

Implementation of Models for Image Processing Applications with Real-Time Constraints.



UGR

Universidad
de Granada

Pablo Guzmán Sánchez

Official PhD Program in Information Technology and Communication.
Department of Computer Architecture and Technology.
University of Granada.

A thesis submitted for the degree of

Philosophiae Doctor (PhD)

2015 April

Editorial: Universidad de Granada. Tesis Doctorales

Autor: Pablo Guzmán Sánchez

ISBN: 978-84-9125-105-7

URI: <http://hdl.handle.net/10481/40186>

Implementación de Modelos de Procesamiento de Imágenes para Aplicaciones con Restricciones de Tiempo Real.



UGR

Universidad
de Granada

Pablo Guzmán Sánchez

Programa Oficial de Doctorado en Tecnologías de la Información y la Comunicación.
Departamento de Arquitectura y Tecnología de Computadores.
Universidad de Granada.

Para optar al grado de

Philosophiae Doctor (PhD)

2015 Abril

DECLARACIÓN

Dr. Eduardo Ros Vidal, Catedrático de Universidad del Departamento de Arquitectura y Tecnología de Computadores de la Universidad de Granada, y Dr. Antonio Javier Alonso Díaz, profesor Titular de Universidad del Departamento de Arquitectura y Tecnología de Computadores de la Universidad de Granada,

CERTIFICAN:

Que la memoria titulada “Implementation of Models for Image Processing Applications with Real-Time Constraints”, ha sido realizada por D. Pablo Guzmán Sánchez bajo nuestra dirección en el Departamento de Arquitectura y Tecnología Computadores de la Universidad de Granada para optar al grado de Doctor por la Universidad de Granada.

El doctorando Pablo Guzmán Sánchez y los directores de la tesis Eduardo Ros Vidal y Antonio Javier Alonso Díaz. Garantizamos, al firmar esta tesis doctoral, que el trabajo ha sido realizado por el doctorando bajo la dirección de los directores de la tesis y hasta donde nuestro conocimiento alcanza, en la realización del trabajo, se han respetado los derechos de otros autores a ser citados, cuando se han utilizado sus resultados o publicaciones.

Granada, 14 de April de 2015

Directores de la Tesis:



Fdo. Eduardo Ros Vidal



Fdo. Antonio Javier Alonso Díaz

Doctorando:



Fdo. Pablo Guzmán Sánchez

AGRADECIMIENTOS

Tras un largo y duro camino, he podido concluir esta etapa de mi vida donde he tenido muchos obstáculos para seguir adelante. He estado acompañado de maravillosas personas y otras no tanto, las cuales me han ayudado a convertir en realidad esta tesis. Empezando por mi director de tesis Eduardo Ros al cual agradezco de corazón al confiar en mí y haberme ayudado en momentos críticos de mi vida. Agradezco también a Paco Pelayo por habernos introducido a Jarno y a mí en el apasionante mundo de la investigación. Sin lugar a dudas, tengo que dar las gracias a mis compañeros Jarno, Rodri, Leo, Fran, Sara, Quique, Matteo, Mauricio, Silvia, Niceto, Nolo, Fergu, Asensio, Richard, Jesus, Encarni, Manolo, Julio, Hector y Paco Illeras por aguantarme tanto tiempo y animarme en los malos momentos. También me gustaría dar las gracias por la ayuda prestada, durante mi estancia en Linköping (Suecia), a Michael Felsberg y Rudolf Mester y en especial a Jordi Altimiras por brindarme la oportunidad de vivir la situación mas dantesca de mi vida al ser atacado por un Gallus Gallus, supuestamente anestesiado, durante una exploración con el equipo de ultrasonidos.

Esta tesis no hubiera sido posible sin la presencia de mi madre, que una vez me escondió el cargador del portátil para que parara de trabajar en la tesis y me fuera a dar un paseo. También doy las gracias al resto de mi familia por apoyarme en esta etapa de mi vida, en especial a mi primo Migue por animarme y haber hecho de psicólogo.

Con motivo de agradecer a todos aquellos que han apoyado, les dedico un fragmento de una gran canción:

“Never gonna give you up, never gonna let you down
Never gonna run around and desert you
Never gonna make you cry, never gonna say goodbye
Never gonna tell a lie and hurt you”

ABSTRACT

IMPLEMENTATION OF MODELS FOR IMAGE PROCESSING APPLICATIONS WITH REAL-TIME CONSTRAINTS

Pablo Guzmán Sánchez, MSc Computer Science
University of Granada, 2015

Supervisor: Eduardo Ros Vidal, Antonio Javier Díaz Alonso

This work has diverse research branches, one of them is the implementation for high performance computing, while the other branches are focused on computer models and bio-inspired algorithms. In this thesis, it has developed different computer vision models for application in automotive, biomedical applications and food industry. Therefore, the main contribution of this work is the applied component, which requires an efficient implementation and processing in specific purpose architectures (for embedded systems) or high performance (to provide results and simulations efficiently, eg shortening time in diagnosis and monitoring of diseases).

This thesis is divided into five sections:

In the first and second section consists in a brief introduction (English and Spanish respectively) to computer vision where it is explained the typical steps that are performed in this area to solve an specific problems. The third section proposes an adaptation of an efficient optical flow model in a hybrid architecture (focal plane and soft-processor), with the objective of developing a real time system that work as driver assistant and advises to the driver about vehicles in overtaking to avoid unexpected maneuvers that could cause an accident. The forth section of this thesis is focused on the field of ultrasound, where it will be exposed a method based on an evolutive algorithm for segmenting the carotid artery to obtain an accurate area estimation and determinate the degree of stenosis of the patient. Additionally it is proposed a model to estimate sub-pixel motion, in combination with other techniques, to measure the wall artery displacement in order to determinate possible pathologies of the subject. The fifth section of this thesis is focused to design an experimental system in real time to analyze images in field of food industry. This system is able to analyze and process images, in order to reject those products that are defective. The last section summarizes the final conclusions of this thesis and exposed future works.

RESUMEN

IMPLEMENTACIÓN DE MODELOS DE PROCESAMIENTO DE IMÁGENES PARA APLICACIONES CON RESTRICCIONES DE TIEMPO REAL

Pablo Guzmán Sánchez, MSc Ingeniería Informática
Universidad de Granada, 2015

Supervisor: Eduardo Ros Vidal, Antonio Javier Díaz Alonso

Este trabajo contiene diversas líneas de investigación, una de ellas es la de implementación de esquemas de computación de altas prestaciones, otra de estas líneas se centra en modelos de computación y algoritmos bio-inspirados. En esta tesis doctoral se han desarrollado diversos modelos de visión para su aplicación en el automóvil, aplicaciones biomédicas y la industria de la alimentación. Por lo tanto, la contribución principal de este trabajo es su componente aplicada, que exige una implementación eficiente y en arquitecturas de procesamiento de propósito específico (para sistemas empotrados) o de altas prestaciones (para disponer de resultados y simulaciones de forma eficiente, por ejemplo acortando tiempos de diagnóstico y seguimiento de patologías).

Esta tesis se estructura principalmente en cinco apartados:

En la primer y segundo apartado, se expone una breve introducción (en Inglés y Español respectivamente) a la visión por computador donde se explican los pasos típicos que se llevan a cabo en este campo para la resolución de problemas concretos. En el tercer apartado se propone una adaptación a un modelo eficiente de flujo óptico en una arquitectura híbrida (plano focal y soft-processor) empotrada con el objetivo de desarrollar un sistema que sirva como asistente de conducción y advierta al conductor de vehículos en adelantamiento en tiempo real para evitar maniobras inoportunas que puedan provocar un accidente. La cuarta parte de esta tesis se centra en el campo de los ultrasonidos, donde se expondrá un método basado en algoritmos evolutivos para la segmentación de la arteria carótida para obtener una precisa estimación del área y de esta manera determinar el grado de estenosis del paciente. Adicionalmente se propone un modelo de estimación de movimiento sub-píxel, en combinación con otras técnicas, para medir el desplazamiento de las arterias para determinar posibles patologías del sujeto. El quinto apartado de esta

tesis se focaliza en diseñar un sistema, en tiempo real, para el análisis de imágenes en el ámbito industrial. Dicho sistema es capaz de analizar y procesar imágenes, con el objetivo de rechazar aquellos productos que están defectuosos. En el último apartado, se resumen las conclusiones finales de esta tesis y se proponen trabajos futuros.

TABLE OF CONTENTS

List of Tables	xvi
List of Figures	xviii
1. INTRODUCTION.....	27
1.1 General.....	27
1.2 Scientific Objectives	31
1.3 Project Framework.....	32
1.3.1 SAPVIA	32
1.3.2 ITREBA	33
1.4 Methods and Tools.....	33
1.5 Organization of Chapters	34
2. INTRODUCCION.....	35
2.1 General.....	35
2.2 Hipótesis y Objetivos Científicos	39
2.3 Marco de Proyectos.....	40
2.3.1 SAPVIA	41
2.3.2 ITREBA	41

2.4 Métodos y Herramientas	42
2.5 Organización de los Capítulos	43
3. ADVANCED DRIVER ASSISTANCE SYSTEMS	45
3.1 Introduction.....	45
3.2 Motivation.....	47
3.3 Background Material and Related Work	49
3.4 Material and Methods	50
3.5 System Description	55
3.5.1 Optical Flow Estimation	55
A. Lucas and Kanade Model for Optical Flow Estimation.....	57
B. Implementation.....	58
C. Pre-Selection of Points of Interest To Speed-Up the Optical Flow Computation.....	61
D. Co-Design Strategy	65
3.5.2 Optical Flow Evaluation	68
3.5.3 Experimental Results	71
3.6 Overtaking Detection	74
A. Estimation of Points of Interest, Optical Flow and Pre-Segmentation	75
B. Binary Matching.....	76
C. Object Detection and Dimension Filtering.....	78
D. Egomotion Filtering	80
E. Increased Vehicle Detection Distance.....	82
3.6.1 System Evaluation and Results.....	84
A. System Performance and Precision Evaluation	84
B. Distance Detection Evaluation	85
C. System feasibility in Low-Light Conditions	87
3.6.2 Discussion	88

3.7 Conclusions.....	89
4. ULTRASOUND IMAGING.....	91
4.1 Introduccion	91
4.2 Artery Segmentation in Ultrasound Imaging.....	93
4.2.1 Motivation.....	93
4.2.2 Background Material and Related Work	95
4.2.3 Material and Methods	95
A. Parametric Active Contour (Snakes).....	95
B. Region-Based Active Contour Model (ACM)	96
C. Segmentation Based on Fuzzy C-Mean Clustering.....	97
D. Active Shape Models (ASM).....	97
4.2.4 System Description	98
A. Evolutionary Scheme	98
B. Feature Extraction	100
C. Ellipse Parameter Estimation	104
D. GPU Implementation	106
4.2.5 System Evaluation and Results.....	109
4.2.6 Discussion.....	113
4.2.7 Conclusions.....	115
4.3 Arterial Wall Tracking.....	115
4.3.1 Motivation.....	115
4.3.2 Background Material and Related Work	116
4.3.3 Material and Methods	117
A. Evaluated Methods.....	117
A.1 Block Matching.....	117
A.2 Optical Flow	118
A.3 Kalman Filter	120
A.4 Similarity Transformation.....	120

A.5 Soft Body Dynamics	121
B. Ground Truth Estimation via Ultrasound Simulation	122
4.3.4 System Description	124
4.3.5 System Evaluation and Results	126
4.3.6 Discussion	132
4.3.7 Conclusions	136
5. INDUSTRIAL INSPECTION	137
5.1 Introduction	137
5.2 Quality Inspection	138
5.2.1 Motivation	138
5.2.2 Background Material and Related Work	139
5.2.3 Material and Methods	140
A. Biscuit Segmentation	141
B. Feature Extraction and Fracture Detection	146
C. Classification	149
5.3 System Evaluation and Results	150
A. Segmentation Evaluation	150
B. Feature Extraction and Classification Evaluation	152
C. GPU Evaluation	154

5.4 Conclusions.....	156
6. CONCLUSIONS AND MAIN CONTRIBUTIONS	157
6.1 Summary.....	157
6.2 Scientific Objectives and Degree of Compliance	158
6.3 Future Work	158
6.4 Scientific Production.....	159
6.5 Main Contributions	160
7. CONCLUSIONES Y CONTRIBUCIONES PRINCIPALES	163
7.1 Sumario	163
7.2 Objetivos Cientificos y su Grado de Cumplimiento	164
7.3 Trabajo Futuro	164
7.4 Producción Científica.....	165
7.5 Contribuciones Principales	166
REFERENCES.....	169
APPENDIX A. TRACKING RESULTS.....	193
APPENDIX B. HSV AND SPHERICAL COLOR SPACES.....	194
APPENDIX C. SEGMENTATION: F-MEASURE VS. PARZEN WINDOW WIDTH.	195
APPENDIX D. SOME EXAMPLES OF THE TRAINING DATA SET.....	196
APPENDIX F. ACRONYMS AND ABBREVIATIONS LIST.....	196

LIST OF TABLES

Table 3.1. Increase/Decrease percentage with respect ten years ago (1991-2012) in accidents, injured people and fatalities involved in an accident of the main countries belonging to the European Union.....	49
Table 3.2. System performance evaluation obtained with a 176x144 spatial resolution.....	60
Table 3.3 Average angular error (AAE) and standard deviation (STD), in Yosemite sequence (without clouds), with the different approaches and densities on Matlab.	69
Table 3.4. Average angular error (AAE) and standard deviation (STD), in Yosemite sequence (without clouds), with the different approaches and densities on Eye-RIS™ system.....	71
Table 3.5. Extrinsic and intrinsic camera parameters used in our tests.	84
Table 3.6. Driver reaction time in relation to the relative speed, taking into account our distance detection (26.7 meters).	86
Table 4.1. Computation time evaluation of the main methods used in the described algorithm on the graphics processing unit (GPU) and parallel CPU implementation. AOS, additive operator splitting; FRS, fast radial symmetry.....	108
Table 4.2. Wall Displacement Simulation Parameters utilized in our experiments.	123
Table 4.3. Evaluated models where the used methods are indicated. Not all the combinations have been used because some of them were nonsensical.	127

Table 5.1. F-Measure results in different color spaces and Parzen Kernels. The results marked in bold correspond to the best results obtained in each approach (highest value per column).....	151
Table 5.2. LOOCV results combining different projections and classification techniques with different Top-Hat histogram sizes.	153
Table 5.3. Obtained LOOCV results combining different projections and classification techniques with different Radon histogram sizes.	154
Table 5.4. Computation time evaluation of the main methods used in our algorithm on GPU and parallel CPU implementation.....	155
Table A1. Longitudinal error results after being evaluated with different methods.	193
Table A2. Radial error results after being evaluated with different methods.	193
Table A3. Position Error results after being evaluated with different methods. .	193
Table A4. Diameter error results after being evaluated with different methods.	193

LIST OF FIGURES

Figure 1.1. Example of the different stages that involve a computer vision application. (a) Original image, (b) Early Cognitive Vision (Point of Interest), (c) Intermediate Cognitive Vision (Descriptors) and (d) High-level Cognitive Vision (Face Detection).	28
Figure 1.2. Illustration of BLIS (BLindspot Information Systems) system, developed by Ford. (Courtesy of Ford Motor Company).	29
Figure 1.3. An example of a Machine Vision inspection system installed in the production chain. (Courtesy of Cognex Corporation).	30
Figura 2.1. Ejemplo de las diferentes etapas que componen una aplicación de visión artificial. (a) Imagen original, (b) Visión Cognitiva Temprana (Puntos de interés), (c) Visión Cognitiva Media (Descriptoros) y (d) Visión Cognitiva de Alto Nivel (Detección facial).	36
Figura 2.2. Ilustración del funcionamiento del sistema BLIS (BLindspot Information Systems) desarrollado por Ford. (Cortesía de Ford Motor Company).37	37
Figura 2.3. Un ejemplo de un sistema de inspección industrial basado en visión artificial instalado en la cadena de producción. (Cortesía de Microscan Systems Inc.).....	39
Figure 3.1. Evolution of road safety, in the last ten years, of the main countries belonging to the European Union. First row illustrate the accidents while second row shows the fatalities and injured people involved in an accident.	48
Figure 3.2. EyeRIS v1.3 smart camera designed by Anafocus.....	51

Figure 3.3. Classical image processing (first row) and SIS image processing pipeline (last row).....	52
Figure 3.4. Internal Q-Eye scheme.	53
Figure 3.5. Complete Eye-RIS™ scheme (image acquisition and processing).	54
Figure 3.6. 5x5 Sparse Integration Block (SIB) representation.	59
Figure 3.7. Neighborhood propagation illustration.....	59
Figure 3.8. Comparison between the gains obtained in different integration blocks, after applying our approaches. Two different implementations are evaluated here, 5x5 (orange bars) and 9x9 (red bars). (Left to right) The first group of columns represents the Sparse Integration Block (SIB) factor gain; the second group shows the obtained gain after apply the optical flow 4:1 propagation. In the third column group figure the gain when the method is optimized in assembler while last column show total gain factor obtained after all the approaches are combined.....	61
Figure 3.9. Roberts Cross convolution filter.....	62
Figure 3.10. (a) Image edge detection in the original image. (b) Image edge detection with a sharpen pre-filtering.....	63

Figure 3.11. (a) System performance using pre-selected points of interest and 5×5 SIB. (b) System performance using pre-selected points of interest and 9×9 SIB. Colored mark illustrates three typical scenarios using different image edge densities. The green mark makes reference to the best performance cases (scenes with low edge density), the blue one, to the most common density values (normal scenes) and the red mark, to the worst cases (scenes with high edge density). The error bar represents the standard deviation of the results for 10 trials.	64
Figure 3.12. Initial Optical-flow Co-Design scheme used in Eye-RIS™ platform.	66
Figure 3.13. Different stages to estimate optical flow in Eye-RIS™ system.	68
Figure 3.12. Optical flow representation. The color corresponds with the direction of the optical flow vector while the magnitude is encoded as the color intensity.	72
Figure 3.13. Optical flow estimation in a traffic sequence. Flow field is overlaid with the original frame. In the first Row (a-c), the optical flow is estimated using 9 × 9 SIB. In the second Row (d-f), the optical flow is estimated using 5×5 SIB.	72
Figure 3.14. Binary edge map processed, with the proposed method in section 3.5.1.C, in the focal plane.	73
Figure 3.15. Optical flow estimation, on edges, in a traffic sequence. The average edge density, in these images, is 42.5%. Flow field is overlaid with the original frame. In the first row (a–c), the optical flow is estimated using 9×9 SIB. In the second Row (d–f), the optical flow is estimated using 5×5 SIB.	74

Figure 3.16. Algorithm dataflow scheme showing the processing structure computing the different steps.....	75
Figure 3.17. (a) Original image of an overtaking scene on a highway; (b) estimated optical flow using the Eye-RIS™ platform; and (c) optical-flow representation. The color corresponds to the direction of the optical-flow vector while the magnitude is encoded as color intensity.....	76
Figure 3.17. (a) Three historical maps and an MOI map; (b) Output map produced.	78
Figure 3.18 Illustration of the camera parameters needed to estimate distances from the camera.	78
Figure 3.19. Illustration describing the parameters involved in obtaining object width.	79
Figure 3.20. (a) Blobs obtained after the binary matching step; (b) Blob obtained when the clustering algorithm is applied.	80
Figure 3.21 Optical-flow scenario when the car turns left and the camera is fixed on the rear view mirror on the left side of the car.....	81
Figure 3.22. Area classification representation, where positive values indicate objects out of the AOI while negative values indicate objects within the AOI.	81
Figure 3.23. Example of projection filtering adaptation on a road scenario.....	82
Figure 3.24. Adaptable frame rate in the scene; sample split period in a sequence.	83
Figure 3.25. Frame rate obtained for different road scenes in which the vertical bar error indicates the deviation.....	85

Figure 3.26. Mean distance detection in relation to the left-hemisphere delay; the vertical bars indicate measurement error.	86
Figure 3.27. Rain scene in low-light condition; (a) Vehicle detected and discarded by the width filter; (b) Vehicle detected as valid after removing the width filter.	88
Figure 3.28. Example of true positive detections with our algorithm on a true scenario.	88
Figure 3.29. Example of false positive detections (first row) and false positives rejected by our approach (second row).	89
Figure 4.1. The ranking results that lead the causes of death in the world.	92
Figure 4.2. Typical manually-driven measurement in ultrasound (US) software (TeleMed) to determinate the diameter of the artery.	94
Figure 4.3. Illustration fitted in the boundaries of an artery by means an ellipse.	100
Figure 4.4. (a) Original image and (b) saliency map after applying fast radial symmetry.	101
Figure 4.5. (a) Original US image; (b) pixel orientation of the original image; (c) SRAD smoothing; (d) pixel orientation of the smoothed (SRAD) image. In b and c, the pixel local orientation is represented with colors with respect to the color reference frame.	103
Figure 4.6. (a) US image smoothed with SRAD; (b) LIP-Sobel gradient; (c) non-max suppression applied to LIP-Sobel gradient estimation; (d) binary representation of the edges.	104
Figure 4.7. Time consumption of the differential evolution algorithm in a (a) parallel CPU and (b) a GPU implementation.	107

Figure 4.8. The factor obtained on the GPU with respect to the parallel CPU implementation.	107
Figure 4.9. The evaluation of the proposed method with different mutation schemes with respect to the ground truth (manually marked) with its respective standard deviations (after 10 trials in each set up).....	110
Figure 4.10. Evaluation of the proposed method with different features mixed (left) without orientation and (right) including the orientation feature...111	
Figure 4.11. The final comparative results of the F-measure, recall and precision with respect to other state-of-the-art methods. ASM, active shape model.112	
Figure 4.12. The Bland–Altman figure comparing the best two obtained results (the proposed one and ASM).	112
Figure 4.13. (a,b) Original US images with the ground truth marked in red; (c,d) fuzzy C-mean; (e,f) active contour model; (g,h) parametric snake; (i,j) active shape model; (k,l) proposed method. Please see the main text for the meaning of each color.	113
Figure 4.14. Some results obtained with the proposed method in different patients.	114
Figure 4.15. Ultrasound images used in the data set test, generated by Field II U.S. simulator.	123
Figure 4.16. Block matching with sub-pixel accuracy by means of the optical flow scheme.....	124
Figure 4.17. Similarity transformation given the motion vectors obtained with the hybrid BM-optical flow method.	124

Figure 4.18. (a) Pipeline of the proposed method with physics simulation (mass-spring) and (b) Illustration of the spring connections in an ultrasound image.....	125
Figure 4.19. Incorporation of Kalman Filter in the previously proposed schemes. Each scheme is encoded with a different color.....	126
Figure 4.20. Longitudinal error results after being evaluated with different methods.	128
Figure 4.21. Radial error results after being evaluated with different methods..	129
Figure 4.22. Position error results after being evaluated with different methods.	129
Figure 4.23. Diameter error results after being evaluated with different methods.	130
Figure 4.24. Bland-Altman figure, where the best obtained results block matching & anisotropic TV-L1 (M5) are compared with (a) block matching (M1), (b) block matching (M5), (c) anisotropic TV-L1 (M3), and (d) anisotropic TV-L1 (M5).	131
Figure 4.25. Diameter evolution over time evaluated in a sequence with long displacements using (a) Lucas-Kanade, (b) block matching & Lucas-Kanade, (c) block matching, (d) anisotropic TV-L1, and (e) block matching & anisotropic TV-L1.....	133
Figure 14.26. Diameter evolution over time evaluated in a sequence with small displacements evaluated with (a) Lucas-Kanade, (b) Block Matching & Lucas-Kanade, (c) Block Matching, (d) Anisotropic TV-L1, and (e) Block Matching & Anisotropic TV-L1.....	134

Figure 4.27. Diameter evolution of the common carotid artery (CCA) in real ultrasound data in different subjects where first row (a) correspond to a healthy patient and the last two rows (b–c) belong to patients with presence of atheroma plaques.	135
Figure 5.1. Inspection system to detect imperfections of the products on the conveyor belt. (Courtesy of Cognex Corp.).....	137
Figure 5.2. A Maria© biscuit on the top of the custard in the production line (Courtesy of Dhul© Company).	138
Figure 5.3. Architecture processing illustration used in this work.	140
Figure 5.4. Data flow scheme showing the whole process through different steps.	141
Figure 5.4. (a) Original image. (b) Clean image after applying Equation (4) with $\alpha = 0.58$, $\beta = 0.55$, $\lambda = 0.9$	143
Figure 5.5. Illustration of spherical color representation $\theta \in 3\pi/2, \pi/2$ respect $\varphi \in 0, \pi/2$ and $\rho = 300$ and HSV color representation $H \in 0,1$ and $S \in 0,1$	143
Figure 5.6. (a) Foreground and (b) Background density of probabilities in the Spherical color domain generated with Gaussian Parzen Kernel. ...	144
Figure 5.7. (a) Original image, (b) Pre-segmented image using Equation (6) in spherical color domain, (c) Results after applying hole filling method.	145
Figure 5.8. Illustration of the proposed method for hole-filling in a binary image.	146
Figure 5.9. (a) Original Image. (b) Logarithmic inverse operation in gray level with $n=100$. (c) Inverse operation in gray level.	147

Figure 5.10. Logarithmic inverse with (a) $n= 50$, (b) $n= 100$, (c) $n= 150$, (d) $n= 200$	147
Figure 5.11. (a) Logarithmic inverse operation in gray level with $n=100$ (2). (b) Top hat with B structure of 7×7 , (c) DoG with $\sigma_1= 3$ $\sigma_2= 6$ and (d) the final result after removing all the non-desired information by means of the output of the previous step (color segmentation).....	148
Figure 5.12. Radon Transformation in a biscuit (a) without the presence of cracks or anomalies with its pertinent (c) normalized histogram (64-bins) output and (b) the Radon Transformation of a biscuit with the presence of cracks and its (d) normalized histogram (64-bins).	149
Figure 5.13. Best obtained results with (a) Parzen (b) Denoised Parzen (c) Ani-TV-L1 (d) Denoised Ani-TV-L1 through Spherical θ, φ color space...152	
Figure. C.1 Plot of the Logarithmic Inverse F-Measure with respect to the Parzen window width ρ in different Kernels (a) HyperCube, (b) HyperSphere, and (c) Gaussian in Spherical color space.	195
Figure. D.1. The first row corresponds to the positive cases (high quality samples) and the rest of the rows belong to some negative cases in the training data set.	196

1. INTRODUCTION

1.1 GENERAL

Machine Vision has been one of the most complex fields in computing. From the 70s, visual perception by computer has been studied and thanks to the new advances in technology; it has been possible to solve machine vision problems that previously were impossible from the computational point of view. It has even developed specific processing systems to execute computer vision applications in real time. The industry is one of the sectors where such kinds of systems have been very successful and thanks to those progresses in computer vision and the technology, it has been possible to develop advanced visual inspection systems for improving the production quality. These systems have been able to replace the human eye, to develop highly repetitive and well-defined tasks in a structured and controlled environment, which 20 years ago those systems were inconceivable. At the beginning those systems were designed to perform simple tasks due to its limitation in capacity and data processing models but in recent years has advanced significantly. Nowadays it can be found hybrid architectures including multi-core processors, DSP (Digital Signal Processor) and GPGPU (General-Purpose Computing on Graphics Processing Units). This field is not limited only to industrial inspection, but covers a huge horizontal spectrum.

Computer vision comes from a branch of artificial intelligence where make use of techniques such as pattern recognition, machine learning, prediction, etc. The purpose is to create a process capable of perceiving, analyzing, "thinking" and interacting in a controlled environment, with the objective of replacing tasks performed by humans. The advantages of using such systems are very wide such as cost reduction, precision, continued unsupervised processing system, avoid monotonous tasks, etc. Due this disipline was created recently, computer vision lacks in many cases of own standards. Generically, cognitive vision can be structured similarly to this in visual perception of primates [KAN91] or humans [NAK95]. This structure consists of three sequential layers, where the previous stage provides information to the next one. Those stages are as follow:

- **Early Cognitive Vision:** It is the first stage and its main rol consists in the extraction of space-time information from one or serveral images to feed back the next stage. The main challenge in this stage is the high degree of processing information due this operations are done in the pixel level. It can be cited for

example: Point of Interest [HAR88][SMI97], Edge detector [CAN83][CAN86], Local orientation pixel [KAS87a][BIG91], etc.

- **Intermediate Cognitive Vision:** Consist in the combination of the extracted features in the previous stage to obtain other more complex. The degree of precision in this stage is crucial; due it is propagated to the next stage. As extracted features it can be cited: Optical Flow [FLE90][GAU02], Disparity [SOL01][MUH02], Local Descriptor [BAY06][LOW99], etc.
- **High-level Cognitive Vision:** It is the last stage in the hierarchy and its responsibility is carry out high-level task as object recognition [DEA13], decision making, scene interpretation [THO94], etc. This stage is in charge to link the extracted vision features from the previous stage and the Artificial Intelligent (AI).

The exposed hierarchy corresponds to a bio-inspired framework where it can be used to solve computer vision problems. However, in many occasion, it is not needed utilize the last stage in order to solve computer vision problems. Figure 1.1 shows an illustrative example where all the described stages are carry out to solve a hipotetical face detection application.

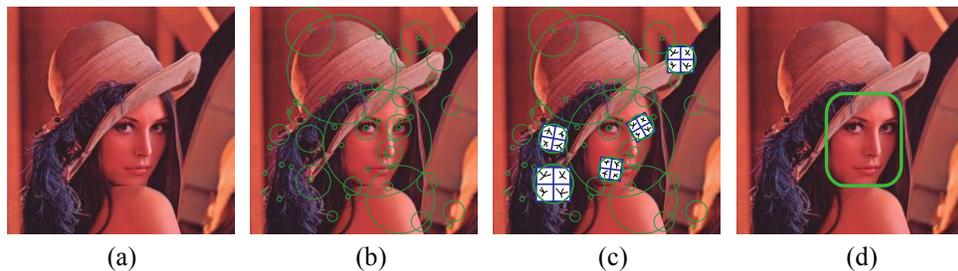


Figure 1.1. Example of the diferent stages that involve a computer visión application. (a) Original image, (b) Early Cognitive Vision (Point of Interest), (c) Intermediate Cognitive Vision (Descriptors) and (d) High-level Cognitive Vision (Face Detection).

Part of this thesis emerges from SAPVIA (Sistema Autónomo Programable de Visión Artificial, Autonomus Programable Artificial Vision System) project with the objective of developing an embedded application in an experimental plataforma created by Anafocus (Innovaciones Microelectrónicas, S.L.). This application is focalized in ADAS (Advanced Driver Assistance Systems) field for helping to the driver and prevents road accidents. Inside the wide spectral range of ADAS applications, we focus in LDW (Lane Departure Warning) systems. This kind of systems must be able to detect

vehicles in the near lane and advice to the driver the danger of moving to this lane, as illustrates Figure 1.2. It is important to remark that the designed system is based on a focal plane able to performance in the analogic plane ultra-fast basic operations and an embedded processor to control tasks and digital processing. The exposed work deal with how co-design existing methods and adapt them to this platform.

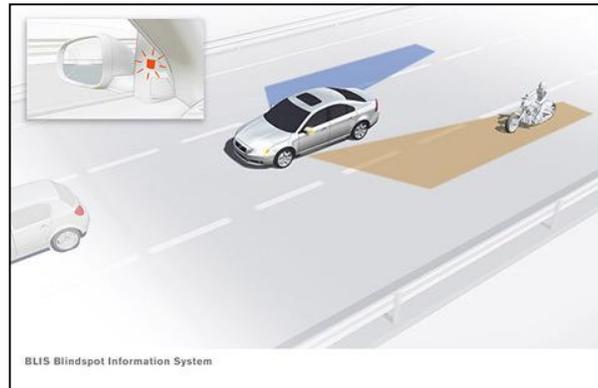


Figure 1.2. Illustration of BLIS (BLindspot Information Systems) system, developed by Ford. (Courtesy of Ford Motor Company).

In addition with the previously described before, this thesis evolves thanks to ITREBA (Procesamiento de Imagen/Video en Tiempo Real para Exploración Biomédica Activa, Active Biomedical Exploration Image/Video Process in Real-Time) project. It expects develop a platform able to capture, process (obtain significant parameters) and show Ultrasound sequences images, where the experts can visualize the extracted parameter in previously registered sequences with the objective of tracking the evolution of the patients and the parameters over the time. One advantage of working with Ultrasound imaging is the non-invasive nature, relative cheap and portable.

In the last decade has increased the cardiovascular diseases, this kind of diseases must be detected in time to avoid premature death. The cardiovascular diseases are one of the three main causes of death in the world. One that must be emphasize is the called “silent killer” o Aneurism that takes millions of lives each years. An Aneurism can cause a rupture of the blood vessel and therefore an immediate death. The non-invasive technique more utilized to determinate the diameter of the blood vessel is Ultrasound imaging and it is used for determining if excess the regular diameter size.

The main problem in Ultrasound imaging is that requires of well-qualified staff to realize the exploration, make an interpretation of the obtained images and search the anomaly in the dynamic exploration process. It is not only obtain an image and process it; the exploration process is complex and the expert must interpret what is displayed to find the areas to realize a diagnosis or follow a disease over the time. This thesis analyzes existing techniques and proposes an alternative method to reduce the degree of specialization to measure, in a semiautomatic way, the diameter of the blood vessel. In addition it developed a method able to measure the dynamic features of the arteries to describe the radial motion and it possible association to specific diseases. The challenge in Ultrasound imaging is the high degree of noise in the signal, resulting quite complex apply classic techniques to determinate the wall vessel motion with high precision.

As commented in the beginning of this section, computer vision has obtained a high impact in the industrial inspection and therefore it is a new challenge to cover in this work. The thesis finalize with the design of a high performance industrial inspection system to detect defective biscuits in the convey belt. As illustrative example of an automatic inspection system, Figure 1.3 shows one of these systems working on a conveyor belt to detect anomalies in plastic water bottles.



Figure 1.3. An example of a Machine Vision inspection system installed in the production chain. (Courtesy of Cognex Corporation).

This work represents a significant contribution in vision systems in industrial field. The combination of different algorithms and its implementation in real-time, in architectures of massively parallel processing, involves an increase of reliability and efficient availability of information.

1.2 SCIENTIFIC OBJECTIVES

In this work, the hypothesis is based on that several classic computer vision algorithms bring an intrinsic limitation caused by the kind of information that is processed. The combination of information generates more precise results for a specific environment representation. The combination of different extracted modalities of a vision system is not a simple sum of them, but requires its integration and cross validation in diverse levels.

In addition, the actual systems incorporate a processing speed that does not cover the real-time requirement for different applications. This affects the use in dynamic environment where the calculated estimations must be done efficiently to facilitate an agile perception-action cycle. An improvement on the vision systems performance involves its adaptation to any work environment where such information should be as requirement the real-time restriction. As work hypothesis, it is suggested the expansion of existent methods and adapt them in specific platforms to be processed in real-time, e.g. with the support of graphic processors (GPUs) or focal planes.

The work hypothesis where are focused the effort of this doctoral thesis is that the computer vision application to applied fields is not imminent. Requires the adaptation of vision models to obtain real-time performance, implementation in architecture with limited performance (embedded system) and making use of robust methods in relation with the signal noise (in the case of the obtained images in Ultrasound). The application of computer vision model in these fields (automotive, bio-medicine and industrial inspection) allows design systems able to provide enough information to simplify the decision-making and make in an efficient way, for the operator (the driver in the case of automotive application and an doctor in the case of ultrasound exploration application) or in automatic way (in the case of automatic industrial inspection). The objectives of this investigation work are listed below:

- Study of motion estimation models, segmentation, feature extraction and classification.
- Development of models for the detection of overtaking vehicles.
- Estimation of the dynamic properties of the artery.

- Segmentation and estimation of the diameter of the artery.
- Study and implementation of a real-time system for industrial inspection.
- Implementation of models in efficient processing platforms.

1.3 PROJECT FRAMEWORK

This thesis has been supported by two research projects, SAPVIA and ITREBA in which this thesis is based. Both projects have different objectives but at the same time those projects share a common denominator, high performance in signal processing architectures. A brief description of each project is carry out in this section.

1.3.1 SAPVIA

SAPVIA (Sistema Autónomo Programable de Visión Artificial, Ref. 1895) emerges as collaboration with the Spin-off Anafocus (Innovaciones Microelectrónicas) to design an ad-hoc application in an experimental embedded artificial vision system called Eye-RIS. This application is involved in the ADAS (Advanced Driver Assistance Systems) field. ADAS systems, based on passive sensors, are considered a huge challenge due the degree of complexity. These systems must work in variable environments with several objects and different scenarios. ADAS systems are designed to help the driver and prevent possible accidents on the road, i.e., work as a support assistance system for the driver. The purpose of the cited virtual assistance is to detect vehicles overtaking in the adjacent lanes and to warn the driver about the risk of changing to another lane. The developed platform by Anafocus is based on a focal plane (analogic) able to realize low-level vision operations (early vision) with a high performance computation. The system is also equipped with a 32 bit soft processor to performance control task and digital processing. The work realized in this project lies in evaluate different computer vision models and apply co-design techniques to adapt it to Eye-RIS system.

1.3.2 ITREBA

The Junta de Andalucía project ITREBA (Procesamiento de Imagen/Video en Tiempo Real para Exploración Biomédica Activa, TIC-5060) consists in the implementation of diverse computer vision techniques applied to Ultrasound imaging in medicine. Ultrasound techniques allow explore and determinate diverse patalogies in non-invasive way. Due this kind of technique requires an active exploration the degree of specialization, therefore only a minority of doctors can utilize this technique because the need of specific formation. The manual extracted parameters during the exploration have a huge inter-operator and inter-exploration variability, this make that this technique must be utilized in a complementary way with others more invasive. For this reason, the fundamental objective in this project is the implementation of a video processing system in real-time to allow, in the frame of vascular explorations with Ultrasounds, different capacities:

- Reduce the inter-operator and inter-exploration variability.
- Significant parameters extraction in a semi-automatic, that will be evaluated with cross validation with respect the one taken manually by the expert.
- Reduction of the degree of specialization.

For this purpose, it will be developed a platform that capture the sequence exploration, process it and visualize the results overlapped in real-time. That is to say, it will be implemented a “augmented visualization” scheme where the specialist can visualize different parameters (overlapped) with the registered exploration sequence. This allows adopt different exploration strategies like evaluate the similitude of the image with respect another stored in a previous exploration. This kind of evaluations can be the measurement of the size or the morphology of an atheroma plaque.

1.4 METHODS AND TOOLS

The followed methodology utilized in this thesis combine the theoretical study, generating models that will be experimentally evaluated through its implementation in computer equipment. It will be follow a theoretical-practice methodology proper of engineering, where it will be covered scientific-technical advances. Besides it will

define testing bench to use a verification pattern and the validation of the developed systems.

The methods implemented in this thesis have been done in different languages and development environment. To study the viability of the methods, it has been utilized MATLAB® environment to determinate the reliability and the robustness of the methods and integrate them into the system. The environment utilized in SAPVIA project was ECLIPSE, using different development languages. In the case of the NIOS II processor, the methods with non-high performance were developed in C while for algorithms with high performance requirement were developed in assembler (ASM). CFPP language was used to develop the algorithms in the focal plane; this language was created by Anafocus and let us develops the Q-Eye unit control. Regarding to ITREBA project, the development environment was Microsoft Visual Studio 2010 with .NET™ technology and OpenCL (Open Computing Language) to develop the massive parallel architecture GPU (Graphics Processing Unit).

1.5 ORGANIZATION OF CHAPTERS

The structure of this thesis is divided in three main chapters, as follow. Chapter 3 describe how work the proposed novel vision system to detect overtaking vehicles. This chapter is divided in three sections, section 3.4 details the utilized architecture to develop the described system, section 3.5 introduces an efficient co-design to estimate optical flow in real time in this architecture and properly evaluated. Section 3.6 details the final system implementation to detect the overtaking vehicles with the pertinent evaluations and finally the conclusions are exposed in section 3.7. Chapter 4 focuses in two applications based on Ultrasound imaging, where section 4.2 is oriented to the segmentation of the artery in transversal section through an evolutionary algorithm and compared with the state of art. Section 4.3 exposed a novel method to determinate the radial motion in the vessel wall in longitudinal Ultrasound imaging. This method makes use of a combination of Block Matching, Variational Optical Flow and semi-elastic methods with the objective to obtain a robust sub-pixel estimation. The exposed method is also exhaustively evaluated with other methods to confirm its reliability. Chapter 5 presents the role of computer vision in the food industry field where it is exposed a system able to analyze and to detect biscuits with presence of fractures to reject out the convey belt. Finally, chapter 6 discusses the obtained results and the obtained scientific contributions of this thesis.

2. INTRODUCCION

2.1 GENERAL

La Visión Artificial ha sido uno de los campos más complejos en la computación por ordenador. Desde la década de los 70 se ha estado estudiando la percepción visual por computador y gracias a los nuevos avances de la tecnología, hoy en día se pueden resolver problemas de visión artificial que antes resultaban imposibles desde el punto de vista computacional. Se han desarrollado incluso sistemas de procesamiento de propósito específico para la visión en tiempo real. La industria ha sido el sector donde, este tipo de sistemas, han tenido una gran acogida. Gracias a los avances en el campo de la visión artificial y la tecnología han hecho posible desarrollar sistemas de inspección visual, mejorando la calidad en la producción. Estos sistemas han sido capaces de reemplazar el ojo humano, para desarrollar tareas muy repetitivas y bien definidas en entornos estructurados y controlados, que hace 20 años eran impensables. Al principio fueron diseñados para realizar tareas sencillas, debido a su limitación en la capacidad y modelos de procesamiento de datos pero en los últimos años se ha avanzado significativamente. Hoy en día, se pueden encontrar arquitecturas híbridas que incluyen procesadores de varios núcleos, DSP (Digital Signal Processor) y GPGPU (General-Purpose Computing on Graphics Processing Units). Este campo no está limitada sólo a la inspección industrial, sino que abarca un inmenso espectro horizontal.

La visión artificial surge como una rama de la inteligencia artificial en las que se emplean técnicas derivadas tales como reconocimiento de patrones, aprendizaje, predicción, etc. La finalidad de la visión artificial es la creación de un proceso capaz de percibir, analizar, “razonar” e interactuar en un entorno controlado, con el objetivo de sustituir tareas realizadas por el ser humano. Las ventajas del uso de este tipo de sistemas son muy amplias tales como reducción de costes, precisión, sistema continuado de procesamiento sin supervisión, eludir tareas monótonas a operarios, etc. Al tratarse de una disciplina con poco tiempo de vida, la visión artificial carece, en muchas ocasiones, de estándares propios. De forma genérica, la visión cognitiva se puede estructurar de forma similar a la presente en la percepción visual de los primates [KAN91] o la del ser humano [NAK95]. Dicha estructuración está compuesta por tres capas secuenciales, donde la etapa anterior provee información a la siguiente. Dicha etapas son las siguientes:

- **Visión Cognitiva Temprana:** Es la primera etapa y su rol principal consiste en extraer información espacio-temporal de una o varias imágenes consecutivas, para realimentar a la siguiente etapa. El principal reto que conlleva esta fase, es el alto volumen de información que se debe procesar dado que las operaciones se realizan a nivel de píxel. Como ejemplos podemos citar: Puntos de Interés [HAR88][SMI97], Detección de bordes [CAN83][CAN86], Orientación local del píxel [KAS87a][BIG91], etc.
- **Visión Cognitiva Media:** Consiste en la combinación de las características extraídas en la etapa anterior para obtener otras más complejas. El grado de precisión de las características obtenidas es una gran responsabilidad de esta etapa, ya que el grado de precisión y robustez es propagado a la etapa posterior. Como características extraídas en esta etapa podemos citar las siguientes: Flujo óptico [FLE90][GAU02], Disparidad [SOL01][MUH02], Descriptores locales [BAY06][LOW99], etc.
- **Visión Cognitiva de Alto Nivel:** Comprende la última etapa en la jerarquía y es la responsable de llevar tareas de alto nivel tales como reconocimiento de objetos [DEA13], interpretación de la escena [THO94], toma de decisiones, etc. Esta etapa es la encargada de enlazar las características de visión extraídas en la etapa anterior y la Inteligencia Artificial (IA).

La jerarquía expuesta corresponde a un marco bio-inspirado en el que nos podemos basar para resolver problemas pertinentes a la visión artificial. No obstante, en muchas ocasiones, no es necesario utilizar la última capa para solucionar tareas relacionadas con la visión. La Figura 2.1 muestra un ejemplo ilustrativo donde se llevan a cabo las etapas previamente descritas para una hipotética aplicación de detección facial.

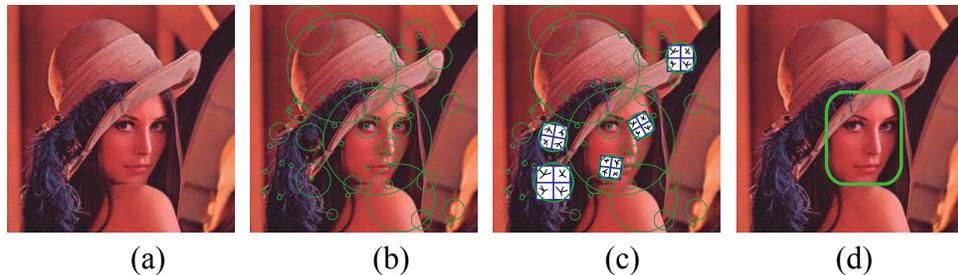


Figura 2.1. Ejemplo de las diferentes etapas que componen una aplicación de visión artificial. (a) Imagen original, (b) Visión Cognitiva Temprana (Puntos de interés), (c) Visión Cognitiva Media (Descriptores) y (d) Visión Cognitiva de Alto Nivel (Detección facial).

Parte de esta tesis surgió a partir de un proyecto denominado SAPVIA (Sistema Autónomo Programable de Visión Artificial) con el objetivo de implementar una aplicación embebida para una plataforma experimental desarrollada por Anafocus (Innovaciones Microelectrónicas, S.L.). Dicha aplicación está focalizada en el ámbito de los ADAS (Advanced Driver Assistance Systems) para la ayuda a la conducción y prevención de accidentes en carretera. Dentro del amplio espectro de los ADAS, nos centraremos en los sistemas LDW (Lane Departure Warning). Este tipo de sistemas deben ser capaces de detectar vehículos en el carril próximo al nuestro y advertir al conductor del peligro que conllevaría un cambio de carril, como ilustra la Figura 2.2. Como elemento a destacar, nuestro sistema se basa en un procesador de plano focal capaz de realizar operaciones básicas de alta velocidad en el plano analógico y de un procesador empotrado para tareas de control y procesamiento digital. El trabajo presentado versa sobre cómo adecuar los algoritmos existentes o los desarrollados mediante técnicas de co-diseño orientadas al presente dispositivo.

