient to communicate her expectations more clearly to the surgeon. It allows the surgeon to educate the patient with more accurate explanations of what can be accomplished, and after, it helps the surgeon plan specific aspects of the procedure so as to achieve the agreed and desired goals. Finally, it facilitates surgical training by allowing trainees to design procedures and understand the results before performing surgery on the patient.

The development of a model fitting method will allow the system to automatically fit a generic deformable model to patient specific 3D breast surface measurements using a physically-based framework. This can be used to quantitatively and reliably assess the aesthetic outcomes of breast reconstructive surgery. In addition, this will also allow the surgeon to quantitatively analyse the degrees of various deformities and asymmetries in the shape of the breast. Finally, a model creation mode will allow the surgeon to interactively adjust the shape of the breast by varying key shape variables analogous to the aesthetic and structural elements surgeons inherently vary manually during breast reconstruction.

Given the triangular mesh data of the patient's breast, the model fitting method could be used to fit a parametric breast model to the data mesh. One aim is to use the reference shape to capture the salient shape features that are innate to the object being modelled and appropriate for comparison with other object instances. Furthermore, the deformations can be used in the reconstruction of fine scale structures and natural irregularities of real world data that the reference shape disregards. The stated reference shape could be obtained from the result of applying global deformations to a geometric primitive. The geometric primitive could be for example defined parametrically, employing a superquadric.

As a final point, it should be noted that has been proposed an European health project and has been accepted for funding. Entitled "PICTURE - Patient Information for the Combined Assessment of Specific Surgical Outcomes in breast cancer", the project combines the analysis of 3D pictures with routine medical examinations which, together, will make it possible to develop techniques for biomechanically modeling the anatomy of the breast, and to understand the effect that the surgery to remove cancerous tissue can have on the patient. With these tools it will be possible

to explore new alternative strategies and to evaluate the consequences of the options available, always respecting the appearance of the breast. This is a three-year project developed in partnership with Philips, the University College London (UK), the Academisch Ziekenhuis Leiden - Leids Universitair Medisch Centrum (the Netherlands) and the Institut National de la Santé et de la Recherche Médicale (France). The opportunity to collaborate with these partners should be seen as a valuable tool to further reinforce the competencies and knowledge of the research group.





# References

- [1] R. Abadir and P. Hornbostel. Breast carcinoma treated by conservative surgery: results of postoperative external radiotherapy with photons only and a nonsplit supraclavicular field. *Am J Clin Oncol*, 24(2):120–123, 2001.
- [2] A. L. Abner, A. Recht, F. A. Vicini, B. Silver, D. Hayes, S. Come, and J. R. Harris. Cosmetic results after surgery, chemotherapy, and radiation therapy for early breast cancer. *International Journal of Radiation Oncology\*Biolog\*Physics*, 21(2):331–338, 1991.
- [3] S. K. Al-Ghazal and R. Blamey. Cosmetic assessment of breast-conserving surgery for primary breast cancer. *The Breast*, 8(4):162–168, 1999.
- [4] S. K. Al-Ghazal, R. W. Blamey, J. Stewart, and A. L. Morgan. The cosmetic outcome in early breast cancer treated with breast conservation. *European Journal of Surgical Oncology*, 25(6):566–570, 1999.
- [5] S. K. Al-Ghazal, L. Fallowfield, and R. W. Blamey. Patient evaluation of cosmetic outcome after conserving surgery for treatment of primary breast cancer. *European Journal of Surgical Oncology*, 25(4):344–346, 1999.
- [6] B. B. Alagoz. Obtaining depth maps from color images by region based stereo matching algorithms. *OncuBilim Algorithm And Systems Labs*, 8(4):1–13, 2008.
- [7] E. L. Allwein, R. E. Schapire, and Y. Singer. Reducing multiclass to binary: A unifying approach for margin classifiers. In *Proc. 17th International Conf. on Machine Learning*, pages 9–16, 2000.
- [8] M. Amichetti, L. Busana, and O. Caffo. Long-term cosmetic outcome and toxicity in patients treated with quadrantectomy and radiation therapy for early-stage breast cancer. *Oncology*, 52(3):177–181, 1991.
- [9] D. Anguelov, P. Srinivasan, D. Koller, S. Thrun, J. Rodgers, and J. Davis. Scape: shape completion and animation of people. *ACM Trans. Graph.*, 24(3):408–416, 2005.
- [10] K. S. Asgeirsson, T. Rasheed, S. J. McCulley, and R. D. Macmillan. Oncological and cosmetic outcomes of oncoplastic breast conserving surgery. *European Journal of Surgical Oncology*, 31(8):817–823, 2005.
- [11] D. Ash, E. Benson, J. Sainsbury, C. Round, and C. Head. Seven-year follow-up on 334 patients treated by breast conserving surgery and short course radical postoperative radiotherapy: A report of the yorkshire breast cancer group. *Clinical Oncology*, 7(2):93–96, 1995.

- [12] A. Bajaj, P. Kon, K. Oberg, and D. Miles. Aesthetic outcomes in patients undergoing breast conservation therapy for the treatment of localized breast cancer. *Plast Reconstr Surg*, 114(6):1442–1449, 2004.
- [13] R. Balaniuk, I. Costa, and J. Mello. Cosmetic breast surgery simulation. In *VIII Symposium on Virtual Reality*, pages 387–396, 2006.
- [14] G. Baroni, M. Riboldi, and P. Patete. System and method for advanced scanning and for deformation simulation of surfaces, wo/2010/073129, July 2010.
- [15] N. Barzigar, A. Roozgard, S. Cheng, and P. Verma. A robust super resolution method for video. In *Proceedings Asilomar Conference on Signals, Systems, and Computers*, pages 1–5, 2012.
- [16] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool. Speeded-up robust features (surf). *Comput. Vis. Image Underst.*, 110(3):346–359, 2008.
- [17] H. Bay, T. Tuytelaars, and L. J. V. Gool. Surf: Speeded up robust features. In *The European Conference on Computer Vision*, pages 404–417, 2006.
- [18] G. F. Beadle, S. Come, I. C. Henderson, B. Silver, S. Hellman, and J. R. Harris. The effect of adjuvant chemotherapy on the cosmetic results after primary radiation treatment for early stage breast cancer. *International Journal of Radiation Oncology Biology Physics*, 10(11):2131–2137, 1984.
- [19] G. F. Beadle, B. Silver, L. Botnick, S. Hellman, and J. R. Harris. Cosmetic results following primary radiation therapy for early breast cancer. *Cancer*, 54(12):2911–2918, 1984.
- [20] C. Bert, K. G. Metheany, K. Doppke, and G. T. Y. Chen. A phantom evaluation of a stereo-vision surface imaging system for radiotherapy patient setup. *Medical Physics*, 32(9):2753–2762, 2005.
- [21] A. Bhatti and S. Nahavandi. Wavelets/multiwavelets in stereo correspondence estimation: a comparative study. In *Digital Image Computing: Techniques and Application*, pages 309–316, 2009.
- [22] C. M. Bishop. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2006.
- [23] J. H. Borger and A. H. Keijser. Conservative breast cancer treatment: analysis of cosmetic results and the role of concomitant adjuvant chemotherapy. *International Journal of Radiation Oncology Biology Physics*, 13(8):1173–1177, 1987.
- [24] K. M. Borgwardt. *Graph Kernels*. PhD thesis, Ludwig-Maximilians University, Germany, 2007.
- [25] P. Boyle and B. Levin. *World Cancer Report*. International Agency for the Research on Cancer, 2009.
- [26] G. Bradski and A. Kaehler. *Learning OpenCV*. O’Reilly Media Inc., 2008.
- [27] M. Brown, R. Szeliski, and S. Winder. Multi-Image Matching Using Multi-Scale Oriented Patches. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 510–517, 2005.

- [28] J. S. Cardoso and M. J. Cardoso. Towards an intelligent medical system for the aesthetic evaluation of breast cancer conservative treatment. *Artificial Intelligence in Medicine*, 40(2):115–126, 2007.
- [29] J. S. Cardoso and J. F. P. da Costa. Learning to classify ordinal data: the data replication method. *Journal of Machine Learning Research*, 8:1393–1429, 2007.
- [30] J. S. Cardoso, J. F. P. da Costa, and M. J. Cardoso. Modelling ordinal relations with SVMs: an application to objective aesthetic evaluation of breast cancer conservative treatment. *Neural Networks*, 18(5–6):808–817, 2005.
- [31] J. S. Cardoso, J. F. P. da Costa, and M. J. Cardoso. SVMs applied to objective aesthetic evaluation of conservative breast cancer treatment. In *Proceedings of the International Joint Conference on Neural Networks*, pages 2481–2486, 2005.
- [32] J. S. Cardoso, H. P. Oliveira, and I. Domingues. Closed shortest path in images: Avoiding unnecessary transformation. 2013. SUBMITTED.
- [33] J. S. Cardoso, R. Sousa, L. F. Teixeira, and M. J. Cardoso. Breast contour detection with stable paths. In A. Fred, J. Filipe, and H. Gamboa, editors, *International Joint Conference on Biomedical Engineering Systems and Technologies*, pages 439–452, 2009.
- [34] J. S. Cardoso, L. F. Teixeira, and M. J. Cardoso. Automatic breast contour detection in digital photographs. In *Proceedings of the International Conference on Health Informatics*, pages 91–98, 2008.
- [35] M. J. Cardoso. Is three better than two? the use of 3d scanners in the assessment of aesthetic results in local breast cancer treatment. *The Breast*, 21(3):227–228, 2012.
- [36] M. J. Cardoso, J. S. Cardoso, A. C. Santos, H. Barros, and M. C. Oliveira. Interobserver agreement and consensus over the esthetic evaluation of conservative treatment for breast cancer. *The Breast*, 15(1):52–57, 2006.
- [37] M. J. Cardoso, J. S. Cardoso, A. C. Santos, C. Vrieling, D. Christie, G. Liljegren, I. Azevedo, J. Johansen, J. Rosa, N. Amaral, R. Saaristo, V. Sacchini, H. Barros, and M. C. Oliveira. Factors determining esthetic outcome after breast cancer conservative treatment. *Breast Journal*, 13(2):140–146, 2007.
- [38] M. J. Cardoso, J. S. Cardoso, C. Vrieling, D. Macmillan, D. Rainsbury, J. Heil, E. Hau, and M. Keshtgar. Recommendations for the aesthetic evaluation of breast cancer conservative treatment. *Breast Cancer Research and Treatment*, 137:629–637, 2012.
- [39] M. J. Cardoso, J. S. Cardoso, T. Wild, W. Krois, and F. Fitzal. Comparing two objective methods for the aesthetic evaluation of breast cancer conservative treatment. *Breast Cancer Research and Treatment*, 116(1):149–152, 2009.
- [40] M. J. Cardoso, A. C. Santos, J. S. Cardoso, H. Barros, and M. C. Oliveira. Choosing observers for evaluation of aesthetic results in breast cancer conservative treatment. *International Journal of Radiation Oncology, Biology and Physics*, 61(3):879–881, 2005.
- [41] G. Catanuto, P. Patete, A. Spano, A. Pennati, G. Baroni, and M. B. Nava. New technologies for the assessment of breast surgical outcomes. *Aesthetic Surgery Journal*, 29(6):505–508, 2009.

- [42] G. Catanuto, A. Spano, A. Pennati, E. Riggio, G. Farinella, G. Impoco, S. Spoto, G. Gallo, and M. Nava. Experimental methodology for digital breast shape analysis and objective surgical outcome evaluation. *Journal of Plastic, Reconstructive & Aesthetic Surgery*, 61(3):314–318, 2008.
- [43] J. Čech and R. Šára. Efficient sampling of disparity space for fast and accurate matching. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, 2007.
- [44] S. K. Cetintas, L. Ozkan, M. Kurt, A. Saran, I. Tasdelen, S. Tolunay, U. Topal, and K. Engin. Factors influencing cosmetic results after breast conserving management (turkish experience). *The Breast*, 11(1):72–80, 2002.
- [45] C.-C. Chang and C.-J. Lin. Libsvm: A library for support vector machines. *ACM Trans. Intell. Syst. Technol.*, 2(3):27:1–27:27, 2011.
- [46] Y.-J. Chang, S.-F. Chen, and J.-D. Huang. A kinect-based system for physical rehabilitation: A pilot study for young adults with motor disabilities. *Research in Developmental Disabilities*, 32(6):2566–2570, 2011.
- [47] V. Cheung, B. J. Frey, and N. Jojic. Video epitomes. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 42–49, 2005.
- [48] H. Chmura Kraemer, V. S. Periyakoil, and A. Noda. Kappa coefficients in medical research. *Statistics in Medicine*, 21(14):2109–2129, 2002.
- [49] M. R. Christiaens, E. van der Schueren, and K. Vantongelen. More detailed documentation of operative procedures in breast conserving treatment: what good will it do us? *European Journal of Surgical Oncology*, 22(4):326–330, 1996.
- [50] D. R. H. Christie, M.-Y. O’Brien, J. A. Christie, T. Kron, S. A. Ferguson, C. S. Hamilton, and J. W. Denham. A comparison of methods of cosmetic assessment in breast conservation treatment. *Breast*, 5:358–367, 1996.
- [51] D. Clarke, A. Martinez, and R. S. Cox. Analysis of cosmetic results and complications in patients with stage i and ii breast cancer treated by biopsy and irradiation. *International Journal of Radiation Oncology Biology Physics*, 9(12):1807–1813, 1983.
- [52] R. A. Cochrane, P. Valasiadou, A. R. M. Wilson, S. K. Al-Ghazal, and R. D. Macmillan. Cosmesis and satisfaction after breast-conserving surgery correlates with the percentage of breast volume excised. *British Journal of Surgery*, 90(12):1505–1509, 2003.
- [53] J. L. Crowley and A. C. Parker. A representation for shape based on peaks and ridges in the difference of low-pass transform. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(2):156–170, 1984.
- [54] Y. Cui, S. Schuon, D. Chan, S. Thrun, and C. Theobalt. 3d shape scanning with a time-of-flight camera. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1173–1180, 2010.
- [55] Y. Cui and D. Stricker. 3d body scanning with one kinect. In *2nd International Conference on 3D Body Scanning Technologies*, pages 121–129, 2011.