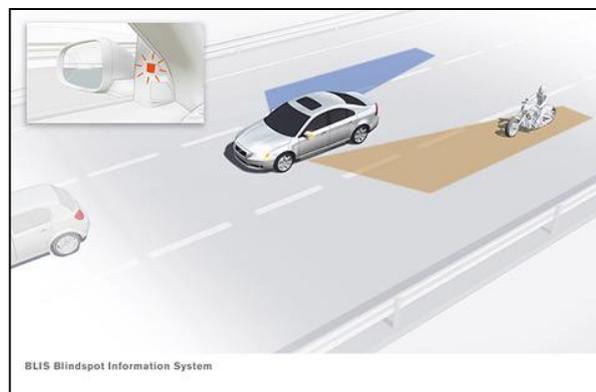


Figura 2.2. Ilustración del funcionamiento del sistema BLIS (BLindspot Information Systems) desarrollado por Ford. (Cortesía de Ford Motor Company).

Adicionalmente a lo descrito anteriormente, la tesis evoluciona gracias al proyecto de excelencia de la Junta de Andalucía ITREBA (Procesamiento de Imagen/Video en Tiempo Real para Exploración Biomédica Activa) en el que se pretende desarrollar una plataforma capaz de capturar, procesar (obtención de parámetros significativos) y mostrar secuencias de ultrasonidos. Donde los especialistas podrán visualizar los parámetros extraídos en las secuencias registradas previamente con el objetivo de realizar seguimientos de pacientes y la evolución de dichos parámetros en diversas monitorizaciones. Una de las ventajas de trabajar con dispositivos de ultrasonidos es que son técnicas de exploración no invasiva, relativamente barata y portátil.

En la última década se ha producido un incremento de enfermedades cardiovasculares que si no son detectadas a tiempo pueden producir la muerte. De hecho las patologías cardiovasculares se encuentran entre las tres principales causas de muerte evitable. Uno de los casos a destacar es el denominado “asesino silencioso” o Aneurisma que se cobra muchas vidas al cabo del año. Una aneurisma puede provocar la ruptura del vaso sanguíneo y producir la muerte instantáneamente. La técnica no invasiva más utilizada se basa en los ultrasonidos, para determinar si el diámetro del vaso excede el tamaño típico.

El problema de las técnicas de ultrasonidos es que requieren de personal altamente especializado para realizar la exploración, interpretar las imágenes obtenidas y buscar el punto o anomalía durante ese proceso de exploración dinámico. Es decir, no se trata de obtener una imagen y luego procesarla, el mismo proceso de exploración es complejo, el especialista debe interpretar lo que está visualizando para encontrar las zonas que quiera revisar (y registrar) en las que puede tomar medidas en imágenes concretas, para hacer un diagnóstico o seguimiento de una patología. Se realizará un estudio de las técnicas existentes y de una propuesta que mejore la precisión para reducir el grado de especialización y poder medir semiautomáticamente el diámetro de los vasos. Adicionalmente implementaremos un sistema capaz de evaluar características dinámicas de las arterias para describir su movimiento radial y una posible asociación con determinadas patologías. El mayor reto en resolver este problema es el alto grado de ruido en la señal, resultando bastante complejo aplicar técnicas clásicas para determinar el movimiento de paredes de vasos con gran precisión.

Como se ha comentado al inicio de esta sección, la visión artificial ha tenido un alto impacto en la inspección industrial y por lo tanto es un reto que pretende afrontar este trabajo, por ello esta tesis concluye con el desarrollo de un sistema de inspección industrial de alto rendimiento para detectar galletas defectuosas en la cadena de producción. Como ejemplo ilustrativo del funcionamiento de sistemas de inspección automatizados, la Figura 2.3 muestra uno de estos sistemas en pleno funcionamiento para la detección de anomalías en los botellines de agua.



Figura 2.3. Un ejemplo de un sistema de inspección industrial basado en visión artificial instalado en la cadena de producción. (Cortesía de Microscan Systems Inc.).

Este trabajo representa un avance significativo en los sistemas de visión en el ámbito industrial. La combinación de información entre diferentes algoritmos existentes y su implementación en tiempo real, en arquitecturas de procesamiento masivo, implican un incremento en la fiabilidad y disponibilidad eficiente de información que aportará el sistema.

2.2 HIPÓTESIS Y OBJETIVOS CIENTÍFICOS

En este trabajo de tesis, la hipótesis consiste en que muchos modelos de visión por computador clásicos conllevan una limitación intrínseca provocada por el tipo de información sobre la que trabajan. Una combinación de información produciría unos resultados más precisos para la representación de un entorno. La combinación de distintas modalidades extraídas de un sistema de visión no es una simple suma de ellas sino que requiere de su integración y validación cruzada a diversos niveles.

Adicionalmente, los sistemas actuales poseen una velocidad de trabajo que no cumplen los requisitos de tiempo real impuestos por distintas aplicaciones, lo que afecta el uso en entornos dinámicos donde el cálculo de las estimaciones se debe hacer de forma eficiente para facilitar un ciclo de percepción-acción ágil. Una mejora de las prestaciones en los sistemas de visión implicaría su adaptación a cualquier entorno de trabajo donde dicha información sea como requisito el tiempo real. Con esta hipótesis de trabajo se plantea extender los modelos y adaptarlos a plataformas de tiempo real, por ejemplo con el apoyo de procesadores gráficos (GPUs) o planos focales.

La hipótesis de trabajo en la que se centran los esfuerzos de esta tesis doctoral es que la aplicación de modelos de visión por computador a campos de aplicación no es

inmediata. Requiere de adaptaciones de modelos de visión para obtener funcionamiento en tiempo real, implementación en arquitecturas de prestaciones limitadas (en el ámbito de los sistemas empotrados) y utilización de modelos robustos frente a ruido de la señal (como en el caso de aplicación de modelos de visión a imágenes obtenidas con ultrasonidos). La aplicación de modelos de visión en estos campos (automóvil , biomédico e inspección industrial) permite diseñar sistemas capaces de suministrar información suficientemente rica para que la posterior toma de decisiones sea sencilla y pueda realizarse de forma eficaz por parte de un operador (el conductor en el caso de aplicación en automóviles y un especialista médico en el caso de aplicación en exploración con ultrasonidos) o de forma automática (en el caso de inspección industrial automatizada). Los objetivos de este trabajo de investigación son los siguientes:

- Estudio de modelos de estimación de movimiento, segmentación, extracción de características y clasificación.
- Desarrollo de modelos para la detección de vehículos en adelantamiento.
- Estimación de las propiedades dinámicas de la arteria.
- Segmentación y estimación del diámetro de la arteria.
- Estudio e implementación de un sistema de inspección industrial en tiempo real.
- Implementación de los modelos en plataformas de procesamiento eficiente.

2.3 MARCO DE PROYECTOS

Esta tesis ha sido soportada por dos proyectos de investigación, SAPVIA e ITREBA en los cuales se basa esta tesis. Ambos proyectos tienen objetivos muy diferentes pero al mismo tiempo comparten un denominador común, procesamiento de la señal en arquitecturas de altas prestaciones. A continuación se realizará una breve descripción de cada uno de los proyectos mencionados.

2.3.1 SAPVIA

SAPVIA (Sistema Autónomo Programable de Visión Artificial, Ref. 1895) surge como una colaboración con el Spin-off Anafocus (Innovaciones Microelectrónicas, S.L.) con el propósito de diseñar una aplicación ad-hoc para un sistema experimental empotrado de visión artificial denominado Eye-IRIS. El ámbito de la aplicación de dicho sistema esta focalizado en el marco de los ADAS (Advanced Driver Assistance Systems). Los Sistemas ADAS, basados en sensores pasivos, están considerados como uno de los mayores retos actuales dentro del campo de la visión artificial, ya que nos encontramos en un ambiente muy variable con diversos objetos y diferentes escenarios. Este tipo de sistemas están orientados para la ayuda a la conducción y prevención de posibles accidentes en la carretera, es decir, funcionan como un asistente de apoyo para el conductor. La finalidad de dicho asistente virtual es la de detectar vehículos en el carril adyacente y advertir al conductor del peligro que conllevaría un cambio de carril. La plataforma desarrollada por Anafocus está basado en un procesador de plano focal (Analógico) capaz de realizar operaciones de baja visión (Visión Temprana) en dicho procesador a alta velocidad. El sistema esta también acompañado de un procesador de 32 bits para tareas de control y procesamiento digital. El trabajo realizado en este proyecto consiste en evaluar diferentes modelos de visión y realizar técnicas de co-diseño para adaptarlo a este dispositivo.

2.3.2 ITREBA

El proyecto de excelencia de la Junta de Andalucía ITREBA (Procesamiento de Imagen/Video en Tiempo Real para Exploración Biomédica Activa, TIC-5060) va dirigido a la implementación de técnicas de visión artificial aplicadas al marco de los ultrasonidos en el ámbito de la medicina. Las técnicas de ultrasonidos permiten explorar e identificar diversas patologías de forma no invasiva, pero al tratarse de una exploración activa requieren de un alto grado de especialización. Por ello sólo una minoría de médicos puede utilizar esta técnica, ya que precisan de formación específica. Los parámetros extraídos manualmente durante la exploración tienen una gran variabilidad inter-operador e inter-exploración, esto hace que esta técnica se utilice de forma complementaria con otras más invasivas. Por todo ello, el objetivo fundamental de este proyecto es la implementación de un sistema de procesamiento de video en tiempo real que permita, en el marco de exploraciones vasculares con ultrasonidos, varias capacidades:

- Reducir la variabilidad inter-operador e inter-exploración.
- Extracción de parámetros significativos de forma automática (que evaluaremos con validación cruzada con respecto a los tomados manualmente por el especialista).
- Reducción del alto grado de especialización exigido.

Para ello se desarrollará una plataforma que reciba la secuencia de exploración, la procese y visualice los resultados superpuestos con la secuencia de exploración en tiempo real. Es decir se implementará un esquema de “visualización aumentada”, en el que el especialista pueda visualizar diversos parámetros (superpuestos) con la secuencia de exploración que está registrada. Esto le permitirá adoptar distintas estrategias de exploración como por ejemplo evaluar la similitud de la imagen (con respecto a otra almacenada en una exploración anterior) mientras está realizando una medida (como tamaño o morfología de una placa de ateroma) en el marco de una monitorización de una patología vascular con reconocimientos periódicos.

2.4 MÉTODOS Y HERRAMIENTAS

La metodología que se va a seguir en esta tesis combinará el estudio teórico, con la generación de modelos que serán validados experimentalmente mediante su programación en equipos informáticos. Se seguirá por tanto una metodología teórico-práctica propia de una ingeniería, en la que se abordan de forma integrada avances científico-técnicos. Además, se definirán bancos de pruebas que sirvan como patrón de verificación y validación de los sistemas desarrollados.

Las implementaciones algorítmicas realizadas en esta tesis se elaboraron en varios lenguajes y entornos de programación. Para estudios de viabilidad de los algoritmos, se hizo uso del entorno MATLAB® donde se evaluaban modelos de visión artificial con el objetivo de determinar la fiabilidad y robustez de los mismos e integrarlos posteriormente a nuestro sistema. En el caso del proyecto SAPVIA se realizaron implementaciones en el entorno ECLIPSE programando en diferentes lenguajes de programación en función de la plataforma a la que iba dirigida. En el caso de la programación para el procesador (soft-core) NIOS II de altera se implemento en C para rutinas que no requerían alto rendimiento, mientras que para algoritmos mas complejos se desarrollaron en ensamblador (ASM). La parte del código pertinente al plano focal se implemento en un lenguaje propio (CFPP) para programar la unidad de control Q-

Eye. A lo que respecta al proyecto ITREBA el entorno de desarrollo utilizado fue Microsoft Visual Studio 2010 conjunto al sistema .NET™ y OpenCL (Open Computing Language) como lenguaje de programación para la arquitectura de procesamiento masivo GPU (Graphics Processing Unit).

2.5 ORGANIZACIÓN DE LOS CAPÍTULOS

La estructura de este trabajo se divide en tres capítulos principales, de la siguiente manera. En el Capítulo 3 se describe el funcionamiento un sistema de visión novel para la detección de vehículos en adelantamiento. Este capítulo está dividido principalmente en tres secciones, la sección 3.4 detalla la arquitectura que se empleará para desarrollar el sistema propuesto, la sección 3.5 introduce un eficiente co-diseño para estimar flujo óptico en tiempo real en esta arquitectura con sus pertinentes evaluaciones. La sección 3.6 detalla la implementación del sistema final para la detección de vehículos en adelantamiento con sus respectivas evaluaciones y finalmente se expondrá las conclusiones finales. El Capítulo 4 se centra en dos aplicaciones en el campo de imágenes por ultrasonidos, donde la sección 4.2 se centra en el marco de la segmentación de la arteria en transversal mediante un algoritmo evolutivo y es comparado con el estado de la técnica. La sección 4.3 se focaliza en el movimiento radial de las paredes de las arterias en imágenes de ultrasonidos longitudinales donde se propone una combinación híbrida de Block Matching y Flujo Óptico variacional y otras técnicas, con el objetivo de obtener una robusta estimación sub-píxel. Esta propuesta es también exhaustivamente evaluada con otros modelos para afianzar su robustez. El Capítulo 5 aborda la visión artificial en el campo de la industria alimenticia donde se expone un sistema capaz de analizar y detectar galletas que presentan fracturas para expulsarlas de la cinta transportadora. Finalmente, en el Capítulo 6 se discuten los resultados obtenidos y las contribuciones científicas de este trabajo.

3. ADVANCED DRIVER ASSISTANCE SYSTEMS

3.1 INTRODUCTION

The automotive industry invests substantial amounts of money in driver-security and driver-assistance systems as well as the governments to decrease the number of injured or fatalities on the road. This issue generates a great expectation to a wide range of beneficiaries where it can be cited for example the automotive industries that try to improve their car designs to be more attractive for the final user incorporating new systems to improve the security on the road, insurance companies also aim to minimize traffic accidents in order to maximize profit as well as the Governments that are direct beneficiaries due to the reduction of costs related with the road accidents (e.g. road repair, sanitary assistance, etc.)

The purpose of this chapter is to design an overtaking detection system (a.k.a. LDW) based on visual motion cues that combines feature extraction, optical flow, solid-objects segmentation and geometry filtering, working with a low-cost compact architecture based on one focal plane and an on-chip embedded processor. The handicap to solve the overtaking detection is the implementation of an efficient motion estimator due to most applications based on technique requires work in real-time; hence, this restriction must be taken into account. Optical flow is the 2-D motion information obtained from a 3-D world. This motion representation is widespread well-known and applied in the science community to solve a wide variety of problems.

It will be showed an efficient approach to estimate the motion velocity vectors with an architecture based on a focal plane processor combined on-chip with a 32 bits NIOS II processor. The presented approach relies on the simplification of the original optical flow model and its efficient implementation in a platform that combines an analog (focal-plane) and digital (NIOS II) processor. The system is fully functional and is organized in different stages where the early processing (focal plane) stage is mainly focus to pre-process the input image stream to reduce the computational cost in the post-processing (NIOS II) stage. It presented the employed co-design techniques and analyzes this novel architecture. The system's performance and accuracy with respect to the different proposed approaches described in the literature will be evaluated. Also it will be discussed the advantages of the described approach as well as the degree of efficiency, which can be obtained from the focal plane processing capabilities of the system.

The second part of this work is focused to describe the proposed approach to solve the overtaking problem. The exposed solution is divided into two stages: firstly analog processing on the focal plane processor dedicated to image conditioning and relevant image-structure selection, and secondly, vehicle tracking and warning-signal generation by optical flow, using a simple digital microcontroller. Our model can detect an approaching vehicle (multiple-lane overtaking scenarios) and warn the driver about the risk of changing lane. Thanks to the use of tightly coupled analog and digital processors the system is able to perform this complex task in real time with very constrained computing resources. The described method has been validated with a sequence of more than 15,000 frames and is effective under different traffic situations as well as weather and illumination conditions.

Our system is based on the Anafocus EYE-RIS™ platform [ROD08] applied to the detection of overtaking vehicles. The challenge of this work is to make use of a multi-purpose machine-vision platform, with a basic set of pre-designed vision primitives, to develop an ad-hoc vision software engine in a standard processor to detect overtaking vehicles. The use of SIMD analog processing capabilities (focal plane) of the platform combined with an optimized computation in the digital processor is the key element that allows the development of the system with significantly constrained resources. For this reason we have had to adapt the properties of the platform to design a suitable model to solve the problem in question, that is, to warn the driver about the risk of changing lane. In addition to this, it is important to note that our system goes beyond previous ones in terms of covering multiple driving situations and addressing the solution of image and motion artifacts not fully covered by previous works. This has motivated the inclusion of novel image-processing techniques as well as fusion with data concerning ego-vehicle kinematics. The final system validates a novel approach and the seldom-used advantage of mixed analog and digital processors and may provide relevant hints for mixed processing and processing optimization to interested readers.

The Sections are organized as follows: Section 3.2 introduces a brief incentive scheme where details the importance of the Advance Driver Assistance Systems (ADAS) while Section 3.3 resumes the state of art in this field. The description of the material utilized, in this work and other works that belong to the state of art, is summarized in Section 3.4. The proposed method, to solve the overtaking detection, is described in depth in Section 3.5 where it will be detailed the implementation and evaluation of an ad hoc optical flow method in EyeRIS™ platform and the description of the introduced model to detect approaching vehicles. Also it is presented a novel

approach to increase the distance detection without extra computation; in Section 3.6.1 the system is evaluated, taking into account the system performance, efficiency, distance detection and light conditions while Section 3.6.3 discuss the obtained results of the evaluated methods and finally Section 3.7 summarizes some conclusions and outlines for proposed future work.

3.2 MOTIVATION

Based on the data collected from CARE (Community Road Accident Database) [CAR14], the amount of accidents on the road (Figure 3.1) has not been decreased significantly from the last ten years. Such information makes us reflect on the importance of the new advances in technology to reduce the accidents. In the last decade, the automotive industry has spent a lot of effort and money on internal safe incorporating airbags, structures able to absorb the greatest possible energy of an impact or intelligent braking systems with the main objective of reducing fatal accidents.

These advances in engineering are reflected in the reduction of 59% in fatalities and 24% of injured people (Germany), as it is illustrated in Figure 3.1. In other words we can say that the figures reported one fatality for every 34 accidents in 1991, while in 2012 a fatal victim are produced for every 83 accidents. On the other hand it is important to remark that the statistics report an accident decrease of 22% in Germany, 15% in Spain, 37% in United Kingdom and a 59% in France while in Italy has an increase of 9%. Table 3.1 shows in more detail the increase or decrease of accidents, injured people and fatalities with respect ten years ago.

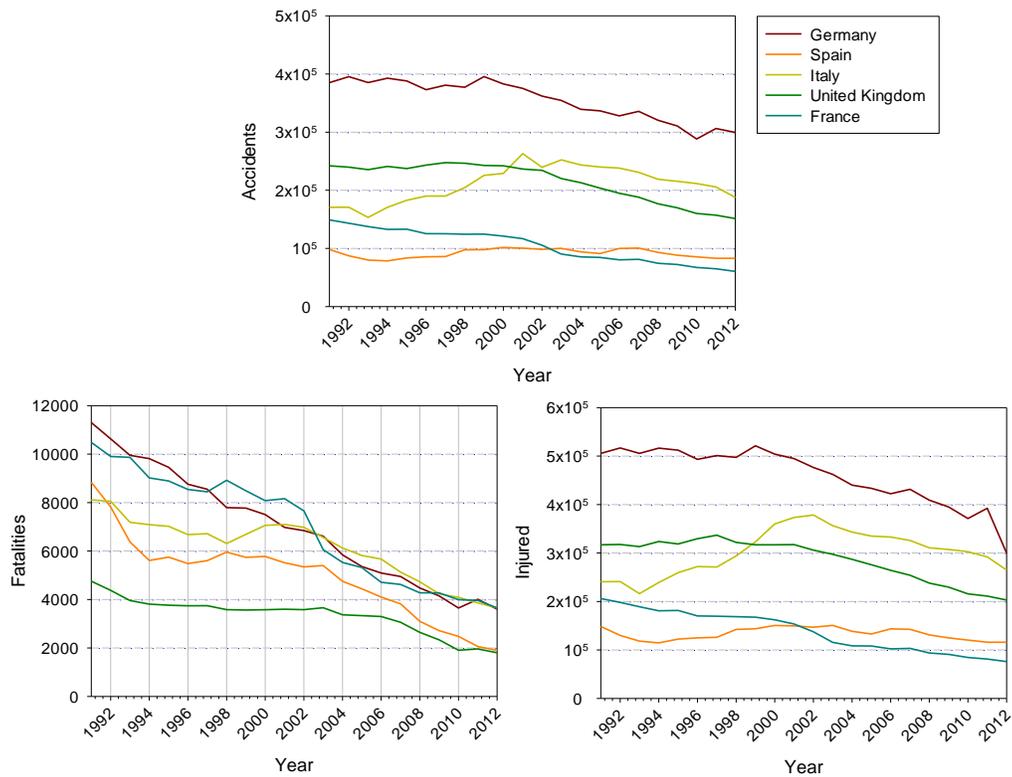


Figure 3.1. Evolution of road safety, in the last ten years, of the main countries belonging to the European Union. First row illustrates the accidents while second row shows the fatalities and injured people involved in an accident.

Nowadays the industry is investing effort to prevent the accident and not only focus in the internal safe of the vehicle, with the objective of reducing the statistics showed before. Advanced Driver Assistance Systems (ADAS) is an innovative technology that would be unthinkable ten years ago. This technology works as a virtual co-pilot, which provides information to the driver in order to improve the quality of driving. The driver assistances are not only limited to provide information but also can take part of the driving task such set the speed limitation, adapt the headlamps, automatic parking system and so on. Actually, the drivers appreciate the importance of road safety as well as the automotive industry. For this reason, science has invested a great effort to provide solutions to improve the road safety, driving comfort and limit energy consumption.

Table 3.1. Increase/Decrease percentage with respect ten years ago (1991-2012) in accidents, injured people and fatalities involved in an accident of the main countries belonging to the European Union.

Country	Accidents	Injured	Fatalities
Germany	-22.20%	-23.81%	-59.04%
Spain	-15.29%	-7.83%	-74.57%
Italy	9.97%	-0.001%	-59.03%
United Kingdom	-37.47%	2.40%	-39.36%
France	-59.40%	-9.2%	-14.15%

3.3 BACKGROUND MATERIAL AND RELATED WORK

The automotive industry has proposed different sensors for driver-assistance systems, ranging from *active sensors* such as LIDAR or RADAR [GAV01] where the price and the information pollution (possible interferences with sensor of similar technology) makes be considered as a suboptimal solution, *passive sensors* such as cameras [DIA08a] [MOT04] [SON07] [SAK06], up to hybrid systems that combine *active and passive sensors* [ALE07]. Based on the interaction of the sensors, the driver assistances can be divided, based on the way that the sensor interacts with the system, into two main classes: *Active* and *Passive* systems. Active systems perform automatic intervention with the vehicle. The functions of active systems range from a simple braking when parking the vehicle before reaches the obstacle (e.g.: vehicle, wall, pedestrian, etc) or avoid collisions in dangerous situations. At high levels of refinement, these systems interact on the direction of steering, braking and engine control to avoid collision with obstacles. An example of these kinds of systems is the automatic parking designed by Lexus. Passive systems are those that are designed to provide information to the driver. These systems without active nature can be seen as a preliminary to the automatic driving of the vehicle, where only warn or suggest to the driver about the risk of a given driving maneuver. An example can be contemplated in the parking system designed by Bosh, where feeds back to the driver about the distance space during the parking maneuver by means of acoustic signals.

In the proposed approach it will be uses a passive image sensor to design a lane-change driver-assistance and blind-spot vehicle passive detection system integrated as part of an advanced driver-assistance system (ADAS). By lane-change driver assistance we mean a system capable of warning the driver about the risk of changing lane when a car is approaching from behind. This problem has been considered by different companies such as Mobileye [MOB09], Volvo [VOL09a] and Ficosa [FIC14], although

they have not reported any specific data about their systems' performances or, for that matter, any benchmarking data. In the research field, different approaches have been proposed for this problem. Liu et al. [LIU07] suggest a feature detection model based on shadow detection and symmetry, while Bertozzi and Broggi introduce a generic obstacle detection [BLA07], making use of stereovision to detect objects pertaining to the road plane. Blanc et al. [BLA07] present a fast-feature-based method that extracts the horizontal edge and peak features in the image to recognize the front of a car by means of Support Vector Machine (SVM) classifiers.

Most of the proposed solutions designed to detect overtaking vehicles are based on motion direction filtering (motion pattern segmentation). Díaz et al. [DIA08b] make use of a classical optical-flow method [LUC81] in a customized digital-signal processor to segment approaching vehicles and then, using a standard processor, track the target with a Kalman filter [WEL02]. Mota et al. [MOT04] expose a bio-inspired model based on the Reichardt correlator [HAS56], where the motion filtering (features correlation) is applied to the vertical structures in the scene. Song and Chen [SON01] suggest a combined method that mixes sparse feature extraction based on corners [HAR88] and a correlation matching technique [SON01] to estimate the motion vector components in order to filter the motion and estimate the horizon line in the scene, thus discarding objects that do not belong to this line. Sakurai et al. [SAK06] use a similar approach to obtain the displacement vectors by means of the Sum of Absolute Differences (SAD) matching algorithm with the help of the vanishing point to address the correlation search.

3.4 MATERIAL AND METHODS

Most of these approaches (described in the previous section) share one thing in common, the use of ad hoc hardware architecture designed to solve the problem. Commercial and non-commercial embedded machine-vision devices present a diverse variety of hardware architectures and technologies. Ficosa [FIC14] and Mobileye [MOB09] opt for an Application-Specific Integrated Circuit (ASIC) technology, which allows low prices (mass production) together with low power consumption compared to Field Programmable Gate Arrays (FPGAs). Díaz et al. [DIA08b] and Mota et al. [MOT04] make use of reconfigurable circuits (FPGA) to develop their system, with the disadvantage of high price but with the possibility of reprogramming the circuit, which cannot be done with ASIC technology. SIMD (Single Instruction Multiple Data) digital architectures are also present in machine-vision systems such as those proposed by

Broggi and Gregoretti [BRO96], which are composed of 256 processor elements designed to detect obstacles on the road [BER96]. Nowadays SIMD architectures are very relevant to machine vision due to their high performance in massive parallel processing as Graphical Processing Units (GPUs), which are able to solve complex mathematical models in real time [ZAC07], but they have the drawback of being expensive, using a lot of power and not being physically portable. On the other hand, SIMD analog architectures (focal plane) [ROD08] [DUD06a] [DUD06b] [FOL08] avoid such inconveniences (physically portable, low price and low power consumption) and provide high performance and fast prototyping in comparison with Digital Signal Processors (DSP), FPGA and standard processors [ZAR08]. This fine-grain architecture must work in cooperation with standard processors to solve coarse-grain tasks. This analog-digital architecture scheme provides an efficient, effective and fast co-design system to solve a wide range of machine-vision tasks.

Analyzing the automotive industry economy, it is well known that the prices cannot exceed more than 50\$ in the production cost. Taking all these factors into account as well as the economical industrial factors, Eye-RIS™ v1.2 platform, designed by Anafocus (Fig. 3.2.), is an excellent candidate for an embedded machine-vision solution to solve the exposed problem in this work as well as cover the automotive industry economical restriction. This vision architecture is a compact system that includes all the elements needed for capturing (sensing) images, enhancing sensor operation, processing the image stream in real-time, interpreting the information contained in such image flow and supporting decision-making based on the outcome of such interpretation.



Figure 3.2. EyeRIS v1.3 smart camera designed by Anafocus

Eye-RIS™ system is a multipurpose platform designed to cover the main low-level machine vision primitives with a competitive price in relation with the offered solutions in the market. The present commercial smart cameras provide a reduced collection of machine vision primitives in relation with the Eye-RIS system. On the other hand, we

can reuse the same architecture to develop another sort of application much faster than an ad-hoc FPGA solution. Other advantages obtained making use of the focal plane are the GOPS (Giga Operations per Second) and power consumption, where the focal plane [ROD08] [18] (250 GOPS) consumes 4mW per GOPS while in a DSP [OLO02][21] (4.3 GOPS) the obtained consumption is 231mW per GOPS. Eye-RIS™ system employs an innovative and proven architecture in which image-processing is accomplished following a hierarchical approach with two main levels (Fig. 3.3):

Early-processing: This level comes right after signal acquisition. The basic tasks at this level are meant to extract useful information from the input image stream. Outputs of this level are reduced sets of data comprising image features such as object locations, shapes, edges, etc.

Post-processing: Here, the amount of data is significantly smaller. Inputs are abstract entities in many cases, and tasks are meant to output complex decisions and to support action-taking. These tasks may involve complex algorithms within long computational flows and may require greater accuracy than early processing.

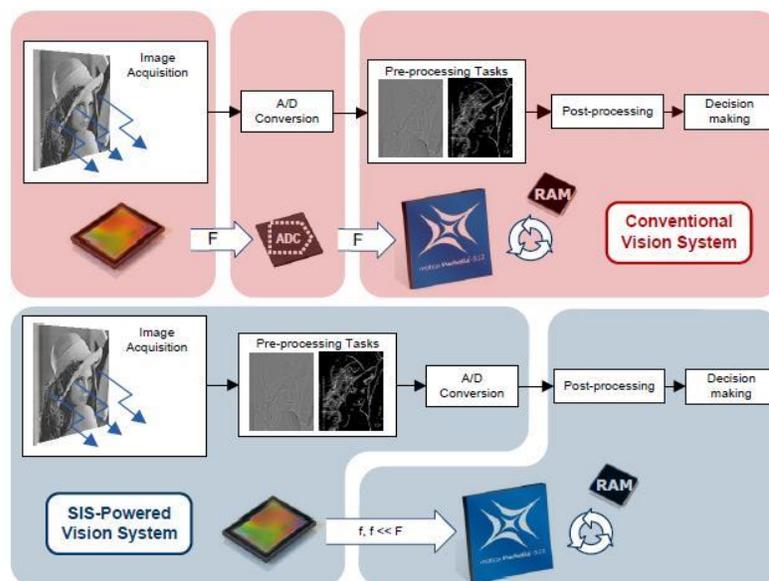


Figure 3.3. Classical image processing (first row) and SIS image processing pipeline (last row).

One unique characteristic of the Eye-RIS™ vision systems compared to other commercial solutions is that image acquisition and early-processing take place at the sensor, which is actually a Smart Image Sensor (SIS). In this device, image acquisition and pre-processing are performed simultaneously in all pixels of the SIS. Consequently,

images do not need to be downloaded from the sensor for the initial stages of the processing. This concept of concurrent sensing-processing extends the Image Sensor concept to the Smart Image Sensor one. The Smart Camera integrates a SIS named Q-Eye (Fig. 3.4), which is a quarter CIF (aka QCIF, 176x144) resolution fully-programmable SIS. It consists of an array of 176 x 144 cells plus a surrounding global circuitry. Each cell comprises multi-mode optical sensors, pixel memories, linear and non-linear analog processors and binary processors. Each cell is interconnected in several ways with its 8 neighboring cells, allowing for highly flexible, programmable, efficient, real-time image acquisition and spatial processing operations. In Smart Image Sensors, each local processor is merged with an optical sensor. This means that each pixel can both sense the corresponding spatial sample of the image and process this data in close interaction and cooperation with other pixels.

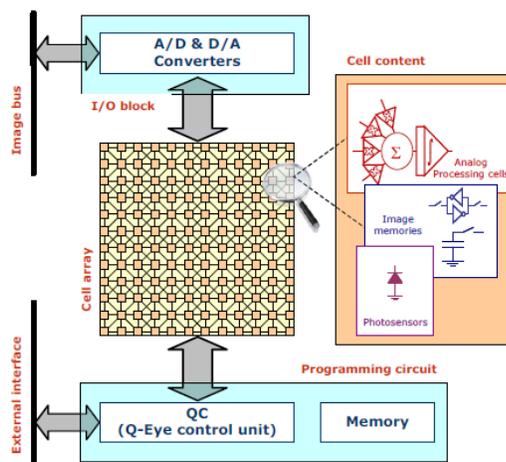


Figure 3.4. Internal Q-Eye scheme.

Eye-RIS™ v1.2 allows ultra-high processing speed beyond 1000 frames per second (fps) thanks to the incorporation of mixed-signal processing at the pixel level (enough light is assumed so that exposure time does not become a bottleneck). Processing speed is also application-dependent. Applications with intensive post-processing algorithms might present slower frame rates, since the performance may be constrained by the processing power of the embedded processor (NIOS II).

On the other hand, the Eye-RIS™ Vision System is not conceived for implementing intensive, iterative gray-level processing tasks. This kind of models can be implemented using the embedded microprocessor but its limited computational power highly limits the complexity of the vision models that can be processed in real time. For this reason,

it is necessary to take advantage of the resources available in the architecture to develop the proposed approaches in this work, to estimate optical flow. The Q-Eye must be seen as a powerful resource for a further processing, i.e. early processing; for this reason, a digital post-processing is needed. After the information of interest is extracted the post-processing level occurs in the NIOS II processor (32-bit RISC working at 70 MHz clock frequency), which is responsible for conducting more complex processing tasks with the reduced data set. Note that this digital processor has a basic design composed of a single arithmetic logic unit (ALU) without a floating-point unit (FPU) and a pipeline of 6 stages. The performance of this soft-core configuration is up to 71 DMIPs (Dhrystone's 2.1 benchmark) and therefore is quite constrained for image processing. In addition, this architecture can also capture information from external sensors through the different communication ports. This external information is extremely useful for achieving our objective.

The presented architecture (Fig. 3.5) has several advantages compared to conventional smart cameras, but imposes some restrictions in programming, due to the analog nature of the SIS Q-Eye, that shall be understood and taken into consideration by application developers. Such restrictions are associated to the internal memories that belong to the Q-Eye sensor. These memories are analogies and suffer of leaks over the time and in each access or operation. This limitation must be taken into account in the co-design stage due it can produce loss in the image quality. Next section the proposed system is described in detail, basing on the architecture presented in this section, in order to detect vehicles in overtaking.

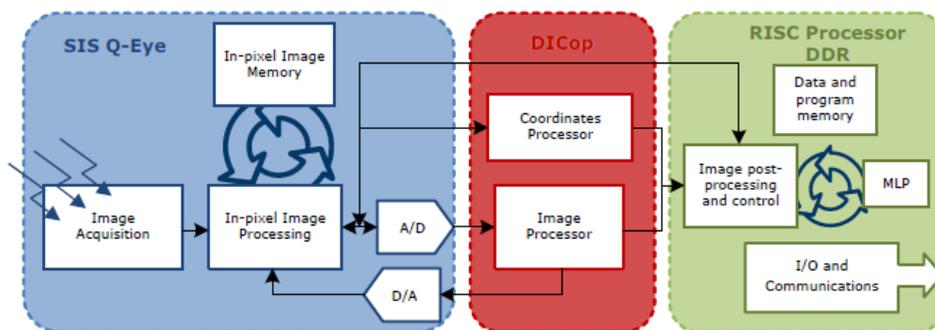


Figure 3.5. Complete Eye-RIS™ scheme (image acquisition and processing).

3.5 SYSTEM DESCRIPTION

Having given a brief outline of the working platform it will be described in detail the proposed model to detect a vehicle approaching from the rear and warn the driver about the risk of changing lane. As commented in Section 3.3, the use of passive sensors (optical sensors) has the advantage of being free of information pollution sensors, unlike RADAR or LIDAR systems. Thanks to the well-structured field of motion of the scene with the overtaking vehicle being clearly distinguished against the landscape, we have based our approach on the analysis of the patterns of motion presented on the image plane. One of the key issues in this work is to demonstrate that mixing analog-digital processors with significantly constrained computing performance and low-power consumption allows this task to be performed efficiently, and to demonstrate the co-design methodology between the analog and digital processor, together with useful generic methods to reduce computing complexity with a minimum degradation of accuracy.

The introduced method, to solve the overtaking detection, is described in depth in this section where it will be detailed, in Section 3.5.1, the implementation and evaluation of an ad hoc optical flow method in EyeRIS™ platform while the explanation of the proposed model to detect approaching vehicles is described in Section 3.5.2 where it is evaluated in different scenarios and it will be also presented a novel approach to increase the distance detection without extra computation.

3.5.1 Optical Flow Estimation

The term Optical Flow refers to the visual phenomenon due to the apparent movement perceived when we move through a scene and/or regarding the objects moving within it. It represents the projection of the 3-D motion presented in the scene to the 2-D plane of the image sensor or the retina. Note that as a consequence of this projection, depth information is partially lost and the estimation of the 3-D scene structure and motion from the available 2-D field is a very complex task. Optical flow has been extensively studied in the computer vision community (see for instance [BAR92]).

Different approaches have been proposed, in the scientific framework, to estimate the optical flow field. The most widely used ones are the gradient based methods. These methods are based on the constant-brightness assumption. An extended model is the

well-known local method described by Lucas and Kanade [LUC81]. Another classical model is the one proposed by Horn and Schunck [HOR81], which introduces a global constraint of smoothness to solve the aperture problem. An actual modification suggested by Brox and Bruhn [BROX04] formulates a new approach to solve the Horn and Schunck model's Achilles heel, the linear smoothness constraint to satisfy the spatial coherence; Brox et al. introduce a non-linear constraint of smoothness which preserves the optical flow boundaries. Another group of methods are based on local phase correlations. Those methods rely on how the effects of displacement in the spatial domain result in the frequency domain. As example of the use of phase information for optical flow it can be cited the work developed by Fleet and Jepson [FLE90] [FLE92]. Correlation techniques are also used in the motion component vector estimation, where block matching methods and similar schemes as the one proposed by Camus [CAM97] are valid alternatives.

In addition to the model choice used to compute the optical flow, its performance and computing resource demands are key elements to develop an embedded system for real-world applications. In the framework of real-time computing approaches Díaz et al. [DIA08b], making use of the Lucas and Kanade [LUC81] approach, developed an embedded system for lane-change decision aid in driving scenarios. Other authors as Mota et al. [MOT04] and Köhler [KOH09] propose bio-inspired models based on Reichardt correlators [HAS56] for the design of low cost approaches. In the framework of analog approaches, authors such as Stocker et al. [STO06] present a focal-plane aVLSI sensor to obtain the optical flow components based on the Horn and Schunck model [HOR81] while Mehta and Etienne-Cummings describe a solution based on a normal flow method [MEH06]. Matching techniques are present in the FPGA world where Niitsuma and Maruyama [NII05] introduce a high performance system able to estimate displacement vectors by means of SAD (Sum of Absolute Differences) matching algorithm.

Following the results of [LIU98] [GAL98] [DIA08b], this work will be focused on Lucas and Kanade's optical flow method [LUC81], which has been highlighted by the mentioned contributions as a good trade-off between accuracy and performance. In this work we will focus to obtain a high computational performance (with low accuracy penalty), taking advantage of the analog and digital processors in Eye-RIS™ system to compute optical flow. It is important to remark that this system is a multipurpose machine vision architecture; hence it is not an ad-hoc embedded system to compute optical flow such [KOH09] [STO06] [MEH06] [NII05] which are designed exclusively for this task.

The description of the optical flow implementation and evaluation is organized as follows: Section A presents an introduction to the optical flow constraint equation of the Lucas and Kanade method used in this work. Section B describes the implementation of an efficient approach in Eye-RISTM architecture while Section C, we suggest an approach based on local features to enhance the performance of the implemented algorithm. Section D details the co-design strategy used to carry out the implementation in Eye-RISTM system. The evaluation of the different approaches is described in Section E and finally; our experimental results are presented in Section F.

A. LUCAS AND KANADE MODEL FOR OPTICAL FLOW ESTIMATION

This section introduces the basics to understand the concept of optical flow and the method used in this work. An ordered sequence of images allows the apparent motion estimation. The optical flow vector can be defined as a temporal variation in the image coordinates across the time, usually denoted as $\vec{v} = (u, v)$, and is computed based on the spatio-temporal derivatives of the pixel luminance. To estimate optical flow, a constraint equation is needed. Hence, it typically formulates the constant-brightness hypothesis. The basis of this assumption is that the pixel brightness remains constant over the movement. Thus, we can model this hypothesis with the following expression:

$$\left(\frac{df(x(t), y(t), t)}{dt}\right) = 0 \quad (3.1)$$

where f represents the luminance values of each pixel in the image. Once the hypothesis is defined since (3.1), is expressed as a derivate of a function with respect to time. Applying the first order Taylor expansion we will obtain the optical flow constraint equation:

$$uf_x + vf_y + f_t = 0 \quad (3.2)$$

where u and v are the optical flow components and the spatio-temporal derivatives are represented by f_x , f_y and f_t respectively. On the basis of the optical flow constraint equation, Lucas and Kanade [LUC81] proposed the minimization of the error Equation (3.2) using the sum of the least squares:

$$E(u, v) = \sum_{i \in B} (f_x(i)u + f_y(i)v + f_t(i))^2 \quad (3.3)$$

The objective of minimize the error Equation (3.3) is to find the displacement

components vector u and v , that minimize the differential error between the previous image warped (making use of the components vector u and v) and the actual image. Hence the Equation (3.3) is minimized by partial derivations respect the optical flow vector $\vec{v} = (u, v)$. The result is presented at Equation (3.4):

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \sum_{i \in B} f_x^2(i) & \sum_{i \in B} f_x(i)f_y(i) \\ \sum_{i \in B} f_x(i)f_y(i) & \sum_{i \in B} f_y^2(i) \end{bmatrix}^{-1} \begin{bmatrix} -\sum_{i \in B} f_x f_t \\ -\sum_{i \in B} f_y f_t \end{bmatrix} \quad (3.4)$$

where u and v are the optical flow components, the spatio-temporal derivates are represented by f_x , f_y and f_t respectively, and the subscript i is the i -th element of the integration block B . Through (3.4), we estimate the optical flow component vectors from a pair of images of a sequence.

The Lucas and Kanade method has been chosen for two main reasons. At first, this method has been ranked with a very good accuracy vs. efficiency trade-off in other literature works [LIU98] [GAL98]. As second reason, it is due to the digital processing restrictions in Eye-RIS™ system as explained in section 3.4 that requires a low complexity model in order to achieve real-time operation.

The next section describes how this model is simplified and optimized (in terms of processing speed) for its implementation in a NIOS II soft-core processor with the focal plane co-processing capability.

B. IMPLEMENTATION

One important problem in optical flow methods is the amount of memory accesses and massive multiplications computed by the model. For this reason, a high optimization becomes necessary to obtain a reasonable system performance. In order to speed up the computation of the Lucas and Kanade model a Sparse Integration Block (SIB) approach is used in (3.4), as show in Figure 3.6. Note that each element of the matrix is composed by two derivatives multiplication (for instance f_y multiply by f_t) and then sparsely added according the mask values. Each zero represents missing data and therefore multiplications that are not performed. This translates in high efficiency at reasonable accuracy requiring affordable computational resources. In our model, 9×9 and 5×5 SIBs are used. In the 9×9 SIB case, allows to reduce the computational load from 410 multiplications per pixel to 130 multiplications. This represents an optimization of

68.29% in terms of computational operations and 69.13% in terms of memory accesses with regard to the original one.

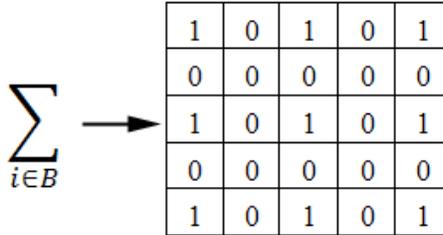


Figure 3.6. 5x5 Sparse Integration Block (SIB) representation.

We apply the principle of vicinity, which assumes that any point in the image will have a similar value to those in its neighborhood. This principle will be used for the optical flow estimation for computation of only a quarter part of the pixel by 4:1 subsampling. Hence, a calculated optical flow vector will be propagated to the neighborhood as shown in Figure 3.7 based on the spatial information coherence that says that close pixels tend to have similar optical flow values. This scheme is more accurate than a 4:1 input image sub-sampled pixel grid strategy because the optical flow estimation takes into account the original spatial-temporal derivatives in the input images. With this approach, we obtain a factor gain up to four, compared with the original one.

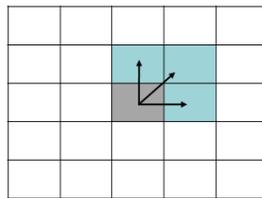


Figure 3.7. Neighborhood propagation illustration.

Once the implementation is detailed, we evaluate the system performance with the different approaches. On one hand, the method was implemented in C with two different SIBs; on the other hand, the same implementation was optimized in assembler with different SIBs. Assembler optimization allows to avoid RAW dependencies, optimizes memory accesses in the pipelined data path, and avoids unnecessary stack accesses usually implemented by the C compiler, absolute registers control, loops unrolling, etc.

Next, a performance study is carried out to evaluate the optimization evolution. Table 3.2, shows that the 5x5 SIB implemented in assembler, reaches a high

performance if we compare it with the 9x9 SIB versions. Obviously, this is due to the memory access increase as well as to the number of operations.

Table 3.2. System performance evaluation obtained with a 176x144 spatial resolution.

Integration Block	Frame Rate (frames per second)
L&K C 9x9 Integration Block	0.3
L&K C 9x9 SIB	0.9
L&K C 9x9 SIB and Propagation	3.6
L&K ASM 9x9 SIB and Propagation	11.9
L&K C 5x5 Integration Block	0.9
L&K C 5x5 SIB	2.3
L&K C 5x5 SIB and Propagation	8.8
L&K ASM 5x5 SIB and Propagation	28.8

Therefore, with these approaches, it can be concluded that the global gain obtained in a 9x9 version amounts to 40.47 the gain factor as show in Figure 3.8. On the other hand, applying these approaches to the 5x5 version, the gain factor is 31.4. The main reason why the 9x9 global gain is higher than 5x5 optimization is because the 5x5 integration stage is significantly reduced (in a factor of 2.7), i.e., the original version uses 25 pixels (in the block) while the SIB version only takes into account 9 pixels; in the meantime the 9x9 is reduced to a 3.2, i.e., the original version analyzes 81 pixels while the SIB version analyzes only 25 pixels. Making use of the neighborhood propagation approach, the optical flow is calculated in a quarter of the sequence, which allows the achievement of such high gains. Figure 3.8 shows the gain factor evolution obtained in each approach and optimization as well as the global gain.

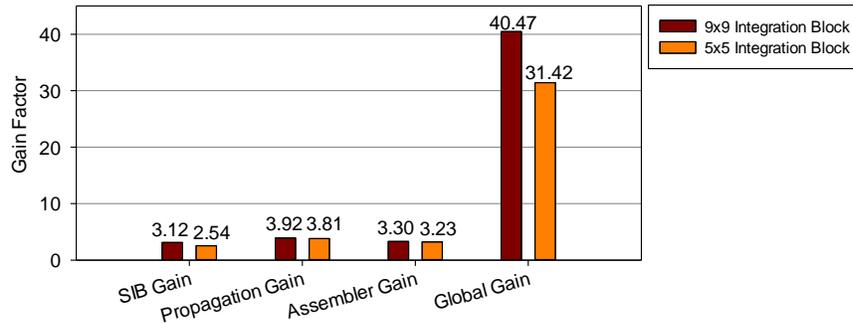


Figure 3.8. Comparison between the gains obtained in different integration blocks, after applying our approaches. Two different implementations are evaluated here, 5x5 (orange bars) and 9x9 (red bars). (Left to right) The first group of columns represents the Sparse Integration Block (SIB) factor gain; the second group shows the obtained gain after apply the optical flow 4:1 propagation. In the third column group figure the gain when the method is optimized in assembler while last column show total gain factor obtained after all the approaches are combined.

It is convenient extrapolate the performance result to a regular PC processor, for instance the Intel Core 2 Duo, to make clear the constraints of the proposed architecture. For a comparative evaluation between NIOS II and an actual processor we make use of Dhrystone 2.1 [WEI84] benchmark. The NIOS II soft-core processor configuration (70 MHz NIOS II/f) used in our test, obtains 71 DMIPS (Dhrystone MIPS) while an Intel Core 2 Duo 2.00 GHz processor obtains 4240 DMIPS (using only one of the processor cores). If we compare both processors an Intel Core 2 Duo obtains a gain factor of 59.7. Furthermore, Intel Core 2 Duo uses a superscalar architecture with two cores and support SIMD instructions (MMX and SSE) while NIOS II is a basic processor with a scalar architecture with a reduced instructions set (add, sub, mul, jmp, etc.). A study about different optimizations feasible on modern processor is shown in [ANG09]. Therefore, it is important to remark that, using the presented processor, we are much more constrained than using standard ones, and these has motivates to employ the proposed model modifications, analog processor utilization and optimizations techniques.

C. PRE-SELECTION OF POINTS OF INTEREST TO SPEED-UP THE OPTICAL FLOW COMPUTATION

The boundaries in an image are areas where optical flow can be more confidently estimated (unless they correspond to 3-D objects where occlusion problems are very common, though this case is less probable). These regions are rich in features; hence, the resulting estimation has more accuracy than in areas with poor contrast structure.

This is so because the Lucas and Kanade model collects weighted spatio-temporal structural information. If the local contrast structure is poor, the optical flow estimation confidence will be low. Instead of computing all the points and discarding unreliable points in a second stage, we can avoid the calculations of low confidence optical flow estimates by discarding these points a priori (using local contrast structure estimates). In order to take advantage of this issue, it will be used the Roberts Cross operator to localize the edges (local contrast maxima). The used kernels are shown in Figure 3.9.

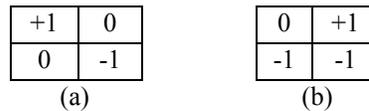


Figure 3.9. Roberts Cross convolution filter.

The sum of the absolute value of each convolution provides edge estimations, where each convolution operation obtains the maximum response when the edge angle reaches 45° (Figure 3.9(a)) 135° (Figure 3.9(b)). The filtering procedure is implemented by applying a Gaussian filter to the edge response and thresholding it with the original signal. The described procedure is indicated in Equation (3.5):

$$f(x) = \begin{cases} 1, & \text{Gaussian Edge Response} < \text{Edge Response} \\ 0, & \text{else} \end{cases} \quad (3.5)$$

Low contrast areas will not provide significant edges. This problem can be solved or reduced by locally performing modifications on the image intensity histograms, for instance by applying a 3×3 Laplacian convolution (aka Sharpen filter) that emphasizes the low-contrast areas. Figure 3.10 shows an image edge detection and a sharpen image edge detection.

It can be observed that a higher edge density is obtained by applying the sharpen filter. In this example, Figure 3.10(a) has 13.49% of density, while Figure 3.10(b) provides a density of 38.20%. Usually, the edge binary map outputs are around 30 up to 40% in high edge density scenes, 50% in the worst case. Using the sharpen filter, this method is more prone to noise. To reduce this noise, we can use the focal plane computational primitives, such as binary dilatation and erosion. Applying successive erosions and dilatations to the binary output includes less noise and becomes sparser before computing the optical flow operations on the NIOS processor. A typical action to remove this noise is applying a 3×3 erosion filter followed by a 3×3 dilatation filter with a simple filled squared mask as structure element to remove single points.

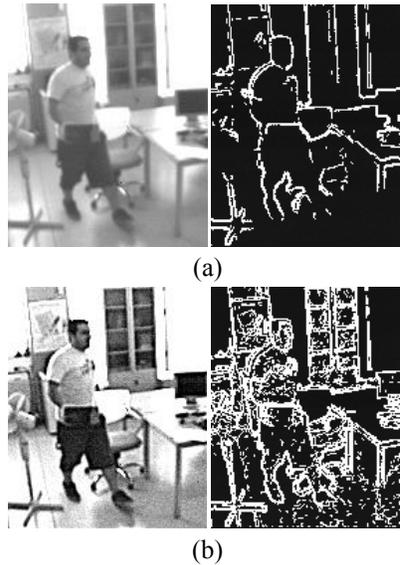
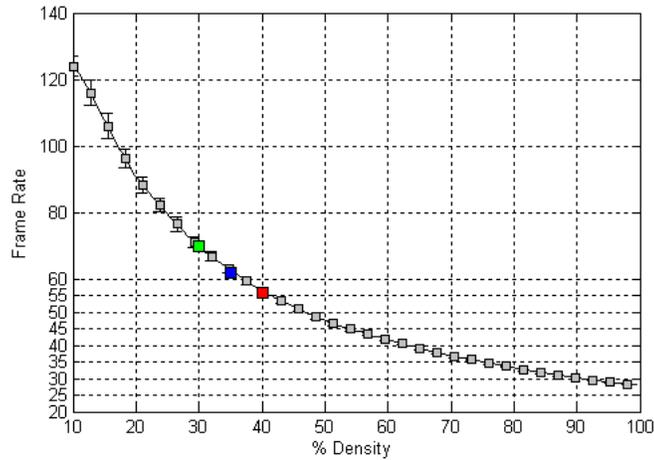
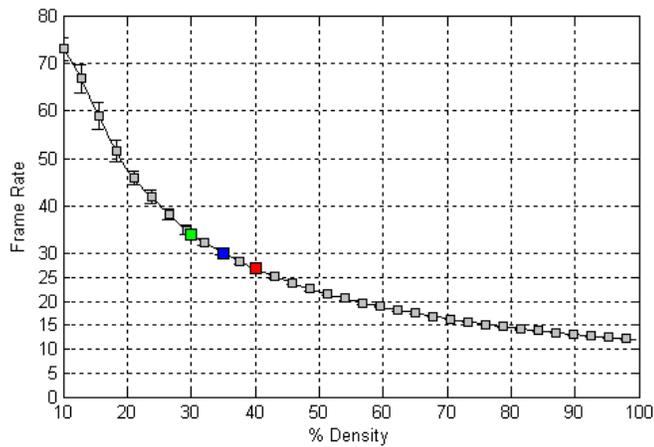


Figure 3.10. (a) Image edge detection in the original image. (b) Image edge detection with a sharpen pre-filtering.

Once we have reached this point, the edge binary map will be integrated with the optical flow estimation with the objective of optimize the computation time. Since the points of interest are previously selected in the focal plane, this process is carried out on the fly without affecting the global performance. Figure 3.11 shows the evaluation of this approach. To carry out this experiment a synthetic density distribution is generated, by filling with 1's from the beginning until the end of the binary mask and adding a 10% in each step. The first chart, Figure 3.11a, refers to the 5×5 SIB and the second one, Figure 3.11b, to the 9×9 SIB. In this experiment, the frame rate is measure while the edge binary mask density increases. This measurement starts from 10% of density until a density of 100%, i.e., the whole image. In both figures, different scene cases that can be found are remarked as well as the standard deviation by means of error bars.



(a)



(b)

Figure 3.11. (a) System performance using pre-selected points of interest and 5×5 SIB. (b) System performance using pre-selected points of interest and 9×9 SIB. Colored mark illustrates three typical scenarios using different image edge densities. The green mark makes reference to the best performance cases (scenes with low edge density), the blue one, to the most common density values (normal scenes) and the red mark, to the worst cases (scenes with high edge density). The error bar represents the standard deviation of the results for 10 trials.

Analyzing the obtained results and taking into consideration that normal scenes have from 30% to 40% of density with a 35% as mean value, it can be deduced that making use of 5×5 SIB, the frame rate can oscillate between 56 (red mark in Figure 3.11a) up to 70 (green mark in Figure 3.11a) frames per second with a typical value of 62 fps (blue mark in Figure 3.11a). In the case of the 9×9 , the frame rate is around 27 (red mark in Figure 3.11b) reaching up to 34 (green mark in Figure 3.11b) frames per second, taking into consideration 30 fps as the usual value (blue mark in Figure 3.11b).

After expose the proposed approach it can be concluded that it is suitable for real time computation beyond video-rate, (25 fps). The main advantage of studying optical flow in points of interest (in our case they are pre-selected edges) is the increase of performance. The gain obtained is around 2 in the worse cases, assuming these cases in scenes with a 50% of density. Considering the best cases, scenes with 30% of density, the gain will arise up to 2.6. As commented before, a common scene usually contains 35% of density. Hence, we can conclude that the mean speed up gain is 2.3. This gain can be used to compute flow at higher frame-rate (and therefore, improve the optical flow accuracy [DIA08b]) or to include new functionalities into the processors towards final concrete applications.

It is important to remark that NIOS II is able to handle 0.044 GOPS to compute optical flow while the focal plane processes 4.1 GOPS to smooth the image, obtain points of interest, and compute the optical flow regularization. To estimate the number of operations used in the focal plane, we carried out an equivalence of a digital processor (NIOS II) to perform the same functionality. Due the amount of operations involved in a previous (focal plane) and final stage (processor) to obtain optical flow, it can be concluded that the optical flow estimation could not be implemented in this architecture without the focal plane assistance.

D. CO-DESIGN STRATEGY

In previous sections, it was detailed the way to estimate optical flow as well as how such optical flow has been improved to speed up the optical flow estimations. In this section, a global description of the co-design implementation to estimate the motion vector components in the Eye-RIS™ system is introduced. As described before, the Q-Eye sensor is a system able to process in the same physical layer where the image is captured (focal plane computation). For this reason, at the same time that the system captures the image, it can be processed and moved to the main memory to apply a post-digital processing.

After the image capture is done, the focal plane processes it with the proposed method to select edges, as described in the previous section, and applies a linear diffusion filter [WEI98] that works as smoothing filter of the captured image and improves the numerical computation of the image derivatives. The estimated edge map and the current and previously captured image on the sensor form part of the optical flow method to be computed in the digital processor. Once the optical flow is computed at the NIOS II processor, a linear diffusion filtering (regularization) is applied to the

optical flow components in the focal plane as indicated in Figure 3.12, to preserve the spatial coherence [HOR81]. Note that when we refer to “regularization”, means the process that performs the local averaging process and improves the spatial coherence based on the smoothness constraint.

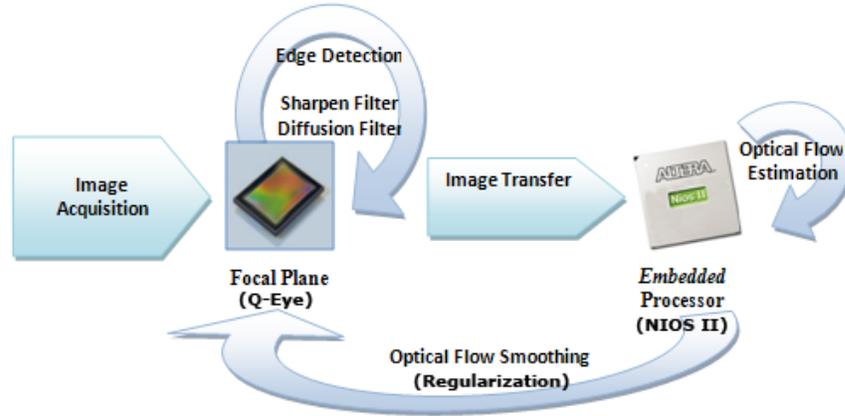


Figure 3.12. Initial Optical-flow Co-Design scheme used in Eye-RIS™ platform.

An important factor to take into account is the exposure time when the image is captured. Adopting a sequential strategy, image acquisition, and NIOS processing being done sequentially (one after the other), is not convenient because it does not take advantage of the pipelined processing capabilities of the system. The focal plane (Q-Eye) is able to work asynchronously with the processor, and the system will be able to capture and process in the focal plane at the same time that we are making use of the digital processor. To accomplish this, the exposure time, focal plane processing, and processor computing time must be taken into account.

Furthermore, analog-based internal memories in the focal plane cannot retain the images for a long time due to transistors leakage. Taking into account that the mean value of an image stored in an internal focal plane memory, decreases around 0.8 LSBs every 40 ms. For this reason, prolonged storage leads generate significant degradation and must be avoid this practices in the co-design. In order to reduce this signal degradation as well as remain a constant sampling period, we must meet the following constraint, as indicated in expression (3.6):

$$PT_{Processor} < ET + PT_{Focal\ Plane} \quad (3.6)$$

where the time of optic flow processing in NIOS II is $PT_{Processor}$, while ET is the capture exposure time and $PT_{Focal\ Plane}$ is the focal plane processing time.

The processing time, in the focal plane, takes approximately 3-40 μs per operation. Hence, this time can be considered negligible if it is compared with the exposure time or the processing time on the digital processor. Although the frame-rate is determined by $PT_{Processor}$, if large movements are presented in the scene, lower exposure times can be used to reduce the displacement. In this case, a slightly more complicated scheme is necessary to reduce, as much as possible, the time that each frame is stored in the analog memories. For this purpose, the optical flow process must be split in different stages. In the first stage, the I_t image is captured by the focal plane at the same time as the partial optical flow estimation (half of the resolution) of a previous captured sequence I_{t-2} and I_{t-1} , is carried out on the processor. The second stage captures I_{t+1} image and processes the unfinished optical flow calculation of the previous stage. The last stage transfers the optical flow vector components to the focal plane to apply the post-processing lineal diffusion filter, which acts as a Gaussian isotropic filter.

Note that we do not have a continuous acquisition process where time between frames is fixed. Contrary, the acquisition process is handled (according to the scheme of Figure 3.13) to avoid the image degradation and preserve, as much as possible, the time interval between the captured images.

Therefore estimating the flow only between pairs of consecutive frames is possible to avoid the problems previously explained. That is to say, the flow is computed between frames I_{t-2} and I_{t-1} and between frames I_t and I_{t+1} but the flow is not computed between frames I_{t-1} and I_t due the time interval can be different. This is because after compute optical flow another further processing algorithms could be applied. Figure 3.13 illustrates this process.

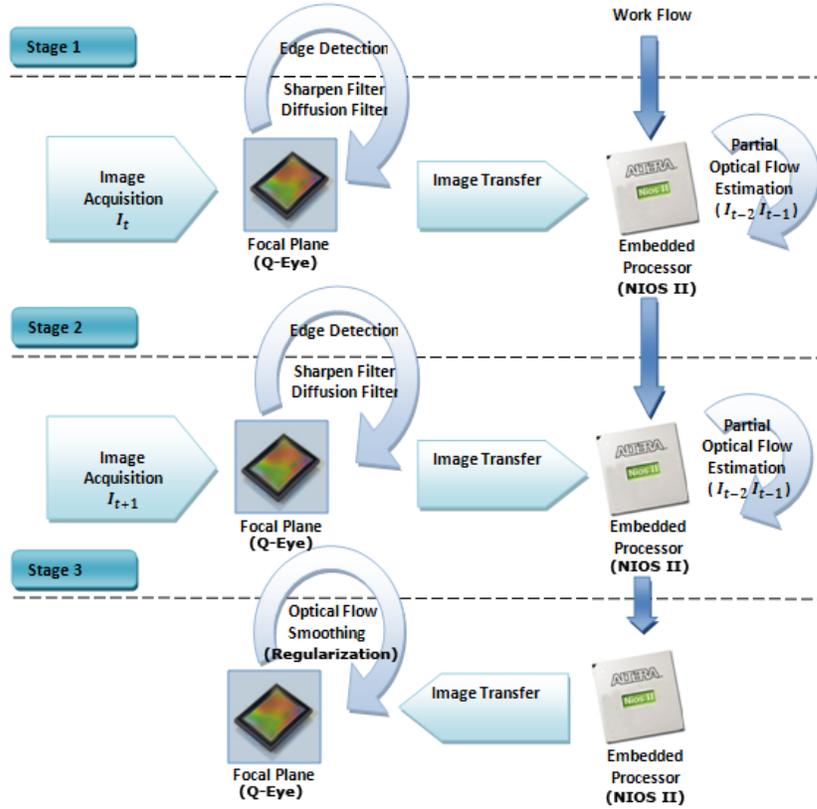


Figure 3.13. Different stages to estimate optical flow in Eye-RIS™ system.

3.5.2 Optical Flow Evaluation

The purpose of this section is to evaluate and validate the suggested approaches described in the previous section. In this paper the error measure (3.7) will be the same that the used by Barron et al. [BAR92], which consists in the angular error estimation between the ground-truth optical flow vector and the estimated one:

$$\psi_E = \arccos(\vec{V}_c \cdot \vec{V}_e) \quad (3.7)$$

where ψ_E is the estimated error, \vec{V}_c the true vector flow from the Ground-truth values, \vec{V}_e the estimated vector flow. Note that this error metric is non-linear and combines information from the angular and magnitude error. Nevertheless it is frequently used and therefore has been used for the sake of comparison with other contributions available at the literature.

Once the angular error estimation procedure is defined, a study of the angular error is carried out on the simplified Lucas and Kanade approach described in this work. As a first step, the original model implementation is evaluated in software (Matlab), with the different proposed approaches and densities (100% and 48.5%) as shown in Table 3.3.

Table 3.3 Average angular error (AAE) and standard deviation (STD), in Yosemite sequence (without clouds), with the different approaches and densities on Matlab.

Integration Block and Used Approach	AAE	STD	AAE	STD
	100%	100%	48.5%	48.5%
5x5 Barron's implementation	11.01°	17.14	10.32°	17.40
5x5	20.68°	21.75	20.38°	20.57
9x9	14.75°	14.85	14.63°	14.16
5x5 SIB	19.64°	20.72	19.34°	19.54
9x9 SIB	14.09°	13.99	13.98°	13.33
5x5 Propagation	20.74°	21.97	20.43°	20.84
9x9 Propagation	14.73°	15.04	14.56°	14.32
5x5 SIB + Propagation	19.71°	20.98	19.39°	19.86
9x9 SIB + Propagation	14.08°	14.21	13.92°	13.51
5x5 SIB Propagation + Regularization $\sigma = 5.29$	12.30°	12.18	12.10°	11.29
9x9 SIB Propagation + Regularization $\sigma = 5.29$	10.51°	8.46	10.51°	8.12
5x5 SIB Propagation + Regularization $\sigma = 7.48$	11.33°	10.26	11.17°	9.53
9x9 SIB Propagation + Regularization $\sigma = 7.48$	9.89°	7.11	9.86°	6.75

To compare the results of the different approaches, it will be compared with the implementation proposed by Barron [BAR92], where the optical flow is estimated making use of a temporal resolution of 5 images and 5×5 integration blocks. Table 3.3 indicates that with a 9×9 integration block, obtains better results that working with a 5×5 integration block and getting an error similar to the Barron's implementation. Making use of large blocks, the model weights better the optical flow components but with the associated problem of the computation time. Comparing the angular error between sparse and non-sparse blocks, the table shows that the error is quite similar between them. Hence, by applying sparse integration blocks (SIB), the performance becomes higher regarding with the original one.

As it was stated in Section 3.5.1.B, the main idea is to propagate the optical flow estimation to the neighborhood. The results obtained (Table 3.3) reveal that making use of this performance optimization; quite similar results are reached if we compare them with the non-propagated version. Hence, it can be said that this approach is totally valid

since the loss in accuracy is insignificant in both sparse and non-sparse integration block approaches.

The last evaluation in Table 3.3, consists in determinate the error after apply the optical flow regularization (spatial coherence). In this way, small errors can be corrected, weighting them with the neighborhood. Due one of the most common smoothing filter used is the Gaussian convolution; this filter will be applied to the estimated optical flow and evaluated. Analyzing the experiment (Table 3.3), it can be deduced that the obtained results are better if the smoothing convolution filter is applied to the optical flow components. The angular error reduction is higher for small integration blocks, both sparse and non-sparse. While using larger masks, the error is lower, if we compare these masks with the smaller ones (5×5). This is due to the fact that, for the information collected in small blocks, the model weighted worse than in the case of masks with larger neighborhoods. Hence, applying the regularization to the result helps to weight again the vicinity, being therefore small mask based approaches more favored. If the obtained results are compared, after the regularization step, with respect the Barron's implementation; the obtained results show quite similar angular errors but on the other hand the standard deviation is reduced more than a half.

All the previous simplified approaches have been evaluated using Matlab code, with double floating point data representation, to illustrate the effect of the successive approaches. To obtain a realistic evaluation, the angular error will be estimated with the results obtained in the Eye-RIS™ system. In order to carry out these measurements, different integration blocks and post smoothing filters (regularization) are taken into account. In this evaluation, it will be assumed as valid all the approaches and simplifications evaluated before. Note that here a new filter, in the regularization step, is used. This filter, lineal diffusion filter [WEI98], implements a low-pass filter that emulates a Gaussian filter using the Resistive Grid module available on the SIS Q-Eye (focal plane). Due to the nature of the linear diffusion and its equivalence with the Gaussian filter [WEI98], the use this filter is more appropriate because it is more precise and exploits the advantages of the focal plane. Table 3.4 shows the measure the average angular error and standard deviation error with different densities (100% and 48.5%). It is important to remark that the optical flow estimation is developed making use of fix point arithmetic approach.

Table 3.4. Average angular error (AAE) and standard deviation (STD), in Yosemite sequence (without clouds), with the different approaches and densities on Eye-RIS™ system.

Integration Block and Used Approach	AAE	STD	AAE	STD
	100%	100%	48.5%	48.5%
5x5 SIB	24.79°	20.12	23.45°	18.55
9x9 SIB	17.10°	14.02	15.88°	12.59
5x5 SIB + Propagation	24.84°	20.09	23.25°	18.41
9x9 SIB + Propagation	17.14°	13.88	15.94°	12.55
5x5 SIB Propagation + Regularization $\sigma = 5.29$	15.61°	12.01	13.12°	10.53
9x9 SIB Propagation + Regularization $\sigma = 5.29$	13.06°	10.11	11.25°	8.86
5x5 SIB Propagation + Regularization $\sigma = 7.48$	14.61°	10.63	12.24°	9.46
9x9 SIB Propagation + Regularization $\sigma = 7.48$	13.09°	9.34	10.44°	7.87

It can be concluded that after performing the angular error measurements of optical flow, the best result obtained is the 9x9 SIB with a lineal diffusion filter with $\sigma = 5.29$ as shown in Table 3.4. The differences with the previous version (Table 3.3) are mainly produced because the method make use of fix point arithmetic of 32 bits in NIOS II, properly rescaled across the different processing operations to keep the relevant information. Note that, though bit-width and representation is significantly different than the purely software version (Matlab version), the results fit quite well to the previous data which validate our fixed-point implementation.

3.5.3 Experimental Results

In this section, the experimental results are presented with real sequences. To evaluate the optical flow results, it has been chosen a traffic sequence where the cars move through the scene. In this sequence, the optical flow has a clear interpretation and therefore, a qualitative evaluation can be done. The original sequence, Ettliger-Tor, can be obtained from [INS10]. The optical flow estimation is carried out in different sequences applying both SIBs (5×5 and 9×9). To interpret the obtained results, the optical flow vector direction is encoded with a color (according to the colored frames of the different images) whereas the vector’s magnitude is expressed by the color intensity as shown in Figure 3.12.

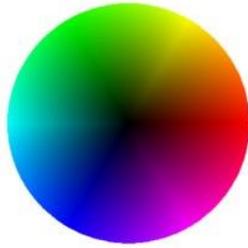


Figure 3.12. Optical flow representation. The color corresponds with the direction of the optical flow vector while the magnitude is encoded as the color intensity.

In the results shown below we can observe that the optical flow increases, as we increase the integration block (5×5 SIB and 9×9 SIB) sparse. To estimate the optical flow in Figure 3.13 and Figure 3.15 a lineal diffusion smoothing is applied to the image, equivalent to a Gaussian filter with $\sigma = 2.5$, while the optical flow regularization corresponds to a Gaussian filter with $\sigma = 3.3$.

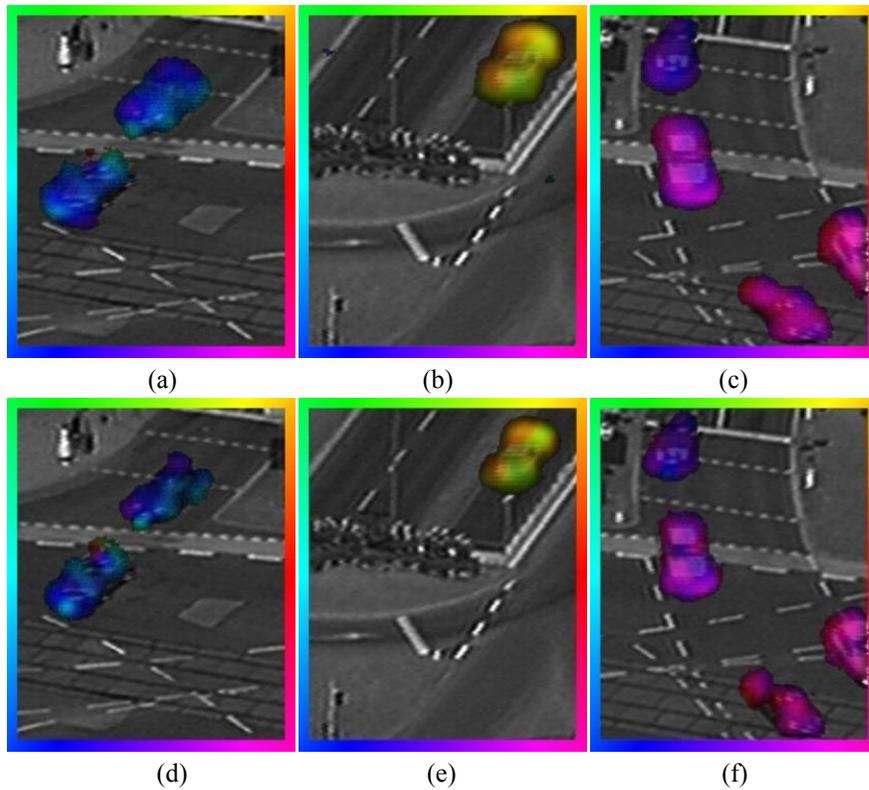


Figure 3.13. Optical flow estimation in a traffic sequence. Flow field is overlaid with the original frame. In the first Row (a-c), the optical flow is estimated using 9×9 SIB. In the second Row (d-f), the optical flow is estimated using 5×5 SIB.

In Section 3.5.1.C, it was proposed a method to obtain the image edge response in a focal plane with the objective of improve the optical flow performance, obtaining a mean gain of 2.3. Once the edge estimation is computed in the focal plane, morphological operations of dilatation and erosion are applied to the binary map (two dilatations and one erosion with a 3×3 kernel) to bring near the optical flow results to the obtained ones without the sparse estimation. Figure 3.14 shows the binary maps results after be processed in the focal plane.

After the points of interest estimation, the motion vectors are calculated in the processor and post smoothed in the Q-Eye. The results of this procedure are shown in Figure 3.15. As can be observed, the output density is slightly lower if we compare it with previous results.

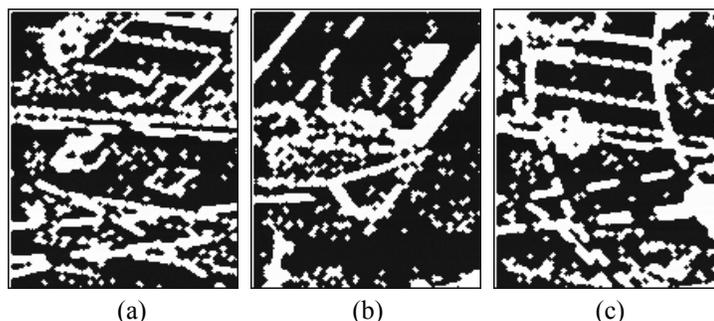


Figure 3.14. Binary edge map processed, with the proposed method in section 3.5.1.C, in the focal plane.

It can be concluded that making use of points of interest, the non-edge detection risk in areas of low contrast must be taken into consideration, due the proposed approach is unable to determinate the optical flow on those regions; as shown in Figure 3.15c,f. In the case of static cameras, other pre-selection schemes can be also used to reduce the input data stream (instead of local contrast structure), for instance local image change ratio (since most of the scenario will be static). When choosing a method to estimate optical flow, the most appropriate approach depends on the target application and scenario. If the application requires a dense optical flow without the risk of areas of low contrast, a 5×5 SIB shall be chosen. When the precision is a crucial factor, pre-selection of point of interest into the estimation of motion vector and using a 9×9 SIB is the best choice. If the time is an essential component, the option that best meets these requirements is using 5×5 SIB with points of interest pre-selection to estimate the optical flow.

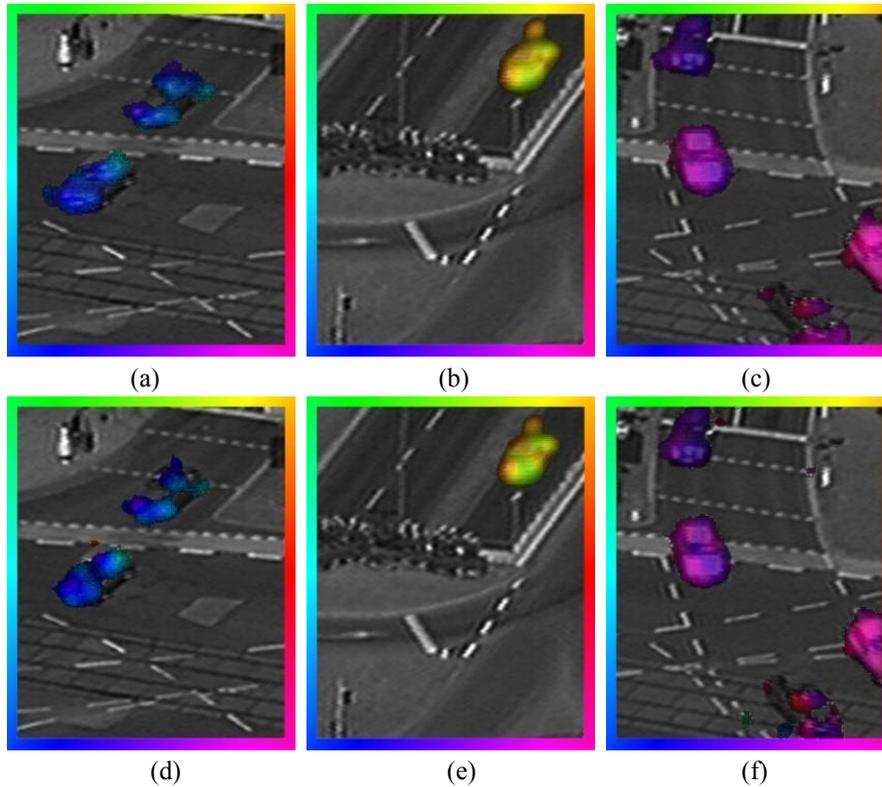


Figure 3.15. Optical flow estimation, on edges, in a traffic sequence. The average edge density, in these images, is 42.5%. Flow field is overlaid with the original frame. In the first row (a–c), the optical flow is estimated using 9×9 SIB. In the second Row (d–f), the optical flow is estimated using 5×5 SIB.

3.6 OVERTAKING DETECTION

Once it has been detailed in depth the ad-hoc implementation to estimate optical flow on Eye-RISTM architecture, this section is focused to describe in detail the followed scheme to detect overtaking cars on the road by mean of motion segmentation. Figure 3.16 illustrates an overview of the introduced scheme in this section. To achieve our objective, information of interest (edges) is extracted in the focal plane, to estimate optical flow at those points and apply a filtering/clustering rule to segment potential vehicles in the digital processor; as described in Section A. After this pre-segmentation, Section B details a basic binary matching process to avoid any illumination and vibration effects that may affect the initial segmentation. Section C exposes the followed metric to reject those detected candidates that do not correspond to standard vehicle dimensions, while Section D details a projection filter to avoid false positives generated by the ego-motion pattern. Finally, in Section E, is proposed an extension of the described method with the objective of increase the distance detection.

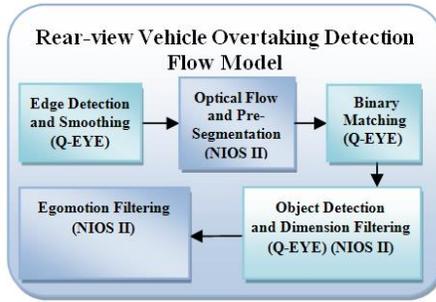


Figure 3.16. Algorithm dataflow scheme showing the processing structure computing the different steps.

A. ESTIMATION OF POINTS OF INTEREST, OPTICAL FLOW AND PRE-SEGMENTATION

In Section 3.5.1 was described in detail an optical-flow method based on points of interest (edges) and its real-time implementation on the Eye-RIS™ platform, clearly showing the benefits of a mixed analog-digital processing strategy. Taking full advantage of the focal-plane processing capabilities and parallel computing that can be carried out with the analog processor (focal plane) and digital processor (NIOS II). After applying different strategies of optimization, the introduced method obtains a gain optimization factor up to 40 in the optical flow implementation.

The first step of the proposed scheme consists estimating the points of interest in the scene. The image is acquired and pre-processed at the same physical layer (the focal plane). Hence at this stage, the estimation of the points of interest and image smoothing are undertaken at the focal plane at the same time (in parallel) as the digital processor makes a partial optical-flow computation of the points of interest of previously captured frames. This stage is replicated to capture the next image of the sequence as well as to compute the remaining optical-flow part of the previously captured frames. After both stages, a regularization operation is applied to the estimated motion vectors to achieve spatial coherence, as described in section 3.5.1.D. Figure 3.17 shows the results obtained with the Eye-RIS™ platform of a highway sequence after applying the described scheme.

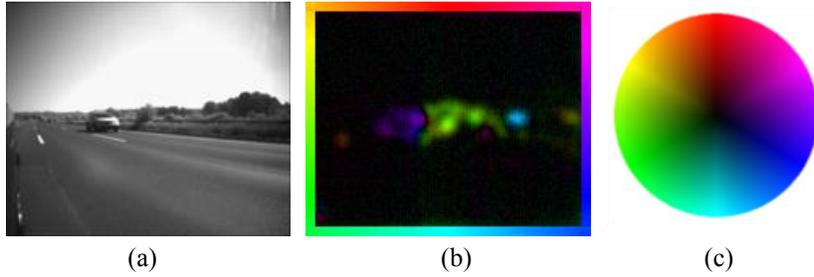


Figure 3.17. (a) Original image of an overtaking scene on a highway; (b) estimated optical flow using the Eye-RIS™ platform; and (c) optical-flow representation. The color corresponds to the direction of the optical-flow vector while the magnitude is encoded as color intensity

Once the optical-flow vectors have been estimated, the next step consists into apply a filtering process to segment potential approaching objects. This proposed filtering process is composed of three basic rules (tree decision scheme) that establish an initial spatial-temporal threshold, a motion direction filtering where the u component must be greater than zero since the sensor is installed on the left side of the car, and finally, a rule to reject those vectors that seem to be generated when the Nyquist theorem is violated. Optical-flow vectors, that suffer this theorem violation, tend to invert the direction with a great magnitude and commonly one of the components of the vector is much bigger than the other. To eliminate those vectors, that present the features explained before, the component of the optical flow v must be less than the absolute value of u . Such filtering rules, which generate the motion of interest (MOI) binary map, are described in equation (3.8),

$$MOI(x, y) \begin{cases} 0, & \text{if } \left(\sum_{i \in B} f_x(x, y) \cdot f_t(x, y) \right)^2 + \left(\sum_{i \in B} f_y(x, y) \cdot f_t(x, y) \right)^2 < \text{Threshold} \\ 0, & \text{if } u(x, y) < 0 \\ 0, & \text{if } u(x, y) < |v(x, y)| \\ 1, & \text{otherwise} \end{cases} \quad (3.8)$$

where f_x , f_y and f_t represent the spatial-temporal derivates, u represents the x optical flow component (horizontal component) and v represents the y optical flow component (vertical motion).

B. BINARY MATCHING

This stage aims to refine the previously obtained binary image to avoid illumination effects, violation of the Nyquist theorem (the video-rate is not sufficient to properly

describe the motion of close and/or fast objects) and vibration effects. To avoid the reported effects it is proposed a binary matching suitable for computation at the focal plane. Note that the utilized architecture is very constrained at the digital part; in order to achieve maximum performance of the system it is essential to take full advantage of the analog part. In addition, this also has the advantage of being more power efficient [ZAR08]. This binary matching consists of a spatial-temporal filtering, which takes into account previous MOI binary images (historical binary maps) obtained during the previous step, to compute a bitwise correlation between them. Assuming that the vehicle to be detected has a defined linear trajectory over time, a right-shift operation is applied to the historical binary maps and correlated (by means of a bitwise AND operator) with the previous historical maps to finally apply the last correlation with the estimated MOI calculated in the previous step. Equation (3.9) defines the described operations,

$$Output = \left(\bigcap_{t=1}^{n-1} \bigcup_{m=1}^2 \left(Flood_{Fill}((MOI \oplus B) \ominus B) \cap Right_Shift(HistMap_t, m) \oplus B \right) \right) \oplus B \quad (3.9)$$

where n is the number of historical maps saved, $Right_Shift(Image, m)$ is a image shifting operation to the right m times, $Flood_Fill$ is a function that fills holes in a binary image, $HistMap_t$ is the historical map at time t and \oplus and \ominus denote dilatation and erosion operations by a structuring element B respectively. In our case, this structuring element B is formed by a square of 3×3 matrix. To optimize this process, the historical maps will be stored in an internal memory (buffer) allocated in the focal plane. The historical buffer size must be taken into account because it is directly proportional to the detection of vehicle delay. In our configuration we have used a historical map buffer size of three. Once the output binary map has been estimated the historical map buffer is updated with the previously calculated MOI binary map. To clarify this process better, Figure 3.17 shows a sequence of three historical binary maps, MOI and the output obtained after applying the procedure described in Equation (3.9).

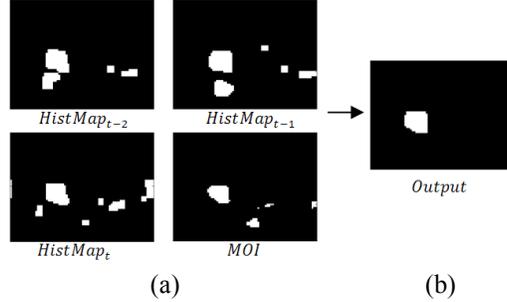


Figure 3.17. (a) Three historical maps and an MOI map; (b) Output map produced.

As commented above, it is important to note that this step takes advantage of focal-plane resources and processing capabilities to accomplish this process. For instance, the focal plane is able to process a bitwise operation between two images in $2.93 \mu s$ while the NIOS II takes $0.18 ms$ to perform the same operation, i.e. the focal plane is up to 64 times faster than NIOS II in this kind of processing. Thus it is important to remark the relevance of the presence of the focal plane in this architecture and its optimized use.

C. OBJECT DETECTION AND DIMENSION FILTERING

Once the refined binary map has been obtained during the previous step the system must locate the detected object and analyze it to determine whether it corresponds to a real vehicle and reject those objects that are outside the area of interest. A common case occurs on bridges, where cars travelling in different directions may be detected by the system. For this reason it is needed the prevention of this kind of situation by assuming a flat road and focusing only on those objects that belong on the road. To this end extra information, from the binary map, must be extracted. The focal plane provides primitives for the extraction of blobs that help us to identify the dimension and position of the blob. This extracted information must be converted from a 2D projection to a 3D coordinated scenario system. It is well known that it is possible to estimate the distance of a point belonging to a plane, as suggested by Gat [GAT05]. To determine the 3D coordinates of an object that belongs on a plane, it is needed to know the intrinsic and extrinsic camera parameters, as illustrated in Figure 5.

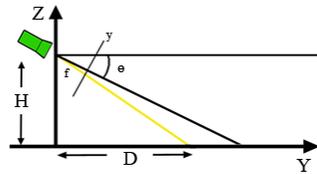


Figure 3.18 Illustration of the camera parameters needed to estimate distances from the camera.

Extrapolating the method described by Gat [GAT05], as suggested by Escalera [ESC01] [29] to take into account the angle tilt, the distance from the bottom of an object to the camera can be described as follow,

$$D = H \cdot \frac{f + y \cdot \tan \theta}{f \cdot \tan \theta - y} \quad (3.10)$$

where D is the distance to a point on the plane, H is the camera height, f is the focal distance, y the vertical coordinate in the image and θ the angle formed between the focal distance and the plane. Observing Equation 3.10, it can be deduced that the maximum estimated distance is the infinity, this case occur when the denominator become to zero. This paper does not concentrate on specifically measuring the distance of a detected object but rather to measure the width and height of such an object and discard those that are not vehicle sized. To arrive at these measurements the proportionality rule can be used, where the width of an object in a 3D world is proportional to the width projected in a 2D image and the distance of such an object is proportional to the focal distance, as shown in Figure 3.19. By solving this equation gives:

$$W = D \cdot \frac{w}{f} \quad (3.11)$$

where W is the real width of an object, D the distance to the object, w the width in pixels in the image and f the focal distance. It is important to note that the estimated Equation (3.11) will be equivalent to estimate the object height. Now those objects detected that do not satisfy vehicle dimensions can be discarded. Furthermore, via Equation 3.11 it can be also discarded those objects that do not belong to the road since measurements made above the horizon line will give negatives values.

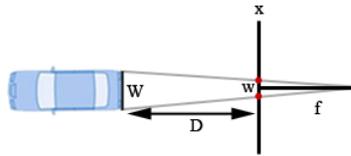


Figure 3.19. Illustration describing the parameters involved in obtaining object width.

In some cases, due to low contrast in certain areas in the image, the object detected may be split into pieces, as shown in Figure 3.20(a). Therefore, before estimating the real dimension of the detected object, it is important to merge the sparse detected blobs

to determine the dimensions, as proposed before in Equation (3.11). This clustering algorithm is based on the Euclidian distance that will merge recursively those blobs that satisfy the distance restriction; in our case the maximum distance was set to 25 pixels. The results after applying this algorithm are illustrated in Figure 3.20.

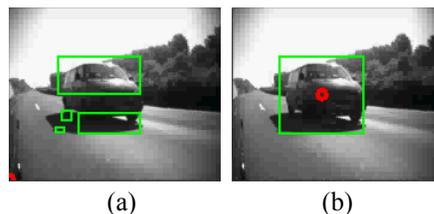


Figure 3.20. (a) Blobs obtained after the binary matching step; (b) Blob obtained when the clustering algorithm is applied.

In addition to the explained in this subsection it is important to remark that those detected vehicles, in non adjacent lanes, could be rejected by mean lateral distance discrimination, as shown in [ESC01]. Due to in certain occasions some drivers commit traffic violations when changing lanes, endangering the road safety, it is preferable do not reject such detected vehicles. In the last row of Figure 3.26, the sequence of vehicle overtaking detection screenshots illustrates how the system is able to detect vehicles in non-adjacent lanes.

D. EGOMOTION FILTERING

Egomotion can be defined as the motion pattern generated by the observer's own displacement in relation to static objects in a scene. When the observer follows a linear path the motion generated is in the opposite direction to that of an object approaching the observer. If the observer does not follow a linear path, for instance when a car turns left with a camera fixed on the left side, part of the scene will generate egomotion with similar features to the ones generated by an approaching vehicle and therefore cause a false positive (see Figure 3.21). For this reason it must be activated a mechanism able to determine when this situation occurs and discard such false positives.

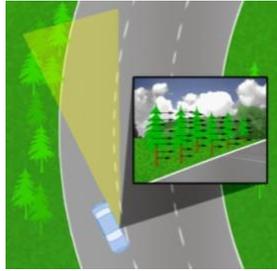


Figure 3.21 Optical-flow scenario when the car turns left and the camera is fixed on the rear view mirror on the left side of the car.

Since the scene is a perspective projection this property can be used to discard false positives generated by egomotion, as explained previously. Based on the line equation it can be described a “linear classifier” that determines whether the centroid of a detected object belongs to the area of interest, as shown in Figure 3.22. The parameters that describe the line, slope and y-intercept are conditioned by the higher vehicle’s centroid along the 2D image space. In Equation (3.12) the output estimates whether an object is in the area of interest (AOI),

$$Sign = \frac{m \cdot x + n - y}{|m \cdot x + n - y|} \quad (3.12)$$

where m is the slope, n is the y-intercept and x and y the Cartesian coordinates of the centroid object.

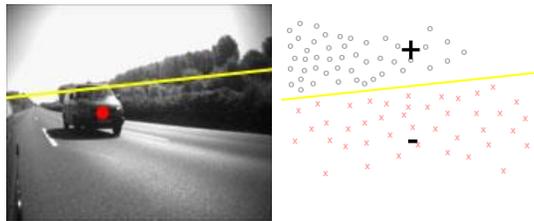


Figure 3.22. Area classification representation, where positive values indicate objects out of the AOI while negative values indicate objects within the AOI.

This approach is not enough in certain circumstances however: when the car turns on a roundabout or takes a sharp bend, for instance, projection filtering cannot discard most of the false positives (note that this issue is not solved by other similar approaches based on image processing [DIA08a][MOT04][SON07]). For this reason it was added an external sensor able to detect such scenarios as taking bends or other turning actions. A simple gyroscope is the most suitable sensor to solve our problem since it provides the angular speed directly. For implementation in a commercial system this sensor

could easily be replaced by a CAN-bus connection to the vehicle network, which reads this information directly from the vehicle sensors. The information is used to adapt the projection filter to the new circumstance, thus limiting the classification area adaptively. Figure 3.23 shows the two points ($p1$ and $p2$) that define the line described in the calibration stage and the newly estimated position of $p1$ ($p1'$) when a vehicle turns to the left. The newly estimated position of $p1$ is defined as follow,

$$p1' = p1 + AngularSpeed \cdot Factor \quad (3.13)$$

where $p1'$ is the new coordinate of $p1$ in the Y axis, $p1$ is the original position, *AngularSpeed* is the information provided from the sensor and *Factor* is a constant gain value. The *Factor* variable in Equation (3.13) is an empirical fixed value where vary depending of intrinsic and extrinsic camera parameters, image resolution and the estimated line to reject possible false negatives. In our tests this value was fixed at 1,000.