- [56] C. D’Aniello, L. Grimaldi, A. Barbato, B. Bosi, and A. Carli. Cosmetic results in 242 patients treated by conservative surgery for breast cancer. *Scandinavian Journal of Plastic and Reconstructive Surgery and Hand Surgery*, 33(4):419–422, 1999.
- [57] B. F. Danoff, R. L. Goodman, J. H. Glick, D. G. Haller, and T. F. Pajak. The effect of adjuvant chemotherapy on cosmesis and complications in patients with breast cancer treated by definitive irradiation. *International Journal of Radiation Oncology\*Biology\*Physics*, 9(11):1625–1630, 1983.
- [58] N. Davidson, S. Khanna, R. Windle, W. Barrie, R. Agrawal, and S. Mitchell. Cosmetic results of early breast carcinoma treated with wide local excision, external beam radiotherapy and iridium-192 boost. *J R Coll Surg Edinb*, 35(3):175–177, 1990.
- [59] A. de la Rochefordiere, A. L. Abner, B. Silver, F. Vicini, A. Recht, and J. R. Harris. Are cosmetic results following conservative surgery and radiation therapy for early breast cancer dependent on technique? *International Journal of Radiation Oncology\*Biology\*Physics*, 23(5):925–931, 1992.
- [60] S. M. Deore, R. Sarin, K. A. Dinshaw, and S. K. Shrivastava. Influence of dose-rate and dose per fraction on clinical outcome of breast cancer treated by external beam irradiation plus iridium-192 implants: Analysis of 289 cases. *International Journal of Radiation Oncology\*Biology\*Physics*, 26(4):601–606, 1993.
- [61] J. A. Dewar, S. Benhamou, E. Benhamou, R. Arriagada, J. Y. Petit, F. Fontaine, and D. Sarrazin. Cosmetic results following lumpectomy axillary dissection and radiotherapy for small breast cancers. *Radiotherapy and Oncology*, 12(4):273–280, 1988.
- [62] A. R. Dominguez and A. K. Nandi. Toward breast cancer diagnosis based on automated segmentation of masses in mammograms. *Pattern Recognition*, 42(6):1138–1148, 2009.
- [63] W. Dong, L. Zhang, G. Shi, and X. Wu. Nonlocal back-projection for adaptive image enlargement. In *International Conference on Image Processing*, pages 349–352, 2009.
- [64] W. Dong, L. Zhang, G. Shi, and X. Wu. Image deblurring and super-resolution by adaptive sparse domain selection and adaptive regularization. *IEEE Transactions on Image Processing*, 20(7):1838–1857, 2011.
- [65] R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern Classification*. Wiley Interscience, 2nd edition, 2000.
- [66] C. Eadie, A. Herd, and S. Stallard. An investigation into digital imaging in assessing cosmetic outcome after breast surgery. *Journal of Audiovisual Media in Medicine*, 23(1):12–16, 2000.
- [67] M. Eder, F. Waldenfels, A. Swobodnik, M. Kloppel, A. Pape, T. Schuster, S. Raith, E. Kitzler, N. Papadopoulos, H. Machens, and L. Kovacs. Objective breast symmetry evaluation using 3-d surface imaging. *The Breast*, 21(2):152–158, 2012.
- [68] A. S. El-Baz, R. Acharya, A. F. Laine, and J. S. Suri. Multi modality state-of-the-art medical image segmentation and registration methodologies (Vol II). Springer, 2011.
- [69] M. Elad and A. Feuer. Restoration of a single superresolution image from several blurred, noisy, and undersampled measured images. *IEEE Transactions on Image Processing*, 6(12):1646–1658, 1997.

- [70] J. H. Elder and S. W. Zucker. Computing contour closure. In *4th European Conference on Computer Vision*, pages 399–412, 1996.
- [71] M. Elter, C. Held, and T. Wittenberg. Contour tracing for segmentation of mammographic masses. *Physics in Medicine and Biology*, 55:5299, 2010.
- [72] B. EORTC. Manual for clinical research and treatment in breast cancer. *Greenich Medical Media, London*, 5th edn., 2004.
- [73] M. A. Fagundes, H. M. Fagundes, C. S. Brito, M. H. Fagundes, A. Daudt, L. A. Bruno, S. J. Azevedo, and L. A. Fagundes. Breast-conserving surgery and definitive radiation: A comparison between quadrantectomy and local excision with special focus on local-regional control and cosmesis. *International Journal of Radiation Oncology\*Biolog\*Physics*, 27(3):553–560, 1993.
- [74] G. Farinella, G. Impoco, G. Gallo, S. Spoto, G. Catanuto, and M. Nava. Objective outcome evaluation of breast surgery. In *Medical Image Computing and Computer-Assisted Intervention*, pages 776–783, 2006.
- [75] G. M. Farinella, G. Impoco, G. Gallo, S. Spoto, and G. Catanuto. Unambiguous analysis of woman breast shape for plastic surgery outcome evaluation. In *Eurographics Italian Chapter Conference*, pages 255–261, 2006.
- [76] S. Farsiu, M. D. Robinson, M. Elad, and P. Milanfar. Fast and robust multiframe super resolution. *IEEE Transactions on Image Processing*, 13(10):1327–1344, 2004.
- [77] O. D. Faugeras. What can be seen in three dimensions with an uncalibrated stereo rig? In *Proceedings of the Second European Conference on Computer Vision*, pages 563–578, 1992.
- [78] B. Fisher, S. Anderson, J. Bryant, R. G. Margolese, M. Deutsch, E. R. Fisher, J. H. Jeong, and N. Wolmark. Twenty-year follow-up of a randomized trial comparing total mastectomy, lumpectomy, and lumpectomy plus irradiation for the treatment of invasive breast cancer. *the New England Journal of Medicine*, 347(16):1233–1241, 2002.
- [79] N. I. Fisher and A. J. Lee. A correlation coefficient for circular data. *Biometrika*, 70(2):327–332, 1983.
- [80] F. Fitzal, W. Krois, H. Trischler, L. Wutzel, O. Riedl, U. Kuhbelbock, B. Wintersteiner, M. Cardoso, P. Dubsky, M. Gnant, R. Jakesz, and T. Wild. The use of a breast symmetry index for objective evaluation of breast cosmesis. *The Breast*, 16(4):429–435, 2007.
- [81] B. Fowble. Conservative surgery and radiation for stage i and ii breast cancer: Identification of a subset of patients with early stage breast cancer for whom breast-conserving therapy may be contraindicated. *The Breast Journal*, 2(1):65–70, 1996.
- [82] S. Fuchs and G. Hirzinger. Extrinsic and depth calibration of tof-cameras. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–6, 2008.
- [83] Y. Furukawa and J. Ponce. Accurate, dense, and robust multiview stereopsis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(8):1362–1376, 2010.
- [84] A. Fusiello and L. Irsara. Quasi-euclidean uncalibrated epipolar rectification. In *19th International Conference on Pattern Recognition*, pages 1–4, 2008.

- [85] A. Fusiello and L. Irsara. Quasi-Euclidean epipolar rectification of uncalibrated images. *Machine Vision and Applications*, 22(4):663–670, 2011.
- [86] G. Galdino, M. Nahabedian, M. Chiaramonte, J. Geng, S. Klatsky, and P. Manson. Clinical applications of three-dimensional photography in breast surgery. *Plast Reconstr Surg.*, 110(1):58–70, 2002.
- [87] J. Gama and P. Brazdil. Cascade generalization. *Mach. Learn.*, 41(3):315–343, 2000.
- [88] V. Garro, P. Zanuttigh, and G. M. Cortelazzo. A new super resolution technique for range data. In *Proceedings of Gruppo Telecomunicazioni Teoria dell’Informazione Meeting*, pages 1–7, 2009.
- [89] A. Geiger, M. Roser, and R. Urtasun. Efficient large-scale stereo matching. In *Asian Conference on Computer Vision*, pages 25–38. Springer, 2010.
- [90] G. H. Golub and C. F. Van Loan. *Matrix computations (3rd ed.)*. Johns Hopkins University Press, Baltimore, MD, USA, 1996.
- [91] J. R. Gray, B. McCormick, L. Cox, and J. Yahalom. Primary breast irradiation in large-breasted or heavy women: Analysis of cosmetic outcome. *International Journal of Radiation Oncology\*Biophysics*, 21(2):347–354, 1991.
- [92] M. Greco, V. Sacchini, R. Agresti, A. Luini, M. D. Vecchio, G. Farante, and R. Raselli. Quadrantectomy is not a disfiguring operation for small breast cancer. *The Breast*, 3(1):3–7, 1994.
- [93] I. Guyon and A. Elisseeff. An introduction to variable and feature selection. *Journal of Machine Learning Research*, 3:1157–1182, 2003.
- [94] B. Han, C. Paulson, and D. Wu. 3d dense reconstruction from 2d video sequence via 3d geometric segmentation. *J. Vis. Comun. Image Represent.*, 22(5):421–431, 2011.
- [95] C. Harris and M. Stephens. A combined corner and edge detector. In *Proc. of Fourth Alvey Vision Conference*, pages 147–151, 1988.
- [96] J. R. Harris, M. B. Levene, G. Svensson, and S. Hellman. Analysis of cosmetic results following primary radiation therapy for stages i and ii carcinoma of the breast. *Int. Journal of Radiation Oncology Biology Physics*, 5(2):257–261, 1979.
- [97] R. I. Hartley. Theory and practice of projective rectification. *Int. J. Comput. Vision*, 35(2):115–127, 1999.
- [98] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2nd edition, 2004.
- [99] F. Hasson, S. Keeney, and H. McKenna. Research guidelines for the delphi survey technique. *Journal of Advanced Nursing*, 32(4):1008–1015, 2000.
- [100] T. Hastie and R. Tibshirani. Classification by pairwise coupling. *Ann. Statist.*, 26(2):451–471, 1998.