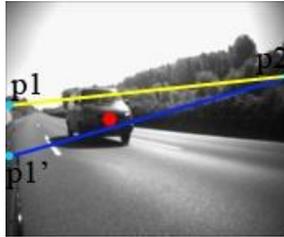


Figure 3.23. Example of projection filtering adaptation on a road scenario.

E. INCREASED VEHICLE DETECTION DISTANCE

The Eye-RIS™ system is able to capture images with low exposure times, thus obtaining high frame rates of up to 10,000 fps. The concomitant problem in road scenes when making use of a high frame rate is that the motion of approaching vehicles at a distance from the camera is hardly noticeable. This can be solved by means of a higher image resolution, but the Eye-RIS™ platform is constrained to QCIF resolution (176×144). Static objects closer to the camera, on the other hand, move faster than distant objects, i.e. the egomotion patterns based on static objects depend on the distance and speed as well as the trajectory. Hence it must be taken this into account the rejection of bad optical-flow estimations of objects close to the camera: the exposure time must be inversely proportional to the speed of the car and the distance to the object. To address both aims, obtaining valid optical-flow estimations from near objects and detecting objects in motion distant from the camera, it is proposed a splitting scheme

of the sequence into different exposure times to satisfy such restrictions. The concept of this introduced solution is shown in Figure 3.24, where the left hemisphere of the sequence computes the optical flow with I_t and I_{t+2} frames while to estimate optical flow in the right hemisphere I_{t+1} and I_{t+2} frames are used. This calls for a frame-rate adaptation mechanism. Note that this mechanism is possible thanks to the complete control of the triggering of image acquisition (as opposed to cameras with a periodic and fixed image-acquisition trigger). This system is capable of controlling the acquisition process according to the scenario and producing a non-periodic, adaptable image-acquisition process that optimizes the processing time and system delays and tunes to the motion range of the scene.

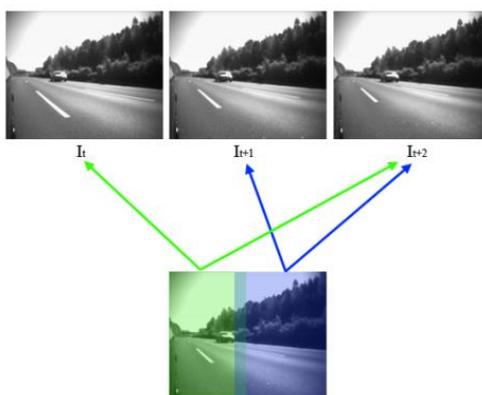


Figure 3.24. Adaptable frame rate in the scene; sample split period in a sequence.

Figure 3.24 shows an overlap area within both hemispheres. The interest of such overlapping is because the sequence may have an uncertainty area in which the left hemisphere may represent incorrect optical-flow estimation due to the vehicle's speed, while the right hemisphere may represent the correct estimations. The opposite may also occur, i.e. the vehicle is moving slowly and so the right hemisphere cannot detect it but the left one can. Hence this overlapping area is the result of the combination, bitwise OR, of both MOI hemispheres (For more details about MOI see Section 3.5.2.B). The choice of the size of the hemisphere and the frame rate adaptation are related to the extrinsic parameters of the camera. In our case these camera parameters are shown in Table 3.5. In the next section this proposed approach is evaluated.

Table 3.5. Extrinsic and intrinsic camera parameters used in our tests.

Camera Parameters	Values
Height	98 cm
Row Tilt	14°
Yaw Tilt	5°
Left Hemisphere Width	96 pixels
Right Hemisphere Width	105 pixels
Overlap Width	32 pixels
Focal Distance	8 mm

3.6.1 System Evaluation and Results

In Section 3.5.2 was described each step of the proposed algorithm with a novel adaptive mechanism to discard false positives generated by the egomotion as well as a complementary approach, detailed in Section 3, to increase the range detection. Now the system is evaluated quantitatively. Firstly, the system performance is estimated to determine the number of frames per second (frame rate) in different scenes that our implementation is able to process with the Eye-RIS™ platform. Precision is another factor that will be studied in this section to determine accuracy by means of an assessment of the exposed system on the highway with diverse elements in the scene. Distance detection is then contrasted with several tests at different speeds on a closed track, including the complementary mechanisms to increase the detection distance described in Section 3.5.2.E, and finally the system is evaluated in low-light conditions, where it is described the pertinent modifications to work under these conditions.

A. SYSTEM PERFORMANCE AND PRECISION EVALUATION

At first it will be evaluated the performance of the described algorithm in the Eye-RIS™ platform and estimated the frame rate that the system is able to process. To measure processing time it will be evaluated in five different road scene sequences, previously recorded with a total of 2,404 images. Figure 3.25 shows the frame rate that the system is able to process in typical road sequences, the rate ranging in the worst case from 22.8 to 31 fps (a mean frame-rate of 26.9 and a standard deviation of 3.1). Having evaluated the results of the performance of the proposed system it can be concluded that the system is capable to work in real-time and is suitable for application in this field. It's important to note that in field tests were carried out obtaining similar results.

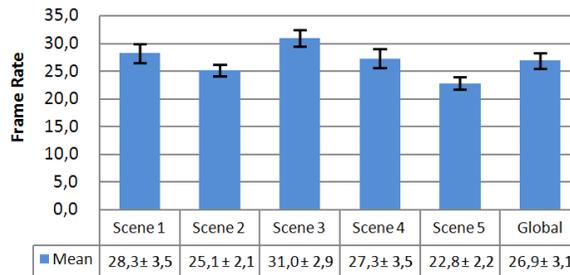


Figure 3.25. Frame rate obtained for different road scenes in which the vertical bar error indicates the deviation.

As far as precision is concerned, the platform was tested on the road to determine the robustness of the proposed algorithm. To accomplish this formal test it was analyzed a continuous video sequence of 15,168 frames while driving on a highway at 80-100 km/h, during which the system detected 76 overtaking cars (positive warnings, or true positives), 14 frames with false positives and no missing vehicles (no false negative). In other words, it can be said that the system is 100% reliable with regard to sensitivity (i.e. all overtaking vehicles were detected). The sensitivity of the system $SENSI = TP/(TP+FN)$, where TP is the number of true positives and FN the number of false negatives.

B. DISTANCE DETECTION EVALUATION

To judge the maximum distance car detection with the described system several tests were undertaken with different relative overtaking speeds (20, 40 and 60 km/h) and different left-hemisphere delays, as explained in Section 3.5.2.E. The measurements were made on a private driving circuit with visual marks on the road to measure distance between objects. Distance measurement at each relative speed was assessed with three different left-hemisphere delays. Each measurement was made three times to obtain a consistent estimation. These assessments were reproduced with the same extrinsic and intrinsic parameters as those used in the road test. Figure 3.26 shows the mean distance detection, taking into account the three speeds in the test evaluation compared to the left-hemisphere delay. The chart reveals that distance detection increases concomitantly with the delay applied to the left hemisphere. The main reason for not evaluating more than three frame delays is because it was found, by trial and error, that the system becomes less stable; i.e. it is more likely to obtain false positives due to the Nyquist theorem violation caused in the optical flow estimation. For this reason, it was decided that the optimal delay value is three.

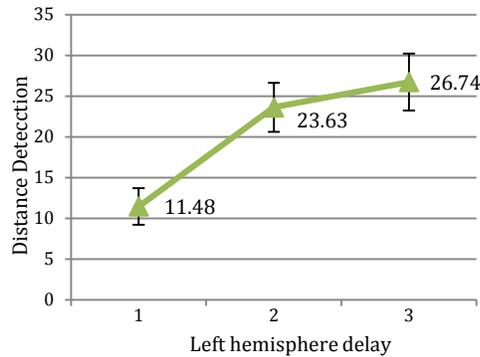


Figure 3.26. Mean distance detection in relation to the left-hemisphere delay; the vertical bars indicate measurement error.

The detection distance establishes driver reaction time to maneuver if necessary. Table 3.6 shows the reaction time in relation to relative speed, taking into account our distance detection with three delays in the left hemisphere. Reaction time decreases with relative speed. According to the research of Green [GRE00], the worst reaction time to overcome is estimated at 1.5 seconds; in our case reaction time is due the steering action taking less time (1.35 sec). With the collected information it can be assumed that the system is able to warn the driver in time.

Table 3.6. Driver reaction time in relation to the relative speed, taking into account our distance detection (26.7 meters).

Relative Speed	Reaction Time
20 Km/h	4.86 s
40 Km/h	2.4 s
60 Km/h	1.6 s

The commercial device to detect overtaking developed by Ford [ICA09], has an overtaking detection range of 20 meters, while that proposed by Volvo [VOL09b] is able to detect within 9.5 meters and the Audi [ICA09] system works over a range of 50 meters. As shown in Figure 3.28, the system is able to detect at a mean range of 26.7 meters for the case of 3-frames hemisphere delay. To extend this range, detection is possible using different optical parameters such as working with higher focal distances for example. In our experiments it was used an 8mm optic. When a longer focal distance (16 mm) was used, however, the system could not cover the area closest to the vehicle and so in order to cover a longer range without losing the near range a dual system with different optical parameters would be required.

C. SYSTEM FEASIBILITY IN LOW-LIGHT CONDITIONS

Driving-assistance systems for vehicle detection in low-light conditions are usually based on the detection of light emitted by headlights, taillights and brake lights. When detecting cars approaching from the rear, O'Malley [OMA08] proposes a solution based on the HSV space and symmetry of the position of lights to detect vehicles. Due to the hypothesis used, this approach is focused primarily on the detection of cars rather than other vehicles such as motorcycles. In the case of frontal detection, Alcantarilla et al [ALC08] describe a method capable of detecting vehicles according to headlight features that are classified by a linear classifier based on SVM and takes into account distance detection to rule out those vehicles above the horizon line.

In our system the proposed algorithm requires no major modifications because it is possible to estimate the optical flow generated by the headlamps without taking into account other features of the car. In our tests the focal plane was responsible for integrating the image in the High Dynamic Range (HDR) using four images with different exposure times (with a total integration time of 12.9 *ms*) to cover a wide gray-level range in all areas. After several tests in low-light conditions it was found that the system worked properly when the lighting was higher than 90 lux. It was also decided to avoid the width filter in low-light conditions (Section 3.5.2.C) because the light emitted by the headlight (on most occasion) reflected on the road, thus distorting the car's proportions, as shown in Figure 3.27(a). It is important to point out that this effect may be controlled by the type of road surface, the illumination of the scene and weather conditions. In order to detect vehicles in these conditions it was used a photo-resistance to determine the degree of illumination in the scene and change the restrictions of our algorithm. After avoiding the vehicle width filter restriction the system works properly, as shown in Figure 3.27(b). At the current state, a qualitative evaluation of the feasibility of the system in such low-visibility conditions is required but, as shown, the results so far are promising and thus our preliminary conclusion is that the system is able to work in low-light conditions as well as normal lighting. This issue will be addressed more fully in the future.

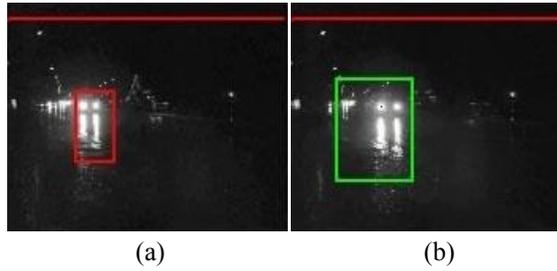


Figure 3.27. Rain scene in low-light condition; (a) Vehicle detected and discarded by the width filter; (b) Vehicle detected as valid after removing the width filter.

3.6.2 Discussion

In this and previous tests it was noticed that the false positives were generated by the wagon-wheel effect [FIN84], where motion direction appears to be opposite to the real one. This effect is generated by the violation of the Nyquist theorem and the system generates a false positive when the wagon-wheel effect is present and the detected area has similar car dimensions (crash barrier, vegetation, shadows etc.). These cases can be ruled out by using further refining schemes, taking into account the egomotion patterns and the speed of our vehicle, though this must be left for future investigation. Figure 3.28 shows some samples of the results in different real overtaking scenarios while Figure 3.29 shows detected false positives and discarded false positives as described in Section 3.5.2.C.



Figure 3.28. Example of true positive detections with our algorithm on a true scenario.

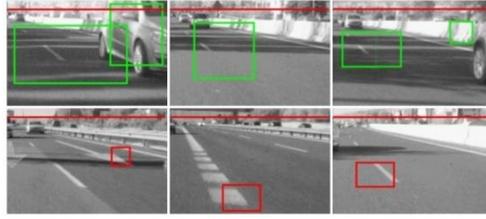


Figure 3.29. Example of false positive detections (first row) and false positives rejected by our approach (second row).

3.7 CONCLUSIONS

Advanced driver-aid systems have begun to play an important role in the automotive industry in recent years. A driver-assistance system has been described and designed to detect approaching vehicles and warn the driver about the risk of changing lane. Hybrid architecture based on a computing focal plane and an on-chip digital processor is used to implement this system. The implementation of the proposed model on this single-chip system is a challenge in itself, only made possible by taking full advantage of the parallel computing resources at the focal-plane and the sequential digital processor. The model has been built on the basis of different stages that are implemented on the focal-plane layer and the digital processing resources, depending on their inherent characteristics.

The described system relies on optical flow to pre-segment approaching objects, after which a final segmentation is carried out on the focal plane and is subsequently evaluated following dimension criteria to discard those that are not vehicle sized. It has been also discussed the effect produced by egomotion when the vehicle turns left on sharp bends or roundabouts and offer a suitable solution based on an external complementary sensor. An inexpensive computational approach to increase the distance of detection is described as well. Finally the whole system is evaluated taking into account its performance, precision, distance detection and the illumination of the scene. It can be concluded that the results are quite promising when bearing in mind the difficulty of this problem. Furthermore, it is worth mentioning that the whole capturing and computing engine is implemented with a single chip, i.e. system-on-a-chip approach. In the evaluation section it is emphasized that all the overtaking cars were detected (no false negatives) making the system very reliable. The false positive cases are affordable because the system will only provide a warning signal if the driver starts a lane-change maneuver.

In future research will be focused on wagon-wheel effect detection to further optimize the accuracy of our approach and also to address the quantitative evaluation of the system in adverse weather conditions as well as night-time scenarios.

4. ULTRASOUND IMAGING

4.1 INTRODUCCION

Vascular ultrasound exploration has been a popular technique widely used in medical sector. One of the advantages that motivates this technology is its non-invasive nature, low price, and the existence of portable systems. Ultrasound, also known as ultrasonography or ecosonography is the procedure to obtain echoes of the high-frequency waves, generated by ultrasound transducers, colliding on a body or object such as tissues, veins, arteries, organs, etc. and translate them into images in order to perform a medical diagnosis. Knowing the speed of propagation of the wave, ultrasonic in this case, the time of flight can be measured and hence estimate the distance of the collided objects. In this type of technology, it is assumed as the average propagation speed to 1540 m/s; that is the one that corresponds to the soft tissues. This assumption produces small errors in the estimation of the distance due the wave travel through different tissues (different propagation media).

The received echoes, originally emitted from the transducer, are the response of the collision over different type of surfaces, producing different types of echoes depending on the surface reflections. In the case of a soft surface, the reflection will be specular (the exit angle formed with a standard to the surface, is equal to the input according to Snell's law). While in the case of a rough surface, it will diffuse reflection type (Scatter). There are also cases where the wave is refracted or absorbed, as is the case of bone-absorbing signal. The relative ratio of the reflected and transmitted power depends on the change in acoustic impedance between two materials. The reflected ultrasound is proportional to the change in acoustic impedance, i.e., as greater is the acoustic impedance change, the proportion of the reflected wave will be higher. Such echoes are registered by the hardware producing black and white anatomical images. This imaging technique can be used to detect and evaluate different diseases related with arterial walls, presence of thrombus in a vein, or other vascular pathologies.

But Ultrasound imaging has also drawbacks, the images produced are very noisy and thus their interpretation requires highly skill personnel. The inter-exploration and inter-specialist variability in the data acquired from Ultrasound explorations is very high. This motivates the development of new models and techniques that facilitate diagnosis from an ultrasound exploration. For instance, tools to semi-automatically segment the artery within an ultrasound image, or even tools to track the artery motion (which is a

dynamic parameter impossible to estimate from a static image). This part of the PhD describes tools developed for these purposes. The other drawback of the ultrasound exploration is the fact that it is done by an expert in a closed loop, i.e., the specialist moves the ultrasound probe according to the interpretation of the images on the screen. Thus semi-automatic segmenting tools need to work in real-time if they aim to be used in this framework. Then, wall motion tracking tools (as the one developed also in this PhD) can be used with recorded sequences, but their processing needs to be done efficiently to allow the specialist to evaluate if the extracted parameter is acceptable or a new recording is necessary (while the patient is still being explored), and this requires it to be done in a few seconds or maximum minutes.

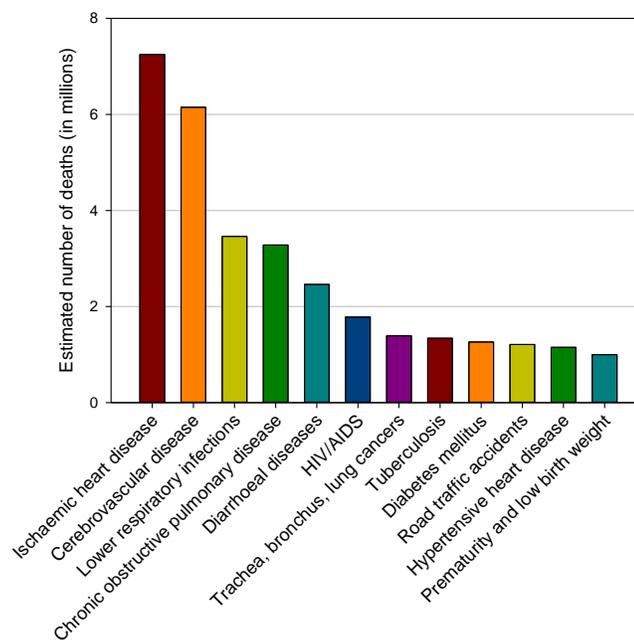


Figure 4.1. The ranking results that lead the causes of death in the world.

It is important to remark that arterial diseases are considered one of the main causes of premature death in the developed world. There are several risk factors associated with the development of arterial diseases, but it has been widely accepted that alcohol, tobacco and an inadequate nutrition are one of the main causes. According to data obtained from the World Health Organization [WHO14], the main diseases related to premature mortality with cardiovascular basis are shown in Figure 4.1 where the first cause of mortality covers 12.8 percent and a total of 25.6 percent of all deaths in relation with cardiovascular problems.

4.2 ARTERY SEGMENTATION IN ULTRASOUND IMAGING

4.2.1 Motivation

In medical imaging, one of the most important topics, which usually turns into a complex task, is image segmentation. Diagnosis based on the measurement of the dimensions of the artery allows experts to identify diseases, such as aneurysm. This disease produces an oversizing in the artery with the risk of a possible rupture. Hence, a good segmentation of the artery based on Ultrasound (US) imaging is important, as it is used for diagnosing different vascular pathologies, such as aneurysm. US imaging represents a crucial medical tool to measure any oversize of the artery given its noninvasive nature, instead of other invasive techniques that make the use of contrast agents. These other techniques are inefficient and expensive as regards their actual cost, in terms of time consumption, and regarding the need for human resources with specific skills. Though US-based explorations require highly specialized personnel, approaches, such as the one described here, aim to create a valid tool, even for not so highly specialized staff.

This chapter is focused on segmenting the outer side of the artery in US imaging, in an easy way, reducing the requirement for specific training, also reducing in this way the inter-intra-specialist variability and, thus, increasing the reliability of the measurements of the diameter of the artery. Semiautomatic measurement schemes, such as the one described in this work, also aim to facilitate the US exploration process, making it suitable even for medical personnel with less specific exploration skills. Figure 4.2 shows an example of a manual segmentation in typical US software. At first, the artery is detected, and the frame is frozen to be analyzed. On this static image, the expert measures, by means of two perpendicular lines, the diameter of the artery. This technique becomes quite rudimentary, because it requires a certain degree of specialization, and even so, the final measurement depends on the current manual skill during the measurement estimation. This leads to high inter-intra-specialist evaluation variability.

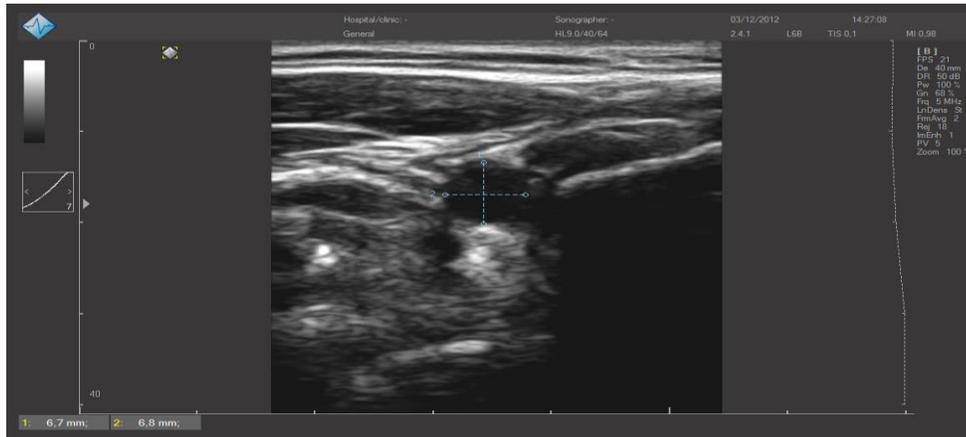


Figure 4.2. Typical manually-driven measurement in ultrasound (US) software (TeleMed) to determinate the diameter of the artery.

The main contribution of this work is the proposal of a novel method based on the geometric nature of the artery, which adopts elliptical shapes. This method is able to segment the artery more accurately than other models proposed in the literature. The optimization of the seed ellipse parameters is done with an evolutionary optimization approach. Due to the high computational cost of the evolutionary models, it will make use of massive parallel architectures and present-day complex algorithms to carry out the segmentation in a reasonable time. The described approach is evaluated and compared with other well-known segmentation techniques, obtaining very promising results in comparison with the state-of-the-art.

The chapter is organized as follows: Section 4.2.2 describe a brief state of art introduction on method based in evolutionary segmentation with ellipses while Section 4.2.3 introduces a detailed explanation of four common segmentation techniques to evaluate the quality of the described approach. Section 4.2.4 is described the proposed technique and detailed the graphics processing unit (GPU) implementation of the suggested method, the obtained performance and a comparison with a parallel CPU implementation. Section 4.2.5 compares the accuracy of the introduced method with respect to the state-of-the-art approaches, while Section 4.2.6 will discuss the obtained results of the evaluated methods and finally Section 4.2.7 summarizes some conclusions.

4.2.2 Background Material and Related Work

Nowadays, many solutions have been proposed in natural imaging to detect ellipses. One of them is the approach introduced by Yao *et al.* [YAO05], where the possibility of making use of a multi-population genetic algorithm, which turns out to be a more robust and faster solution than classic methods, such as the Randomized Hough transform (RHT) [MCL98], was demonstrated. Similar works have been carried out by Lutton and Martinez [LUT94] or Mainzer [MAI02]. In these works, genetic algorithms were also used to detect ellipses in natural images. This kind of method cannot be extrapolated to the US domain because of the existence of several artifacts, such as not-well-defined edges, noise and the presence of high discontinuities in the edges, *etc.* It is also important to remark that those approaches are only based on the binary domain and focused on natural images. In the US framework, an approach also based on ellipse estimation has been recently introduced by Moursi and Sakka [MOU09]. This method first applies a segmentation to obtain an early approximation of the ellipse parameters and is subordinated to the pre-segmentation stage; if this stage fails, the rest of the segmentation stages will also fail. In Section (4.2.3) are detailed also other popular techniques utilized in medical segmentation field and evaluated with the proposed one.

4.2.3 Material and Methods

In this subsection, the most common segmentation methods in medicine will be briefly described and compared with the submitted approach.

A. PARAMETRIC ACTIVE CONTOUR (SNAKES)

Kass *et al.* [KAS87b] introduce a model that consists in parameterizing the curve based on the topology of the image and the internal features of the curve. Given a curve $v(s) = [x(s), y(s)]$, $s \in [0, 1]$, the objective is to fit such a curve to the nearest contour. Hence, Equation (4.1) is minimized.

$$F_{Total} = \int_0^1 F_{int}(v(s)) + F_{ext}(v(s)) + F_{const}(v(s)) ds + F_{Balloon}(v(s)) \quad (4.1)$$

where $F_{int}(v(s))$ is the internal force, $F_{ext}(v(s))$ the external force, $F_{const}(v(s))$ another external force (e.g., one given by the user) and $F_{Balloon}(v(s))$ an external force

that provides expansion or contraction to the contour. The internal force can be defined as in Equation (4.2).

$$F_{int} = F_{cont} + F_{curv} = \alpha(s) \frac{dv}{ds} - \beta(s) \frac{d^2v}{ds^2} \quad (4.2)$$

The first term, F_{cont} , minimizes the length of the arc and α controls the elasticity of the contour. On the other hand, the second term, F_{curv} , minimizes the curve, and β handles the flexibility and smoothness of the contour.

The original approach [KAS87b] was prone to problems when the initialization was far away from the solution. The use of the external force $F_{ext}(v(s))$ aims to minimize this problem. As Xu and Prince [XUU98] propose, expanding the external force field equation with a model that produces vector diffusion Equation (4.3) called gradient vector flow.

$$F_{ext} = \int \int \mu(u_x^2 + u_y^2 + v_x^2 + v_y^2) + (f_x^2 + f_y^2)((u - f_x)^2 + (v - f_y)^2) dx dy \quad (4.3)$$

where μ is a non-negative weight that provides the degree of smoothness, f_x and f_y are the initial external force vectors and (u, v) the new estimated vectors. The last term is motivated by the approach of L. Cohen and I. Cohen [COH93], which included a new extra external force able to inflate or deflate the shape of the contour, and it is defined in Equation (4.4), where $\gamma \in [-1, 1]$, causing the negative values to be a deflation of the shape and the positive ones an inflation. This method has been used, for instance, to approximate the Gallbladder's shape [CIE10], to segment tumors in livers [CVA05], segmentation of masses from breast US images [JUM10] or to segment skin cancer [TAN09].

$$F_{Balloon}(s) = \gamma \vec{N}(s) \quad (4.4)$$

B. REGION-BASED ACTIVE CONTOUR MODEL (ACM)

Zhang et al. [ZHA10a] proposed a novel method that allows a selective segmentation by means of a combination of the geodesic active contour [CAS97] and the Chan–Vese methods [CHA01]. This model makes use of the Signed Pressure Force (SPF), i.e., it modulates the expansion or contraction force of the region of interest, as indicated in Equation (4.5).

$$SPF = \frac{u_0 - \frac{c1 + c2}{2}}{\max\left\{\left|u_0 - \frac{c1 + c2}{2}\right|\right\}} \quad (4.5)$$

where u_0 is the original image and $c1$ and $c2$ are the means of the internal and external regions of the level set. SPF works as a constraint of the region. Hence, Zhang et al. [ZHA10a] removed the weight by the SPF function in geodesic active contour Equation (4.6), where $g(|\nabla I|)$ is the diffusivity function, ϕ the active contour, f the balloon force, which allows one to control the shrinking or expanding of the contour, and α a constant to control the diffusion on weak edges.

$$\frac{\partial \phi}{\partial t} = g(|\nabla I|) \cdot |\nabla \phi| \cdot \left(\operatorname{div} \left(\frac{\nabla \phi}{|\nabla \phi|} \right) + f \right) + \alpha \cdot \nabla g(|\nabla I|)^T \cdot \nabla \phi \quad (4.6)$$

$$\frac{\partial \phi}{\partial t} = SPF \cdot |\nabla \phi| \quad (4.7)$$

The Chan–Vese method has no problem with the leaks; for that reason, $\nabla g(|\nabla I|)^T \cdot \nabla \phi$ is removed, because it is no longer needed. Furthermore, the diffusion term in the original Equation (4.6) was substituted by a Gaussian convolution. Hence, the final equation after the modifications explained above is simplified as indicated in Equation (4.7).

C. SEGMENTATION BASED ON FUZZY C-MEAN CLUSTERING

Fuzzy C-mean (FCM) [DUN73] has become a popular method for segmentation in medical imaging. Abdel-Dayem and El-Sakka [ABD07] make use of this method to segment the carotid in US imaging, where the features that make up the clustering data are formed by the intensity, mean and standard deviation of a 5×5 block. Once the feature vector is extracted per each pixel, the FCM is used to segment the image into three classes. Finally, the user selects the desired segmented area, and by means of morphological reconstruction [VIN93], the selected area will be extracted from the rest.

D. ACTIVE SHAPE MODELS (ASM)

Incorporating shape prior knowledge has been one of the most recent advances in segmentation in the last decade. Cootes et al. [COO95] proposed a statistical method

for the shape that can be deformed with respect to the mean through the most relevant eigenvectors (P) and a parameter, b , as indicated in Equation (4.8).

$$x = \bar{x} + P \cdot b \quad (4.8)$$

The objective of this method is to estimate the parameters to locate the desired object to be segmented by means of matching each landmark with the previously trained normalized gradient profiles and solving linear equations to estimate the desired parameters in a multi-scale strategy. Those parameters are the translation in the x and y axis, scale, rotation and deformation Equation (4.9). Due to the robustness of this method, has been considered in many fields for instance in face tracking [ZHA08] or Tibia bone segmentation in US images [HEE01]. In this evaluation test, seven different kinds of artery topologies were trained in four scales and twelve landmarks per scale, which is enough to cover the common cases.

$$v = \{t_x, t_y, s, \theta, b\} \quad (4.9)$$

4.2.4 System Description

As an alternative to the described methods to segment the artery based on US images, the proposed approach is based on a well-known stochastic technique. The description of the introduced method is organized as follows: Section A presents an introduction to the evolutionary utilized scheme focused to solve the segmentation issue. Section B describes the extracted features from the US images with the objective of address the evolutionary algorithm through the solution space while Section C, explains the ad-hoc objective function and finally Section D describes the implementation of the suggested approach on GPU and its performance evaluation.

A. EVOLUTIONARY SCHEME

Storn and Price [STO97] designed a method based on an evolution scheme to find optimal parameters in a subset of the solutions space, called Differential Evolution (DE). This method has demonstrated that working on low-dimensional problems [HAN06], a good performance is obtained. The main objective of the proposed segmentation method is to find the parameters of a given ellipse that better fit in the boundaries of the artery. This optimization problem is addressed using DE to obtain the parameters (x_c, y_c, a, b, θ) that define ellipse Equation (4.10), as indicated in Figure 4.3.

$$\begin{aligned}x(t) &= x_c + a \cdot \cos t \cdot \cos \theta - b \cdot \sin t \cdot \sin \theta \\y(t) &= y_c + a \cdot \cos t \cdot \sin \theta - b \cdot \sin t \cdot \cos \theta\end{aligned}\tag{4.10}$$

where x_c and y_c are the center coordinates of the ellipse, a and b the minor and major axis, θ the angle with respect to the x-axis and the major axis of the ellipse and t a parameter that varies within the range $[0 - \pi]$.

The pseudo-code of DE is shown in Algorithm 4.1, where the `InitPopulation` is a function that initializes the population in a bounded space defined by the maximum and minimum values of the parameters (`MaxPar` and `MinPar`). The fitness function (`GetFitness`) defines how well the parameters of the agents fit with the desired function (see Section 4.2.4.C for more information about the fitness function). To provide an exploration option to the population in each generation, a mutation operator (`Mutate`) is included. The new mutated agent is defined in Equation (4.11); where agents of the same generation $(\vec{a}, \vec{b}, \vec{c})$ are used to estimate the new one, with a desired mutation factor (F).

$$\vec{u} = \vec{a} \cdot F (\vec{b} - \vec{c})\tag{4.11}$$

Algorithm 4.1. Differential evolution (DE) pseudocode implementation.

```

1 Population = InitPopulation(MaxPar,MinPar);
2 FitPop = GetFitness(Population);
3 BestAgent = GetBestAgent(Population);
4 while (NumIter < NumIterMax)
5 MutPop = Mutate(Population,BestAgent,F);
6 CrPop = Cross(Population,MutPop,CR);
7 FitCr = GetFitness(CrPop);
8 Population = Replace(Population,CRPop,FitCr,FitPop);
9 BestAgent = GetBestAgent(Population);
10 NumIter = NumIter + 1;
11 end while
12 return BestAgent

```

Different approaches of this operator will be evaluated in Section 4.2.5. Another important exploration mechanism is the one provided by the cross-operator (*Cross*), where the mutated agents will be mixed with the current generation with a probability factor, CR (*Cross Rate*).

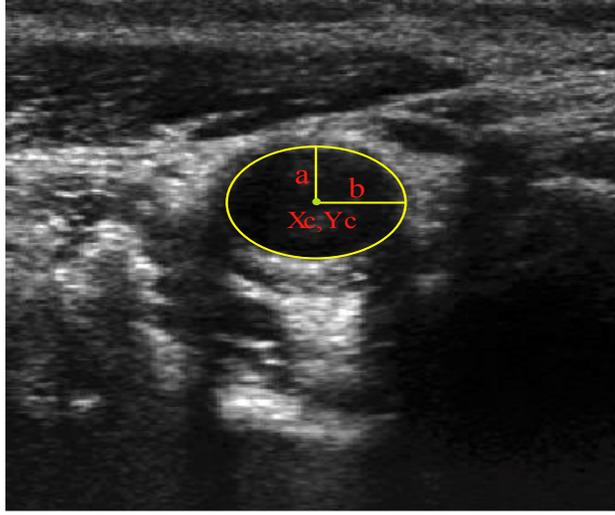


Figure 4.3. Illustration fitted in the boundaries of an artery by means an ellipse.

At this point, the main parameter optimization method used in this paper has been defined. The next subsection will detail the features extracted in the US image used by the DE model to find the desired ellipse parameters that define the best contour of the artery. This search will be led by an *ad hoc* multi-objective function, as explained in Subsection 4.2.4.C, to solve this problem.

B. FEATURE EXTRACTION

To address the DE method over the solution space and estimate the parameters that obtain the best ellipse fitting over the artery, it is necessary to extract features that make it easier to find the desired solution. At first, the proposed approach must estimate the potential x_c and y_c parameters, which define the center of the ellipse by means of the method described by Loy and Zelinsky [LOY03], known as fast radial symmetry (FRS). This algorithm is based on a voting system, which determines the positive and negative affected pixels, $p_{\pm ve}$, through the norm of the gradient, $g(p)$, at point p using a ratio of size n , the orientation projection image, $O(p_{\pm ve}(p))$, and the magnitude projection image, $M(p_{\pm ve}(p))$.

$$p_{\pm ve}(p) = p \pm \text{round}\left(\frac{g(p)}{\|g(p)\|} n\right) \quad (4.12)$$

$$O_n(p_{\pm ve}(p)) = O_n(p_{\pm ve}(p)) \pm 1 \quad (4.13)$$

$$M_n(p_{\pm ve}(p)) = M_n(p_{\pm ve}(p)) \pm \|g(p)\| \quad (4.14)$$

To determinate the saliency map, S , Equation (4.16), the result of Equation (4.15) is convolved with a Gaussian to remove any undesired noise. Parameter α in Equation (4.15) acts as a gain factor of the saliency map to allow the enhancement of low contrast voting gradients.

$$F_n(p) = \|O_n(p)\|^{(\alpha)} M_n(p) \quad (4.15)$$

$$S = \sum_{n \in N} F_n * G_\sigma \quad (4.16)$$

As opposed in the original version of this method, the normalization of Equations (4.13) and (4.14) and the integration of S Equation (4.16) are not performed. These changes are done so as to obtain a faster implementation, as explained in Subsection 4.2.4.D. Figure 4.4b shows an example of the output obtained after applying the FRS algorithm to an US image.

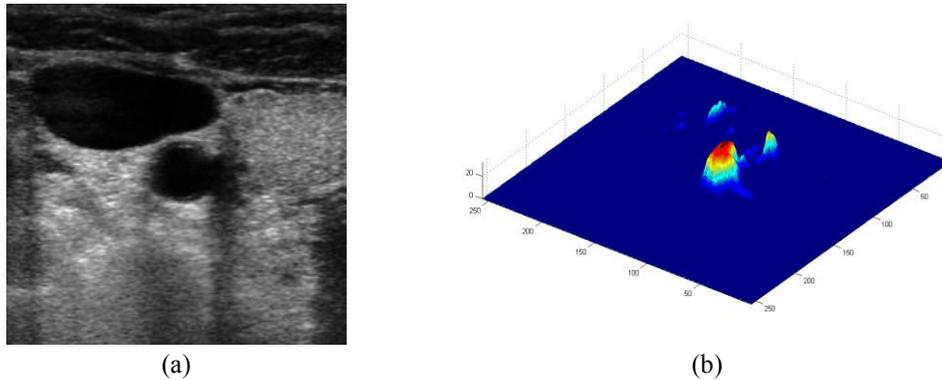


Figure 4.4. (a) Original image and (b) saliency map after applying fast radial symmetry.

At this point, one of the functions that will drive the DE method to optimize parameters x_c and y_c has been described. The guidance functions to obtain a , b and θ parameters in expression Equation (4.10) will be based on edge features, in this case, on the pixel orientation, gradient and binarized edges. The pixel orientation is estimated by means of the eigenvector with a high eigenvalue of the first order structure tensor J

Equation (4.17), where I_x , I_y and I_{xy} denote the spatial derivatives and G_σ a Gaussian convolution for each spatial derivative.

$$J = G_\sigma * \begin{bmatrix} I_x^2 & I_{xy}^2 \\ I_{xy}^2 & I_y^2 \end{bmatrix} \quad (4.17)$$

One of the main problems of working with US images is the amount of speckle noise. Yongjian and Acton [YON02] described a solution to reduce the speckle noise, named Speckle Reducing Anisotropic Diffusion (SRAD), where a non-linear diffusion method to estimate the strength of the edges in US images through instantaneous Coefficient of Variation (ICOV) Equation (4.18) is suggested. To obtain the degree of exchange in the diffusion scheme, Yongjian and Acton use mechanism Equation (4.19), where the diffusion becomes zero in areas with highly contrasted edges and the opposite case when the area is flat. Parameter \tilde{q} determines the level of speckle noise, and therefore, it allows the control of the diffusion over time in Image (I).

$$q(x) = \sqrt{\frac{\left| \frac{1}{2} |\nabla I(x)|^2 - \frac{1}{16} (\nabla^2 I(x))^2 \right|}{\left(I(x) + \frac{1}{4} \nabla^2 I(x) \right)^2}} \quad (4.18)$$

$$c(q(x)) = \frac{1}{\left(1 + \frac{|q(x)^2 - \tilde{q}(x)^2|}{\tilde{q}(x)^2(1 + \tilde{q}(x)^2)} \right)} \quad (4.19)$$

Such noise will interfere in the pixel orientation estimation, as shown in Figure 4.5b. To avoid this artifact, the SRAD method will be applied with the objective of removing the noise and the texture generated by the tissue to obtain a clear pixel orientation, as illustrated in Figure 4.5d.

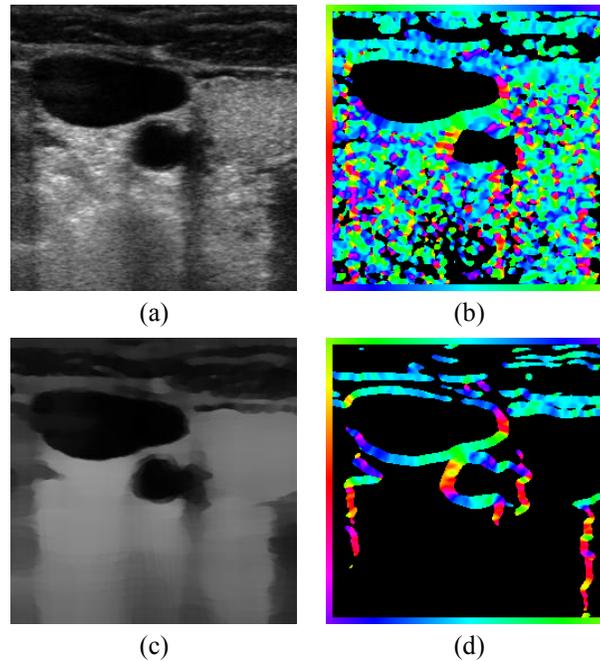


Figure 4.5. (a) Original US image; (b) pixel orientation of the original image; (c) SRAD smoothing; (d) pixel orientation of the smoothed (SRAD) image. In b and c, the pixel local orientation is represented with colors with respect to the color reference frame.

However, this information is not enough yet to address the search. Given the large solution space provided by the pixel orientation, edge information will be incorporated to define the echogenicity. A change of different tissue structures, such as muscle to fat or artery layers to blood, provides a high degree of echogenicity. For this reason, the Logarithmic Image Processing (LIP) edge detector [PAL06] is used. This detector makes use of the logarithm derivatives [DEN95] with the objective of enhancing such areas as much as possible. Figure 4.6a shows the US image filtered by the SRAD method, while Figure 4.6b illustrates the obtained result after applying the LIP-Sobel gradient. The non-max suppression [CAN86] results are shown in Figure 4.6c and finally; the binary map (Figure 4.6d) after thresholding the non-max suppression output is included.

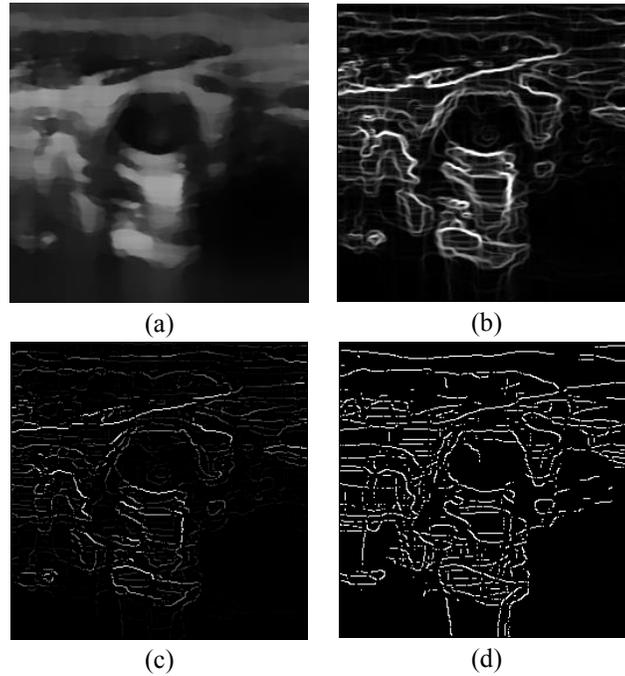


Figure 4.6. (a) US image smoothed with SRAD; (b) LIP-Sobel gradient; (c) non-max suppression applied to LIP-Sobel gradient estimation; (d) binary representation of the edges.

C. ELLIPSE PARAMETER ESTIMATION

Once the features that will drive the evolutionary method to obtain the parameters (x_c, y_c, a, b, θ) that define ellipse Equation (4.10) have been defined, the fitness function of the DE algorithm will be described in this section. The proposed method intends to minimize the angle difference between the norm vectors of the ellipse and the pixel orientation of the image and, at the same time, to maximize the gradient and the binary edge of the artery and maximize the centricity through FRS. To solve this problem, the method exposed by Hajela and Lin [HAJ92] is adopted, where the approach follows a strategy based on weights for each different objective to integrate them into a unique fitness function. The proposed implementation makes use of a penalty scheme that penalizes points of the ellipse where clear orientation, gradient or edge Equations (4.21)–(4.23) do not exist. This penalization will be adaptive, by means of Equation (4.20), where the degree of penalization is defined by the Euclidean norm of the axis of the ellipse and its ratio aspect. A weight factor will be also included to this penalization. After an empirical evaluation, it was demonstrated that by setting those values to $W_o = -90$, $W_g = -1$ and $W_b = -1$, the method obtains reasonable results.

$$D_p = \frac{\sqrt{a^2 + b^2}}{ratio} \quad (4.20)$$

$$f_o(\vec{u}_t, \vec{v}_t, D_r) = \begin{cases} \text{if } \vec{v}_t \neq \text{NAN} & \frac{1}{A_v(\vec{u}_t, \vec{v}_t) + \varepsilon} \\ \text{else} & W_o \cdot D_p \end{cases} \quad (4.21)$$

$$f_g(G_v, D_r) = \begin{cases} \text{if } G_v > \text{Threshold} & G_v \\ \text{else} & W_g \cdot D_p \end{cases} \quad (4.22)$$

$$f_b(B_v, D_r) = \begin{cases} \text{if } B_v = 1 & B_v \\ \text{else} & W_b \cdot D_p \end{cases} \quad (4.23)$$

$$f_{FRS}(x_c, y_c) = FRS(x_c, y_c) \quad (4.24)$$

The first objective function Equation (4.21) is the orientation error, f_o , which estimates the generated angle between the norms of the ellipse vector, (\vec{u}_t) , and the pixel orientation, (\vec{v}_t) , vector by means of function A_v , which defines the angle between both vectors. The second objective function, f_g , Equation (4.22), aims to find shapes with a higher echogenicity given by the LIP-Sobel gradient, G_v . To avoid the gradient generated by small variations in the intensity of the image, a threshold scheme is utilized. The function is set to the gradient if the threshold is higher than a defined value; otherwise, a penalization is applied to avoid the discontinuity in the shape of the ellipse and the gradient map. The third objective function, f_b , Equation (4.23), acts in a similar way as the previous one, but in this case, a binary map, B_v , is used. Finally, the fourth objective function, f_{FRS} , Equation (4.24), will guide the ellipse into the center of the artery using the FRS map. Once all the objective functions have been defined, the fitness function can be expressed as shown in Equation (4.25), where the weights, $\alpha_1, \alpha_2, \alpha_3, \alpha_4$, are incorporated for each respective objective function. After empirical tests, it was found that those values fixed to $\alpha_1 = 100, \alpha_2 = 2, \alpha_3 = 100, \alpha_4 = 30$ of the exposed method achieve satisfactory results.

$$\begin{aligned}
Fitness = Perimeter^{-1} & \left(\alpha_1 \sum_{t=0}^{\pi} f_o(\vec{u}_t, \vec{v}_t, D_r) + \alpha_2 \sum_{t=0}^{\pi} f_g(G_v(x(t), y(t)), D_r) \right. \\
& \left. + \alpha_3 \sum_{t=0}^{\pi} f_b(B_v(x(t), y(t)), D_r) + \alpha_4 f_{FRS}(x_c, y_c) \right) \quad (4.25)
\end{aligned}$$

D. GPU IMPLEMENTATION

One of the main problems when making use of evolutionary schemes is the computation cost, because it requires many iterations (generations) and a considerable number of individual members in each population. To avoid this limitation, a well-known parallel computing architecture, GPU (graphics processing unit), is utilized. The GPU used to evaluate the proposed implementation is NVidia GTX 580 (Fermi architecture; NVidia Corporation, Santa Clara, Calif., US), and it is equipped with 512 processing cores. In this section, the same implementation is also compared with an Intel i7 CPU 950 (Intel Corporation, Santa Clara, Calif., US) that incorporates eight cores (four cores with two logic-cores per each physical one). The technology used to develop the methods, in both architectures, is Open Computing Language (OpenCL), motivated by its degree of versatility to be executed in different platforms. At this point, the implementation will be briefly described and the differences with other methods in literature will be listed.