- [101] E. Hau, L. H. Browne, S. Khanna, S. Cail, G. Cert, Y. Chin, C. Clark, S. Inder, A. Szwajcer, and P. H. Graham. Radiotherapy breast boost with reduced whole-breast dose is associated with improved cosmesis: The results of a comprehensive assessment from the st. george and wollongong randomized breast boost trial. *International Journal of Radiation Oncology\*Biology\*Physics*, 82(2):682–689, 2012.
- [102] M. Heath, K. W. Bowyer, and D. Kopans. Current Status of the Digital Database for Screening Mammography. *Digital Mammography - MAMMO*, pages 457–460, 1998.
- [103] J. Heil, A. Carolus, J. Dahlkamp, M. Golatta, C. Domschke, F. Schuetz, M. Blumenstein, G. Rauch, and C. Sohn. Objective assessment of aesthetic outcome after breast conserving therapy: Subjective third party panel rating and objective bcct.core software evaluation. *The Breast*, 21(1):61–65, 2012.
- [104] J. Heil, J. Dahlkamp, M. Golatta, J. Rom, C. Domschke, G. Rauch, M. Cardoso, and C. Sohn. Aesthetics in breast conserving therapy: Do objectively measured results match patients’ evaluations? *Annals of Surgical Oncology*, 18(1):134–138, 2011.
- [105] P. Henry, M. Krainin, E. Herbst, X. Ren, and D. Fox. RGB-D mapping: Using kinect-style depth cameras for dense 3D modeling of indoor environments. *International Journal of Robotics Research*, 31(5):647–663, 2012.
- [106] H. Henseler, B. Khambay, A. Bowman, J. Smith, J. Siebert, S. Oehler, X. Ju, A. Ayoub, and A. Ray. Investigation into accuracy and reproducibility of a 3d breast imaging system using multiple stereo cameras. *Journal of Plastic, Reconstructive and Aesthetic Surgery*, 64(5):577–582, 2011.
- [107] D. C. Herrera, J. Kannala, and J. Heikkila. Accurate and practical calibration of a depth and color camera pair. In *Proceedings of the 14th international conference on Computer analysis of images and patterns*, pages 437–445, 2011.
- [108] D. C. Herrera, J. Kannala, and J. Heikkila. Joint depth and color camera calibration with distortion correction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(10):2058–2064, 2012.
- [109] H. Hirschmüller. Stereo processing by semiglobal matching and mutual information. *IEEE transactions on pattern analysis and machine intelligence*, 30(2):328–41, 2008.
- [110] U. Hoeller, A. Kuhlmei, A. Bajrovic, K. Grader, J. Berger, S. Tribius, F. Fehlauer, and W. Alberti. Cosmesis from the patient’s and the doctor’s view. *International Journal of Radiation Oncology, Biology and Physics*, 57(2):345–354, 2003.
- [111] M. Irani and S. Peleg. Improving resolution by image registration. *Graph. Models Image Process.*, 53(3):231–239, 1991.
- [112] F. Isgro and E. Trucco. Projective rectification without epipolar geometry. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 94–99, 1999.
- [113] N. Isogai, K. Sai, H. Kammishi, M. Watatani, H. Inui, and H. Shiozaki. Quantitative analysis of the reconstructed breast using a 3-dimensional laser light scanner. *Ann Plast Surg*, 56(3):237–242, 2006.



- [114] S. Izadi, D. Kim, O. Hilliges, D. Molyneaux, R. Newcombe, P. Kohli, J. Shotton, S. Hodges, D. Freeman, A. Davison, and A. Fitzgibbon. Kinectfusion: real-time 3d reconstruction and interaction using a moving depth camera. In *Proceedings of the 24th annual ACM symposium on User interface software and technology*, pages 559–568, 2011.
- [115] A. Jemal, F. Bray, M. M. Center, J. Ferlay, E. Ward, and D. Forman. Global cancer statistics. *CA: A Cancer Journal for Clinicians*, 61(2):69–90, 2011.
- [116] I. H. Jermyn and H. Ishikawa. Globally optimal regions and boundaries as minimum ratio weight cycles. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(10):1075–1088, 2001.
- [117] H. Jin, S. Soatto, and A. J. Yezzi. Multi-view stereo reconstruction of dense shape and complex appearance. *Int. J. Comput. Vision*, 63(3):175–189, 2005.
- [118] J. Johansen, J. Overgaard, and M. Overgaard. Effect of adjuvant systemic treatment on cosmetic outcome and late normal-tissue reactions after breast conservation. *Acta Oncologica*, 46(4):525–533, 2007.
- [119] J. Jones and D. Hunter. Consensus methods for medical and health services research. *British Medical Journal*, 311(7001):376–380, 1995.
- [120] H. Kaija, S. Rauni, I. Jorma, and H. Matti. Consistency of patient- and doctor-assessed cosmetic outcome after conservative treatment of breast cancer. *Breast Cancer Research and Treatment*, 45(3):225–228, 1997.
- [121] N. Kaur, J.-Y. Petit, M. Rietjens, F. Maffini, A. Luini, G. Gatti, P. C. Rey, C. Urban, and F. De Lorenzi. Comparative study of surgical margins in oncoplastic surgery and quadrantectomy in breast cancer. *Annals of Surgical Oncology*, 12(7):539–545, 2005.
- [122] M. Kawale, J. Lee, S. Y. Leung, M. C. Fingeret, G. P. Reece, M. A. Crosby, E. K. Beahm, M. K. Markey, and F. A. Merchant. 3d symmetry measure invariant to subject pose during image acquisition. *Breast Cancer: Basic and Clinical Research*, 5:131–142, 2011.
- [123] Y. Ke and R. Sukthankar. PCA-SIFT: A more distinctive representation for local image descriptors. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2:506–513, 2004.
- [124] D. Keren, S. Peleg, and R. Brada. Image sequence enhancement using sub-pixel displacements. In *Proceedings of the Computer Society Conference on Computer Vision and Pattern Recognition*, pages 742–746, 1988.
- [125] D. T. Kien. *A review of 3D reconstruction from video sequences*. Ph.d. thesis, University of Amsterdam, The Netherlands, 2005.
- [126] M. S. Kim, G. P. Reece, E. K. Beahm, M. J. Miller, E. Neely Atkinson, and M. K. Markey. Objective assessment of aesthetic outcomes of breast cancer treatment: Measuring ptosis from clinical photographs. *Comput. Biol. Med.*, 37(1):49–59, 2007.
- [127] N. F. Kim and J. S. Park. Segmentation of object regions using depth information. In *Proceedings of the International Conference on Image Processing*, pages 231–234, 2004.