The differential evolution method has a 100% parallel nature and is very suitable for parallel architectures, such as GPUs [DEV10][ZHU11][ZHU10], The adopted scheme is similar to the one described in [DEV10], with the difference being that the XorShift pseudorandom number generator [MAR03] is used instead of generating random vectors in the host in each generation, with the objective of speeding up the performance. In the evaluation of the proposed method, it is important to remark that unlike other approaches; the advantage of the parallelism on a CPU was taken into account (making use of all cores in parallel) to evaluate, in the most equitable way, the performance of both architectures (GPU and CPU). It has been demonstrated [MAL08] that the number of agents (the size of the DE population of solutions) matters; a small number of agents can produce an early convergence. Therefore, an assessment is carried out by taking into consideration the number of agents compared to the number of generations. Figure 4.7 shows the performance evaluation considering the number of agents and number of generations in the CPU (Figure 4.7a) and GPU (Figure 4.7b). In Figure 4.8, the gain factor obtained on the GPU with respect to the CPU is illustrated, and it shows a significant gain in performance (up to 54 times faster). After some experimental tests,

it was noticed that the use of 4,096 agents and 200 generations is enough to ensure the convergence to an optimized solution.

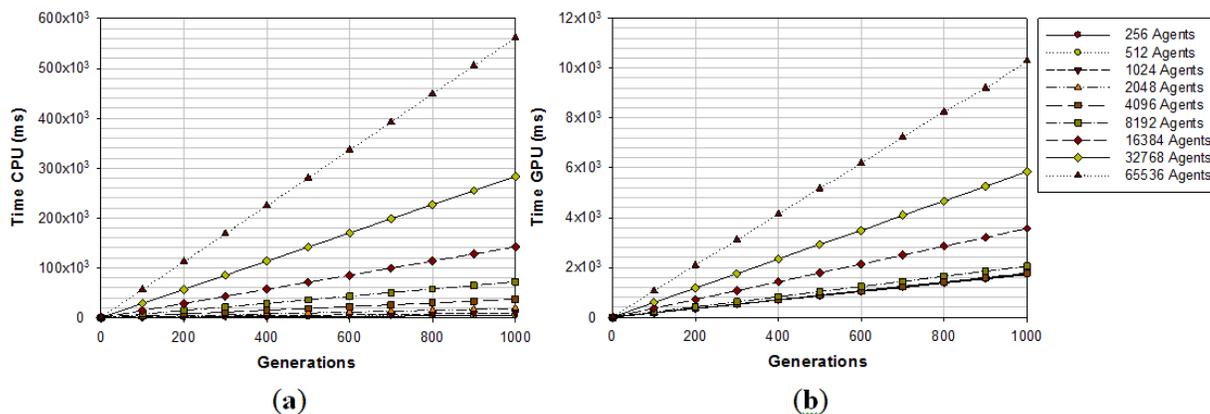


Figure 4.7. Time consumption of the differential evolution algorithm in a (a) parallel CPU and (b) a GPU implementation.

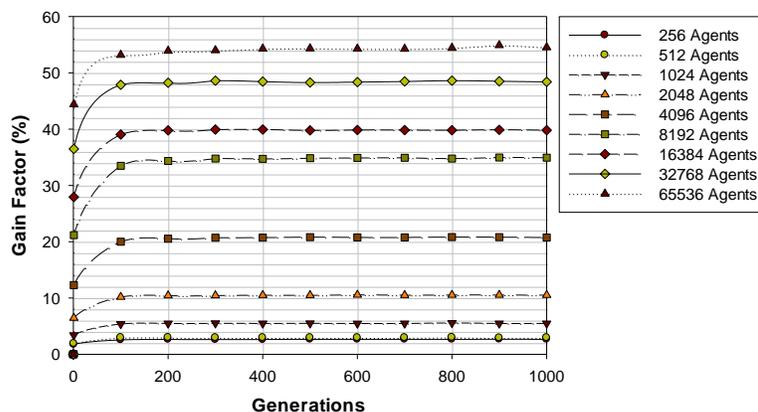


Figure 4.8. The factor obtained on the GPU with respect to the parallel CPU implementation.

Once DE is evaluated in terms of computational speed on a GPU and parallel CPU, the feature extraction implementation in the US images (256×256) on the GPU is detailed. The parallel implementation of the SRAD and FRS methods will be explained in more detail, but the rest of the methods will not be described here, because their implementation is trivial and does not represent any kind of novelty. Nevertheless, all the performance rates of the specific modules are benchmarked in Table 4.1.

Table 4.1. Computation time evaluation of the main methods used in the described algorithm on the graphics processing unit (GPU) and parallel CPU implementation. AOS, additive operator splitting; FRS, fast radial symmetry.

Method	GPU Time	CPU (8 cores) Time
SRAD-AOS 5 iterations	7.65 ms	13.58 ms
FRS	3.78 ms	-
Pixel Orientation	1.14 ms	6.95 ms
Non-Max Suppression	0.16 ms	2.29 ms
LIP-Sobel Gradient	0.11 ms	1.42 ms

SRAD, as well as non-linear diffusion methods in general, is very expensive in computation terms. Weickert [WEI08][WEI98] proposed a semi-implicit scheme to speed up this kind of method: denominated additive operator splitting (AOS). In a recent work, Cao *et al.* [CAO09] exposed a SRAD GPU implementation based on an AOS scheme. To solve the tridiagonal matrix system, which is involved in the AOS method, Cao *et al.* utilized the cyclic reduction (CR) [HOC65] method. In this work, the parallel cyclic reduction (PCR) [HOC81] technique was chosen, due to the fact that it has an impact on the processing gain factor [ZHA10b] in comparison with CR. This approach is a mix of processing techniques, where each system is executed in parallel, and the elements of the system are solved into an iterative parallel scheme.

On the other hand, the FRS algorithm presents a sequential processing that becomes quite difficult to parallelize on a GPU. To solve this problem, Glavtchev *et al.* [GLA11] proposed a smart solution whose computation time becomes invariant to voting size ratio Equation (4.12). To solve the voting integration bottleneck, Glavtchev *et al.* delegated this task to Open Graphics Library (OpenGL) by means of 3D primitives. Since this solution is totally focused on being solved on a GPU, the equivalent CPU implementation cannot be performed. Regarding the LIP-Sobel method and post-processing stages (close to the LIP-Canny model), the developed approach was done in a similar way as in the work presented by Palomar *et al.* [PAL10], but in the last stage, only applying non-max suppression [CAN86] and omitting the hysteresis stage.

Finally, to conclude this section, the global time consumption is estimated. Choosing a population of 4,096 agents and 200 generations in the evolutionary scheme and the image processing analysis (detailed in Section 4.2.4.B), the whole computation time takes 300 ms per image. It is important to remark that in this processing time, the image

transferences to the GPU and other minor algorithms, such as casting, range change, non-max suppression, thresholding, *etc.*, are also included.

4.2.5 System Evaluation and Results

In this evaluation, the US images were acquired from two different US devices with the objective of evaluating the methods in a non-constrained platform. Those devices are a Siemens Antares (128 lines; Siemens AG, Munich, Germany) and a TeleMed Echo Blaster (64 lines; Telemed UAB, Vilnius, Lithuania). One of the most significant differences between these pieces of US equipment is their resolution (double resolution in Siemens with respect to the TeleMed device). To determine the accuracy of the methods, the well-known F-measure estimator Equations (4.26)–(4.28) are used. This estimator evaluates a benchmark dataset of 40 US images of different patients and areas, with their respective ground truth (where an expert marked it manually, point by point).

$$recall = \frac{\#correctly\ classified\ foreground\ pixels}{\#foreground\ pixels\ in\ GT} \quad (4.26)$$

$$precision = \frac{\#correctly\ classified\ foreground\ pixels}{\#pixels\ classified\ as\ foreground} \quad (4.27)$$

$$F - Measure = 2 \cdot \frac{recall \cdot precision}{recall + precision} \quad (4.28)$$

The stochastic method differential evolution was described in Section 4.2.4.A, which states that the mutation operator aids the model to explore the space solution through Equation (4.11). To determine the most appropriate strategy mutation, different mutation models Equations (4.29)–(4.32) with the exposed metrics Equations (4.26)–(4.28) are evaluated.

- (1) DE/Best/1:

$$\vec{u}_i = \vec{v}_{Best} + F \cdot (\vec{v}_i - \vec{v}_{rand1}) \quad (4.29)$$

- (2) DE/Current to Best/1:

$$\vec{u}_i = \vec{v}_i + F \cdot (\vec{v}_{Best} - \vec{v}_i) + F \cdot (\vec{v}_{rand1} - \vec{v}_{rand2}) \quad (4.30)$$

(3) DE/Current to Best/2:

$$\vec{u}_i = \vec{v}_i + F \cdot (\vec{v}_{Best} - \vec{v}_i) + F \cdot (\vec{v}_{Best} - \vec{v}_{rand1}) \quad (4.31)$$

(4) DE/Rand/1

$$\vec{u}_i = \vec{v}_i + F \cdot (\vec{v}_{rand1} - \vec{v}_{rand2}) \quad (4.32)$$

where \vec{u}_i denotes the mutated vector, \vec{v}_{Best} the best agent in the current generation, \vec{v}_i the current agent and \vec{v}_{rand1} and \vec{v}_{rand2} the random vectors in the evaluated generation. Figure 4.9 shows the obtained segmentation results with different mutation operators. To perform this test, the experiment has been repeated 10 times (per each operator) and the mean of the obtained results has been computed. It can be observed that the best mutation operation can have an impact of about 3% in precision and almost 1% in the F-measure with respect to the worst case.

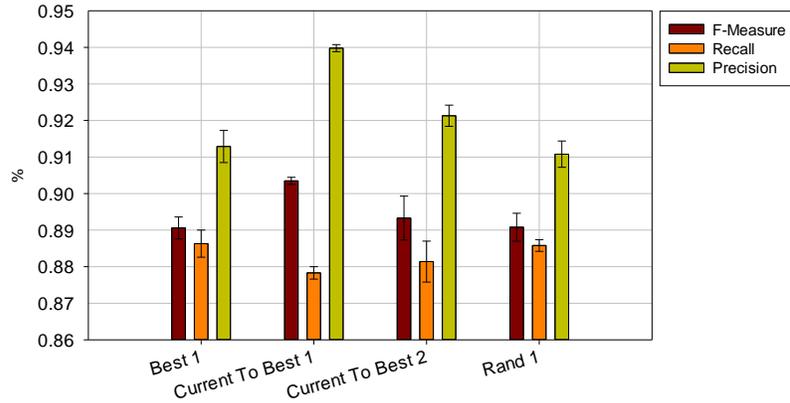


Figure 4.9. The evaluation of the proposed method with different mutation schemes with respect to the ground truth (manually marked) with its respective standard deviations (after 10 trials in each set up).

It is clearly demonstrated (Figure 4.9) that the scheme Current to Best 1 Equation (4.30) is the best option, providing also a small standard deviation, while Rand 1 and Current to Best 2 schemes provide the worst results. To validate the proposed extracted features and to quantify the repercussion of each one, the impact of the combination of the features of each image is evaluated (Figure 4.10). It can be observed that the use of the binary mask and the edge information is not enough to find the desired parameters. The orientation feature improves the accuracy of the method up to 30% more than

making use of the other features. This means that it becomes the most important characteristic in the described method. The rest of the features aid in refining the accuracy of the model and reduce the standard deviation in the results.

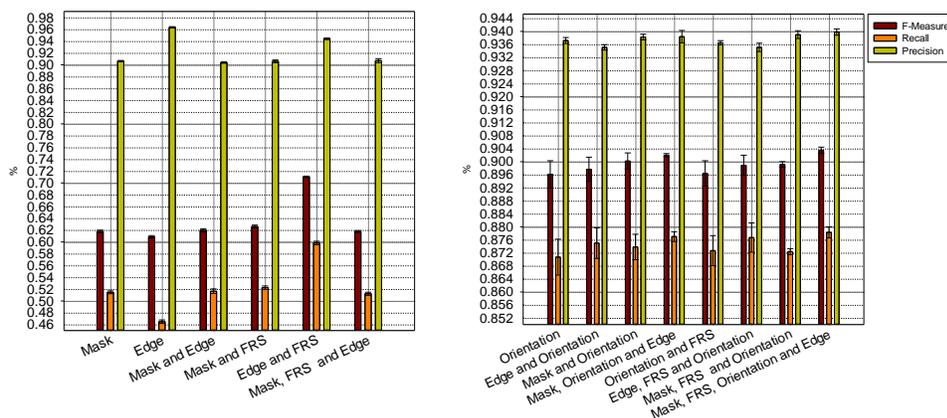


Figure 4.10. Evaluation of the proposed method with different features mixed (**left**) without orientation and (**right**) including the orientation feature.

Finally, a last evaluation is carried out, where the results of the alternative methods (Section 4.2.3) and the proposed approach (Section 4.2.4) are compared. Figure 4.11 reveals that the presented method obtains the best results, and not only in the F-measure metric, but also in the recall and precision. In the Discussion Section, some examples of the results of this comparative evaluation are shown. Furthermore, the suggested method has an improvement in the F-measure of 3%, 5% in recall and 2% in precision with respect to active shape models (ASMs). In the case of the comparison with respect to the parametric snake method, the obtained improvement is about 4% in the F-measure, 7% in recall and 2% in precision.

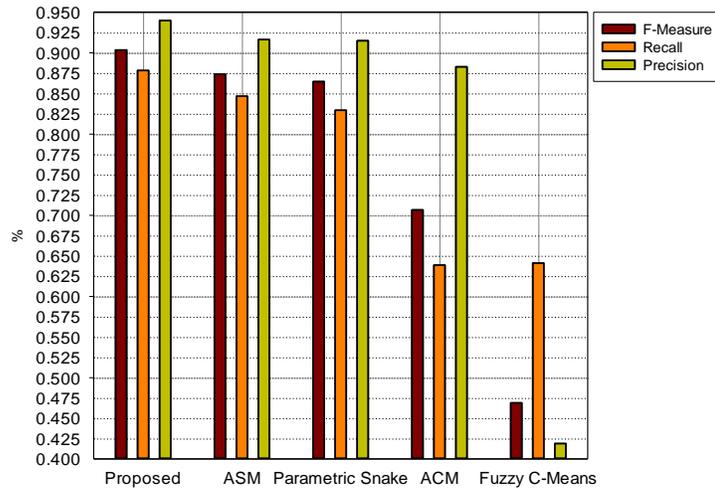


Figure 4.11. The final comparative results of the F-measure, recall and precision with respect to other state-of-the-art methods. ASM, active shape model.

To obtain a more detailed evaluation between the proposed method and the second best in the ranking, both methods are compared by means of a Bland–Altman plot. Figure 4.12 shows the difference between the proposed method and the second best results (ASM) and their averages. The middle line indicates the average difference of both methods, whereas the upper and lower lines represent the 95% limits of agreement with the 16.97% with respect to the mean difference. It can be concluded that Figure 4.12 shows an overall good agreement of the amplitudes between the presented method and the reference one (ASM).

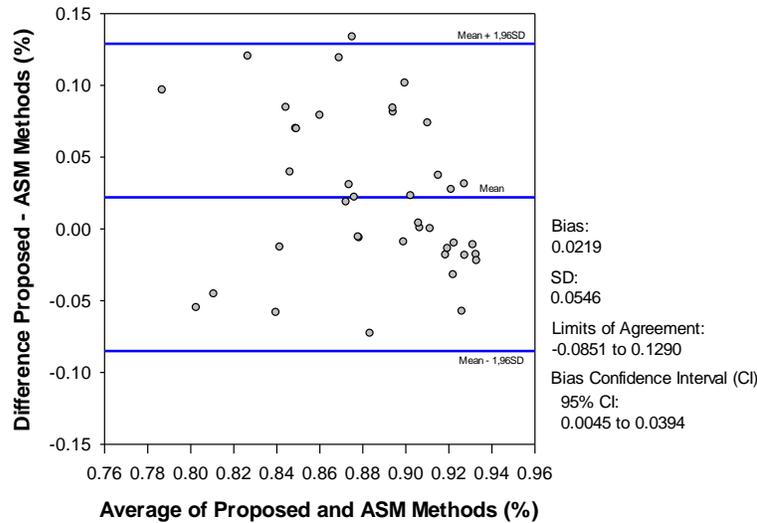


Figure 4.12. The Bland–Altman figure comparing the best two obtained results (the proposed one and ASM).

4.2.6 Discussion

After the evaluation performed in the previous section, some results obtained with the assessed segmentation methods are exhibited, and the pros and cons of each model are discussed. Figure 4.13a,b shows the ground truth (red color) that defines the area of the artery. The other one, Figure 4.13c–l, illustrates the results obtained by the different methods, where the green color illustrates the false positives, the false negatives are marked in blue color and cyan indicates the correct segmentation. Fuzzy C-mean (Figure 4.13c,d) and ACM (Figure 4.13e,f) produce leaks in the segmentation, because these methods are not appropriate when the US image produces shadows (a quite common artifact). On the other hand, methods that preserve the shape (parametric snake, ASM and the proposed one) are more robust for handling the leaks, as is illustrated in Figure 4.13g,l, where such an artifact is not produced.

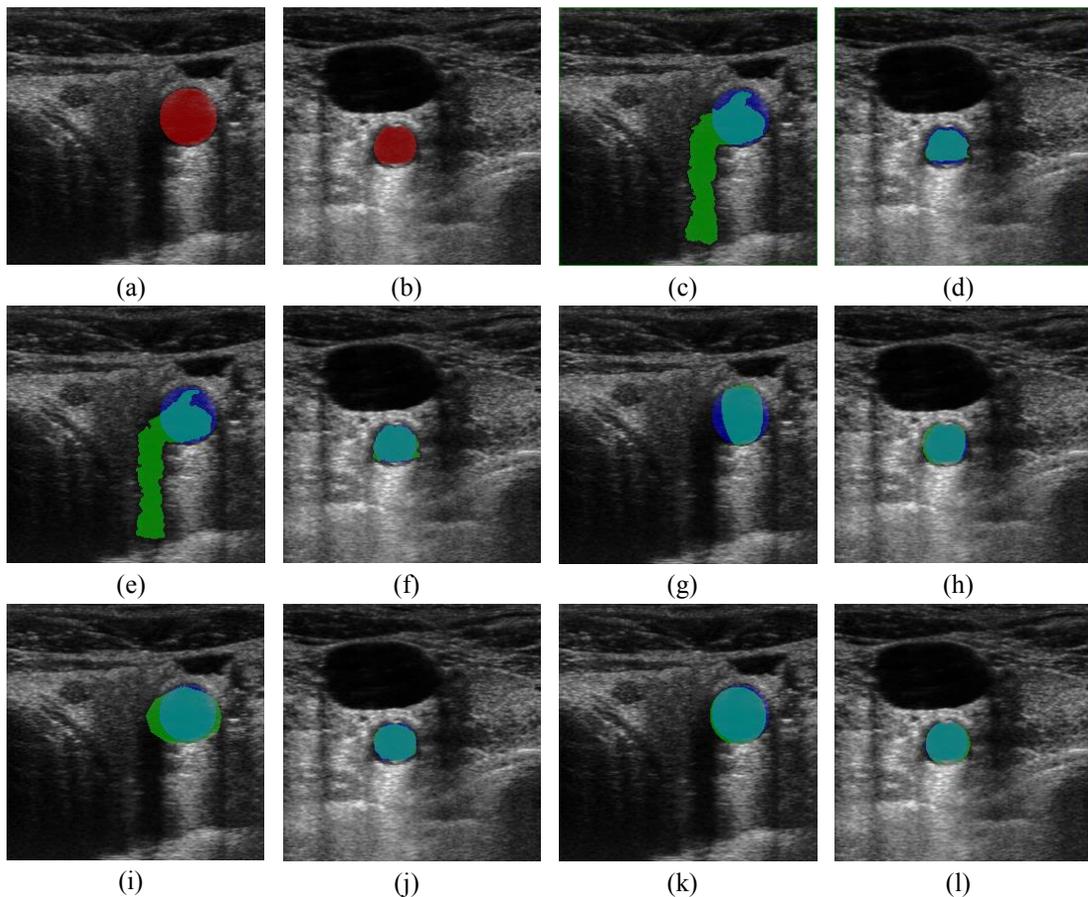


Figure 4.13. (a,b) Original US images with the ground truth marked in red; (c,d) fuzzy C-mean; (e,f) active contour model; (g,h) parametric snake; (i,j) active shape model; (k,l) proposed method. Please see the main text for the meaning of each color.

Comparing the group of methods that preserve the shape, it is important to remark that the parametric snake model provides different results in relation with its parameter settings (thus, it is sensitive to its internal configuration parameters). Another inconvenience occurs in the initialization of the method, which must be done close to the final solution. Those problems vanish with the ASM method, where the variation of the shape is defined in the training stage (avoiding parameter sensitivity), and its space search is longer than parametric snake, through its multi-scale scheme. The suggested method does not require the training stage (unlike ASM), because it is based on the geometric nature of the artery, the ellipse.

As demonstrated in Section 4.2.5, the proposed approach reaches a good accuracy (obtaining an improvement with respect to ASM of 16.97%, as shown in the Bland–Altman figure (Figure 4.11), and obtains an excellent GPU performance (up to 54 times faster than the parallel CPU implementation), as demonstrated in Section 4.2.4.D. Finally, to conclude the discussion chapter, some accomplished results with the introduced method in different patients and topologies of the artery are illustrated (Figure 4.14).

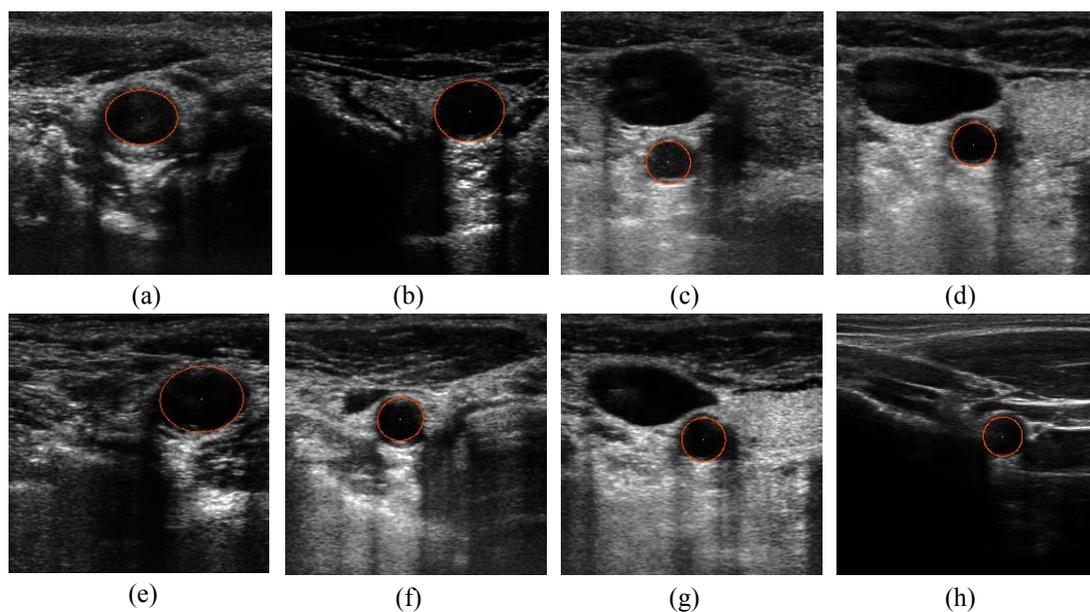


Figure 4.14. Some results obtained with the proposed method in different patients.

4.2.7 Conclusions

In this work, a method based on an evolutionary approach for optimizing different kinds of features to fit an ellipse that best defines the edges of the artery has been proposed. It has been demonstrated that it can be computed efficiently making intensive use of a GPU platform. Its high accuracy in relation with other state-of-the-art methods is also highlighted.

The submitted approach outperforms other methods, not only in terms of accuracy, but also because in comparison with the ASM method, the suggested method does not require any kind of previous training stage. Another advantage that must be remarked upon is that the proposed method supports large search spaces, unlike ASM or parametric snakes, which need to be initialized close to the final solution.

4.3 ARTERIAL WALL TRACKING

4.3.1 Motivation

Estimating the variation of motion in the artery for vascular characterization [BON00] is a new technique that helps doctors to detect specific diseases. Other non-invasive techniques such as Ankle Brachial Pressure Index (ABPI) [YAO69] or Augmentation Index (AIx) have been used to estimate parameters (blood pressure) that are associated with peripheral vascular diseases. For example, Mortensen *et al.* [MOR09] demonstrated the relation of AIx and the Marfan syndrome, the role that involves AIx in the hypertension field [SHI08] and the increase of the arterial stiffness in human subjects with Type 1 diabetes mellitus [WILK00]. The way to estimate the pressure parameters becomes very limited due to the fact that such measures cannot be estimated in other parts of the body besides the carotid artery where we have an easy access with US. Arterial pressure and arterial wall motion are related since estimating the pressure requires measuring the variation of the diameter of the artery, as it is indicated in Equation (4.34). The importance of the wall motion artery's characterization has been also discussed by several authors who have demonstrated that radial [BON00, RAM03, KAN98, BON96] and longitudinal [ZAN12, ZAN10] [CIN06] motion are promising indicators to be associated with certain diseases or pathologies.

Existing commercial solutions such as Tissue Doppler Imaging (TDI) focus on the velocity measurement of the myocardial motion using Doppler principles. This technique has been extended to other applications in echocardiography [BON00, RAM03], to determine the mechanical properties of vessels by means of TDI. The main problems with using TDI are that the motion vector measurement can only be done in parallel to the direction of the ultrasound beam, TDI measures absolute tissue velocity, and it is not able to distinguish all passive motion [HOO06].

The chapter is organized as follows: Section 4.3.2 describe a brief state of art introduction on method artery tracking. In Section 4.3.3 introduces a detailed explanation of the evaluated models, the process to generate the ground truth estimation and the combination of the methods that will be evaluated and compared in Section 4.3.4. In Section 4.3.5, the obtained results of the evaluated methods will be discussed and finally, Section 4.3.6 summarizes some conclusions and outlines for future work.

4.3.2 Background Material and Related Work

Different solutions have been proposed to characterize the wall artery motion directly from ultrasound images in order to complement the information about motion patterns extracted from B-mode US. Image intensity correlation techniques have been widely used in ultrasound due to their robustness under noisy environments. Golemati *et al.* [GOL12] compared the displacement error produced in block matching [DUF95] and optical flow [LUC81, HOR81] methods over a simulated dataset. The matching feature is also an important factor, where Soleimani [SOL11] demonstrates that by including the gradient in the local search, the method results improve. The inclusion of a Kalman filter [KAL60] to update the reference block and the displacement vector [ZAH13, GAS01] has been also evaluated. This method becomes useful when the registered data is corrupted by significant amounts of noise, but in cases where the information is not corrupted at all, the filter does not improve the accuracy of the system rather produces over-smoothing. Other authors [BEU11, VAP11] go one step further and not only measure the displacement of the wall, but also include the Pulse Wave Velocity (PWV) to estimate the pressure by mean Moens-Korteweg Equation (4.33), that relates the PWV to the elasticity of the arterial wall:

$$PWV = \sqrt{\frac{A_0 \partial P}{\rho \partial A}} \quad (4.33)$$

where A_0 is the arterial diameter in the diastole, ρ is the density of blood, ∂A is the variation of the diameter over time (determined using cross-correlation in [BEU11,VAP11]) with respect to the artery in resting position, and ∂P is the difference in the pressure with respect to the end-diastolic pressure $P(0)$. The pressure produced in the artery can be estimated using PWW , as shown in Equation (4.34):

$$\partial P = \frac{\rho \cdot PWW^2}{A_0} \partial A \quad (4.34)$$

Compared to previous works, our method supports sub-pixel accuracy and incorporate collective motion information to define the wall artery motion. The major contribution of this paper is the evaluation of different methods and how they can be integrated to better address our problem of estimating the change in diameter ∂A . In this work, in order to enhance existing motion tracking methods, a combination of similarity transformation, non-rigid deformations, statistical filtering, and hybrid motion estimation techniques are proposed. In this way, it will be possible to estimate useful parameters instead of using more expensive and invasive methods that put the patient's well-being at risk.

4.3.3 Material and Methods

In this section, the methodological ‘building blocks’ used in this Thesis will be first briefly described. Then, the performance of the methods in different analysis pipelines will be evaluated.

A. EVALUATED METHODS

A.1 Block Matching

The block matching (BM) technique has been a very popular method in the ultrasound field because it provides a robust estimation of the motion by means of comparing the similarity between blocks of different images. One of the uses of motion estimation via BM technique is the one proposed by Basarab [BAS08], where the elastography map is estimated to show hidden objects such as cysts or cancer tumors in ultrasound imaging. This work uses a multiscale scheme to avoid errors in the motion estimation and to obtain a low sub-pixel resolution. It is important to remark that normalized cross correlation (NCC) block matching method is one of the most popular techniques

utilized in ultrasound tracking [ZAN12, ZAN10, CIN06, GAS102, CIN05]. In our evaluation, it was decided to make use of Lewis method [LEW95] due to the fact that the obtained performance is much superior to the original one (approximately 15 times faster). Lewis method consists of a modification of the NCC technique where the similarity is given by Equation (4.35):

$$\frac{\sum_{(i,j) \in W} (I_1(i,j) - \bar{I}_1) \cdot (I_2(x+i, y+j) - \bar{I}_2)}{\sqrt{\sum_{(i,j) \in W} (I_1(i,j) - \bar{I}_1)^2 \cdot \sum_{(i,j) \in W} (I_2(x+i, y+j) - \bar{I}_2)^2}} \quad (4.35)$$

where I_1 is the reference block with size W , I_2 is the image where the correlation is carried out, \bar{I}_1 and \bar{I}_2 are their respective means. The numerator in Equation (4.35) can be seen as a convolution (as in Equation (4.36)) between two images, $f(i,j) = I_1(i,j) - \bar{I}_1$ and $t(i,j) = I_2(i,j) - \bar{I}_2$, and can be efficiently solved by means of a Fourier convolution:

$$\sum_{(i,j) \in W} f(i,j) \cdot t(x+i, y+j) \quad (4.36)$$

On the other hand, the denominator of Equation (4.35) must be solved efficiently. Lewis [LEW95] proposed the use of the integral of the image technique to compute it efficiently and reduce the computation cost.

A.2 Optical Flow

The temporal variation in an ordered sequence of images allows the estimation of the optical flow 2D vector, usually denoted as $\vec{v} = (u, v)$, and is computed based on the constant-brightness hypothesis, which assumes that the pixel brightness remains constant over time. This leads to the formulation of the famous optical flow constraint Equation (4.37):

$$uf_x + vf_y + f_t = 0 \quad (4.37)$$

where u and v are the optical flow components and the spatio-temporal derivatives are represented by f_x , f_y and f_t respectively. It is important to remark that in the two considered optical flow methods (described in this Section), the texture domain of the image was used, as proposed by Wedel *et al.* [WED08], so as to avoid problems with the brightness assumption.

A.2.1. OPTICAL FLOW: LUCAS AND KANADE

On the basis of the optical flow constraint equation, Lucas and Kanade [LUC81] proposed the minimization of the error Equation (4.38) using the sum of the least squares:

$$E(u, v) = \sum_{i \in B} (f_x(i)u + f_y(i)v + f_t(i))^2 \quad (4.38)$$

Equation (4.38) is minimized by taking the partial derivatives with respect to the optical flow vector \vec{v} . The resulting \vec{v} minimizes the differential error between the previous image and the current image and is given is presented in Equation (4.39):

$$\vec{v} = \begin{bmatrix} \sum_{i \in B} f_x^2(i) & \sum_{i \in B} f_x(i)f_y(i) \\ \sum_{i \in B} f_x(i)f_y(i) & \sum_{i \in B} f_y^2(i) \end{bmatrix}^{-1} \begin{bmatrix} -\sum_{i \in B} f_x(i)f_t(i) \\ -\sum_{i \in B} f_y(i)f_t(i) \end{bmatrix} \quad (4.39)$$

where \vec{v} is the optical flow vector, the spatio-temporal derivatives are represented by f_x, f_y and f_t respectively, and the subscript i is the i -th element of the integration block B .

A.2.2 OPTICAL FLOW: ANISOTROPIC TV-L1

The anisotropic optical flow Equation (4.40), introduced by Werlberger *et al.* [WER09], is an extension of the popular method TV-L1 optical flow [ZAC07], which is based on a regularized propagation technique similar to the one proposed by Horn and Schunck [HOR81]. Werlberger introduces an anisotropic diffusor that does not propagate values through the edges, with better preserves image structure. In this case, the original formulation was changed in order to reduce the computation cost in the last term of Equation (4.40), where possible artifacts (e.g., occlusions) can be rectified over the time incorporating feedback from previous optical flow:

$$\min_{\vec{u}, \vec{v}} \sup_{|\vec{p}_d| \leq 1} \left\{ \int_{\Omega} \sum_{d=1}^2 \left[(D^{1/2} \nabla u_d) \cdot \vec{p}_d - \varepsilon \frac{|\vec{p}_d|^2}{2} + \frac{1}{2\theta} (u_d - v_d)^2 \right] + \delta |\rho(\vec{v}(\vec{x}))| \right. \\ \left. + \frac{\lambda}{2} \int_{\Omega} (u'_d - u_d)^2 d\vec{x} \right\} \quad (4.40)$$

where $D^{1/2} = \exp(-\alpha |\nabla I|^\beta) \vec{n} \vec{n}^T + \vec{n}^\perp \vec{n}^{\perp T}$, $\vec{n} = \frac{\nabla I}{|\nabla I|}$ as the normal vector, and \vec{n}^\perp the tangent vector of a given point. u_d is the optical flow vector and u'_d is the previous warped optical flow [BRO04, ANA89].

A.3 Kalman Filter

In noisy systems, the Kalman Filter [KAL60] has been proposed due to its robustness and efficiency. This method is based on a statistical approach to determine the current estimation of a linear system from a collection of previous observations over time as described in Equation (4.41):

$$\hat{x}_k = A \hat{x}_{k-1} + B u_{k-1} + w_{k-1} \quad (4.41)$$

where \hat{x}_k is the predicted estimation in the current time k , \hat{x}_{k-1} is the previous observation, A and B describe the transition and control matrix respectively, u_{k-1} is the control signal, and w_{k-1} , the process noise of the system.

A.4 Similarity Transformation

Incorporating shape prior knowledge has become common practice in segmentation methods in the last decades. Cootes *et al.* [COO95] introduce a statistical method able to deform a contour by means of weighting relevant eigenvectors (P) by shape parameters (b) with the objective of adapting the contour to a desired object in the image, as shown in Equation (4.42):

$$x = \bar{x} + P \cdot b \quad (4.42)$$

The objective of this method is estimating the shape parameters as well as the pose parameters (translation in x and y-axis, scale, rotation) that locate the desired object to be segmented in Equation (4.43):

$$v = \{t_x, t_y, s, \theta, b\} \quad (4.43)$$

These parameters are found by means of matching each landmark with the previous trained normalized gradient profiles and solving linear equations [HAM03]. In our work, we will not adopt the statistically-based deformations, thus the local deformation terms can be discarded. In other words, we neglect the shape variability encoded in b and focus on estimating the remaining pose parameters only. Therefore, Equation (4.42) is no longer needed in our approach and we only utilize the weighted similarity transformation proposed in [HAM03]. Such a similarity transformation is obtained by means of the weighted sum minimization in Equation (4.44):

$$E = (x1 - M(s, \theta)[x2] - t)^T W (x1 - M(s, \theta)[x2] - t) \quad (4.44)$$

where:

$$M(s, \theta) \begin{bmatrix} x \\ y \end{bmatrix} = \begin{pmatrix} (s \cdot \cos\theta) \cdot x - (s \cdot \sin\theta) \cdot y \\ (s \cdot \sin\theta) \cdot x + (s \cdot \cos\theta) \cdot y \end{pmatrix}, \quad (4.45)$$

$$t = (t_x, t_y, \dots, t_x, t_y)^T$$

$x1$ is the origin point and $x2$ the translated point, s the scale, θ the rotation, t the translation vector, and W is a diagonal matrix of weights for each point.

A.5 Soft Body Dynamics

Soft body models have been widely used in computer science to carry out realistic physical simulations of motion and deformable objects. Rather than a statistically-based deformation model in Equation (4.43), this work will focus on the popular mass-spring model, which is based on a mesh of nodes (masses) and connected by means of elastic links (springs). The basis of this method relies on Hooke's law, to simulate the spring force, and the second Newton's law to simulate the dynamics by time integration. In this work, a simplification of the idea introduced by Hamarneh *et al.* [HAM03] will be adopted. The authors describe a system in Equation (4.46) that involves forces generated by the springs' system, in Equation (4.47), in a controlled environment (Equation (4.48)) allowing for speedup/slowdown of the velocity of the mesh's nodes:

$$f_i = f_i^{Hooke} + f_i^{viscous} \quad (4.46)$$

$$f_i^{Hooke} = -k_s(\|x_i - x_j\| - r_{ij}) \frac{x_i - x_j}{\|x_i - x_j\|} - \left(k_d(v_i - v_j)^T \frac{x_i - x_j}{\|x_i - x_j\|} \right) \frac{x_i - x_j}{\|x_i - x_j\|} \quad (4.47)$$

$$f_i^{viscous} = -k_v v_i \quad (4.48)$$

where f_i is the final estimated force, k_s the Hooke's spring constant, x_i the Cartesian coordinate of the i -th node, r_{ij} the rest length associated to a link between two nodes, k_d the damping constant, k_v the viscosity coefficient, and v_i as the velocity of the i -th node. Once the nodes' force is obtained, it is possible to estimate the acceleration, velocity, and position of each node by means of the iterative scheme in Equation (4.49):

$$\begin{aligned} a_i &= \frac{f_i}{m_i} \\ v_i &= v_i^{old} + a_i \Delta t \\ x_i &= x_i^{old} + v_i \Delta t \end{aligned} \quad (4.49)$$

where a_i is the acceleration, f_i the force described in Equation (4.46), v_i the velocity, x_i the position, and Δt the time interval.

B. GROUND TRUTH ESTIMATION VIA ULTRASOUND SIMULATION

To evaluate the proposed methods, the Field II Ultrasound MATLAB library [JEN96] was used to generate nine sequences simulating the wall displacement of the common carotid artery. Each sequence involves a complete cycle of the cardiac system with a frequency of 25 Hz per cycle. These simulations were generated with 1024 physical elements, a transducer center frequency of 5 MHz, 100 MHz of sampling frequency, and 64 active elements. The sequences are based on three different topologies as shown in Figure 4.15, where diverse amplitudes of motion are used, as explained later in this section. Although these synthetic images are clearer than real US artery images, having the ground-truth motion allows quantitative evaluation metrics to compare different methods. We include also the simulation parameters used to produce these synthetic sequences (Table 4.2) to facilitate the reproduction of our results.

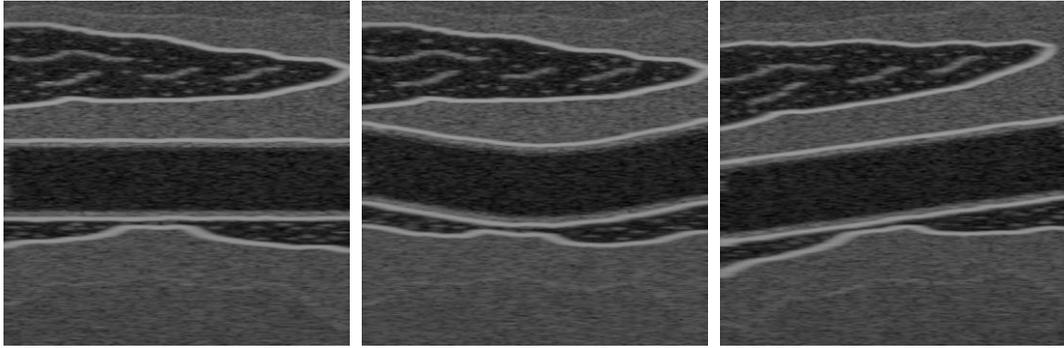


Figure 4.15. Ultrasound images used in the data set test, generated by Field II U.S. simulator.

Following the steps of Stoitsis *et al.* [STO08], that describe a mathematical mechanical deformation model of the arterial wall as a separable method in space and time, this paper will simulate the radial and longitudinal displacement of the artery by means of Equations (4.50) and (4.51), which is a simplification of the original method.

$$rt(t) = \prod (t_0, t_1) \cdot \sin^2 \frac{\pi \cdot t}{c \cdot T} + \prod (t_1, t_2) \cdot (a + b \cdot t) \quad (4.50)$$

$$\prod (t_i, t_j) = f \cdot (1 + \tanh(d \cdot (t - t_i))) \cdot (1 + \tanh(d \cdot (t_j - t))) \quad (4.51)$$

where a and f determinate the amplitude of the waveform, b defines the slope in the second part of the curve, c and T are coefficients that control the initial part of the curve, d determines the wall artery speed, t_1 and t_2 correspond to the duration of the first and second pulse of the waveform and t is the time variable. In the generated data set, the chosen parameters are shown in Table 4.2, where different values of f let us control the amplitude of the artery displacement. For each generated artery topology, three values of f will be applied to produce different radial and longitudinal displacements.

Table 4.2. Wall Displacement Simulation Parameters utilized in our experiments.

Parameter	Value
a	15.14
b	-0.64
c	1.5
t_1	0.25t
t_2	0.65t
d	1.22
f	0.06, 0.12, 0.25
T	1

4.3.4 System Description

Following the introduction of the methodological building blocks in the previous section, the described combination of methods to be evaluated in Section 3 will be explained in detail. To estimate the motion of the wall of the artery, two optical flow methods and the block matching approach proposed by Lewis [LEW95] will be evaluated.

One of the main problems in classical block matching techniques is the sub-pixel accuracy. To handle this problem in an efficient way, other authors [CHA10, CHI07] describe a combination of optical flow and block matching, in order to increase the motion vector precision. This approach relies on estimating firstly the motion vector by means of the block matching technique, and then applying a warping [BRO04, ANA89] to the block and computing the optical flow to estimate the sub-pixel information (as shown in Figure 4.16).



Figure 4.16. Block matching with sub-pixel accuracy by means of the optical flow scheme.

Most of the wall artery tracking papers take into account the tracking individually [GOL12, GAS10] but do not consider all the tracking points as a set of data that define a semi-rigid object in motion. In this paper we use all tracked points when estimating the similarity transformation parameters (rotation, scale, and translation), as explained in Section A.6. This method obtains all the motion vectors for all point and estimates the transformation parameters by means of Equation (4.45). One of the advantages of this method is the robustness against noise, allowing the computation of the parameters that define the transformation even with some wrong (or outlier) motion vectors. Figure 4.17 describes the steps that define this proposed scheme.

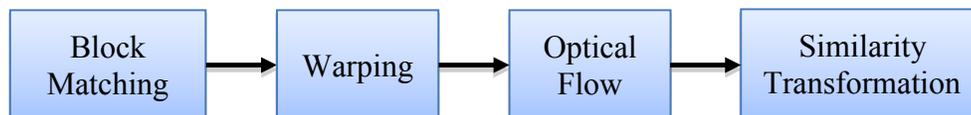
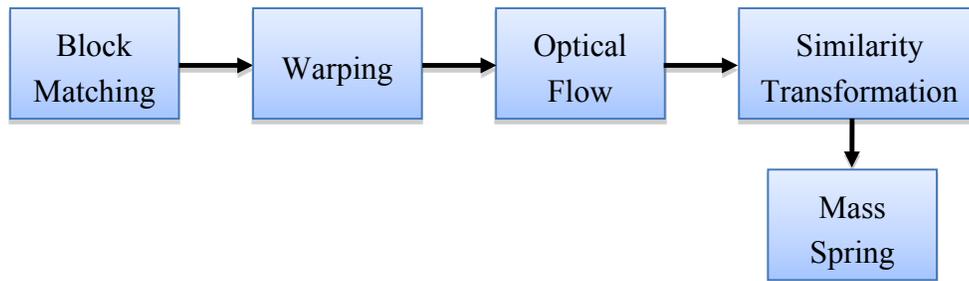
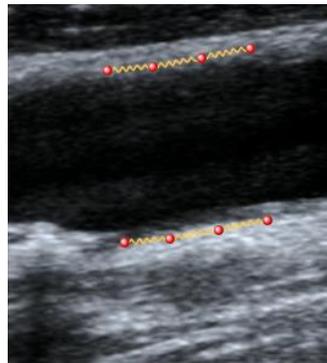


Figure 4.17. Similarity transformation given the motion vectors obtained with the hybrid BM-optical flow method.

When the similarity transformation is applied, it produces an error in the position transformation. This artifact happened because the method considers the systolic-diastolic-systolic transition as a scale-translation transformation over time. To reduce this error, rather than using the statistical shape deformation model of Equation (4.42), a physically-based simulation method (Soft Body, Section A.7) could be included to the upper and lower set of points (individually) that track the artery, as indicated in Figure 4.18a. The pipeline of this approach is illustrated in Figure 4.18b.



(a)



(b)

Figure 4.18. (a) Pipeline of the proposed method with physics simulation (mass-spring) and (b) Illustration of the spring connections in an ultrasound image.

Finally, the KF (Kalman Filter, Section A.5) will be incorporated into this evaluation. This filter was previously used by Gastouniotti *et al.* [GAS10] to increase the accuracy of tracking the wall artery tracking. In this particular case, the KF will be incorporated into different proposed schemes as described by Figure 4.19 with the objective of being evaluated a posteriori in Section 4.3.5 as well as the other described schemes. It is important to note that in our evaluations a simple updating scheme was

added to update the reference block (the initial tracked block) in each frame of the sequence, as shown in Equation (4.52):

$$B_{new_ref} = \alpha \cdot B_{old_ref} + (1 - \alpha) \cdot B_{estimated} \quad (4.52)$$

where B_{new_ref} is the new estimated reference block, B_{old_ref} the old referenced block, $B_{estimated}$ is the estimated displaced block, and α is the parameter that controls the amount of information that remains from the old reference block; we empirically set α to 0.98.

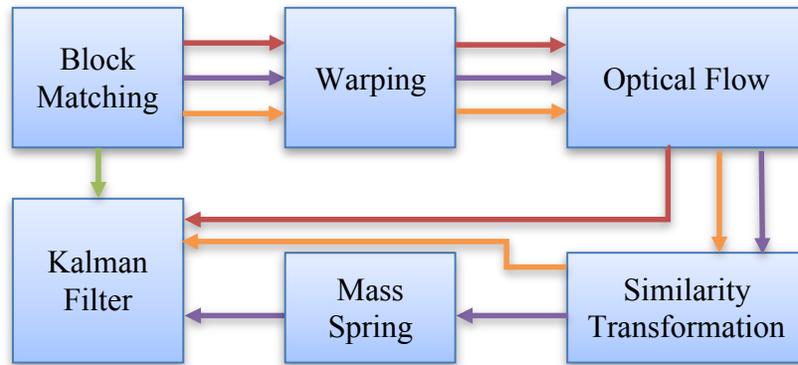


Figure 4.19. Incorporation of Kalman Filter in the previously proposed schemes. Each scheme is encoded with a different color.

4.3.5 System Evaluation and Results

As described in Section 4.3.3.B, we generated a set of nine sequences that simulate the wall artery motion with different amplitudes and topologies. The displacement of the wall of the artery, in our evaluation set, varies from 2.50 up to 9 pixels. To evaluate the proposed methods in Section 4.3.4, a similar metric as the one used by Golemati *et al.* [GOL12] will be used. In our evaluation, each sequence contained 84 frames and involved three cardiac cycles. To obtain an impartial evaluation, six different positions (a total of 54 evaluations per method) will be evaluated in each sequence and the Cartesian coordinate error Equations (4.53)–(4.55) and the diameter error over the time Equation (4.56) will be measured by mean of root-mean-squared error (RMSE):

$$\bar{\epsilon}_{long} = \sqrt{\frac{1}{N \cdot M} \sum_{j=0}^{M-1} \sum_{i=0}^{N-1} (x(i, j) - x'(i, j))^2} \quad (4.53)$$

$$\bar{\epsilon}_{rad} = \sqrt{\frac{1}{N \cdot M} \sum_{j=0}^{M-1} \sum_{i=0}^{N-1} (y(i, j) - y'(i, j))^2} \quad (4.54)$$

$$\bar{\epsilon}_p = \sqrt{\frac{1}{N \cdot M} \sum_{j=0}^{M-1} \sum_{i=0}^{N-1} (\|p(i, j) - p'(i, j)\|^2)} \quad (4.55)$$

$$\bar{\epsilon}_d = \sqrt{\frac{1}{\frac{N}{2} \cdot M} \sum_{j=0}^{M-1} \sum_{i=0}^{\frac{N}{2}-1} (d(i, j) - d'(i, j))^2} \quad (4.56)$$

where $x(i, j)$, $y(i, j)$ and $p(i, j)$ are the i -th x , y and (x, y) coordinates in the j -th frame of the sequence with its respective ground truth $x'(i, j)$, $y'(i, j)$, $p'(i, j)$ and $d(i, j)$ is the estimated diameter of the artery and its ground truth $d'(i, j)$. N is the number of points, in our case, six points, and M is the number of frames per sequence. Our evaluation uses five motion models: Lucas and Kanade, Anisotropic Huber L1, BM (Lewis method), BM + Lucas & Kanade, and BM + Anisotropic Huber TV-L1. For each motion model, the methods described in Section 4.3.4 will be used. These methods include the similarity transformation (ST), the Kalman filter (KF), and the mass-spring (MS) physics based model. The motion methods will also be evaluated individually as shown in Table 4.3.

Table 4.3. Evaluated models where the used methods are indicated. Not all the combinations have been used because some of them were nonsensical.

Methods	ST	KF	MS
M1			
M2		x	
M3	x		
M4	x	x	
M5	x		x
M6	x	x	x

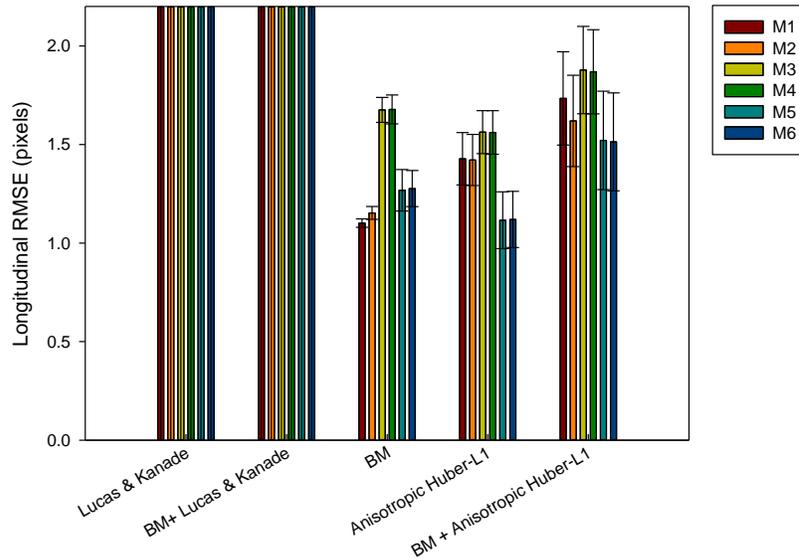


Figure 4.20. Longitudinal error results after being evaluated with different methods.

At first, the estimated longitudinal error will be evaluated by means of Equation (4.53), in our evaluation set (Section 4.3.3.B). As can be appreciated in Figure 4.20, the methods that make use of the Lucas and Kanade algorithm generate the highest number of errors. This is due to the fact that this method does not handle properly the aperture problem compared to the other motion evaluated methods. M1 with block matching and M5 method with anisotropic TV-L1 motion estimation produce the best results in the longitudinal motion estimation. In some plots the errors values are above the maximum values of the plot. We have reduced the plot range to better discriminate among the other approaches.

In the radial motion evaluation (Figure 4.21) computed by means of Equation (4.54), the methods with Lucas and Kanade obtain again worse results. But unlike the longitudinal evaluation, the mix of block matching and anisotropic TV-L1 obtain the best results and a lower deviation with respect to the best results of the block matching approaches, achieving almost 50% less error. It is important to remark that in both evaluations (longitudinal and radial), the inclusion of the mass-spring method (M5–M6) helps, in general, to reduce the position error.

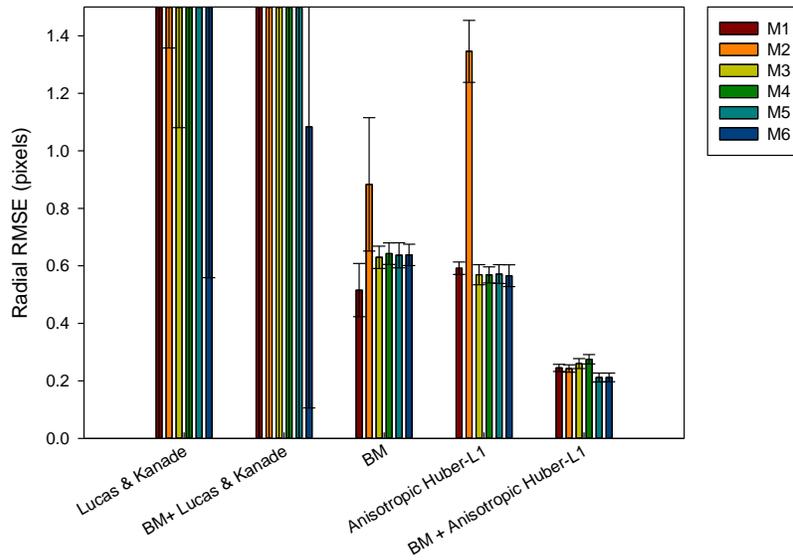


Figure 4.21. Radial error results after being evaluated with different methods.

A global position evaluation (Figure 4.22) is carried out by Equation (4.55), where the radial and longitudinal displacement is taken into account. In general, the block matching (M1 version) and anisotropic TV-L1 methods achieve the best results with the difference that BM generates less standard deviation in the error among different evaluated sequences.

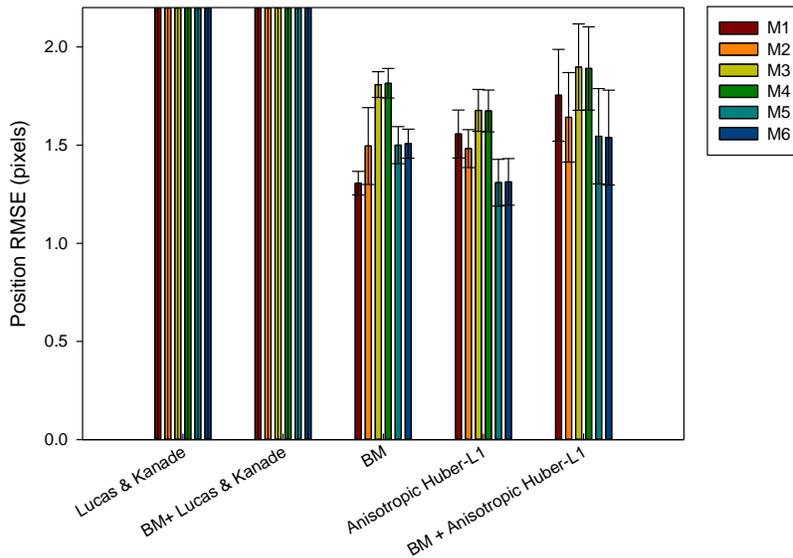


Figure 4.22. Position error results after being evaluated with different methods.

It is important to obtain a method capable of achieving good position precision. In our case, it is not critical to estimate the evolution of the diameter of the artery over time with a model that generates a small deviation on the elastomer's position (± 1 pixel). To evaluate the most significant parameter, the diameter of the elastomer over time, estimations obtained from Equation (4.56) will be evaluated in our assessment set (Figure 4.23). The results reveal that the best method is the combination of block matching and anisotropic TV-L1 (M5). If we compare it with the previous results that evaluate the position error, the best solution (BM-M1) generates 3.1 times higher error than the new best solution obtained. The combination of optical flow and block matching obtain almost two times higher precision than the methods working individually.

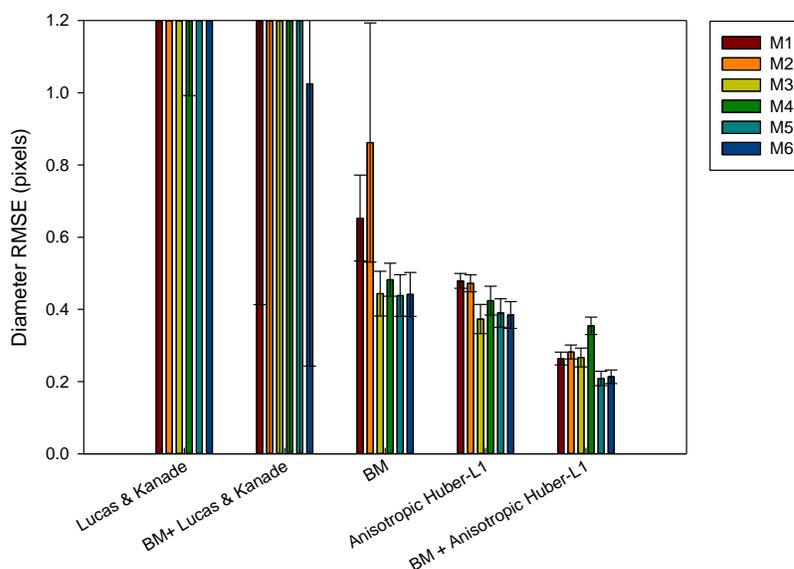


Figure 4.23. Diameter error results after being evaluated with different methods.

After evaluating the different described approaches, it can be observed that the inclusion of the Kalman filter (M2, M4 and M6) provides worse results with respect to the other approaches. The methods that include the similarity transformation reduce the error up to 14% in some proposed approaches. After incorporating the soft body model, an increase of 13% in the precision is obtained in the last method (BM and anisotropic TV-L1). Finally, it can be concluded that a combination of optical flow and block matching and the M5 scheme becomes the most precise technique to estimate the desired parameter.

To show a more detailed evaluation between the best method and other results, a Bland-Altman plot is produced. Figure 4.24 shows the difference between the best

methods with other evaluated methods and their averages. The middle line indicates the average difference of both methods, whereas the upper and lower lines represent 95% limits of agreement with 15.80% (Figure 4.24a), 49.78% (Figure 4.24b), 33.66% (Figure 4.24c), and 78.39% (Figure 4.24d) of window (defined by 1.96 times the standard deviation with respect to the mean difference) displacement with respect to the origin coordinate. It can be concluded from Figure 4.24 that there is an overall good agreement of the amplitudes between the BM and anisotropic TV-L1 (M5) method and the reference ones. To facilitate the evaluation, further tabulated results are listed in the Appendix A. In the next section, the obtained results will be discussed and the method will be evaluated in real cases with different human subjects *in vivo*, with the objective of validating this technique.

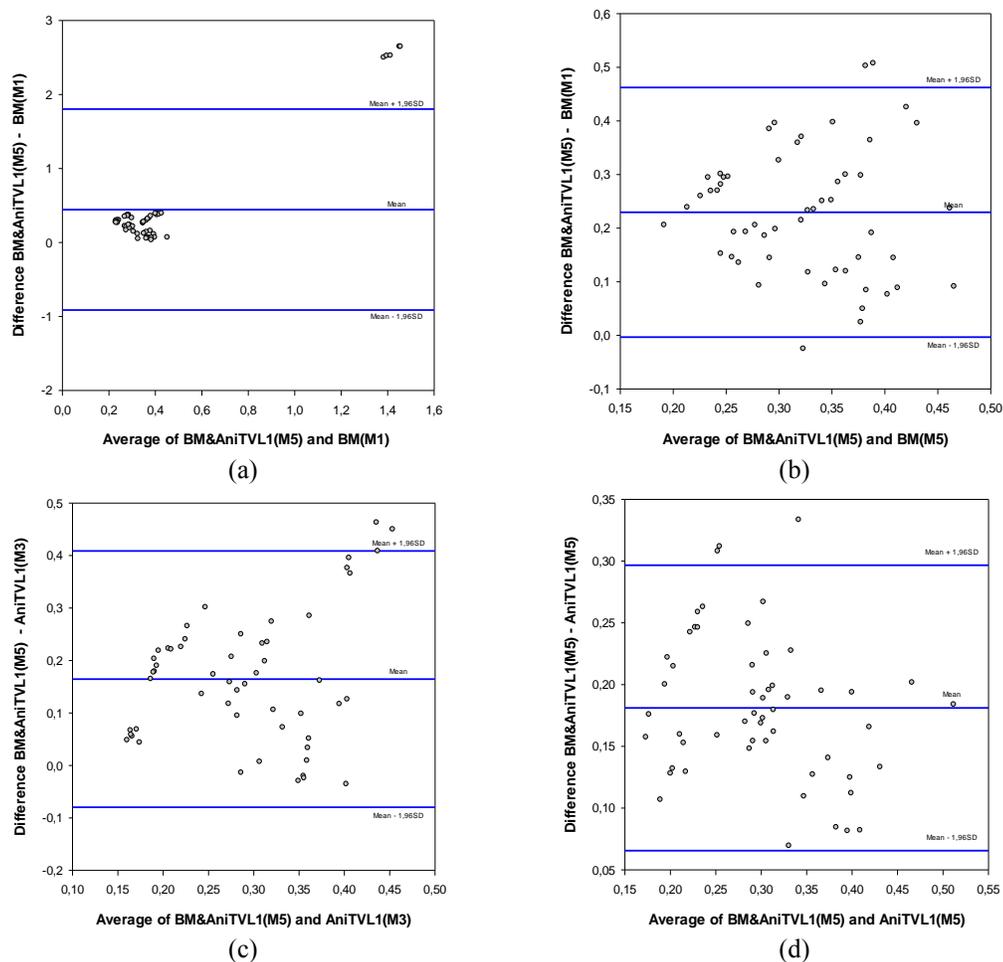


Figure 4.24. Bland-Altman figure, where the best obtained results block matching & anisotropic TV-L1 (M5) are compared with (a) block matching (M1), (b) block matching (M5), (c) anisotropic TV-L1 (M3), and (d) anisotropic TV-L1 (M5).

4.3.6 Discussion

In this section, the proposed approaches and the impact of each method in different evaluations will be discussed. Figures 4.25 and 4.26 show the described evaluations in two different sequences, in order to illustrate the response of the methods under different motion variations.

Figure 4.25 corresponds to a simulation that produces a maximum displacement of 18 pixels (the variation in displacement between the upper and lower arterial wall), while Figure 4.26 generates a maximum wall displacement of 5 pixels. As was discussed in the previous section, the Lucas and Kanade method is not the most appropriate algorithm to register the motion in US images as illustrated in Figures 4.25a,b and 4.26a,b. When displacement vectors are long, the optical flow based techniques are not the most convenient ones, because these methods have a maximum limit to determine the motion vector (thus improving the working range would require multi-scale schemes such as [BAR12]) as illustrated in Figure 4.25d.

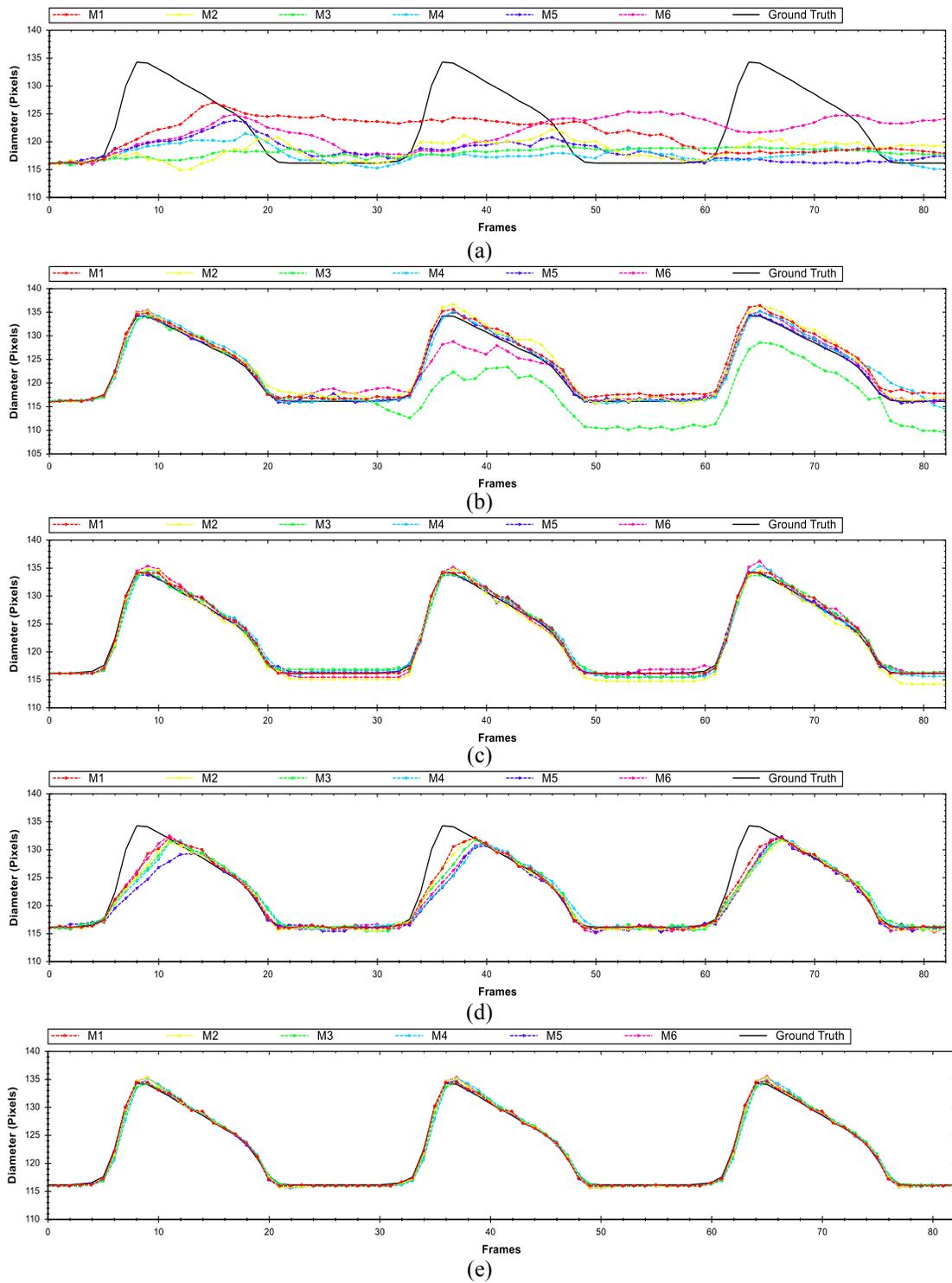


Figure 4.25. Diameter evolution over time evaluated in a sequence with long displacements using (a) Lucas-Kanade, (b) block matching & Lucas-Kanade, (c) block matching, (d) anisotropic TV-L1, and (e) block matching & anisotropic TV-L1.

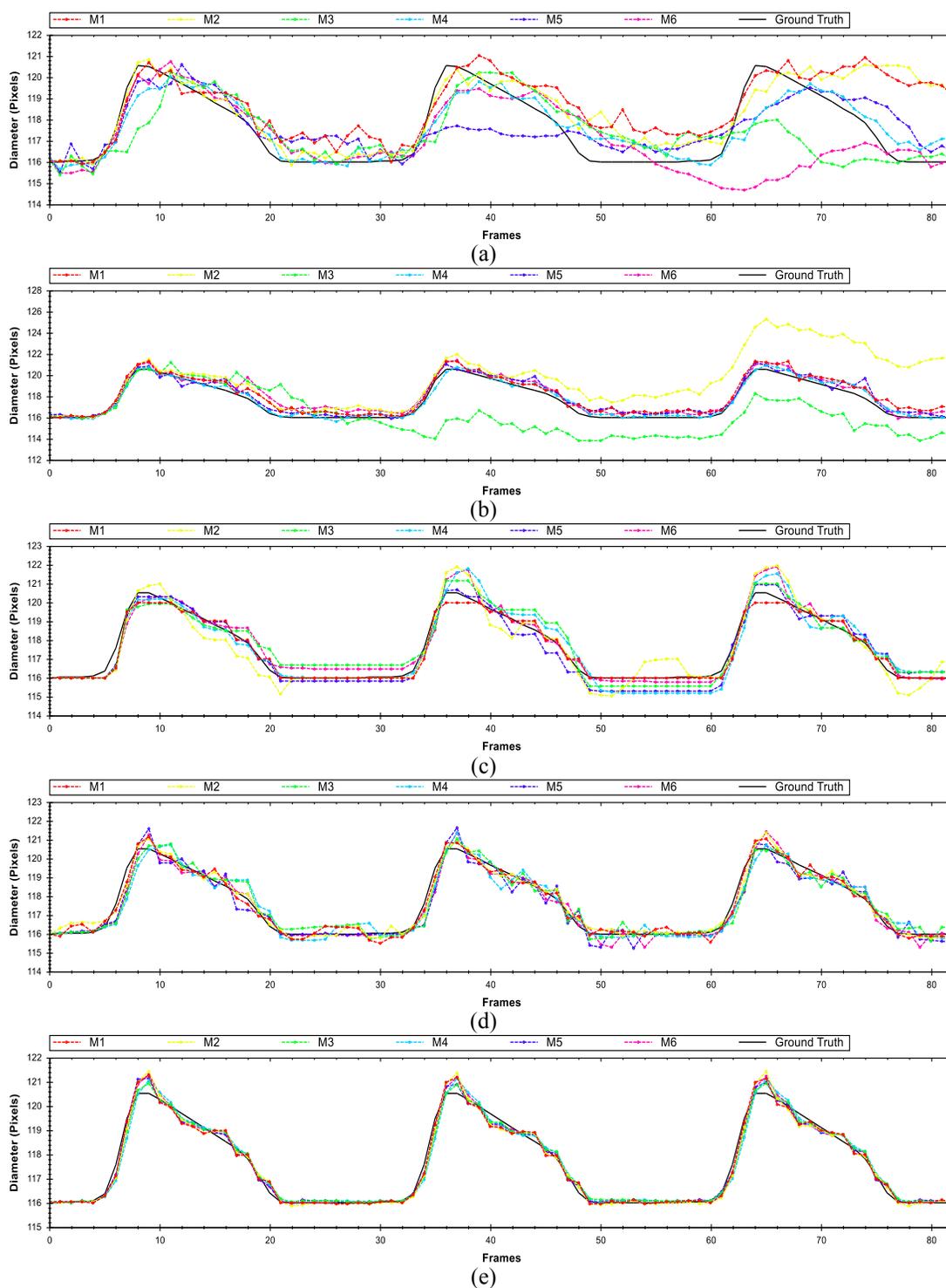


Figure 14.26. Diameter evolution over time evaluated in a sequence with small displacements evaluated with (a) Lucas-Kanade, (b) Block Matching & Lucas-Kanade, (c) Block Matching, (d) Anisotropic TV-L1, and (e) Block Matching & Anisotropic TV-L1.

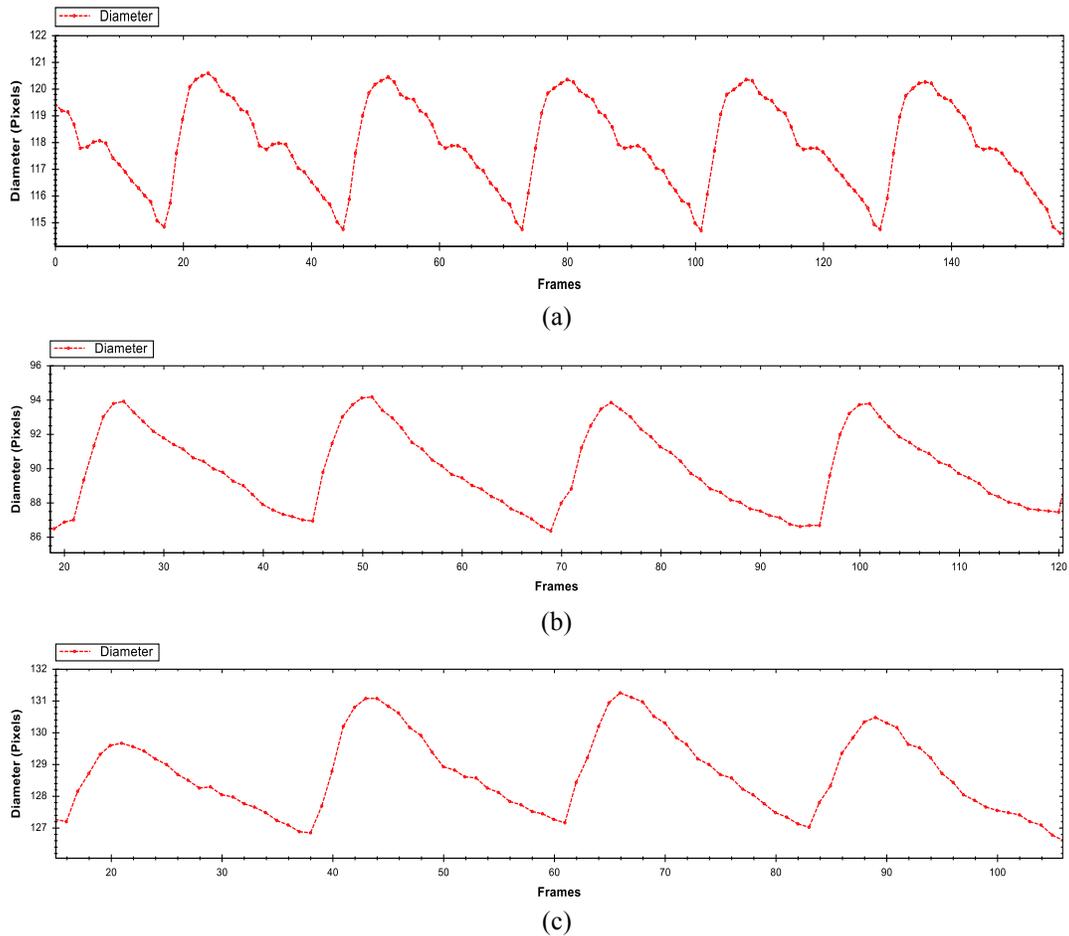


Figure 4.27. Diameter evolution of the common carotid artery (CCA) in real ultrasound data in different subjects where first row (a) correspond to a healthy patient and the last two rows (b–c) belong to patients with presence of atheroma plaques.

Moreover, Block Matching techniques do not have this limitation, but produce rough results and allow no sub-pixel precision (Figures 4.25c and 4.26c). A good solution is the combination of correlation and optical flow techniques to avoid this displacement limitation and obtain sub-pixel precision (Figures 4.25c and 4.26c), acquiring twice more precision than with only block matching method (approx. 0.25 pixel error), but with the inconvenience of increased computation time. The inclusion of Kalman filters does not significantly increase or decrease the results, but it is interesting to include it in hypothetical cases when the system has a severe disturbance and noise. The main problem with the use of this filter is that, depending on the settings of the parameters, the signal may be over smoothed and shifted in relation with the desired one.

At this point, the proposed methods have been evaluated on synthetic US B-Mode imaging. To demonstrate and validate that the best approach is able to work in real sequences, it is evaluated in different subjects in the common carotid artery as shown in Figure 4.27. In this brief evaluation, a healthy patient (Figure 4.27a) was involved, whose diameter motion curves showed clearly the dicrotic peak (attributed to the elastic recoil of the aorta and aortic valve) while the other two patients (with presence of atheroma plaque) the dicrotic peak is absent (Figure 4.27b,c).

4.3.7 Conclusions

The objective of this work was the design, evaluation and comparison of methods able to characterize arterial wall motion. A set of methods has been evaluated with the objective of determining which approach better handles our problem, the estimation of the diameter of the artery over time. The motion methods were selected according to the obtained results in other works, with the goal of comparing our approach against these other methods and evaluating its accuracy. It has been demonstrated that our described combination of methods based on similarity transformation, non-rigid deformations, statistical filtering, and hybrid motion estimation techniques enhance existing state of the art approaches, up to 2.5 times more accurate than state of art techniques.

Synthetic US sequences with different patterns of motion were generated to allow quantitative comparative analysis of different methods and combination of techniques. Our experiments involve a total of 1620 evaluations within nine simulated sequences of 84 frames each and four error metrics. In fact, the assessment that appropriate integration of different techniques has a clear impact on the final performance represents an important contribution of this work. Another advantage that must be remarked is that the proposed methods supports large displacement vectors unlike optical flow techniques that are limited in working range and require multi-scale schemes.

5. INDUSTRIAL INSPECTION

5.1 INTRODUCCION

Since the industrial evolution promoted by General Motors in the 60s, with the introduction of Programmable Logic Controllers (PLCs), the industry stacked in terms of new technology innovations. At the beginning of the 80s, a new technology started to be introduced slowly and without being noticed in the industry. Such new technology is nowadays named machine vision. Those systems are able to replace the human eye (with a low price investment) providing robustness, reliability, and fast processing. For this reason, the industry has been one of the sectors where this technology has been widely accepted. In industry, quality control has become an essential factor where it is required robust, fast and reliable systems. Currently there are many applications, making use of machine vision, focused in quality control in order to detect product defects, surface imperfections, detection of non-installed parts, etc.



Figure 5.1. Inspection system to detect imperfections of the products on the conveyor belt. (Courtesy of *Cognex Corp.*).

In the early stage of this technology, the performed tasks were very basic due to the bottleneck in the data processing performance and therefore the developed tasks were quite rudimentary. Nowadays, thanks to the advance of the technology, it has been possible to develop visual inspection systems able to realize complex machine vision tasks and hence, to improve the quality of the inspection in the line production as shows Figure 5.1. These advances are composed of hybrid architecture that includes multi-core processing, DSP (Digital Signal Processor) and GPGPU (General-Purpose Computing on Graphics Processing Units) with the objective of providing high image processing performance.

5.2 QUALITY INSPECTION

5.2.1 Motivation

In this work, a machine vision system able to inspect biscuits in the line production, as illustrates Figure 5.2, to detect those ones that present fractures and reject them from the transport belt is presented. The proposed method consists of a collection of steps that involves biscuit segmentation, crack detection, feature extraction, and classification. Each defined step is also evaluated in deep to determinate its accuracy and demonstrates that the described approach is efficient at solving the problem. In order to speed up the computation time, our approach is implemented on a GPU architecture obtaining a higher processing speed in comparison with a CPU implementation. The great challenge of using GPU architectures is the adaptation of sequential methods to be executed in parallel. Therefore, this paper is also focused on using the appropriate methods that better fit this parallel architecture in order to solve this problem.



Figure 5.2. A Maria© biscuit on the top of the custard in the production line (Courtesy of Dhul© Company).

The chapter is organized as follows: Section 5.2.2 introduces a detailed description of the proposed model to detect and extract the features that define the fractures in a biscuit; also, the classification stage to determinate the quality of the product is described; in Section 5.2.3, the described method is evaluated in depth in order to determine its accuracy and measures of the computation cost in different hardware architectures are also provided; and finally, Section 5.2.4 summarizes some conclusions and outlines for future works.

5.2.2 Background Material and Related Work

Currently, there are many applications making use of machine vision systems in the quality control and optimization area. These applications are focused on detecting defects in products, surface imperfections, incorrect installed parts in the assembly, etc. In the textile quality control area, Mark et al [MAK06] make use of a frame grabber (Matrox) to detect imperfections in woven fabrics using a combination of Gabor wavelet network and morphological operations. Other types of techniques, based on 3D reconstructions, are also used to inspect the surface. One example is the one exposed by Abbott et al [ABB00], where they handle tridimensional reconstruction techniques, by laser, to estimate the optimal cutting position of wood studying the surface. Other topics related with surface inspection are present, e.g., the exposed solution by Vitrià et al [VIT07]. The proposed system is able to inspect the surface of the cork stoppers and classify them into different categories in order to determine the quality of the cork. The features that describe the quality are based on the presented crack in the surface. The author segments the image by means of an adaptive threshold technique, analyzes the blob features (area, length, perimeter, convex perimeter, compactness, roughness, elongation, etc.) and catalogs the cork stopper by means of a Bayesian classification.

Machine vision systems are also present in the food industry, for example the ad-hoc architecture, based on a FPGA, designed by Peterson [PEA09]. This high-speed inspection system determines the quality of grains through the study of valleys in the intensity signal. Making use of the same platform, Haff et al [HAF10] create a high-speed pistachio classifier analyzing the color histogram. Another example is the work described by Riquelme et al [SUN00], where olives are classified in order to determine their quality by means of color segmentation and morphological features. Also, more complicated quality inspections have been carried out, as is the case described by Sun [SUN00, SUN03]. This inspection system has the objective of analyzing the topping percentage and dispersion of the ingredients on pizzas by means of segmentation techniques and statistical distributions over the pizza base. Other example is the work proposed by Blasco et al. [BLA09] where designed a complete vision system to sort automatically the pomegranate in the conveyor belt by means of statistical color information. In relation with the presented work we must cite the exposed work of Senni et al. [SEN14], in this article the author describes a computer vision system based on infrared cameras in order to detect foreign object in biscuits using template matching. With the purpose of extending more information about this food inspection, we suggest

to the reader this collection of books [SUN12, SUN11, GRA03] where a wide range of visual food inspection subjects can be obtained.

After the brief overview about the importance and usefulness of machine vision in the industry, we will introduce a specific problem and demonstrate how to solve it with the objective of showing how useful are the architectures based on massive parallel processing (GPU) and comparing our approach with classical parallel and sequential (CPU) processing.

5.2.3 Material and Methods

Our system consists of a CMOS camera uEye 1240ML, lens of 8 mm of focal distance and a distance from the camera to the biscuit of 25.5 mm. The environment where was taken the dataset samples was illuminated with halogen lamps (5500° Kelvin) giving 1206 Lux. The processing unit utilized to process the images was an NVidia© (580 GTX) GPU. Nowadays, GPUs are an essential hardware tool in computer vision due to the high computational performance. The main problem of efficiently making use of this kind of architectures is taking advantage of their inherent parallel processing. On one hand, they can lead to outstanding performance in algorithms with a parallel nature. But in the case of sequential algorithms, they need to be parallelized (if possible at all) in order to take any advantage. This problem of taking full advantage of parallel processing architectures has become a hot topic in the computer vision research field. One of main motivations of this work was to demonstrate the possibility of making use of parallel processing of algorithms that have a sequential nature.

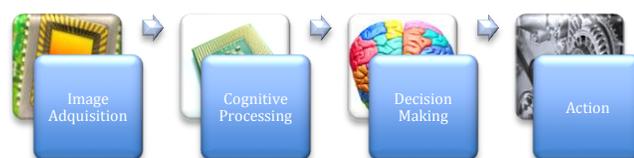


Figure 5.3. Architecture processing illustration used in this work.

In this paper, an industrial problem will be solved with a parallel architecture following the scheme shown in Figure 5.3. This process involves a sensor to capture the image that will be processed in the GPU to extract features and determine, in another stage, whether the biscuit passes the quality control or not. Once, the discarding decision has been taken, the product needs to be dropped out from the production line, the pertinent actions need to be carried out, i.e., sending a signal to the PLC to reject the product.

This work will be focused in the “cognitive processing” and the decision making stage as illustrated in Figure 5.4. The first step is focused on the biscuit segmentation in order to mark the area of interest and to remove information that does not belong to it. The second step involves the detection of possible fractures in the biscuits. After enhancing the fracture related features, the third step will be extracting the features in the area that has been previously segmented and finally, to use such information to determine the quality of the product by means of a classifier.

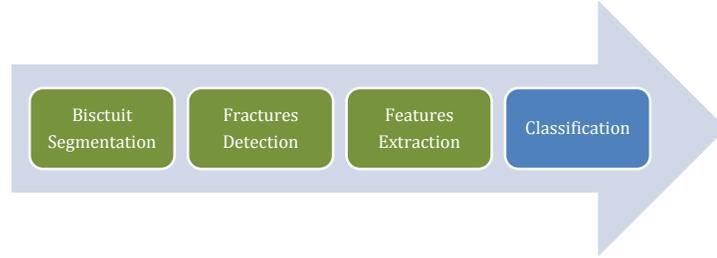


Figure 5.4. Data flow scheme showing the whole process through different steps.

A. BISCUIT SEGMENTATION

The first step in the presented scheme is to perform the biscuit segmentation inside of the recipient in order to drop out external non-related information that can mislead the feature extraction step. To carry out this step, color segmentation techniques will be used. This topic has been very common in computer vision, for instance, in skin segmentation to detect faces [VEZ03] and up to robotic applications segmenting desired objects such as a ball [WAS02]. In the color segmentation field, a huge variety of methods to classify the desired color can be found; in this case, we aim for a method inherently parallel to better suit the GPU processing architecture. Non-parametric methods become fast in the classification task since they can be built using a LUT (Look Up Table) with the density of probabilities. This makes these approaches very attractive in computational time terms and GPU-friendly. A fast approach to estimate the density function is the one introduced by Parzen [PAR62] that allows the use of different kernels. The method described here will make use of the most common Parzen kernels [DEV82]: Hyper Sphere (5.1), Hyper Cube (5.2), and Gaussian (5.3),

$$P(X|w_i) = \frac{1}{N_i} \cdot \sum_{m=1}^{N_i} \begin{cases} v_p^{-1} & \text{if } \{X|\delta_E(X, Z_i^m) \leq \rho_i\} \\ 0 & \text{if } \{X|\delta_E(X, Z_i^m) > \rho_i\} \end{cases} \quad (5.1)$$

$$P(X|w_i) = \frac{1}{N_i} \cdot \sum_{m=1}^{N_i} \begin{cases} (2\rho_i)^{-d} & \text{if } \{X|\delta_T(X, Z_i^m) \leq \rho_i\} \\ 0 & \text{if } \{X|\delta_T(X, Z_i^m) > \rho_i\} \end{cases} \quad (5.2)$$

$$P(X|w_i) = \frac{1}{N_i \cdot \prod_{i=1}^d \rho_i \cdot (2\pi)^{\frac{d}{2}}} \sum_{m=1}^{N_i} \prod_{i=1}^d \exp \left\{ -\frac{1}{2} \left(\frac{X_i - Z_i^m}{\rho_i} \right)^2 \right\} \quad (5.3)$$

where N_i is the number of prototypes, v_p is the volume of the sphere, $\delta_E(X, z_i^m)$ denotes the Euclidean distance, ρ_i is the *Parzen* window width, $\delta_T(X, z_i^m)$ is the Chebyshev distance, and d the number of dimensions.

To improve the quality of color segmentation, a modified ROF model [RUD92] will be used but with the incorporation of the weighted function as shown in (5.4), to preserve diffusion over the edges. In our case, we use a similar weighted function as in Unger et al [UNG08], but with the extension of multiple channels (5.5). To speed up this model, it has been decided to adopt a similar approach as the one introduced by Chambolle [CHA04], with the inclusion of an adaptive diffusion scheme [ZHU08], where formulates the total variation minimization as a primal dual algorithm.

$$\int_{\Omega} |\nabla u| g(x) + \frac{\lambda}{2} (u - f)^2 d\Omega \quad (5.4)$$

$$g(x) = e^{-\beta \left(\frac{\sum_{i=1}^{NC} |\nabla u_i|}{NC} \right)^\alpha} \quad (5.5)$$

In Equation (5.4), u is the desired cleaned image, f the original image, and $\lambda \in [0,1]$ the regulation parameter, while in Equation (5.5), α and β denote parameters that define the weighted function, NC denotes the number of channels (in our case $NC = 3$), and $|\nabla u_i|$ is the norm of channel i . After applying Equation (5.4) to the image, the result is a denoised image as illustrated by Figure 5.4.

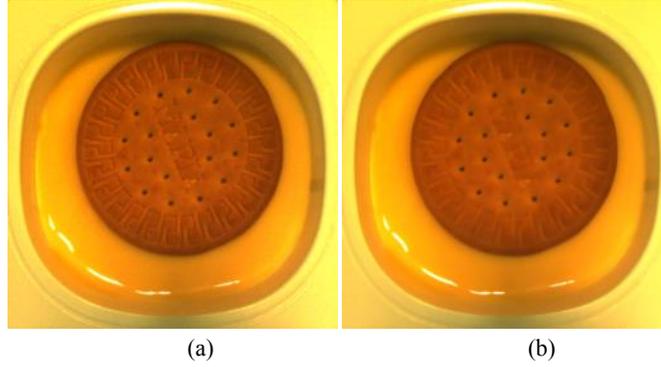


Figure 5.4. (a) Original image. (b) Clean image after applying Equation (4) with $\alpha = 0.58$, $\beta = 0.55$, $\lambda = 0.9$.

Once the image has been denoised, the next step in the introduced method is the segmentation. To carry out the binary segmentation, we will use the method by Under et al [UNG08] with a modification, in order to include the anisotropic diffusion [WER09] as indicated in Equation (5.6),

$$\min_{u,v} \sup_{|\vec{p}| \leq 1} \left\{ \int_{\Omega} \left(D^{1/2} \nabla u \right) \cdot \vec{p} - \varepsilon \frac{|\vec{p}|^2}{2} + \frac{1}{2\theta} (u - v)^2 + |v - f| d\Omega \right\} \quad (5.6)$$

$$f = -\log(\text{prob}(\text{ForeGround})) + \log(\text{prob}(\text{BackGround})) \quad (5.7)$$

where $D^{1/2} = \exp(-\alpha|\nabla I|^\beta) \vec{n} \vec{n}^T + \vec{n}^+ \vec{n}^{+T}$, $\vec{n} = \frac{\nabla I}{|\nabla I|}$ as the normal vector, \vec{n}^+ the tangent vector of a given point, u is the segmented image, v an intermediate variable, \vec{p} is the re-projection vector, and f the logarithmic likelihood of the background (tub and cream) and foreground (the biscuit). In order to evaluate the submitted segmentation method and to determine the more appropriate color space, in Section 5.4 we will evaluate the methods in three different color spaces RGB, Spherical (Fig 5.5.a), and HSV (Fig 5.5.b). For more information about HSV and Spherical color spaces, we have included the equations of these transformations in Appendix A.

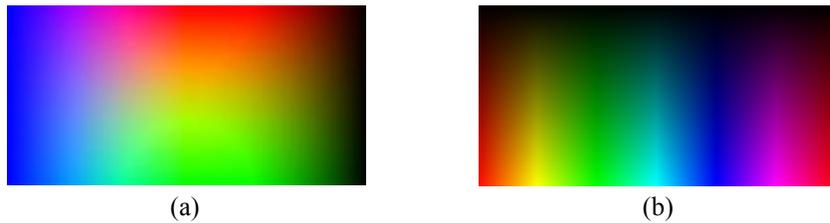


Figure 5.5. Illustration of spherical color representation $\theta \in \left[\frac{3\pi}{2}, \frac{\pi}{2} \right]$ respect $\phi \in \left[0, \frac{\pi}{2} \right]$ and $\rho = 300$ and HSV color representation $H \in [0,1]$ and $S \in [0,1]$.

The advantage of using Spherical color space is that the visible color spectrum can be represented in two dimensions rejecting the intensity. This simplification allows us to pre-build a two dimension Look-Up-Table (LUT), where the input of the LUT is defined by the color space coordinates in the matrix (LUT) and the output is the density of probabilities that define if a color is considered as Background (Tub and Cream) or Foreground (Biscuit). Making use of the LUT, let us save in memory reservation in the GPU and computation time to estimate the likelihoods. Scaling the "2D-LUT" in a 360x360 floating point matrix, the memory used is 0.49 MB. In the case of making use of a 3D-LUT of 360x360x360, the total used memory rises up to 177.97 MB, hence the 2D-LUT version obtains a reduction factor of 363.20. With the purpose of generating a multimodal probabilistic map, the Mean shift [COM02] method is used to cluster the selected color and generate its likelihoods as illustrated by Figure 5.6, where it an example of the generated 2D-LUT (Foreground and Background) is shown.

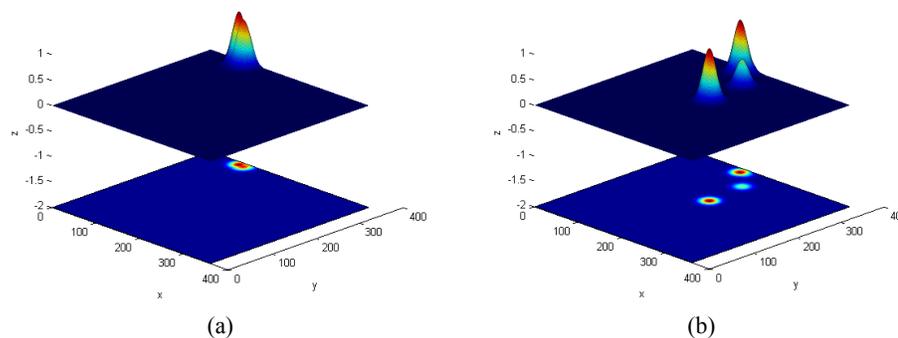


Figure 5.6. (a) Foreground and (b) Background density of probabilities in the Spherical color domain generated with Gaussian Parzen Kernel.