- [128] Y. M. Kim, D. Chan, C. Theobalt, and S. Thrun. Design and calibration of a multi-view tof sensor fusion system. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–7, 2008.
- [129] A. Klaus, M. Sormann, and K. Karner. Segment-Based Stereo Matching Using Belief Propagation and a Self-Adapting Dissimilarity Measure. In *18th International Conference on Pattern Recognition*, pages 15–18, 2006.
- [130] H. Kobatake and S. Hashimoto. Convergence index filter for vector fields. *IEEE Transactions on Image Processing*, 8(8):1029–1038, 1999.
- [131] R. Koch, M. Pollefeys, and L. Van Gool. Realistic surface reconstruction of 3d scenes from uncalibrated image sequences. *The Journal of Visualization and Computer Animation*, 11(3):115–127, 2000.
- [132] V. Kolmogorov and R. Zabih. Computing visual correspondence with occlusions via graph cuts. In *International Conference on Computer Vision*, pages 508–515, 2001.
- [133] V. Kolmogorov and R. Zabih. Multi-camera scene reconstruction via graph cuts. In *Proceedings of the 7th European Conference on Computer Vision*, pages 82–96, 2002.
- [134] L. Kovacs, M. Eder, R. Hallweck, A. Zimmermann, M. Settles, A. Schneider, K. Udoscic, K. Schwenzer-Zimmerer, N. A. Papadopoulos, and E. Biemer. New aspects of breast volume measurement using 3-dimensional surface imaging. *Ann Plast Surg*, 57(6):602–610, 2006.
- [135] L. Kovacs, M. Eder, R. Hollweck, A. Zimmermann, M. Settles, A. Schneider, M. Endlich, A. Mueller, K. Schwenzer-Zimmerer, N. A. Papadopoulos, and E. Biemer. Comparison between breast volume measurement using 3d surface imaging and classical techniques. *The Breast*, 16(2):137–145, 2007.
- [136] L. Kovacs, A. Yassouridis, A. Zimmermann, G. Brockmann, A. Wohnl, M. Blaschke, M. Eder, K. Schwenzer-Zimmerer, R. Rosenberg, N. A. Papadopoulos, and E. Biemer. Optimization of 3-dimensional imaging of the breast region with 3-dimensional laser scanners. *Ann Plast Surg*, 56(3):229–236, 2006.
- [137] L. Krishnan, A. L. Stanton, C. A. Collins, V. E. Liston, and W. R. Jewell. Form or function? part 2. objective cosmetic and functional correlates of quality of life in women treated with breast-conserving surgical procedures and radiotherapy. *Cancer*, 91(12):2282–2287, 2001.
- [138] J. M. Kurtz. Impact of radiotherapy on breast cosmesis. *The Breast*, 4(3):163–169, 1995.
- [139] J. Lee, S. Chen, G. P. Reece, M. A. Crosby, E. K. Beahm, and M. K. Markey. A novel quantitative measure of breast curvature based on catenary. *IEEE Transactions on Biomedical Engineering*, 59(4):1115–1124, 2012.
- [140] J. Lee, M. Kawale, F. A. Merchant, J. Weston, M. C. Fingeret, D. Ladewig, G. P. Reece, M. A. Crosby, E. K. Beahm, and M. K. Markey. Validation of stereophotogrammetry of the human torso. *Breast Cancer: Basic and Clinical Research*, 5:15–25, 2011.
- [141] C. Leung. *Efficient Methods for 3D Reconstruction From Multiple Images*. Ph.d. thesis, University of Queensland, Australia, 2006.

- [142] A. Levinshtein, S. Dickinson, and C. Sminchisescu. Multiscale symmetric part detection and grouping. In *12th IEEE International Conference on Computer Vision*, pages 2162–2169, 2009.
- [143] A. Levinshtein, C. Sminchisescu, and S. Dickinson. Optimal contour closure by superpixel grouping. In *11th European conference on Computer vision*, pages 480–493, 2010.
- [144] M. Lhuillier and L. Quan. A quasi-dense approach to surface reconstruction from uncalibrated images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(3):418–433, 2005.
- [145] J. Li, E. Li, Y. Chen, L. Xu, and Y. Zhang. Bundled depth-map merging for multi-view stereo. In *Proceedings of the Conference on Computer Vision and Pattern Recognition*, pages 2769–2776, 2010.
- [146] X. Li, Y. Hu, X. Gao, D. Tao, and B. Ning. A multi-frame image super-resolution method. *Signal Process.*, 90(2):405–414, 2010.
- [147] G. Liljegren, L. Holmberg, and G. Westman. The cosmetic outcome in early breast cancer treated with sector resection with or without radiotherapy. *European Journal of Cancer*, 29(15):2083–2089, 1993.
- [148] E. V. Limbergen, A. Rijnders, E. V. Schueren, T. Lerut, and R. Christiaens. Cosmetic evaluation of breast conserving treatment for mammary cancer. 2. a quantitative analysis of the influence of radiation dose, fractionation schedules and surgical treatment techniques on cosmetic results. *Radiotherapy and Oncology*, 16(4):253–267, 1989.
- [149] E. V. Limbergen, E. V. Schueren, and K. . V. Tongelen. Cosmetic evaluation of breast conserving treatment for mammary cancer. 1. proposal of a quantitative scoring system. *Radiotherapy and Oncology*, 16(3):159–167, 1989.
- [150] I. Lindsey, J. W. Serpell, W. R. Johnson, and A. Rodger. Cosmesis following complete local excision of breast cancer. *Australian and New Zealand Journal of Surgery*, 67(7):428–432, 1997.
- [151] H. C. Longuet-Higgins. A computer algorithm for reconstructing a scene from two projections. *Nature*, 293:133–135, 1981.
- [152] C. Loop and Z. Zhang. Computing rectifying homographies for stereo vision. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 125–131, 1999.
- [153] A. Losken, I. Fishman, D. D. Denson, H. R. Moyer, and G. W. Carlson. An objective evaluation of breast symmetry and shape differences using 3-dimensional images. *Ann Plast Surg*, 55(6):571–575, 2005.
- [154] A. Losken, H. Seify, D. D. Denson, A. A. Paredes, and G. W. Carlson. Validating three-dimensional imaging of the breast. *Ann Plast Surg*, 54(5):471–476, 2005.
- [155] D. G. Lowe. Object recognition from local scale-invariant features. In *Proceedings of the Seventh IEEE International Conference on Computer Vision*, pages 1150–1157, 1999.
- [156] D. G. Lowe. Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision*, 60(2):91–110, 2004.

- [157] L. Lucchese and G. Cortelazzo. A noise-robust frequency domain technique for estimating planar roto-translations. *IEEE Transactions on Signal Processing*, 48(6):1769–1786, 2000.
- [158] A. Luini, G. Gatti, V. Galimberti, S. Zurrida, M. Intra, O. Gentilini, G. Paganelli, G. Viale, R. Orecchia, P. Veronesi, and U. Veronesi. Conservative treatment of breast cancer: Its evolution. *Breast Cancer Research and Treatment*, 94:195–198, 2005.
- [159] A. T. Magalhães, H. P. Oliveira, S. Costa, J. S. Cardoso, and M. J. Cardoso. Value of Photographic Side-views in the Objective Evaluation of the Aesthetic Results of Breast Cancer Conservative Treatment. In *33rd Annual San Antonio Breast Cancer Symposium*, 2010.
- [160] A. T. Magalhães, H. P. Oliveira, A. J. Moura, J. S. Cardoso, and M. J. Cardoso. Aesthetic Evaluation after Mastectomy and Immediate Breast Reconstruction using the BCCT.core Software. In *3rd International Meeting of Oncoplastic and Reconstructive Breast Surgery*, 2010.
- [161] A. T. Magalhães, H. P. Oliveira, J. Soares, A. J. Moura, J. S. Cardoso, and M. J. Cardoso. 3D Model for Aesthetic Objective Evaluation after Breast Cancer Surgery using Infrared Laser Projector. In *3rd International Meeting of Oncoplastic and Reconstructive Breast Surgery*, 2010.
- [162] B. Marcel, M. Briot, and R. Murrieta-Cid. Calcul de translation et rotation par la transformation de fourier. *Journal Traitement du Signal*, 14(2):135–149, 1997.
- [163] A. Martelli. Edge detection using heuristic search methods. *Computer Graphics and Image Processing*, 1(2):169–182, 1972.
- [164] A. Martelli. An application of heuristic search methods to edge and contour detection. *Communications ACM*, 19(2):73–83, 1976.
- [165] D. Martin, C. Fowlkes, D. Tal, and J. Malik. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In *Proceedings of the 8th International Conference on Computer Vision*, volume 2, pages 416–423, July 2001.
- [166] J. Matas. Robust wide-baseline stereo from maximally stable extremal regions. *Image and Vision Computing*, 22(10):761–767, 2004.
- [167] K. Mikolajczyk and C. Schmid. Scale and affine invariant interest point detectors. *International Journal of Computer Vision*, 60(1):63–86, 2004.
- [168] K. Mikolajczyk and C. Schmid. Performance evaluation of local descriptors. *IEEE transactions on pattern analysis and machine intelligence*, 27(10):1615–30, 2005.
- [169] P. A. V. Miranda, A. X. Falcao, and T. V. Spina. Riverbed: A novel user-steered image segmentation method based on optimum boundary tracking. *IEEE Transactions on Image Processing*, 21(6):3042–3052, 2012.
- [170] L. Moisan and B. Stival. A Probabilistic Criterion to Detect Rigid Point Matches Between Two Images and Estimate the Fundamental Matrix. *International Journal of Computer Vision*, 57(3):201–218, 2004.
- [171] P. Monasse. Quasi-Euclidean Epipolar Rectification. *Image Processing On Line*, 2011.