The presence of holes or fractures in the biscuit (with low intensity) can produce mistakes in the segmentation as illustrated in Figure 5.7. This motivates the use of morphological operations to fill those holes, but this operation is not trivial on the GPU. Removing sparse points in a binary image is an easy task, since it can be done using the Opening morph operant (Erosion and Dilatation), but filling the holes is not trivial on the GPU. In the sequential programming field, Robison and Whelan [ROB04] proposed a fast method, considered one of the fastest methods that exist nowadays. On the GPU field, we have only found the method introduced by Karas [KAR10], adapting the sequential reconstruction [VINC93] scheme, and the extension of Teodoro et al. [TEO12], that adopts a modification of the previous method to speed up the convergence [VINC93].

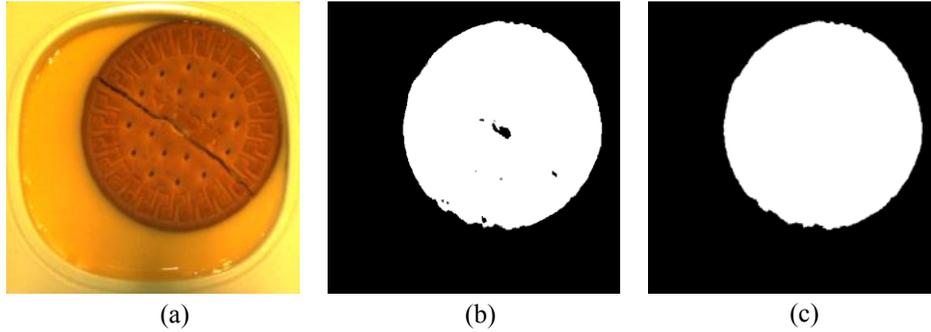


Figure 5.7. (a) Original image, (b) Pre-segmented image using Equation (6) in spherical color domain, (c) Results after applying hole filling method.

In our case, we are working in the binary domain and those algorithms are designed to work in gray scale and therefore, they are not suitable to solve our problem. To carry out the hole-filling algorithm on the GPU, the utilized approach will be based on the CCL (Connected-Component Labeling) method [KAL11]. One of the main reasons for this is because it does not need any kind of atomic access. The first step of the method consists in realizing a bitwise “not” operation into the binary image and applying the CCL algorithm, following the strategy proposed in [KAL11]. The second step consists on collecting the blob indexes belonging to the boundary of the binary image and marking them as 1 into a LUT vector, with the same size of the image and previously initialized to zero, at the position determined by the blob index. The third step involves the use of the LUT vector to be consulted while the exposed method scans the previous processed labeled image to write into another image the value inside the LUT vector; e.g., if a label is read and contains a value equal to 5, the 5th position in the LUT will be read and the pertinent value is written into the output image. The last step of the explained algorithm consists of the application of a bitwise “not” operation to the image obtained in the previous step. To better understand the process of the method, Figure 5.8 shows a short example of the different steps of the algorithm.

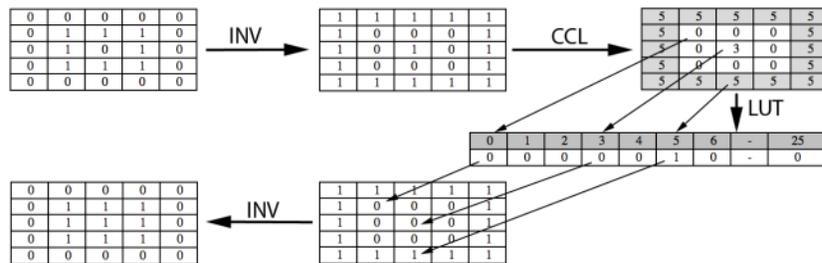


Figure 5.8. Illustration of the proposed method for hole-filling in a binary image.

Once the biscuit has been segmented, with the goal of selecting relevant information from the studied object, suitable features will be extracted in order to classify the biscuit following a quality criterion. In our case, such quality criteria consist in the rejection of broken biscuits.

B. FEATURE EXTRACTION AND FRACTURE DETECTION

Once the desired object in the image (in our case the biscuit) has been segmented, the objective is to enhance and detect potential cracks (fractures) in the captured image. The detection of fractures has been widely studied in the machine vision field due to the importance of creating an automatic and autonomous system able to replace the human eye. As a brief overview about this topic, we can refer to methods such as the introduced by Subirat [SUB06], where 2D wavelets are used in a multi-scale scheme to locate cracks on the pavement. Sinha and Fieguth [SIN06] make use of a two-step algorithm, local-global, where the first step involves statistical properties (local) of the image to extract candidates and the second step applies cleaning and linking operations (global) to merge the cracks. A general method to detect fractures comes from Han et al [HAN10] that introduced a feature estimation based on the Hough transformation [HOU62, DUD72] of different gradient responses, based on 20x20 pixels blocks, with the incorporation of an SVM [COR95] (Support Vector Machine) classifier to determinate whether a block is a crack or not. In the biscuit crack detection field, we have only found the method exposed by Nashat et al [NAS11, NAS14] in which they extract the potential fractures from the absolute difference of the red and blue channel and apply the Hough transformation to obtain a pyramidal feature extraction, to train and classify with SVM, in order to determinate the areas with fractures.

In the exposed approach, we will make use of the logarithmic inverse (5.8) on the gray scale of the image to enhance the crack candidates. Areas with low intensity are potential zones where the cracks are defined due to the “absorption” of the light. An example of the enhancement of potential cracks is illustrated in Figure 5.9 where it can be observed how the logarithmic inverse of the image enhances better than an inverse operation (i.e. Subtraction of the maximum value in an n-bit image, in our case 255, and the original image).

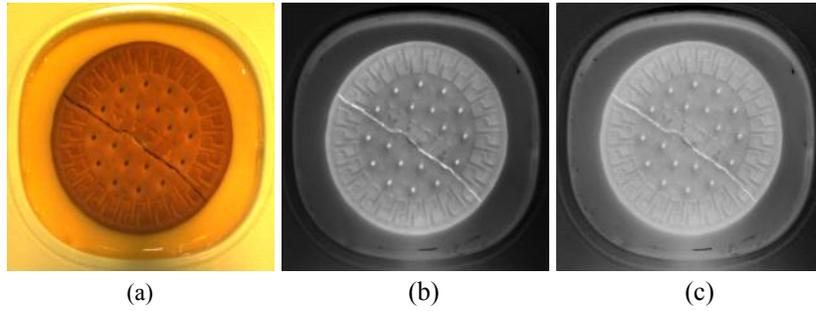


Figure 5.9. (a) Original Image. (b) Logarithmic inverse operation in gray level with $n=100$. (c) Inverse operation in gray level.

$$\text{Log Inv} = n \cdot \log(m) - n \cdot \log(\text{Img} + 1) \quad (5.8)$$

In Equation (5.8), n is the parameter that controls the flexion of the generated curve. To better understand how the n parameter affects to the results in (5.8), Figure 5.10 illustrates the output with different values of n .

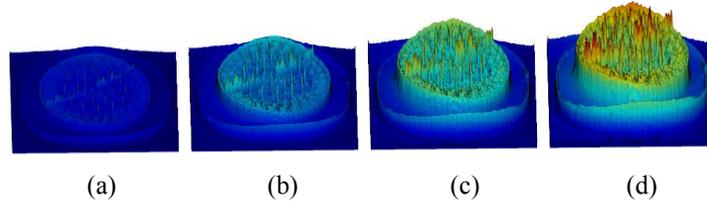


Figure 5.10. Logarithmic inverse with (a) $n= 50$, (b) $n= 100$, (c) $n= 150$, (d) $n= 200$.

After applying Equation (5.8) to the image (Figure 5.11.a), the Top-hat (5.9) morphological operation is carried out to enhance the cracks. This operation has the objective of highlighting the valleys, as shown in Figure 5.11.b, and of removing high responses such as the ones present in biscuit and tub edges. In Equation (5.9), \oplus and \ominus denote dilatation and erosion operations by a structuring element B . At this point, we could have used a threshold but instead of this, a DoG (Difference of Gaussian) filter will be applied before the threshold step to enhance the edge and remove undesired frequencies, acting as a band-pass filter (Figure 5.11.c).

$$\text{Top hat} = (\text{Img} \ominus B) \oplus B - \text{Img} \quad (5.9)$$

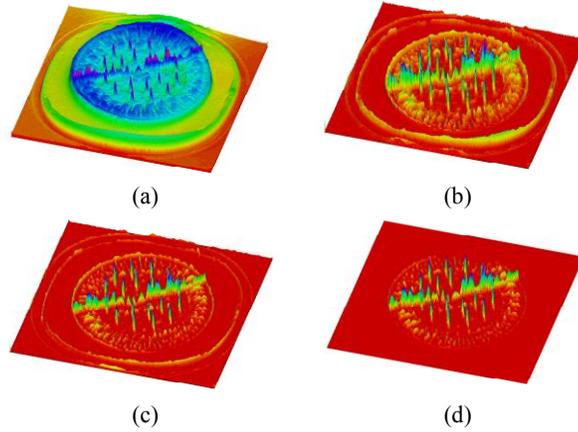


Figure 5.11. (a) Logarithmic inverse operation in gray level with $n=100$ (2). (b) Top hat with B structure of 7×7 , (c) DoG with $\sigma_1=3$ $\sigma_2=6$ and (d) the final result after removing all the non-desired information by means of the output of the previous step (color segmentation).

The Hough transform has been quite popular in order to detect cracks [NAS11, HAN10, SIN06], but the cited methods require a previous thresholding step to work in the binary domain. To avoid such step, our approach will use the bi-dimensional Radon transform [FID85]. This method (5.10) consists in the integral transformation of a function over a set of straight lines L with respect to the arc length t ,

$$R(\theta, s) = \int_{-\infty}^{\infty} f((t \sin \theta + s \cos \theta), (s \sin \theta - t \cos \theta)) dt \quad (5.10)$$

where s is the minimum distance from the line to the origin and θ is the angle formed by the straight line with respect the x-axis. One of the most important drawbacks of this method is the computation cost, in order to avoid this problem our approach adopts the steps of Mendl et al. [UNG08] that introduces an efficient implementation on GPU. In order to extract the features vector that describes the biscuit, we will compute the histogram of the Radon Transformation. Figure 5.12.a-c shows the obtained Radon transform of a correct biscuit and the estimated normalized histogram while Figure 5.12.b-d illustrates the related outputs of a biscuit with the presence of a fracture.

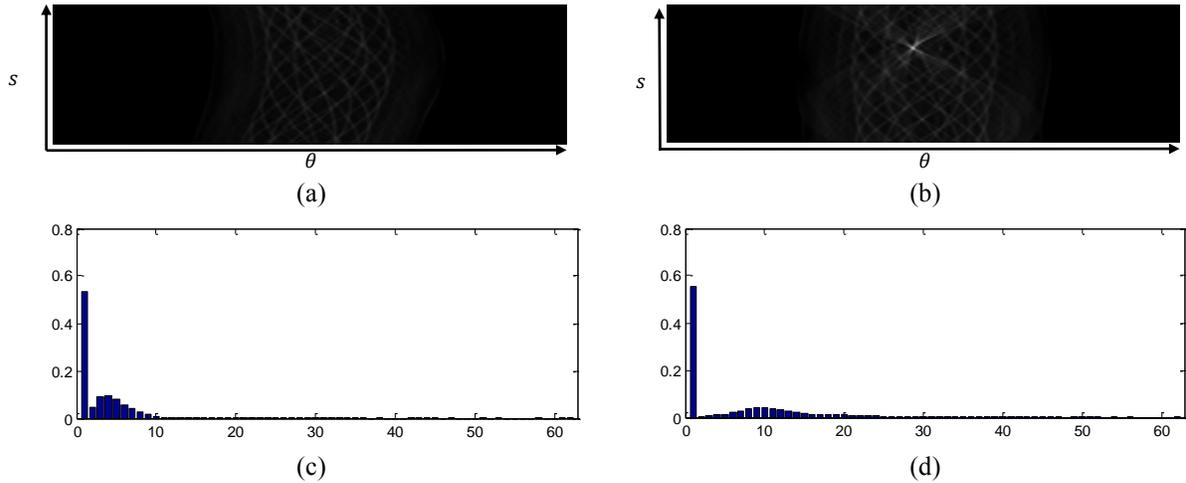


Figure 5.12. Radon Transformation in a biscuit (a) without the presence of cracks or anomalies with its pertinent (c) normalized histogram (64-bins) output and (b) the Radon Transformation of a biscuit with the presence of cracks and its (d) normalized histogram (64-bins).

C. CLASSIFICATION

Support Vector Machine (SVM) [COR95] is a widely used classification technique that has received great attention [YUA05, RUA13] in many different fields. This method consists of a transformation of the data dimension to a higher dimension space by means of a kernel function in order to facilitate the linear separation. Once the transformation to a higher dimension representation space is done, SVM aims to find a hyperplane that separates two classes, minimizing the classification error and maximizing the margin of separation between the hyperplane and the closest points. In order to contrast the obtained results using SVM techniques, we have also included in the evaluation the decision tree method C4.5 [QUI93].

Once we have described the classifiers used in this work, the projection techniques utilized in our evaluation will be detailed. Linear Discriminant Analysis (LDA) [FIS36] has been an important technique in pattern recognition; this method aims to minimize the dispersion of the inter-class distance while maximizing the intra-class one. Also, an extension of the LDA technique that consists of a non-linear generalization of LDA will be evaluated. The cited LDA extension, Kernel Discriminant Analysis (KDA) [MIK99], is based on the kernel trick that is based on the generation of a mapping function that allows the transformation of a given space into another one. Besides, the Principal Component Analysis (PCA) [JOL86] technique has been considered to be

taken into account. In this method, the individuality of the classes is not taken into account and the projection to another space is carried out by means of the Eigen-vectors with higher Eigen-values of the covariance matrix. As described before with the LDA extension based on kernels (KDA), there is also an equivalent to PCA, called KPCA [SCH98]. This method will be included in our evaluation as well.

5.3 SYSTEM EVALUATION AND RESULTS

Section 5.3 described in detail the proposed steps to perform the classification of an undesired product following a quality criterion, in this case fractures, by means of feature extraction in the image. In this section, the exposed method will be evaluated globally and also specifically as regards the different stages, i.e., segmentation, feature extraction, and classification.

A. SEGMENTATION EVALUATION

In order to evaluate the accuracy of the Parzen kernels and the Total Variation method with different color spaces, we will make use of the F-Measure estimator (5.13), as detailed in [BRU11]. To generate the evaluation dataset (ground truth), a collection of five biscuits have been segmented by hand in order to estimate the F-Measure value.

$$recall = \frac{\#correctly\ classified\ foreground\ pixels}{\#foreground\ pixels\ in\ GT} \quad (5.11)$$

$$precision = \frac{\#correctly\ classified\ foreground\ pixels}{\#pixels\ classified\ as\ foreground} \quad (5.12)$$

$$F - Measure = 2 \cdot \frac{recall \cdot precision}{recall + precision} \cdot 100 \quad (5.13)$$

Once different Parzen Kernels (Gaussian, Sphere and Cube) in standalone and making use of the Total Variation method in different color spaces (RGB, HSV and Spherical) have been applied, the obtained results in this evaluation are shown in Table 5.1. It is clear that the inclusion of intensity information in color segmentation is not very appropriate because the methods that split the intensity channel obtain better results. In the case of the Spherical domain, it also leads to good results but avoiding such information (intensity), the accuracy increases. Another point to be remarked is

that the inclusion of the denoising stage leads to improvements that in some cases represent accuracy increases up to 8%.

Table 5.1. F-Measure results in different color spaces and Parzen Kernels. The results marked in bold correspond to the best results obtained in each approach (highest value per column).

Color Spaces	Methods	Parzen %	Denoised Parzen %	Ani-TV-L1 Seg. %	Denoised Ani-TV-L1 Seg %.
RGB	Hyper Cube Kernel	96.28	97.12	97.42	97.51
	Hyper Sphere Kernel	96.70	97.57	97.47	97.79
	Gaussian Kernel	96.72	97.57	97.47	97.90
HSV	Hyper Cube Kernel	87.01	93.87	93.87	95.87
	Hyper Sphere Kernel	88.00	94.10	94.10	95.80
	Gaussian Kernel	88.20	94.10	94.10	95.80
HS	Hyper Cube Kernel	90.42	98.60	98.51	98.60
	Hyper Sphere Kernel	93.70	98.58	98.50	98.58
	Gaussian Kernel	93.71	98.59	98.50	98.59
Spherical (ρ, θ, φ)	Hyper Cube Kernel	96.73	98.07	97.95	98.07
	Hyper Sphere Kernel	96.88	97.98	97.86	97.98
	Gaussian Kernel	95.83	97.97	97.88	97.97
Spherical (θ, φ)	Hyper Cube Kernel	98.06	98.47	98.68	98.72
	Hyper Sphere Kernel	98.03	98.46	98.66	98.72
	Gaussian Kernel	98.02	98.47	98.59	98.74

It has been demonstrated that working in the spherical color domain, the obtained results are much better than working in RGB or HSV color space. Also, different Parzen windows kernels have been compared, concluding that the best kernel is the Gaussian one in conjunction with the Anisotropic Total Variation segmentation method. Besides, it has been verified that by denoising the image, the accuracy of the method increases. Figure 5.13 shows the best-obtained results after evaluating the biscuit segmentation with different approaches.

In order to better understand the impact of each method in the segmentation, in the Appendix, the reader can observe the obtained F-Measure with different window sizes and methods. It is important to note that by making use of the Gaussian Kernel, the segmentation becomes more stable, providing more flexibility when selecting the windows width as it is illustrated in the Appendix.

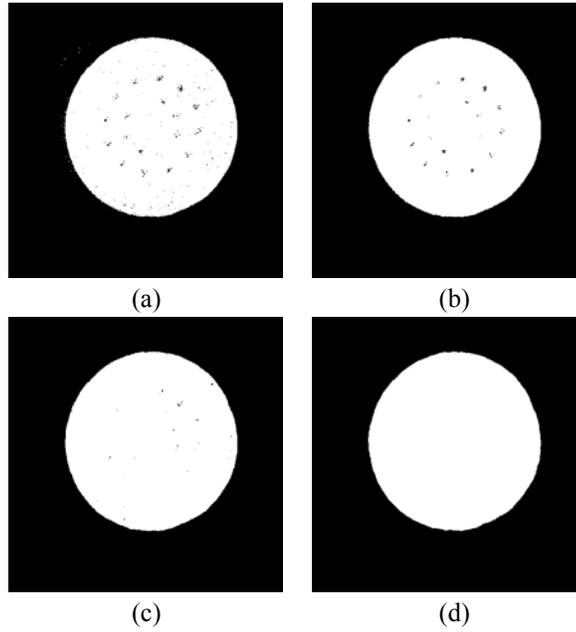


Figure 5.13. Best obtained results with (a) Parzen (b) Denoised Parzen (c) Ani-TV-L1 (d) Denoised Ani-TV-L1 through Spherical (θ, φ) color space.

B. FEATURE EXTRACTION AND CLASSIFICATION EVALUATION

Once the best segmentation method has been evaluated and determined, the accuracy of the extracted features will be evaluated. For this evaluation, the Leave-One-Out Cross-Validation (LOOCV) index will be the one used to measure the efficiency with the inclusion of the Radon transformation versus the Top-hat feature extraction over 50 different patterns. LOOCV is used when a low number of samples are available; it takes only one sample data as test and the rest of the samples as training data. This is repeated as many times as samples are available, each time taking a different sample as test data. The total generated error (5.14) consists of the integration of the classification errors with the original data. The main reason for using this evaluation technique is because the obtained error helps us to determine the features that perform better classification scores, in our case; it determines the histogram size and justifies the use of Radon transformation in the suggested method.

$$Error = \frac{100}{N} \sum_{i=1}^N Error_i \quad (5.14)$$

As it was commented in Section 5.3.3, the popular SVM and C45 techniques will be used with three different projection methods (LDA, KDA, and KPCA). At first, the histogram of the Top-Hat features is evaluated over different histogram sizes and combination of techniques, as illustrated in Table 5.2. In general, it can be observed that the increment of the histogram size improves the accuracy globally. The best obtained results belong to the combination of the KPCA/PCA projection and SVM using Gaussian kernels.

Table 5.2. LOOCV results combining different projections and classification techniques with different Top-Hat histogram sizes.

Classifier	Histogram 128 bins (Error %)	Histogram 64 bins (Error %)	Histogram 32 bins (Error %)	Histogram 16 bins (Error %)	Histogram 8 bins (Error %)	Histogram 4 bins (Error %)
SVM	8.7	11.91	12.20	10.78	12.78	23.03
C45	12.53	12.41	12.32	12.28	12.45	12.45
PCA-SVM	0.24	0.24	0.24	0.24	0.24	0.24
PCA-C45	12.41	12.86	13.16	12.49	12.41	12.45
LDA-SVM	32.15	34.27	25.94	17.11	23.65	25.94
LDA-C45	12.74	12.95	12.61	12.49	12.36	12.53
KDA-SVM	2.20	1.99	6.16	4.45	12.11	12.36
KDA-C45	12.24	12.24	12.24	12.24	12.24	12.24
KPCA-SVM	0.24	0.24	0.24	0.24	4.20	1.19
KPCA-C45	12.24	12.24	12.36	12.49	12.66	13.07

The same evaluation is carried out with the histogram of the Radon transformation as shown in Table 5.3. As in the previous results, the increase of the histogram size represents a global increment in accuracy but the best obtained result give 0% in the LOOV metric, in the case of the combination of the KDA and SVM techniques. It is also important to remark that the use of Radon transformation to generate the histogram performs much better than the Top-hat histogram. In the case of SVM without projection, the performance is up to 72 times more accurate or in the case of SVM with KPCA, accomplishing up to 6 times higher accuracy.

Table 5.3. Obtained LOOCV results combining different projections and classification techniques with different Radon histogram sizes.

Classifier	Histogram 128 bins (Error %)	Histogram 64 bins (Error %)	Histogram 32 bins (Error %)	Histogram 16 bins (Error %)	Histogram 8 bins (Error %)	Histogram 4 bins (Error %)
SVM	0.12	0.24	2.12	2.12	4.49	6.66
C45	12.28	12.28	12.28	12.32	12.41	16.15
PCA-SVM	0.12	0.24	2.12	2.12	4.33	6.66
PCA-C45	12.86	13.24	12.82	12.32	12.57	12.28
LDA-SVM	24.19	31.77	20.78	10.62	13.45	30.57
LDA-C45	12.99	12.70	12.49	12.32	12.45	13.41
KDA-SVM	0	0.08	0.12	0.16	0.16	0.58
KDA-C45	12.28	12.24	12.24	12.24	12.24	12.36
KPCA-SVM	0.04	0.24	0.12	0.12	0.24	0.29
KPCA-C45	12.24	12.28	12.24	12.24	12.24	12.24

C. GPU EVALUATION.

One of the main problems in computer vision is the high computation cost of many models. In the case of variational methods (5.4, 5.6), they have an iterative nature and require more computation time than other techniques. To avoid this limitation, a GPU is used to evaluate the suggested implementation, in this work an NVidia© GTX 580 (Fermi architecture; NVidia Corporation, Santa Clara, Calif., US) equipped with 512 processing cores was used. To contrast the obtained results on GPU, the same implementation is also compared with an Intel i7 CPU 950 (Intel Corporation, Santa Clara, Calif., US) with eight cores (four cores with two logic-cores per each physical one). We have declined to use Open Computing Language (OpenCL) due to its degree of versatility to be executed in different platforms. It is important to point out that, unlike in other articles, the parallelism of the CPU (using all cores) has been taken into account to assess the performance of both architectures (GPU and CPU).

Table 5.4. Computation time evaluation of the main methods used in our algorithm on GPU and parallel CPU implementation.

Method	GPU Time (ms)	8-cores CPU Time (ms)	1-core CPU Time (ms)
TV-L2 Denosing (5 iterations, RGB)	10.61	73.48	462.42
Ani-TV-L1 Huber Segmentation (20 iterations)	9.4	93.96	544.74
Morphological Reconstruction	2.18	4.67	23.03
Top-Hat (6x6 Block size)	1.61	11.06	61.15
DoG	1.96	10.19	58.55
Log-Inv	0.10	0.37	1.82
Radon Transformation	5.31	48.03	1958.10
RGB2Spherical	0.14	1.61	8.60

In order to compare the proposed Morphological reconstruction method, in the binary domain, other methods have been also evaluated. Those methods are the ones based on the implemented function in Matlab (imfill) and the DownHill [ROB04] method where the measured computation times are 58.78 ms and 75.33 ms respectively, i.e. 27 and 34 times faster correspondingly. These results reveal the importance of using parallel architectures to speed up the processing time.

The total computation time to determine the patterns that defines the quality of the product is 44.43 ms (GPU), 545.12 ms (8-cores CPU), and 3477.51 ms (1-core CPU). In other words, the system is able to analyze up to 22 images per second on GPU, allowing working with transport belts at the speed of 2 m/s and obtaining a speed gain factor of 12.26 with respect to the 8-cores CPU execution and 78.26 in relation with the 1-core CPU execution. These results demonstrate the importance of the GPUs to execute heavy computational loads in real-time. In the case that further performance is needed, it could be viable by means of current GPUs. For example, the Nvidia Titan Black GPU provides a performance 3.49 times faster as regards computation speed and 1.7 times faster with respect to the memory speed in comparison with the GPU used in this work. With these performance improvements, we would expect to further increase the computation speed up to ≈ 4 -6 times faster, i.e., ≈ 88 -132 images per second.

5.4 CONCLUSIONS

Industrial visual inspection is playing an important role in the present years. In this work, we have presented and evaluated a method able to segment and determine the presence of cracks or missing parts of the biscuit. A massive parallel architecture was used to implement this system to demonstrate the importance of this kind of architectures in visual inspection. Our system relies on a collection of steps where at first, biscuit color segmentation is performed with the objective of centering the attention in the biscuit area. The second step of the proposed method consists in the potential crack detections in the biscuit, focusing the detections only in the previous segmentation. The third step involves the extraction of features by means of the Radon histogram to finally classify and decide whether the product is suitable or should be rejected according to the defined quality criteria.

The suggested steps of the exposed method have been evaluated, in the case of the segmentation step. The evaluation has compared different color spaces with different Parzen kernels and the incorporation of Anisotropic Total Variation formulation to improve the segmentation (5% of improvement with respect the worse segmentation) giving a total of 60 different evaluations. Also, it has been proposed a novel method to fill holes in the binary domain on GPU. The accuracy in the classification has been evaluated by means of the Radon Histogram, in comparison with another type of feature. In this evaluation, two kinds of classification methods over four kinds of projection techniques have been tested, giving a total of 120 evaluations over 50 different samples. It has been demonstrated that including projection techniques significantly improves the accuracy, giving 0% of error in the classification. It can be concluded that the incorporation of variational methods improves the segmentation and the use of the Radon transform is an appropriate domain to extract features. Finally, it is important to remark the importance of the incorporation of the GPU to industrial inspection, where a performance speed up of 12.26 with respect the CPU implementation is obtained, i.e., with the GPU implementation the system is able to analyze 22 images per second while with the parallel CPU implementation, it is only able to process around 2 images per second.

In a future research, this work is to be extended to another kind of products and utilizing stereo vision or structure from motion techniques to incorporate additional information to the system.

6. CONCLUSIONS AND MAIN CONTRIBUTIONS

6.1 SUMMARY

The main purpose of this work was to investigate the most appropriate computer vision techniques and their implementations in real-time, in different architectures, to solve different problems in the industry. Some of the computer vision models that have been evaluated and implemented are shown below:

- Optical Flow (Lucas & Kanade and Anisotropic TV-L1).
- Block Matching (Lewis Method).
- Local Features (Fast Radial Symmetry, Pixel Orientation, LIP-SOBEL, DoG, TopHat, etc).
- Image Denoising (TV-L2 and SRAD).
- Domain Transformation (Radon and Fourier Transformation).
- Segmentation (Anisotropic TV-L1 Huber, Parzen and Active Shape Models).
- Morphological Reconstruction.

Those techniques have been previously evaluated with other methods, not cited in this work, before the implementation to ensure that the obtained robustness was fulfilled the application specifications and determinate those methods that can be parallelized on massive parallel architectures. The implementation of computer vision methods is a challenge; not all techniques have a parallel nature computation and require different designs to be implemented in a parallel architecture. One example of this problem is the SRAD method, which includes the AOS diffusion scheme that has not a parallel nature or the case of Fast Radial Symmetry where the voting scheme can be substituted with OpenGL techniques to make use of the advantages in the GPU.

6.2 SCIENTIFIC OBJECTIVES AND DEGREE OF COMPLIANCE

After the exposition of this work, we will briefly revisit the proposed objectives exposed in Chapter 1.2 to highlight the achievements and how they are related with the work objectives. It is important to remark that all the objectives have been reached. The objectives in this research work are:

- Study of motion estimation models, segmentation, feature extraction and classification. (Chapters 3, 4 and 5).
- Development of models for the detection of overtaking vehicles. (Chapter 3.5.2).
- Estimation of the dynamic properties of the artery. (Chapter 4.3).
- Segmentation and estimation of the diameter of the artery. (Chapter 4.2).
- Study and implementation of a real-time system for industrial inspection. (Chapter 5).
- Implementation of models in efficient processing platforms. (Chapters 3, 4 and 5).

6.3 FUTURE WORK

Future work will address the study of how to include more relevant cues in the optical flow in order to increase the speed and solve large displacement problems. Also, include more advanced methods to segment the obtained motion. Other point that could be of interest is the inclusion of Active Shape Model features into the Differential Evolution method such Block Matching into the objective function with the objective of improving the segmentation and the search in the solution space. In addition it is also of interest exploring the possibility of including Stereo Vision or structure from motion in the industrial inspection field with the objective of increase the collected information.

6.4 SCIENTIFIC PRODUCTION

Throughout of this thesis, it has been studied the following: (1) how to estimate Optical Flow efficiently with in a hybrid (analogic and digital) architecture [GUZ10]; (2) detection of overtaking vehicles in real-time with a low-cost architecture [GUZ11]; (3) arterial segmentation by means of an evolutive method and local features [GUZ14a]; (4) sub-pixel arterial wall tracking in Ultrasound images [GUZ14b] and (5) visual industrial inspection to detect defective biscuits in the chain production [GUZ14c]. Published papers (including a pending publication) are:

Journal Publications:

- **Guzmán, P.**; Díaz, J.; Agís, R.; Ros, E. Optical Flow in a Smart Sensor Based on Hybrid Analog-Digital Architecture. *Sensors* 2010, 10, 2975-2994.
- **Guzmán, P.**, Díaz, J., Ralli, J., Agís, R., & Ros, E. (2011). Low-cost sensor to detect overtaking based on optical flow. *Machine Vision and Applications*, 1-13.
- Ralli, J., Díaz, J., **Guzmán, P.**, & Ros, E. (2012). Experimental Study of Image Representation Spaces in Variational Disparity Calculation. *EURASIP Journal on Advances in Signal Processing*, 2012(1), 254.
- **Guzman, P.**, Ros, R., & Ros, E. (2014, February). Artery Segmentation in Ultrasound Images Based on an Evolutionary Scheme. In *Informatics* (Vol. 1, No. 1, pp. 52-71).
- **Guzman, P.**, Hamarneh, G., Ros, R., & Ros, E. (2014). Arterial Mechanical Motion Estimation Based on a Semi-Rigid Body Deformation Approach. *Sensors*, 14(6), 9429-9450.
- **Guzman, P.**, Agis R., Ros E. Machine Vision Inspection of Defective Biscuits in Real-Time. (Pending of Journal review).

Conferences:

- Agís Rodrigo, Díaz Javier, Ortigosa Pilar, **Guzmán Pablo**, Ros Eduardo: Optical Flow Reliability Model Approximated with RBF. *IWANN* (2) 2011: 90-97.

Books:

- Prieto, A. Periféricos Avanzados. Chapter: Smart Cameras: Sensores ópticos avanzados para aplicaciones industriales de visión. **Pablo Guzmán Sánchez**, Rodrigo Agís Melero, Ester Martín Garzón. ISBN: 978-84-15452-03-4.

6.5 MAIN CONTRIBUTIONS

At this point, we highlight the main contributions of this work.

- We have adapted an optical flow method and efficiently implemented in an embedded architecture with low computational resources, in combination with the focal plane processing at the sensor level.
- We have adopted a co-design strategy to develop a novel system, based on a low cost camera, able to detect vehicles overtaking in real time. It has also been tested in real-world conditions obtaining very promising results.
- We have develop a method to segment the cross section of the artery in Ultrasound images using an evolutionary scheme, based on local features, achieving better results than other alternative segmentation techniques.
- We have designed a dual motion estimation method (Block Matching and Optical Flow) to track the arterial wall in longitudinal Ultrasound imaging, obtaining more accuracy than current alternative techniques.
- We have developed using a co-design strategy a robust vision inspection system to work in the industrial field with the purpose of inspecting the quality and detecting defective biscuits in the chain production. The exposed scheme has demonstrated to be a very effective approach by combining transforming image domain and advanced classification techniques.
- We have efficiently used massive parallel processing architectures by adapting vision models to the specific features of these architectures. Using this strategy we have obtained outstanding performance results compared to the direct implementation on general purpose architectures.

- We have developed a computer vision "Framework" on GPU that includes 360 kernels with a total of 13,000 lines of code. This framework allows an easy co-design of computer vision systems and re-using the code to develop new algorithms.

7. CONCLUSIONES Y CONTRIBUCIONES PRINCIPALES

7.1 SUMARIO

El principal proposito de esta tesis fue investigar las técnicas de vision por computador mas propicias y su implementacion en tiempo real, en diferentes arquitecturas para resolver diferentes problemas en el ambito industrial. A continuacion se mostrará algunos de los modelos de vision que han sido evaluados e implementados.

- Flujo Óptico (Lucas & Kanade y Anisotrópico TV-L1).
- Block Matching (Lewis Method).
- Características locales (Fast Radial Symmetry, Orientacion del Píxel, LIP-SOBEL, DoG, TopHat, etc).
- Limpieza de imágenes (TV-L2 y SRAD).
- Transformación a otros dominios (Radon y la transformada de Fourier).
- Segmentación (Anisotrópico TV-L1 Huber, Parzen y Active Shape Models).
- Reconstrucción Morfológica.

Estas técnicas han sido previamente evaluadas con otros modelos, no citados en este trabajo, antes de la implementacion con el objetivo de asegurarnos la robustez y determinar aquellos modelos que se pueden paralelizar en arquitecturas de computacion masiva paralela. La implementación de modelos de vision por computador supone un reto, debido a que no todas las técnicas tienen una naturaleza de computacion paralela y se requiere que rediseñar para que se pueda computar en arquitecturas paralelas. Un ejemplo de este problema es el que presenta el modelo SRAD que incluye el esquema de difusión AOS y no presenta una naturaleza paralela o en el caso de Fast Radial Symmetry utiliza un esquema de votacion bastante costoso y puede ser sustituido con técnicas de OpenGL para hacer uso de las ventajas de la GPU.

7.2 OBJETIVOS CIENTIFICOS Y SU GRADO DE CUMPLIMIENTO

Tras haber expuesto el trabajo realizado en esta tesis, elaboraremos un breve análisis de los objetivos propuestos en el capítulo 2.2 para remarcar en que puntos se han cubierto. Cabe destacar que dichos objetivos han sido cubiertos completamente. Los objetivos mencionados en este trabajo de investigación son los siguientes:

- Estudio de modelos de estimación de movimiento, segmentación, extracción de características y clasificación. (Capítulos 3, 4 y 5).
- Desarrollo de modelos para la detección de vehículos en adelantamiento. (Capítulo 3.5.2)
- Estimación de las propiedades dinámicas de la arteria. (Capítulo 4.3)
- Segmentación y estimación del diámetro de la arteria. (Capítulos 4.2)
- Estudio e implementación de un sistema de inspección industrial en tiempo real. (Capítulo 5).
- Implementación de los modelos en plataformas de procesamiento eficiente. (Capítulos 3, 4 y 5).

7.3 TRABAJO FUTURO

El objetivo para trabajo futuro consiste en incluir nuevas características para acelerar el procesamiento y resolver grandes desplazamientos. También se planeará en incluir métodos avanzados para segmentar el movimiento. Otro punto que puede ser de interés es incluir partes del modelo de Active Shape Model en el algoritmo de Evolución Diferencial, como por ejemplo Block Matching dentro de la función objetivo para mejorar la segmentación y la búsqueda en el espacio de soluciones. Adicionalmente podría ser interesante explorar la posibilidad de incluir visión stereo o Structure from Motion en el campo de la inspección industrial con el objetivo de incrementar información.

7.4 PRODUCCIÓN CIENTÍFICA

A lo largo de esta tesis, se ha estudiado lo siguiente: (1) como estimar eficientemente Flujo Óptico en una arquitectura híbrida (Analógica y Digital) [GUZ10]; (2) detección de vehículos en adelantamiento en tiempo real con una arquitectura de bajo coste [GUZ11]; (3) segmentación de la arteria mediante características locales y un algoritmo evolutivo [GUZ14a]. Tracking sub-pixel de la pared arterial [GUZ14b] y (5) Inspección industrial visual para detectar galletas defectuosas en la cadena de producción [GUZ14c]. Las publicaciones, incluyendo las pendientes de revisión, son las siguientes:

Publicaciones:

- **Guzmán, P.**; Díaz, J.; Agís, R.; Ros, E. Optical Flow in a Smart Sensor Based on Hybrid Analog-Digital Architecture. *Sensors* 2010, 10, 2975-2994.
- **Guzmán, P.**, Díaz, J., Ralli, J., Agís, R., & Ros, E. (2011). Low-cost sensor to detect overtaking based on optical flow. *Machine Vision and Applications*, 1-13.
- Ralli, J., Díaz, J., **Guzmán, P.**, & Ros, E. (2012). Experimental Study of Image Representation Spaces in Variational Disparity Calculation. *EURASIP Journal on Advances in Signal Processing*, 2012(1), 254.
- **Guzman, P.**, Ros, R., & Ros, E. (2014, February). Artery Segmentation in Ultrasound Images Based on an Evolutionary Scheme. In *Informatics* (Vol. 1, No. 1, pp. 52-71).
- **Guzman, P.**, Hamarneh, G., Ros, R., & Ros, E. (2014). Arterial Mechanical Motion Estimation Based on a Semi-Rigid Body Deformation Approach. *Sensors*, 14(6), 9429-9450.
- **Guzman, P.**, Agis R., Ros E. Machine Vision Inspection of Defective Biscuits in Real-Time. (Pendiente de revisión en la revista Elsevier, *Food Engineering*).

Conferencias:

- Agís Rodrigo, Díaz Javier, Ortigosa Pilar, **Guzmán Pablo**, Ros Eduardo: Optical Flow Reliability Model Approximated with RBF. *IWANN* (2) 2011: 90-97.

Libros:

- Prieto, A. Periféricos Avanzados. Capitulo: Smart Cameras: Sensores ópticos avanzados para aplicaciones industriales de visión. **Pablo Guzmán Sánchez**, Rodrigo Agís Melero, Ester Martín Garzón. ISBN: 978-84-15452-03-4.

7.5 CONTRIBUCIONES PRINCIPALES

En este apartado se presentarán las principales contribuciones llevadas a cabo en este trabajo.

- Se ha demostrado la capacidad de desarrollar un metodo eficiente de Flujo Óptico en una arquitectura empotrada con pocos recursos computaciones en combinacion con un plano focal.
- Mediante el uso de una cámara inteligente de bajo coste, se ha demostrado que es posible diseñar un sistema capaz de detectar vehículos en adelantamiento en tiempo real. Tambien se ha probado en condiciones reales obteniendo unos resultados muy prometedores.
- Se ha validado la segmentación de arterias en sección transversal en Ultrasonidos utilizando un esquema evolutivo, dirigido mediante las características locales, obteniendo mejores resultados que otras técnicas populares de segmentación.
- Se ha diseñado un método dual de estimación de movimiento (Block Matching y Flujo Óptico) para estimar el desplazamiento de la pared de la arteria en sección longitudinal en Ultrasonidos, obteniendo mejores resultados que técnicas actuales.
- Se ha probado que factible co-diseñar un sistema robusto de inspección automática por visión artificial para trabajar en el ámbito de la industria, con el propósito de determinar la calidad y detectar galletas con defectos en la línea de producción. El esquema propuesto ha sido evaluado y se ha demostrado que tiene una gran robustez combinando el dominio de la imagen y haciendo uso de avanzados técnicas de clasificación.
- Se ha hecho uso de arquitecturas de procesamiento masivo de forma eficiente, adaptando modelos de visión a las características de dichas arquitecturas.

Mediante esta estrategia hemos obtenido un rendimiento notable en comparación con la implementación en arquitecturas de propósito general.

- Se ha desarrollado un marco de trabajo basado en visión por computador en GPU que incluye 360 kernels con un total de 13.000 líneas de código. El propósito de este marco de trabajo es facilitar el co-diseño de algoritmos de visión y reutilizar código para otros futuros algoritmos.

REFERENCES

- [ABB00] Abbott, A.L., D.L. Schmoltdt, and P.A. Araman. 2000. A next generation processing system for edging and trimming. In: Proc. of the Twenty Eighth Annual Hardwood Symp., Davis, West Virginia, 2000.
- [ABD07] Abdel-Dayem, A.; El-Sakka, M. Fuzzy c-Means Clustering for Segmenting Carotid Artery Ultrasound Images. In Proceedings of the International Conference on Image Analysis and Recognition (ICIAR 2007), Montreal, Canada, 22–24 August 2007; pp. 933–948.
- [ALC08] Alcantarilla, P.F., Bergasa, L.M., Jimenez, P., Sotelo, M.A., Parra, I., Fernandez, D.: Night Time Vehicle Detection for Driving Assistance LightBeam Controller. IEEE Intelligent Vehicles Symposium. Eindhoven, The Netherlands, June 4-6. (2008).
- [ALE07] Alessandretti, G., Broggi, A., Cerri, P.: Vehicle and guard rail detection using radar and vision data fusion. IEEE Transactions on Intelligent Transportation Systems, Vol. 8, No. 1, pp. 95–105 (2007).
- [ANA89] Anandan, P. A computational framework and an algorithm for the measurement of visual motion. *Int. J. Comput. Vis.* 1989, 2, 283–310.
- [ANG09] Anguita, M.; Díaz, J.; Ros, E.; Fernández-Baldomero, F.J. Optimization strategies for high-performance computing of optical-flow in general-purpose processors. *IEEE Trans. Circ. Syst. Video T.* 2009, 19, 1475-1488.
- [BAR12] Barranco, F.; Díaz, J.; Pino, B.; Ros, E. A multi-resolution approach for massively-parallel hardware-friendly optical flow estimation. *J. Vis. Commun. Image Represent.* 2012, 23, 1272–1283.

- [BAR92]** Barron, J.L.; Fleet, D.J.; Beauchemin, S.; Burkitt, T. Performance of optical flow techniques. In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Champaign, IL, USA, June 1992; pp. 236-242.
- [BAS08]** Basarab, A.; Liebgott, H.; Morestin, F.; Lyshchik, A.; Higashi, T.; Asato, R.; Delachartre, P. A method for vector displacement estimation with ultrasound imaging and its application for thyroid nodular disease. *Med. Image Anal.* 2008, 12, 259–274.
- [BAY06]** Bay, H., Tuytelaars, T., & Van Gool, L. (2006). Surf: Speeded up robust features. In *Computer Vision–ECCV 2006* (pp. 404-417). Springer Berlin Heidelberg.
- [BER96]** Bertozzi, M., Broggi, A.: Real-time lane and obstacle detection on the gold system. *IEEE Intelligent Vehicle Symposium*, pp. 213–218 (1996).
- [BEU11]** Beulen, B.W.; Bijnens, N.; Koutsouridis, G.G.; Brands, P.J.; Rutten, M.C.; van de Vosse, F.N. Toward noninvasive blood pressure assessment in arteries by using ultrasound. *Ultrasound Med. Biol.* 2011, 37, 788–797.
- [BIG91]** Bigun, J., Granlund, G. H., & Wiklund, J. (1991). Multidimensional orientation estimation with applications to texture analysis and optical flow. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 13(8), 775-790.
- [BLA07]** Blanc, N., Steux, B., Hinz, T.: LaRASideCam - a fast and robust vision-based blindspot detection system. In *Proc. IEEE Intelligent Vehicles Symp., Istanbul, Turkey, Jun.13-15, pp.480-485. (2007).*
- [BLA09]** Blasco, J., Cubero, S., Gómez-Sanchís, J., Mira, P., & Moltó, E. (2009). Development of a machine for the automatic sorting of pomegranate (*Punica granatum*) arils based on computer vision. *Journal of Food Engineering*, 90(1), 27-34.

- [BON00]** Bonnefous, O.; Luizy, F.; Kownator, S. Arterial Wall Motion imaging: A new ultrasound approach to vascular characterization. *Medicamundi* 2000, 44, 37–43.
- [BON96]** Bonnefous, O.; Montaudon, M.; Sananes, J.C.; Denis, E. Non invasive echographic techniques for arterial wall characterization. In *Proceedings of the IEEE Ultrasonics Symposium, San Antonio, TX, USA, 3–6 November 1996; Volume 2*, pp. 1059–1064.
- [BRO04]** Brox, T.; Bruhn, A.; Papenberg, N.; Weickert, J. High accuracy optical flow estimation based on a theory for warping. In *Computer Vision-ECCV 2004*; Springer: Berlin/Heidelberg, Germany, 2004; pp. 25–36.
- [BRO96]** Broggi, A., Conte, G., Gregoretti, F., Sansoè, C., Reyneri, L. M.: The Evolution of the PAPRICA System. *Integrated Computer-Aided Engineering Journal - Special Issue on Massively Parallel Computing*, 4(1) (1996).
- [BROX04]** Brox, T.; Bruhn, A.; Papenberg, N.; Weickert, J. High accuracy optical flow estimation based on a theory for warping. *Proceedings of 8th European Conference on Computer Vision*; Pajdla, T., Matas, J., Eds.; Springer: Prague, Czech Republic, 2004; pp. 25-36.
- [BRU11]** Brutzer, Sebastian; Höferlin, Benjamin; Heidemann, Gunther: Evaluation of Background Subtraction Techniques for Video Surveillance. In: *Computer Vision and Pattern Recognition (CVPR)*, 1937-1944, 2011.
- [CAM97]** Camus, T. Real-time quantized optical flow. *Real-Time Imaging* 1997, 3, 71-86.