- [172] I. Moreira, I. Amaral, I. Domingues, A. Cardoso, M. J. Cardoso, and J. S. Cardoso. Inbreast: Towards a full field digital mammographic database. *Academic Radiology*, 19(2):236–248, 2012.
- [173] S. M. Moss, L. Nyström, H. Jonsson, E. Paci, E. Lynge, S. Njor, and M. Broeders. The impact of mammographic screening on breast cancer mortality in europe: a review of trend studies. *Journal of Medical Screening*, 19(1):26–32, 2012.
- [174] H. R. Moyer, G. W. Carlson, T. M. Styblo, and A. Losken. Three-dimensional digital evaluation of breast symmetry after breast conservation therapy. *Journal of the American College of Surgeons*, 207(2):227–232, 2008.
- [175] M. Noguchi, Y. Saito, Y. Mizukami, A. Nonomura, N. Ohta, N. Koyasaki, T. Taniya, and I. Miyazaki. Breast deformity, its correction, and assessment of breast conserving surgery. *Breast Cancer Research and Treatment*, 18(2):111–118, 1991.
- [176] H. P. Oliveira and J. S. Cardoso. Image retargeting using stable paths. In *Proceedings of the Fourth International Conference on Computer Vision Theory and Applications*, pages 40–47, 2009.
- [177] H. P. Oliveira, J. S. Cardoso, A. Magalhães, and M. J. Cardoso. Simultaneous detection of prominent points on breast cancer conservative treatment images. In *Proceedings of The IEEE International Conference on Image Processing*, pages 2841–2844, 2012.
- [178] H. P. Oliveira, J. S. Cardoso, A. Magalhães, and M. J. Cardoso. A 3d low-cost solution for the aesthetic evaluation of breast cancer conservative treatment. 2013. SUBMITTED.
- [179] H. P. Oliveira, J. S. Cardoso, A. Magalhães, and M. J. Cardoso. Methods for the aesthetic evaluation of breast cancer conservation treatment: A technological review. *Current Medical Imaging Reviews*, 9(1):32–46, 2013.
- [180] H. P. Oliveira, A. Magalhaes, M. J. Cardoso, and J. S. Cardoso. An accurate and interpretable model for bcct.core. In *Proceedings of the 32<sup>nd</sup> Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 6158–6161, 2010.
- [181] H. P. Oliveira, A. Magalhaes, M. J. Cardoso, and J. S. Cardoso. Improving the bcct.core model with lateral information. In *Proceedings of the 10<sup>th</sup> IEEE International Conference on Information Technology and Applications in Biomedicine*, pages 1–4, 2010.
- [182] H. P. Oliveira, P. Patete, G. Baroni, and J. S. Cardoso. Development of a bcct quantitative 3d evaluation system through low-cost solutions. In *Proceedings of the 2<sup>nd</sup> International Conference on 3D Body Scanning Technologies*, pages 16–27, 2011.
- [183] H. P. Oliveira, M. D. Silva, A. Magalhães, M. J. Cardoso, and J. S. Cardoso. Is kinect depth data accurate for the aesthetic evaluation after breast cancer surgeries? In *Proceedings of Iberian Conference on Pattern Recognition and Image Analysis*, pages 261–268, 2013.
- [184] I. A. Olivotto, M. A. Rose, R. T. Osteen, S. Love, B. Cady, B. Silver, A. Recht, and J. R. Harris. Late cosmetic outcome after conservative surgery and radiotherapy: Analysis of causes of cosmetic failure. *International Journal of Radiation Oncology\*Biophysics*, 17(4):747–753, 1989.

- [185] M. Ozuysal, M. Calonder, V. Lepetit, and P. Fua. Fast keypoint recognition using random ferns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(3):448–461, 2010.
- [186] C. Papathanassiou and M. Petrou. Super-resolution: an overview. *IEEE International Geoscience and Remote Sensing Symposium*, 8:5655–5658, 2005.
- [187] S. C. Park, M. K. Park, and M. G. Kang. Super-resolution image reconstruction: a technical overview. *IEEE Signal Processing Magazine*, 20(3):21–36, 2003.
- [188] P. Patete, M. Riboldi, and G. Baron. Integrated 3d body scanner with real-time compensation of breathing movement artifacts. In *Proceedings of the 1<sup>st</sup> International Conference on 3D Body Scanning Technologies*, pages 185–190, 2010.
- [189] P. Patete, M. Riboldi, M. Spadea, G. Catanuto, A. Spano, M. Nava, and G. Baroni. Motion compensation in hand-held laser scanning for surface modeling in plastic and reconstructive surgery. *Annals of Biomedical Engineering*, 37(9):1877–1885, 2009.
- [190] M. P. Patterson, R. D. Pezner, L. R. Hill, N. L. Vora, K. R. Desai, and J. A. Lipsett. Patient self-evaluation of cosmetic outcome of breast-preserving cancer treatment. *International Journal of Radiation Oncology\*Biology\*Physics*, 11(10):1849–1852, 1985.
- [191] S. Peleg, D. Keren, and L. Schweitzer. Improving image resolution using subpixel motion. *Pattern Recogn. Lett.*, 5(3):223–226, 1987.
- [192] R. D. Pezner, J. A. Lipsett, N. L. Vora, and K. R. Desai. Limited usefulness of observer-based cosmesis scales employed to evaluate patients treated conservatively for breast cancer. *International Journal of Radiation Oncology Biology Physics*, 11(6):1117–1119, 1985.
- [193] R. D. Pezner, M. P. Patterson, L. R. Hill, N. Vora, K. R. Desai, J. O. Archambeau, and J. A. Lipsett. Breast retraction assessment: an objective evaluation of cosmetic results of patients treated conservatively for breast cancer. *International Journal of Radiation Oncology Biology Physics*, 11(3):575–578, 1985.
- [194] D. L. Pham, C. Xu, and J. L. Prince. Current methods in medical image segmentation. *Annual Review of Biomedical Engineering*, 2(1):315–337, 2000.
- [195] L. C. Pickup. *Machine Learning in Multi-frame Image Super-resolution*. Ph.d. thesis, University of Oxford, United Kingdom, 2008.
- [196] B. Pierquin, J. Huart, M. Raynal, Y. Otmezguine, E. Calitchi, J. J. Mazon, G. Ganem, J. P. L. Bourgeois, G. Marinello, and M. Julien. Conservative treatment for breast cancer: long-term results (15 years). *Radiotherapy Oncology*, 20(1):16–23, 1991.
- [197] M. Pollefeys. Visual 3D Modelling from Images (accessed 09/11/2012). <http://www.cs.unc.edu/marc/tutorial/>, 2002.
- [198] M. Pollefeys, R. Koch, M. Vergauwen, and L. Van Gool. Automated reconstruction of 3d scenes from sequences of images. *Journal of Photogrammetry and Remote Sensing*, 55(4):251–267, 2009.
- [199] J. R. Quinlan. Induction of decision trees. *Mach. Learn.*, 1(1):81–106, 1986.

- [200] G. R. Ray and V. J. Fish. Biopsy and definitive radiation therapy in stage i and ii adenocarcinoma of the female breast: Analysis of cosmesis and the role of electron beam supplementation. *International Journal of Radiation Oncology\*Biology\*Physics*, 9(6):813–818, 1983.
- [201] G. R. Ray, V. J. Fish, J. B. Marmor, W. Rogoway, P. Kushlan, C. Arnold, R. H. Lee, and F. Marzoni. Impact of adjuvant chemotherapy on cosmesis and complications in stages i and ii carcinoma of the breast treated by biopsy and radiation therapy. *International Journal of Radiation Oncology\*Biology\*Physics*, 10(6):837–841, 1984.
- [202] F. Remondino and S. El-Hakim. Image-based 3d modelling: A review. *The Photogrammetric Record*, 21(115):269–291, 2006.
- [203] X. Ren, C. C. Fowlkes, and J. Malik. Cue integration for Figure/Ground labeling. In *Advances in Neural Information Processing Systems*, pages 1121–1128, 2005.
- [204] X. Ren, C. C. Fowlkes, and J. Malik. Scale-invariant contour completion using conditional random fields. In *Tenth IEEE International Conference on Computer Vision*, pages 1214–1221, 2005.
- [205] A. Roelstraete, M. V. Lancker, A. D. Schryver, and G. Storme. Adjuvant radiation after conservative surgery for early breast cancer. local control and cosmetic outcome. *Am J Clin Oncol*, 16(4):284–290, 1993.
- [206] M. Rose, I. Olivotto, B. Cady, C. Koufman, R. Osteen, B. Silver, A. Recht, and J. Harris. Conservative surgery and radiation therapy for early breast cancer. long-term cosmetic results. *Arch Surg*, 124(2):153–157, 1989.
- [207] E. Rosten and T. Drummond. Fusing points and lines for high performance tracking. In *Tenth IEEE International Conference on Computer Vision*, pages 1508–1515, 2005.
- [208] E. Rosten and T. Drummond. Machine learning for high-speed corner detection. In *Proceedings of the 9th European conference on Computer Vision*, pages 430–443, 2006.
- [209] Y. Rubner, C. Tomasi, and L. J. Guibas. The earth mover’s distance as a metric for image retrieval. *International Journal of Computer Vision*, 40(2):99–121, 2000.
- [210] V. Sacchini, A. Luini, S. Tana, L. Lozza, V. Galimberti, M. Merson, R. Agresti, P. Veronesi, and M. Greco. Quantitative and qualitative cosmetic evaluation after conservative treatment for breast cancer. *European Journal Cancer*, 27(11):1395–1400, 1991.
- [211] R. Šára. Finding the largest unambiguous component of stereo matching. In *Proceedings of the 7th European Conference on Computer Vision*, pages 900–914, 2002.
- [212] R. Šára. Robust correspondence recognition for computer vision. In *Proceedings of the International Conference on Computational Statistics*, pages 119–131, 2006.
- [213] R. Sarin, K. A. Dinshaw, S. K. Shrivastava, V. Sharma, and S. M. Deore. Therapeutic factors influencing the cosmetic outcome and late complications in the conservative management of early breast cancer. *International Journal of Radiation Oncology\*Biology\*Physics*, 27(2):285–292, 1993.

- [214] D. Scaramuzza, A. Harati, and R. Siegwart. Extrinsic self calibration of a camera and a 3d laser range finder from natural scenes. In *Proceedings of the International Conference on Intelligent Robots and Systems*, pages 4164–4169, 2007.
- [215] M. Silva. *Kinect Based System Applied to Breast Cancer Conservative Treatment*. M.sc. thesis, Universidade do Porto, Faculdade de Engenharia, 2012.
- [216] J. Smisek, M. Jancosek, and T. Pajdla. 3d with kinect. In *Proceedings of the IEEE International Conference on Computer Vision Workshops*, pages 1154–1160, 2011.
- [217] S. M. Smith and J. M. Brady. SUSAN—A new approach to low level image processing. *International journal of computer vision*, 23(1):45–78, 1997.
- [218] K. C. Sneeuw, N. K. Aaronson, J. R. Yarnold, M. Broderick, J. R. G. Ross, and A. Goddard. Cosmetic and functional outcomes of breast conserving treatment for early stage breast cancer. 1. comparison of patients’ ratings, observers’ ratings, and objective assessments. *Radiotherapy and Oncology*, 25(3):153–159, 1992.
- [219] J. Soares. *Uncalibrated Stereo Vision applied to Breast Cancer Treatment Aesthetic Assessment*. M.sc. thesis, Universidade do Porto, Faculdade de Engenharia, 2011.
- [220] E. Song, S. Xu, X. Xu, J. Zeng, Y. Lan, S. Zhang, and C.-C. Hung. Hybrid segmentation of mass in mammograms using template matching and dynamic programming. *Academic Radiology*, 17(11):1414–1424, 2010.
- [221] R. Sousa, J. S. Cardoso, J. F. P. da Costa, and M. J. Cardoso. Breast contour detection with shape priors. In *Proceedings of the 15<sup>th</sup> IEEE International Conference on Image Processing*, pages 1440–1443, 2008.
- [222] R. Sousa, H. P. Oliveira, and J. S. Cardoso. Feature selection with complexity measure in a quadratic programming setting. In *Proceedings of Iberian Conference on Pattern Recognition and Image Analysis*, pages 524–531, 2011.
- [223] J. S. Stahl and S. Wang. Globally optimal grouping for symmetric closed boundaries by combining boundary and region information. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(3):395–411, 2008.
- [224] B. S. I. Standards. *Photography - Electronic still-picture cameras - Resolution measurements*. International Organization for Standardization (ISO), Geneva, Switzerland, 2000.
- [225] R. Steeves, P. Phromratanapongse, W. Wolberg, and D. Tormey. Cosmesis and local control after irradiation in women treated conservatively for breast cancer. *Arch Surg*, 124(12):1369–1373, 1989.
- [226] J. Suckling, J. Parker, D. Dance, S. Astley, I. Hutt, C. Boggis, I. Ricketts, E. Stamatakis, N. Cerneaz, S. Kok, P. Taylor, D. Betal, and J. Savage. The mammographic images analysis society digital mammogram database. *Experta Medica International Congress Series*, 1069:375–378, 1994.
- [227] C. Sun and S. Pallottino. Circular shortest path in images. *Pattern Recognition*, 36(3):709–719, 2003.
- [228] R. Szeliski. *Computer Vision: Algorithms and Applications*. Springer-Verlag New York, Inc., New York, NY, USA, 2010.