- [CAN83] Canny, J. F. (1983). Finding edges and lines in images. Massachusetts Inst. of Tech. Report, 1.
- [CAN86] Canny, J. (1986). A computational approach to edge detection. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, (6), 679-698.
- [CAO09] Cao, T.; Wang, B.; Liu, D.C. Optimized GPU Framework for Semi-Implicit AOS Scheme Based Speckle Reducing Nolinear Diffusion. In *Proceedings of the SPIE Medical Imaging, Lake Buena Vista, FL, USA, 8–12 February 2009*.
- [CAR14] CARE: Community Road Accident Database.(Accessed Feb. 4, 2014). [Online].
- [CAS97] Caselles, V.; Kimmel, R.; Sapiro, G. Geodesic active contours. *Int. J. Comput. Vis.* 1997, 22, 61–79.
- [CHA01] Chan, T.F.; Vese, L.A. Active contours without edges. *IEEE T. Image Process.* 2001, 10, 266–277.
- [CHA04] Chambolle, A. (2004). An algorithm for total variation minimization and applications. *Journal of Mathematical imaging and vision*, 20(1-2), 89-97.
- [CHA10] Chan, S.H.; Vo, D.T.; Nguyen, T.Q. Subpixel motion estimation without interpolation. In *Proceedings of the 2010 IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP), Dallas, TX, USA, 14–19 March 2010*; pp. 722–725.

- [CHI07] Chi, Y.M.; Tran, T.D.; Etienne-Cummings, R. Optical flow approximation of sub-pixel accurate block matching for video coding. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2007), Honolulu, HI, USA, 15–20 April 2007; pp. 1017–1020.
- [CIE10] Ciecholewski, M. Gallbladder Boundary Segmentation from Ultrasound Images Using Active Contour Model. In *Intelligent Data Engineering and Automated Learning–IDEAL*; Springer: Berlin/Heidelberg, Germany, 2010; pp. 63–69.
- [CIN05] Cinthio, M.; Ahlgren, A.R.; Jansson, T.; Eriksson, A.; Persson, H.W.; Lindstrom, K. Evaluation of an ultrasonic echo-tracking method for measurements of arterial wall movements in two dimensions. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* 2005, 52, 1300–1311.
- [CIN06] Cinthio, M.; Ahlgren, A.R.; Bergkvist, J.; Jansson, T.; Persson, H.W.; Lindstrom, K. Longitudinal movements and resulting shear strain of the arterial wall. *Am. J. Physiol.* 2006, 291, H394–402.
- [COH93] Cohen, L.D.; Cohen, I. Finite-element methods for active contour models and balloons for 2-D and 3-D images. *IEEE T. Pattern Anal.* 1993, 15, 1131–1147.
- [COM02] Comaniciu, D., & Meer, P. (2002). Mean shift: A robust approach toward feature space analysis. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(5), 603-619.
- [COO95] Cootes, T.F.; Taylor, C.J.; Cooper, D.H.; Graham, J. Active shape models—Their training and application. *Comput. Vis. Image Underst.* 1995, 61, 38–59.
- [COR95] Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine learning*, 20(3), 273-297.

- [CVA05]** Cvancarova, M.; Albregtsen, F.; Brabrand, K.; Samset, E. Segmentation of Ultrasound Images of Liver Tumors Applying Snake Algorithms and GVF. In Proceedings of the 19th International Congress and Exhibition: Computer Assisted Radiology and Surgery (CARS 2005), Berlin, Germany, 22–25 June 2005; pp. 218–223.
- [DAV75]** Davis, L. S.: A survey of edge detection techniques , Computer Graphics and Image Processing, vol. 4, no. 3, pp. 248-270, September (1975).
- [DEA13]** Dean, T., Ruzon, M. A., Segal, M., Shlens, J., Vijayanarasimhan, S., & Yagnik, J. (2013, June). Fast, accurate detection of 100,000 object classes on a single machine. In Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on (pp. 1814-1821). IEEE.
- [DEN95]** Deng, G.; Cahill, L.W.; Tobin, G.R. The study of logarithmic image processing model and its application to image enhancement. *IEEE T. Image Process.* 1995, 506–512.
- [DEV10]** De Veroneses, L.; Krohling, R. Differential Evolution Algorithm on the GPU with C-CUDA. In Proceedings of the IEEE Congress on Evolutionary Computation (CEC), Barcelona, Spain, 18–23 July 2010; pp. 1–7.
- [DEV82]** P.A. Devijver, J. Kittler; Pattern Recognition: A Statistical Approach. Prentice-Hall, 1982.
- [DIA08a]** Díaz, J., Ros, E., Rotter A., Muehlenberg, M.: Lane change Decision Aid system based on motion driven vehicle tracking, *IEEE Transactions on Vehicular Technology*, 57(5), pp. 2736 – 2746 (2008).

- [DIA08b]** Díaz, J.; Ros, E.; Agís, R.; Bernier, J.L. Superpipelined high-performance optical-flow computation architecture. *Comput. Vis. Image Understand.* 2008, 112, 262-273.
- [DUD06a]** Dudek, P., Carey, S.J.: A General-Purpose 128x128 SIMD Processor Array with Integrated Image Sensor. *Electronics Letters*, vol.42, no.12, pp.678-679, June. (2006).
- [DUD06b]** Dudek, P., W.Barr, D.R., Lopich, A., Carey, S.J.: Demonstration of real-time image processing on the SCAMP-3 vision system. *IEEE International Workshop on Cellular Neural Networks and their Applications, CNNA 2006*, pp.13-13, Istanbul, August. (2006).
- [DUD72]** Duda, R. O., & Hart, P. E. (1972). Use of the Hough transformation to detect lines and curves in pictures. *Communications of the ACM*, 15(1), 11-15.
- [DUF95]** Dufaux, F.; Moscheni, F. Motion estimation techniques for digital TV: A review and a new contribution. *Proc. IEEE 1995*, 83, 858–876.
- [DUN73]** Dunn, J.C. A fuzzy relative of the isodata process and its use in detecting compact well-separated clusters, *J. Cybernet.* 1973, 3, 32–57.
- [ESC01]** A. de la Escalera. *Visión por computador: fundamentos y métodos.* Prentice Hall, 2001.
- [FIC14]** Ficosa Digital Blind Spot Detector (accessed Jan. 23, 2014). [Online]. Available: www.ficosa.com
- [FID85]** Fiddy, M. A. (1985). The Radon transform and some of its applications. *Journal of Modern Optics*, 32(1), 3-4.

- [FIN84]** Finlay, D. J., Dodwell, P. C., Caelli, T. M.: The wagon-wheel effect. *Perception* 13: 237–248. (1984).
- [FIS36]** Fisher, R. A. (1936). The use of multiple measurements in taxonomic problems. *Annals of eugenics*, 7(2), 179-188.
- [FLE90]** Fleet, D. J., & Jepson, A. D. (1990). Computation of component image velocity from local phase information. *International Journal of Computer Vision*, 5(1), 77-104.
- [FLE92]** Fleet, D.J. *Measurement of Image Velocity*; Kluwer Academic: Norwell, MA, USA, 1992.
- [FOL08]** Foldesy, P., Zarandy, A., Rekeczky, C.: Configurable 3D-integrated focal-plane cellular sensor-processor array architecture. *International Journal of Circuit Theory and applications* 36: (5-6) pp. 573-588. (2008).
- [GAL98]** Galvin, B.; McCane, B.; Novins, K.; Mason, D.; Mills, S. Recovering motion fields: an evaluation of eight optical flow algorithms. In *Proceedings of British Machine Vision Conference*, Southampton, UK, October 1998; pp. 195-204.
- [GAS10]** Gastounioti, A.; Golemati, S.; Stoitsis, J.; Nikita, K.S. Kalman-filter-based block matching for arterial wall motion estimation from B-mode ultrasound. In *Proceedings of the 2010 IEEE International Conference on Imaging Systems and Techniques (IST)*, Thessaloniki, Greece, 1–2 July 2010; pp. 234–239.
- [GAT05]** Gat, I., Benady, M., Shashua, A.: A monocular vision advance warning system for the automotive aftermarket. *SAE Transactions*, vol.114, 7, pp. 403-410. (2005).
- [GAU02]** Gautama, T., & Van Hulle, M. A. (2002). A phase-based approach to the estimation of the optical flow field using spatial filtering. *Neural Networks, IEEE Transactions on*, 13(5), 1127-1136.

- [GAV01]** Gavrilá, D., Kunert M., Lages, U.: A multi-sensor approach for the protection of vulnerable traffic participants- the PROTECTOR project, IEEE Instrumentation and Measurement Technology Conference, Volume 3, pp. 2044-2048, Budapest, Hungary (2001).
- [GLA11]** Glavtchev, V.; Muyan-Ozcelik, P.; Ota, J.M.; Owens, J.D. Feature-Based Speed Limit Sign Detection Using a Graphics Processing Unit. In Proceedings of the IEEE on Intelligent Vehicles Symposium (IV), Baden-Baden, Germany, 5–9 June 2011; pp. 195–200.
- [GOL12]** Golemati, S.; Stoitsis, J.S.; Gastounioti, A.; Dimopoulos, A.C.; Koropouli, V.; Nikita, K.S. Comparison of block matching and differential methods for motion analysis of the carotid artery wall from ultrasound images. IEEE Trans. Inf. Technol. Biomed. 2012, 16, 852–858.
- [GRA03]** Graves, M. (2003). Machine vision for the inspection of natural products (p. 5). London: Springer.
- [GRE00]** Green, M.: How long does it take to stop? Methodological analysis of driver perception-brake times. Transp. Hum. Factors, vol. 2, no. 3, pp. 195–216, Sep. (2000).
- [GUZ10]** Guzmán, P., Díaz, J., Agís, R., Ros, E.: Optical Flow in a Smart Sensor Based on Hybrid Analog-Digital Architecture. Sensors (Special issue on Motion Sensors), 10, 2975-2994. (2010).
- [GUZ11]** Guzmán, P., Díaz, J., Ralli, J., Agís, R., & Ros, E. (2011). Low-cost sensor to detect overtaking based on optical flow. Machine Vision and Applications, 1-13.
- [GUZ12]** Prieto, A. Periféricos Avanzados. Smart Cameras: Sensores ópticos avanzados para aplicaciones industriales de visión. Pablo Guzmán

Sánchez, Rodrigo Agís Melero, Ester Martín Garzón. ISBN: 978-84-15452-03-4.

- [GUZ14a]** Guzman, P., Ros, R., & Ros, E. (2014, February). Artery Segmentation in Ultrasound Images Based on an Evolutionary Scheme. In *Informatics* (Vol. 1, No. 1, pp. 52-71). Multidisciplinary Digital Publishing Institute.
- [GUZ14b]** Guzman, P., Hamarneh, G., Ros, R., & Ros, E. (2014). Arterial Mechanical Motion Estimation Based on a Semi-Rigid Body Deformation Approach. *Sensors*, 14(6), 9429-9450.
- [GUZ14c]** Guzman, P., Agis R., Ros E. Machine Vision Inspection of Defective Biscuits in Real-Time. (Pending of review).
- [HAF10]** Haff, R.P. Pearson, T.C. Toyofuku. Sorting of In-Shell Pistachio Nuts from Kernels Using Color Imaging. *Applied engineering in agriculture*. 2010, v. 26, no. 4, p. 633-638.
- [HAJ92]** Hajela, P; Lin, C.Y; Genetic search strategies in multicriterion optimal design. *Struct. Optim.* 1992, 4, 99–107.
- [HAM03]** Hamarneh, G.; McInerney, T. Physics-Based Shape Deformations for Medical Image Analysis. In *Proceedings of the SPIE-IS&T Electronic Imaging: Image Processing: Algorithms and Systems*, Santa Clara, CA, USA, 20 January 2003; Volume 5014, pp. 354–362.
- [HAN06]** Hansen, N. Compilation of results on the 2005 CEC benchmark function set, Tech. Rep., Institute of Computational Science ETH Zurich, 2006.
- [HAN10]** Han Hu, Quanquan Gu, Jie Zhou, HTF: a novel feature for general crack detection. The 17th IEEE International Conference on Image Processing (ICIP), Hong Kong, China, pp.1633-1636, 2010.

- [HAR88]** Harris, C., Stephens, M.: A combined corner and edge detector. Proc.of 4th Alvery Vision Conference on Manchester, pp. 189-192. (1988).
- [HAS56]** Hassenstein, B., Reichardt, W.: Systemtheoretische Analyse der Zeit-Reihenfolgen und Vorzeichenauswertung bei der ewegungsperzeption des Rüsselkäfers Chlorophanus, Zeitschrift für Naturforschung 11b 513–524. (1956).
- [HAS56]** Hassenstein, B.; Reichardt, W. Systemtheoretische Analyse der Zeit-Reihenfolgen und Vorzeichenauswertung bei der Bewegungspereption des Rüsselkäfers Chlorophanus. Z. Naturforsch. 1956, 11b, 513-524.
- [HEE01]** He, P., & Zheng, J. (2001). Segmentation of tibia bone in ultrasound images using active shape models. In Engineering in Medicine and Biology Society, 2001. Proceedings of the 23rd Annual International Conference of the IEEE (Vol. 3, pp. 2712-2715). IEEE.
- [HOC65]** Hockney, R.W. A fast direct solution of Poisson’s equation using Fourier analysis. *J. ACM* 1965, 12, 95–113.
- [HOC81]** Hockney, R.W.; Jesshope, C.R. *Parallel Computers*; Adam Hilger: London, UK, 1981.
- [HOO06]** Ho, C.Y.; Solomon, S.D. A clinician’s guide to tissue Doppler imaging. *Circulation* 2006, 113, e396–e398.
- [HOR81]** Horn B.; Schunck, B. Determining optical flow. *Artif. Intell.* 1981, 17, 185-203.
- [HOU62]** Hough, P. V. (1962). U.S. Patent No. 3,069,654. Washington, DC: U.S. Patent and Trademark Office.

- [ICA09]** I-Car: Blind spot object detection systems. (Accessed Jan. 23, 2009).[Online].Available:
<http://www.i-car.com/pdf/advantage/online/2009/101309.pdf>
- [INS10]** Institut für Algorithmen und Kognitive Systeme. Ettliger-Tor Sequence.
- [JEN96]** Jensen, J.A. Field: A program for simulating ultrasound systems. In Proceedings of the 10th Nordicbaltic Conference on Biomedical Imaging, Tampere, Finland, 9–13 June 1996; Volume 4, Supplement 1, PART 1, pp. 351–353.
- [JOL86]** Jolliffe, I. T. (1986). Principal component analysis. Springer Series in Statistics, Berlin: Springer, 1986, 1.
- [JUM10]** Jumaat, A. K., Rahman, W. E. Z. W. A., Ibrahim, A., & Mahmud, R. (2010). Segmentation of masses from breast ultrasound images using parametric active contour algorithm. *Procedia-Social and Behavioral Sciences*, 8, 640-647.
- [KAL11]** O. Kalentev, A. Rai, S. Kemnitz, and R. Schneider, "Connected component labeling on a 2D grid using CUDA", presented at J. Parallel Distrib. Comput., 2011, pp.615-620.
- [KAL60]** Kalman, R.E. A new approach to linear filtering and prediction problems. *J. Basic Eng.* 1960, 82, 35–45.
- [KAN91]** E. R. Kandell, J.H. Schwartz, and T.M. Jessell. Principles of Neural Science. Elsevier, 1991. 5, 7, 17, 19, 97, 176, 182.
- [KAN98]** Kanters, S.D.J.M.; Elgersma, O.E.H.; Banga, J.D.; van Leeuwen, M.S.; Algra, A. Reproducibility of measurements of intima-media thickness and distensibility in the common carotid artery. *Eur. J. Vasc. Endovasc. Surg.* 1998, 16, 28–35.

- [KAR10]** P. Karas. Efficient Computation of Morphological Greyscale Reconstruction. In MEMICS, volume 16 of OASICS, pages 54–61. Schloss Dagstuhl – Leibniz-Zentrum fuer Informatik, Germany, 2010.
- [KAS87a]** Kass, M., & Witkin, A. (1987). Analyzing oriented patterns. *Computer vision, graphics, and image processing*, 37(3), 362-385.
- [KAS87b]** Kass, M.; Witkin, A.; Terzopoulos D. Snakes: Active contour models. *Int. J. Comput. Vis.* 1987, 1, 321–331.
- [KOH09]** Köhler, T.; Röchter, F.; Lindemann, J.P.; Möller, R. Bio-inspired motion detection in an FPGA-based smart camera module. *Bioinspir. Biomim.* 2009, 4, 15008.
- [LEW95]** Lewis, J.P. Fast normalized cross-correlation. *Vis. Interface* 1995, 10, 120–123.
- [LIU07]** Liu, W., Wen, X., Duan, B., Yuan, H., Wang, N.: Rear Vehicle Detection and Tracking for Lane Change Assist. *Proceedings of the 2007 IEEE Intelligent Vehicles Symposium*. pp.252-257 (2007).
- [LIU98]** Liu, H., Hong, T., Herman, M., Camus, T., Chellapa, R.: Accuracy vs. efficiency trade-off in optical flow algorithms. *Computer Vision Image Understand.* 72, 271–286 (1998).
- [LOW99]** Lowe, D. G. (1999). Object recognition from local scale-invariant features. In *Computer vision, 1999. The proceedings of the seventh IEEE international conference on* (Vol. 2, pp. 1150-1157). IEEE.
- [LOY03]** Loy, G.; Zelinsky, A. Fast radial symmetry for detecting points of interest. *IEEE T. Pattern Anal.* 2003, 25, 959–973.

- [LUC81] Lucas, B. and Kanade, T.: An iterative image registration technique with an application to stereo vision. In Proc. Seventh International Joint Conference on Artificial Intelligence, Vancouver, Canada, pp. 674–679 (1981).
- [LUT94] Lutton, E.; Martinez, P. A Genetic Algorithm for the Detection of 2D Geometric Primitives in Images. In Proceedings of the 12th International Conference on Pattern Recognition, Jerusalem, Israel, 9–13 October 1994; pp. 526–528.
- [MAI02] Mainzer, T. Genetic Algorithm for Shape Detection. Technical Report No. DCSE/TR-2002-06, University of West Bohemia, Pilsen, Czech Republic, 2002.
- [MAK06] K.L. Mak, P. Peng, Detecting defects in textile fabrics with optimal Gabor filters, *Int. J. Comput. Sci.* 1 1306–4428. (2006).
- [MAL08] Mallipeddi, R.; Suganthan, P.N. Empirical Study on the Effect of Population Size on Differential Evolution Algorithm. In Proceedings of the IEEE Congress on Evolutionary Computation (CEC), Hong Kong, China, 1–6 June 2008; pp. 3664–3671.
- [MAR03] Marsaglia, G. Xorshift RNGs. *J. Stat. Softw.* 2003, 8, 1–6.
- [MCL98] McLaughlin, R.A. Randomized Hough transform: Improved ellipse detection with comparison. *Pattern Recogn. Lett.* 1998, 19, 299–305.
- [MEH06] Mehta, S.; Etienne-Cummings, R. A simplified normal optical flow measurement CMOS camera. *IEEE Trans. Circ. Syst. I* 2006, 53, 1223-1234.
- [MEN13] Mendl, C. B., Eliuk, S., Noga, M., & Boulanger, P. (2013). Comprehensive Analysis of High-Performance Computing Methods for Filtered Back-Projection. *Electronic Letters on Computer Vision and Image Analysis*, 12(1), 1-16.

- [MIK99]** Mika, S., Ratsch, G., Weston, J., Scholkopf, B., & Muller, K. (1999, August). Fisher discriminant analysis with kernels. In *Neural Networks for Signal Processing IX, 1999. Proceedings of the 1999 IEEE Signal Processing Society Workshop.* (pp. 41-48). IEEE.
- [MOB09]** N.V. Mobileye: Blind Spot Detection and Lane Change Assist (BSD/LCA) (accessed Jan. 23, 2009). [Online]. Available: <http://www.mobileye.com/default.asp?PageID=226>
- [MOR09]** Mortensen, K.; Aydin, M.A.; Rybczynski, M.; Baulmann, J.; Schahidi, N.A.; Kean, G.; Kühne, K.; Bernhardt, A.M.; Franzen, O.; Mir, T.; et al. Augmentation index relates to progression of aortic disease in adults with Marfan syndrome. *Am. J. Hypertens.* 2009, 22, 971–979.
- [MOT04]** Mota, S., Ros, E., Ortigosa, E. M., Pelayo, F. J.: Bio-Inspired motion detection for blind spot overtaking monitor, *Int. Journal of Robotics and Automation*, 19(4), pp. 190–196 (2004).
- [MOU09]** Moursi, S.G.; Sakka, M.R.E. Semi-automatic snake based segmentation of carotid artery ultrasound images. *Commun. Arab Comput. Soc. (ACS)* 2009, 2, 1–32.
- [MUH02]** Mühlmann, K., Maier, D., Hesser, J., & Männer, R. (2002). Calculating dense disparity maps from color stereo images, an efficient implementation. *International Journal of Computer Vision*, 47(1-3), 79-88.
- [NAK95]** Nakayama, K., He, Z. J., & Shimojo, S. (1995). Visual surface representation: A critical link between lower-level and higher-level vision. *Visual cognition: An invitation to cognitive science*, 2, 1-70.

- [NAS11]** Nashat, S., Abdullah, A., & Abdullah, M. Z. (2011, May). A robust crack detection method for non-uniform distributions of coloured and textured image. In *Imaging Systems and Techniques (IST)*, 2011 IEEE International Conference on (pp. 98-103). IEEE.
- [NAS14]** Nashat, S., Abdullah, A., & Abdullah, M. Z. (2014). Machine vision for crack inspection of biscuits featuring pyramid detection scheme. *Journal of Food Engineering*, 120, 233-247.
- [NII05]** Niitsuma, H.; Maruyama, T. High speed computation of the optical flow. In *Proceedings of the 13th International Conference on Image Analysis and Processing (ICIAP '05)*, Cagliari, Italy, September 2005; pp. 287-295.
- [OLO02]** Olofsson, A.; Lange, F.A. 4.32 GOPS 1W general-purpose DSP with an enhanced instruction set for wireless communication. *IEEE Int. Solid-State Circuits Conf. Dig. Tech. Pap.* 2002, 2, 36-89.
- [OMA08]** O'Malley, R., Glavin, M., Jones, E.: Vehicle Detection at Night Based on Tail-Light Detection. 1st International Symposium on Vehicular Computing Systems, Trinity College Dublin, July. (2008).
- [PAL06]** Palomares, J.M.; Gonzalez, J.; Ros, E.; Prieto, A. General logarithmic image processing. *IEEE T. Image Process.* 2006, 15, 3602–3608.
- [PAL10]** Palomar, R.; Palomares, J.M.; Castillo, J.M.; Olivares, J.; Gómez-Luna, J. Parallelizing and Optimizing LIP-Canny Using Nvidia Cuda. In *Trends in Applied Intelligent Systems*; Springer: Berlin/Heidelberg, Germany, 2010; pp. 389–398.
- [PAR62]** Parzen, E. On estimation of a probability density function and mode. *Annals of Mathematical Statistics* 33: 1065–1076. 1962.

- [PEA09]** Pearson, T. C. Hardware based image processing for high-speed inspection of grains. *Computers and Electronics in Agric.* 69(1): 12-18. 2009.
- [QUI93]** Quinlan, J. R. *C4.5: Programs for Machine Learning*. Morgan Kaufmann Publishers, 1993.
- [RAM03]** Ramnarine, K.V.; Hartshorne, T.; Sensier, Y.; Naylor, M.; Walker, J.; Naylor, A.R.; Panerai, R.B.; Evans, D.H. Tissue Doppler imaging of carotid plaque wall motion: A pilot study. *Cardiovasc. Ultrasound* 2003, 1, doi:10.1186/1476-7120-1-17.
- [RIQ08]** Riquelme, M. T., Barreiro, P., Ruiz-Altisent, M., & Valero, C. (2008). Olive classification according to external damage using image analysis. *Journal of Food Engineering*, 87(3), 371-379.
- [ROB04]** K Robinson, PF Whelan. Efficient morphological reconstruction: A downhill filter. *Pattern Recognition Letters* 25 (15), 1759-1767, 2004.
- [ROD08]** Rodríguez-Vazquez, A. The eye-RIS CMOS vision system. *Analog Circuit Design: Sensors, Actuators and Power Drivers*; Springer: Dordrecht, The Netherlands, 2008; pp. 15-32.
- [RUA13]** Ruan, N., Chen, Y., & Gao, D. Y. (2013, October). An efficient classification using support vector machines. In *Science and Information Conference (SAI), 2013* (pp. 585-589). IEEE.
- [RUD92]** Rudin, L., Osher, S., and Fatemi, E. Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259-268. 1992.
- [SAK06]** Sakurai, K., Kyo, S., Okazaki S.: Implementation of Overtaking Vehicle Detection Using the IMAPCAR Highly Parallel Image Processor. In *Proc. of ITS Congress*. (2006).

- [SCH98]** Schölkopf, B., Smola, A., & Müller, K. R. (1998). Nonlinear component analysis as a kernel eigenvalue problem. *Neural computation*, 10(5), 1299-1319.
- [SEN14]** Senni, L., Ricci, M., Palazzi, A., Burrascano, P., Pennisi, P., & Ghirelli, F. (2014). On-line automatic detection of foreign bodies in biscuits by infrared thermography and image processing. *Journal of Food Engineering*, 128, 146-156.
- [SHI08]** Shimizu, M.; Kario, K. Review: Role of the augmentation index in hypertension. *Ther. Adv. Cardiovasc. Dis.* 2008, 2, 25–35.
- [SIN06]** S.K. Sinha, P.W. Feiguth, Segmentation of buried concrete pipe images, *Automation in Construction* 15 (1) (2006) 47–57.
- [SMI97]** Smith, S. M., & Brady, J. M. (1997). SUSAN—A new approach to low level image processing. *International journal of computer vision*, 23(1), 45-78.
- [SOI99]** Soille, P., *Morphological Image Analysis: Principles and Applications*, Springer-Verlag, 1999, pp. 173-174.
- [SOL01]** Solari, F., Sabatini, S. P., & Bisio, G. M. (2001). Fast technique for phase-based disparity estimation with no explicit calculation of phase. *Electronics Letters*, 37(23), 1382-1383.
- [SOL11]** Soleimani, E.; Dizaji, M.M.; Saberi, H. Carotid artery wall motion estimation from consecutive ultrasonic images: Comparison between block-matching and maximum-gradient algorithms. *J. Tehran Univ. Heart Cent.* 2011, 6, 72–78.
- [SON01]** Song, K.T., Huang, J. H.: Fast Optical Flow Estimation and Its Application to Real-time Obstacle Avoidance. In *Proc. Of 2001 IEEE ICRA*, Seoul, Korea, pp.2891-2896. (2001).

- [SON07] Song, K.T., Chen, H. Y.: Lateral driving assistance using optical flow and scene analysis. Proceedings of the 2007 IEEE Intelligent Vehicles Symposium. pp.624–9 (2007).
- [STO06] Stocker, A.A. Analog integrated 2-D optical flow sensor. Analog Integr. Circuit. Signal 2006, 46, 121-138.
- [STO08] Stoitsis, J.; Golemati, S.; Koropouli, V.; Nikita, K.S. Simulating dynamic B-mode ultrasound image data of the common carotid artery. In Proceedings of the IEEE International Workshop on Imaging Systems and Techniques (IST 2008), Crete, Greece, 10–12 September 2008; pp. 144–148.
- [STO97] Storn, R.; Price, K. Differential evolution—A simple and efficient heuristic for global optimization over continuous spaces. *J. Glob. Optim.* 1997, 11, 341–359.
- [SUB06] P. Subirats, J. Dumoulin, V. Legeay, and D. Barba, "Automation of Pavement Surface Crack Detection using the Continuous Wavelet Transform", in Proc. ICIP, 2006, pp. 3037-3040.
- [SUN00] Sun, D. W. (2000). Inspecting pizza topping percentage and distribution by a computer vision method. *Journal of Food Engineering*, 44(4), 245-249.
- [SUN03] Sun, D. W., & Brosnan, T. (2003). Pizza quality evaluation using computer vision—Part 2: Pizza topping analysis. *Journal of Food Engineering*, 57(1), 91-95.
- [SUN11] Sun, D. W. (2011). *Computer vision technology for food quality evaluation*. Academic Press.
- [SUN12] Sun, D. W. (Ed.). (2012). *Computer vision technology in the food and beverage industries*. Elsevier.
- [TAN09] Tang, J. A multi-direction GVF snake for the segmentation of skin cancer images. *Pattern Recogn.* 2009, 42, 1172–1179.

- [TEO12]** Teodoro, G. and Pan, T. and Kurc, T.M. and Cooper, L. and Kong, J. and Saltz, J.H. A Fast Parallel Implementation of Queue-based Morphological Reconstruction using GPUs.Center for Comprehensive Informatics, Emory University, Atlanta, GA 30322. Emory University, Center for Comprehensive Informatics Technical Report CCI-TR-2012-2, 2012.
- [THO94]** Thomanek, F., Dickmanns, E. D., & Dickmanns, D. (1994, October). Multiple object recognition and scene interpretation for autonomous road vehicle guidance. In Intelligent Vehicles' 94 Symposium, Proceedings of the (pp. 231-236). IEEE.
- [UNG08]** Unger, M., Pock, T., Trobin, W., Cremers, D., & Bischof, H. (2008, September). TVSeg-Interactive Total Variation Based Image Segmentation. In BMVC (pp. 1-10).
- [VAP11]** Vappou, J.; Luo, J.; Okajima, K.; Di Tullio, M.R.; Konofagou, E.E. Non-invasive measurement of local pulse pressure by pulse wave-based ultrasound manometry (PWUM). *Physiol. Meas.* 2011, 32, 1653–1662.
- [VEZ03]** Vladimir Vezhnevets, VassiliSazonov, Alla Andreeva. A Survey on Pixel-Based Skin Color Detection Techniques. Graphics and Media Laboratory, Faculty of Computational Mathematics and Cybernetics, Moscow State University, Moscow, Russia. 2003.
- [VIN93]** Vincent, L. Morphological grayscale reconstruction in image analysis: Applications and efficient algorithms. *IEEE T. Image Process.* 1993, 2,176–201.
- [VIT07]** Vitria, J., Bressan, M., & Radeva, P. (2007). Bayesian classification of cork stoppers using class-conditional independent component analysis. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, 37(1), 32-38.

- [VOL09a]** Volvo BLIS System.
(Accessed Jan. 23, 2009). [Online]. Available:
<http://www.mynrma.com.au/blis.asp>
- [VOL09b]** Volvo Club: Volvo S60. Press Information.
(Accessed Jan. 23, 2009).[Online]. Available:
http://www.volvoclub.org.uk/press/pdf/presskits/S60_MY2005.pdf
- [WAS02]** Z. Wasik, and A. Saffiotti, Robust color segmentation for the robocup domain. In Proc. of the Int. Conf. on Pattern Recognition (ICPR'02), Quebec City, Quebec, CA, 2002.
- [WED08]** Wedel, A.; Pock, T.; Zach, C.; Bischof, H.; Cremers, D. An improved algorithm for TV-L1 optical flow. In Proceedings of the International Dagstuhl Seminar, Dagstuhl Castle, Germany, 13–18 July 2008; pp. 23–45.
- [WEI08]** Weichert, J. *Anisotropic Diffusion in Image Processing*; BG Teubner: Stuttgart, Germany, 2008; pp. 278–290.
- [WEI84]** Weicker, R.P. Dhrystone: a synthetic systems programming benchmark. CACM 1984, 27, 1013-1030.
- [WEI98]** Weickert, J.; Romeny, B.H.; Viergever, M.A. Efficient and reliable schemes for nonlinear diffusion filtering. *IEEE T. Image Process.* 1998, 7, 398–409.
- [WEL02]** Welch, G., Bishop, G.: An introduction to the Kalman filter. Dept. Comput. Sci., Univ. North Carolina Chapel Hill, Chapel Hill, NC, Tech. Rep. TR 95-041. (2002).
- [WER09]** Werlberger, M., Trobin, W., Pock, T., Wedel, A., Cremers, D., & Bischof, H. (2009, September). Anisotropic Huber-L1 Optical Flow. In BMVC (Vol. 1, No. 2, p. 3).

- [WHO14]** World Health Organization. (Accessed March. 5, 2014). [Online]. Available:

<http://www.who.int/features/qa/18/en/>
- [WILK00]** Wilkinson, I.B.; MacCallum, H.; Rooijmans, D.F.; Murray, G.D.; Cockcroft, J.R.; McKnight, J.A.; Webb, D.J. Increased augmentation index and systolic stress in type 1 diabetes mellitus. *QJM* 2000, 93, 441–448.
- [XUU98]** Xu, C.; Prince, J.L. Snakes, shapes, and gradient vector flow. *IEEE T. Image Process.* 1998, 7, 359–369.
- [YAO05]** Yao, J.; Kharma, N.; Grogono, P. A multi-population genetic algorithm for robust and fast ellipse detection. *Pattern Anal. Appl.* 2005, 8, 149–162.
- [YAO69]** Yao, S.T.; Hobbs, J.T.; Irvine, W.T. Ankle systolic pressure measurements in arterial disease affecting the lower extremities. *Br. J. Surg.* 1969, 56, 676–679.
- [YON02]** Yongjian, Y.; Acton, S.T. Speckle reducing anisotropic diffusion. *IEEE T. Image Process.* 2002, 11, 1260–1270.
- [YUA05]** Yuan, Y., & Huang, T. (2005). A polynomial smooth support vector machine for classification. In *Advanced Data Mining and Applications* (pp. 157-164). Springer Berlin Heidelberg.
- [ZAC07]** Zach, C.; Pock, T.; Bischof, H. A duality based approach for realtime TV-L1 optical flow. In *Pattern Recognition*; Springer: Berlin/Heidelberg, Germany, 2007; pp. 214–223.
- [ZAH13]** Zahnd, G.; Orkisz, M.; Sérusclat, A.; Moulin, P.; Vray, D. Evaluation of a Kalman-based block matching method to assess the bi-dimensional motion of the carotid artery wall in B-mode ultrasound sequences. *Med. Image Anal.* 2013, 17, 573–585.

- [ZAN10]** Zahnd, G.; Boussel, L.; Marion, A.; Durand, M.; Moulin, P.; Sérusclat, A.; Vray, D. Measurement of two-dimensional movement parameters of the carotid artery wall for early detection of arteriosclerosis: A preliminary clinical study. *Ultrasound Med. Biol.* 2010, 37, 1421–1429.
- [ZAN12]** Zahnd, G.; Vray, D.; Sérusclat, A.; Alibay, D.; Bartold, M.; Brown, A.; Durand, M.; Jamieson, L.M.; Kapellas, K.; Maple-Brown, L.J.; et al. Longitudinal displacement of the carotid wall and cardiovascular risk factors: Associations with aging, adiposity, blood pressure and periodontal disease independent of cross-sectional distensibility and intima-media thickness. *Ultrasound Med. Biol.* 2012, 38, 1705–1715.
- [ZAR08]** Zarándy, Á., Rekeczky, C., Földesy, P.: Analysis of 2D operators on topographic and non-topographic processor architectures. *CNNA 2008. 11th international workshop on cellular neural networks and their applications.* Santiago de Compostela, IEEE. (2008).
- [ZHA10a]** Zhang, K.; Zhang, L.; Song, H.; Zhou, W. Active contours with selective local or global segmentation: A new formulation and level set method. *Image Vis. Comput.* 2010, 28, 668–676.
- [ZHA10b]** Zhang, Y.; Cohen, J.; Owens, J.D. Fast Tridiagonal Solvers on the GPU. In *Proceedings of the 15th ACM SIGPLAN Symposium Principles and Practice of Parallel Programming (PPoPP '10)*, Bangalore, India, 9–14 January 2010; pp. 127–136.
- [ZHU08]** Zhu, M. and Chan, T. An efficient primal-dual hybrid gradient algorithm for total variation image restoration. *CAM Report 08-34*, UCLA. 2008.
- [ZHU10]** Zhu, W.; Li, Y. GPU-Accelerated Differential Evolutionary Markov Chain Monte Carlo Method for Multi-Objective Optimization over Continuous Space. In *Proceedings of the 2nd*

Workshop on Bio-Inspired Algorithms for Distributed Systems (BADs), New York, NY, USA, 7–11 June 2010; pp. 1–8.

[ZHU11]

Zhu, W. Massively parallel differential evolution-pattern search optimization with graphics hardware acceleration: An investigation on bound constrained optimization problems. *J. Glob. Optim.* 2011, 50, 417–437.

APPENDIX A. TRACKING RESULTS

Table A1. Longitudinal error results after being evaluated with different methods.

Methods	M1	M2	M3	M4	M5	M6
Lucas & Kanade	62.87	16.95	59.98	73.91	75.99	42.52
BM+ Lucas & Kanade	52.79	45.05	57.41	39.51	29.76	19.48
BM	1.10	1.15	1.68	1.68	1.27	1.28
Anisotropic Huber-L1	1.43	1.42	1.56	1.56	1.12	1.12
BM + Anisotropic Huber-L1	1.73	1.62	1.88	1.87	1.52	1.51

Table A2. Radial error results after being evaluated with different methods.

Methods	M1	M2	M3	M4	M5	M6
Lucas & Kanade	7.89	2.37	4.96	8.50	4.93	3.33
BM+ Lucas & Kanade	4.69	4.42	3.96	2.77	1.84	1.08
BM	0.52	0.88	0.63	0.64	0.64	0.64
Anisotropic Huber-L1	0.59	1.35	0.57	0.57	0.57	0.57
BM + Anisotropic Huber-L1	0.25	0.24	0.26	0.28	0.21	0.21

Table A3. Position Error results after being evaluated with different methods.

Methods	M1	M2	M3	M4	M5	M6
Lucas & Kanade	63.75	17.38	60.66	75.02	76.37	42.99
BM+ Lucas & Kanade	53.35	45.51	57.63	39.68	29.86	19.55
BM	1.30	1.49	1.80	1.81	1.49	1.50
Anisotropic Huber-L1	1.55	1.48	1.67	1.67	1.30	1.31
BM + Anisotropic Huber-L1	1.75	1.64	1.89	1.89	1.54	1.53

Table A4. Diameter error results after being evaluated with different methods.

Methods	M1	M2	M3	M4	M5	M6
Lucas & Kanade	4.36	3.08	3.79	3.70	2.52	2.89
BM+ Lucas & Kanade	9.30	5.79	2.52	2.25	1.44	1.02
BM	0.65	0.86	0.44	0.48	0.43	0.44
Anisotropic Huber-L1	0.47	0.47	0.37	0.42	0.38	0.38
BM + Anisotropic Huber-L1	0.26	0.28	0.26	0.35	0.20	0.21

APPENDIX B. HSV AND SPHERICAL COLOR SPACES

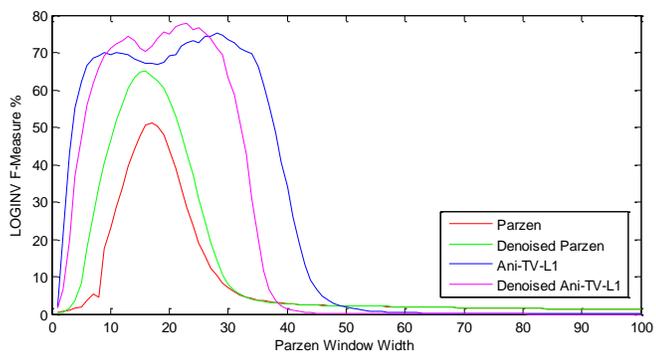
RGB-Spherical color conversion:

$$Spherical(\rho, \theta, \varphi) = \begin{cases} \rho = \sqrt{R^2 + G^2 + B^2} \\ \theta = \tan^{-1} \frac{G}{R} \\ \varphi = \cos^{-1} \frac{B}{\rho} \end{cases} \quad (B.1)$$

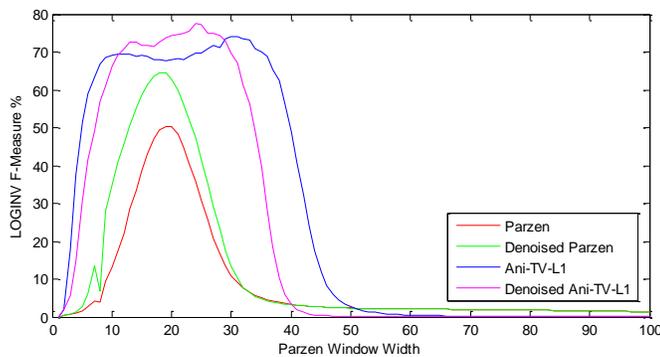
RGB-HSV color conversion:

$$\begin{aligned} VMax &= \max(R, G, B) \\ VMin &= \min(R, G, B) \\ \Delta &= VMax - VMin \\ H &= \begin{cases} 60 \cdot \left(\frac{G - B}{\Delta} \text{mod} 6 \right) \text{ if } VMax = R \\ 60 \cdot \left(\frac{B - R}{\Delta} + 2 \right) \text{ if } VMax = G \\ 60 \cdot \left(\frac{R - G}{\Delta} + 4 \right) \text{ if } VMax = B \end{cases} \\ S &= \begin{cases} 0 & \text{if } \Delta = 0 \\ \frac{\Delta}{VMax} & \text{if } \Delta \neq 0 \end{cases} \\ V &= VMax \end{aligned} \quad (B.2)$$

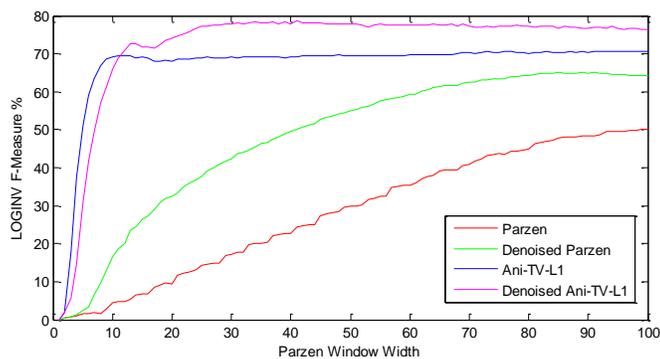
APPENDIX C. SEGMENTATION: F-MEASURE VS. PARZEN WINDOW WIDTH



(a)



(b)



(c)

Figure. C.1 Plot of the Logarithmic Inverse F-Measure with respect to the Parzen window width ρ in different Kernels (a) HyperCube, (b) HyperSphere, and (c) Gaussian in Spherical color space.

APPENDIX D. SOME EXAMPLES OF THE TRAINING DATA SET

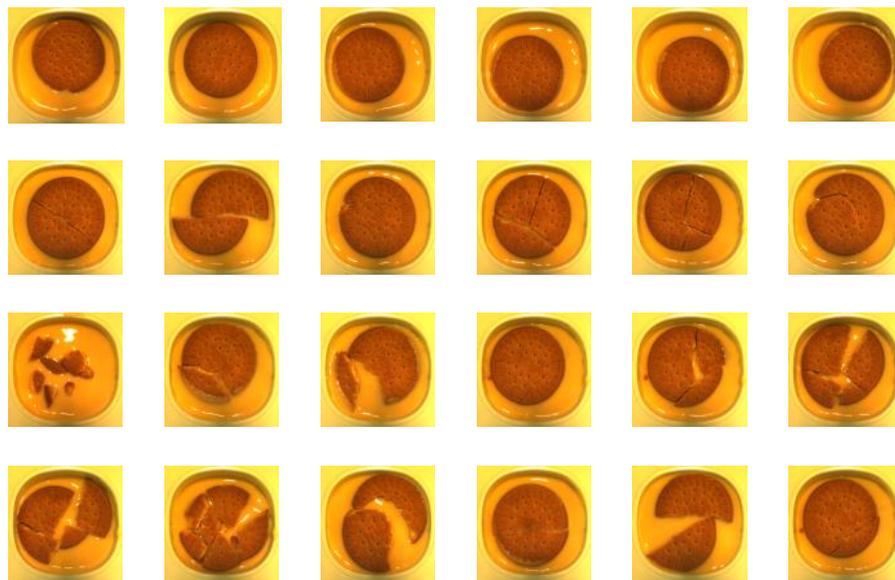


Figure. D.1. The first row corresponds to the positive cases (high quality samples) and the rest of the rows belong to some negative cases in the training data set.

APPENDIX F. ACRONYMS AND ABBREVIATIONS LIST.

AAE: Average Angular Error

ABPI: Ankle Brachial Pressure Index

ACM: Region-Based Active Contour Model

ADAS: Advanced Driver Assistance System

AI: Artificial Intelligence

AIx: Augmentation Index

ALU: Arithmetic Logic Unit

AOI: Area Of Interest

AOS: Additive Operator Splitting

ASIC: Application-Specific Integrated Circuit

ASM: Active Shape Models

BLIS: BLindspot Information Systems

BM: Block Matching

CARE: Community Road Accident Database

CCA: Common Carotid Artery

CR: Cyclic Reduction

DE: Differential Evolution

DMIPS: Dhrystone MIPS

DSP: Digital Signal Processor

FCM: Fuzzy C-mean

FPGA: Field Programmable Gate Arrays

FPS: Frame Per Second

FPU: Floating-Point Unit

FRS: Fast Radial Symmetry

GOPS: *Giga Operations per Second*

GPGPU: *General-Purpose Computing on Graphics Processing Units*

GPU: *Graphics Processing Units*

HDR: *High Dynamic Range*

ICOV: *Instantaneous Coefficient Of Variation*

ITREBA: *Procesamiento de Imagen/Video en Tiempo Real para Exploración Biomédica Activa*

KF: *Kalman Filter*

LDW: *Lane Departure Warning*

LIDAR: *Laser Imaging Detection and Ranging*

LIP: *Logarithmic Image Processing*

MIPS: *Million of Instructions Per Second*

MMX: *MultiMedia eXtension*

MOI: *Motion Of Interest*

MS: *Mass-Spring*

NCC: *Normalized Cross Correlation*

OpenCL: *Open Computing Language*

PCR: *Parallel Cyclic Reduction*

PWV: *Pulse Wave Velocity*

QCIF: *Quarter Common Intermediate Format*

RADAR: *Radio Detection and Ranging*

RHT: *Randomized Hough Transform*

RISC: *Reduced Instruction Set Computer*

RMSE: *Root Mean Squared Error*

SAD: *Sum of Absolute Differences*

SAPVIA: *Sistema Autónomo Programable de Visión Artificial*

SIB: *Sparse Integration Block*

SIMD: *Single Instruction, Multiple Data*

SIS: *Smart Image Sensor*

SPF: *Signed Pressure Force*

SRAD: *Speckle Reducing Anisotropic Diffusion*

SSE: *Streaming SIMD Extensions*

ST: *Similarity Transformation*

STD: *Standard Deviation*

SVM: *Support Vector Machine*

TDI: *Tissue Doppler Imaging*

TV: *Total Variation*

US: *UltraSound*