- [229] M. E. Taylor, C. A. Perez, K. J. Halverson, R. R. Kuske, G. W. Philpott, D. M. Garcia, J. E. Mortimer, R. J. Myerson, D. Radford, and C. Rush. Factors influencing cosmetic results after conservation therapy for breast cancer. *International Journal of Radiation Oncology\*Biology\*Physics*, 31(4):753–764, 1995.
- [230] O. M. Tepper, N. S. Karp, K. Small, J. Unger, L. Rudolph, A. Pritchard, and M. Choi. Three-dimensional imaging provides valuable clinical data to aid in unilateral tissue expander-implant breast reconstruction. *The Breast Journal*, 14(6):543–50, 2008.
- [231] O. M. Tepper, K. Small, L. Rudolph, M. Choi, and N. Karp. Virtual 3-dimensional modeling as a valuable adjunct to aesthetic and reconstructive breast surgery. *The American Journal of Surgery*, 192(4):548–551, 2006.
- [232] O. M. Tepper, J. G. Unger, K. H. Small, D. Feldman, N. Kumar, M. Choi, and N. S. Karp. Mammometrics: the standardization of aesthetic and reconstructive breast surgery. *Plast Reconstr Surg*, 125(1):393–400, 2010.
- [233] S. Timp and N. Karssemeijer. A new 2D segmentation method based on dynamic programming applied to computer aided detection in mammography. *Medical Physics*, 31(5):958–971, 2004.
- [234] J. Tong, J. Zhou, L. Liu, Z. Pan, and H. Yan. Scanning 3d full human bodies using kinects. *IEEE Transactions on Visualization and Computer Graphics*, 18(4):643–650, 2012.
- [235] E. Touboul, Y. Belkacemi, J. P. Lefranc, S. Uzan, M. Ozsahin, D. Korbas, L. Buffat, J. Balosso, F. Pene, and J. Blondon. Early breast cancer: influence of type of boost (electrons vs iridium-192 implant) on local control and cosmesis after conservative surgery and radiation therapy. *Radiotherapy and Oncology*, 34(2):105–113, 1995.
- [236] S. A. Friedman, R. Osteen, and J. R. Harris. Factors influencing cosmetic outcome of conservative surgery and radiotherapy for breast cancer. *The Surgical Clinics of North America*, 70(4):901–916, 1990.
- [237] R. Tsai. A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses. *IEEE Journal of Robotics and Automation*, 3(4):323–344, 1987.
- [238] L. I. Tsouskas and I. S. Fentiman. Breast compliance: a new method for evaluation of cosmetic outcome after conservative treatment of early breast cancer. *Breast Cancer Research and Treatment*, 15:185–190, 1990.
- [239] I. Turesson and G. Notter. The influence of fraction size in radiotherapy on the late normal tissue reaction–i: Comparison of the effects of daily and once-a-week fractionation on human skin. *International Journal of Radiation Oncology\*Biology\*Physics*, 10(5):593–598, 1984.
- [240] I. Turesson and G. Notter. The influence of fraction size in radiotherapy on the late normal tissue reaction–ii: Comparison of the effects of daily and twice-a-week fractionation on human skin. *International Journal of Radiation Oncology\*Biology\*Physics*, 10(5):599–606, 1984.
- [241] T. Tuytelaars and K. Mikolajczyk. Local invariant feature detectors: a survey. *Found. Trends. Comput. Graph. Vis.*, 3(3):177–280, 2008.

- [242] P. Vandewalle. *Super-Resolution from Unregistered Aliased Images*. PhD thesis, Ecole Polytechnique Federale de Lausanne, Switzerland, 2006.
- [243] P. Vandewalle, S. Süsstrunk, and M. Vetterli. A frequency domain approach to registration of aliased images with application to super-resolution. *EURASIP J. Appl. Signal Process.*, pages 233–233, 2006.
- [244] U. Veronesi. Conservative treatment of breast cancer: A trial in progress at the cancer institute of milan. *Cancer*, 39(6):2822–2826, 1977.
- [245] U. Veronesi, N. Cascinelli, L. Mariani, M. Greco, R. Saccozzi, A. Luini, M. Aguilar, and E. Marubini. Twenty-year follow-up of a randomized study comparing breast-conserving surgery with radical mastectomy for early breast cancer. *the New England Journal of Medicine*, 347(16):1227–1232, 2002.
- [246] N. Villaroman, D. Rowe, and B. Swan. Teaching natural user interaction using openni and the microsoft kinect sensor. In *Proceedings of the conference on Information technology education*, pages 227–232, 2011.
- [247] G. Vogiatzis, C. Hernández Esteban, P. H. S. Torr, and R. Cipolla. Multiview stereo via volumetric graph-cuts and occlusion robust photo-consistency. *IEEE Trans. Pattern Anal. Mach. Intell.*, 29(12):2241–2246, 2007.
- [248] C. Vrieling, L. Collette, E. Bartelink, J. H. Borger, S. J. Brenninkmeyer, J. C. Horiot, M. Pierart, P. M. Poortmans, H. Struikmans, E. V. der Schueren, J. A. V. Dongen, E. V. Limbergen, and H. Bartelink. Validation of the methods of cosmetic assessment after breast-conserving therapy in the EORTC “boost versus no boost” trial. *International Journal of Radiation Oncology Biology Physics*, 45(3):667–676, 1999.
- [249] C. Vrieling, L. Collette, A. Fourquet, W. J. Hoogenraad, J.-C. Horiot, J. J. Jager, M. Pierart, P. M. Poortmans, H. Struikmans, B. Maat, E. V. Limbergen, and H. Bartelink. The influence of patient, tumor and treatment factors on the cosmetic results after breast-conserving therapy in the EORTC “boost versus no boost” trial. *Radiotherapy and Oncology*, 55(3):219–232, 2000.
- [250] R. Wang, J. Choi, and G. Medioni. Accurate full body scanning from a single fixed 3d camera. In *International Conference on 3D Imaging, Modeling, Processing, Visualization, and Transmission*, pages 1–8, 2012.
- [251] S. Wang, T. Kubota, J. M. Siskind, and J. Wang. Salient closed boundary extraction with ratio contour. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(4):546–561, 2005.
- [252] Z. Wang and J. Road. A region based stereo matching algorithm using cooperative optimization. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, 2008.
- [253] D. E. Wazer, T. DiPetrillo, R. Schmidt-Ullrich, L. Weld, T. J. Smith, D. J. Marchant, and N. J. Robert. Factors influencing cosmetic outcome and complication risk after conservative surgery and radiotherapy for early-stage breast carcinoma. *Journal of Clinical Oncology*, 10(3):356–63, 1992.

- [254] A. Weiss, D. Hirshberg, and M. Black. Home 3d body scans from noisy image and range data. In *IEEE International Conference on Computer Vision*, pages 1951–1958, 2011.
- [255] L. R. Williams and A. R. Hanson. Perceptual completion of occluded surfaces. *Computer Vision and Image Understanding*, 64(1):1–20, 1996.
- [256] J. Yang, J. Wright, T. Huang, and Y. Ma. Image super-resolution via sparse representation. *IEEE Transactions on Image Processing*, 19(11):2861–2873, 2010.
- [257] Q. Yang, R. Yang, J. Davis, and D. Nister. Spatial-depth super resolution for range images. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, 2007.
- [258] F. Yi and I. Moon. Image segmentation: A survey of graph-cut methods. In *International Conference on Systems and Informatics*, pages 1936–1941, 2012.
- [259] A. Yla-Jaaskia and F. Ade. Grouping symmetrical structures for object segmentation and description. *Computer Vision and Image Understanding*, 63(3):399–417, 1996.
- [260] J. Zgajnar. Digital evaluation of breast cosmesis after breast-conserving treatment: End of the beginning. *The Breast*, 16(5):441–442, 2007.
- [261] C. Zhang and Z. Zhang. Calibration between depth and color sensors for commodity depth cameras. In *Proceedings of the IEEE International Conference on Multimedia and Expo*, pages 1–6, 2011.
- [262] Q. Zhang and R. Pless. Extrinsic calibration of a camera and laser range finder (improves camera calibration). In *Proceedings of the International Conference on Intelligent Robots and Systems*, pages 2301–2306, 2004.
- [263] Z. Zhang. Flexible camera calibration by viewing a plane from unknown orientations. In *Proceedings of the Seventh IEEE International Conference on Computer Vision*, pages 666–673, 1999.
- [264] Z. Zhang. A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11):1330–1334, 2000.
- [265] J. Zhou and B. Li. Image rectification for stereoscopic visualization. *J. Opt. Soc. Am. A*, 25(11):2721–2733, 2008.
- [266] J. Zhu, L. Wang, R. Yang, and J. Davis. Fusion of time-of-flight depth and stereo for high accuracy depth maps. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, 2008.
- [267] Q. Zhu, G. Song, and J. Shi. Untangling cycles for contour grouping. In *11th IEEE International Conference on Computer Vision*, pages 1–8, 2007.
- [268] Q. Zhu, H. Wang, and W. Tian. A practical new approach to 3D scene recovery. *Signal Processing*, 89(11):2152–2158, 2009.
- [269] M. Zollhofer, M. Martinek, G. Greiner, M. Stamminger, and J. SuBmuth. Automatic reconstruction of personalized avatars from 3d face scans. *Comput. Animat. Virtual Worlds*, 22(2-3):195–202, 2011.