

João Paulo Ribeiro de Oliveira

Caries Detection in Panoramic Dental X-ray Images



Universidade da Beira Interior

Departamento de Informática

August 2009

João Paulo Ribeiro de Oliveira

Caries Detection in Panoramic Dental X-ray Images



*Thesis submitted to the Department of Computer Science
for the fulfilment of the requirements for the degree of Master in Science
made under the supervision of Doctor Hugo C. Proença, Assistant Professor
at the Department of Computer Science of University of Beira Interior, Covilhã, Portugal*

Universidade da Beira Interior
Departamento de Informática
August 2009

Acknowledgements

I leave this section the special thank you to everyone who in some way contributed to the final result of this document.

First of all a special thanks to Professor Dr. Hugo Proença, for all the support and all the wisdom available to the achieving of this dissertation. Note also that always showed the willingness to answer questions and discuss the different decisions to make for the different steps of our dissertation. To Professor Dr. Hugo Proença a special and honorable thanks.

Special thanks also to Dr. Rui Conceição because without him it would not be possible the completion of this dissertation. Also thanks the availability provided to obtain an opinion from a specialist about certain decisions taken in our dissertation.

To my closest family, my mother Ana Bela Oliveira, my brothers José Miguel Oliveira and Luís Gonçalo Oliveira and to my grandmother Beatriz Ferreira. For all the support and strength throughout these years of university.

To my special friends Joel Carvalho, Orlando Pereira and Sílvio Filipe.

A word of appreciation to the research group of Soft Computing and Image Analysis Lab (Socia Lab.) for all the meetings where we discuss the work of all the members.

I thank my colleagues in the Socia Lab., Release and NMCG throughout the companionship and mutual provided in different times of the day. Leave a special word of appreciation to Gil Melfe.

The last but not the least to my beautiful girlfriend, Ana Teresa de Jesus Pereira. Which throughout all these years was my shelter and in the same time my fountain of strength. Thank you.

Thanks to all the people who were not particularly mentioned but still had an important role in my education.

Contents

Acknowledgements	iii
Contents	v
List of Figures	ix
List of Tables	xiii
Acronym	xv
1 Introduction	1
1.1 Main Goal	2
1.2 Motivation	3
1.3 Computer Vision	3
1.4 Block Diagram	5
1.5 Organization	6
2 State of the Art	9
2.1 Dental X-ray	9
2.1.1 X-ray	10
2.1.2 Stomatologic	11
2.1.3 Main Applications	12
2.2 Applications of Dental X-ray in Computer Science	13
2.2.1 Clinical Environments	13
2.2.2 Biometrics	14

2.2.3	Teeth Segmentation	15
2.2.4	Active Contours	17
2.2.4.1	Snakes	18
2.2.4.2	Geometric	19
2.2.4.3	Level Set Methods	19
2.2.4.4	Geodesic	19
2.2.4.5	Without Edges	20
2.3	Conclusion	22
3	Data-Set Images	23
3.1	Dental X-ray camera	23
3.2	Properties	23
3.2.1	Morphologic Properties	24
3.3	Images Statistics	25
3.3.1	Number of Teeth per Image	26
3.3.2	Average Size	27
3.3.3	Data Summary	28
3.4	Conclusion	29
4	ROI Definition and Jaws Partition	31
4.1	ROI Definition	31
4.2	Jaws Partition	33
4.3	Conclusion	35
5	Teeth Gap Valley Detection	37
5.1	Line Sum Intensities - Method 1	37
5.2	Center Distance with the Active Contours - Method Two	41
5.2.1	Morphological Operators	42
5.2.2	Divide the jaws	45
5.2.3	Polar Coordinates	45
5.2.4	Active Contours	46

5.2.4.1	Without Edges	48
5.2.5	Teeth Division	50
5.3	Conclusion	58
6	Tooth Segmentation and Dental Caries Detection	61
6.1	Tooth Segmentation	62
6.2	Dental Caries Detection	66
6.2.1	Dental Caries	67
6.2.2	Feature Extraction	67
6.2.2.1	Statistical Features based on the image properties . .	68
6.2.2.2	Region based features	70
6.2.2.3	Boundary or Border Tooth Features	74
6.2.2.4	Region Texture Features	80
6.2.3	PCA	83
6.2.3.1	PCA applied to our Features	84
6.3	Conclusion	84
7	Results	87
7.1	ROI Definition	88
7.2	Jaws Partition	90
7.3	Teeth Gap Valley Detection	91
7.3.1	Line Sum Intensities - Method 1	91
7.3.2	Center Distance with the Active Contours - Method 2	92
7.4	Tooth Segmentation	93
7.5	Dental Caries Detection	94
7.5.1	Data-Set Normalization	96
7.5.2	Results with our Tooth Segmentation	97
7.5.2.1	Artificial Neural Network	98
7.5.2.2	Support Vector Machine (SVM)	98
7.5.2.3	Naive Bayes	99
7.5.3	Results with the perfect Tooth Segmentation	100

7.5.3.1	Artificial Neural Network	100
7.5.3.2	Support Vector Machine (SVM)	102
7.5.3.3	Naive Bayes	102
7.5.4	Discussion	103
7.6	Conclusion	104
8	Conclusion	105
8.1	Contribute	105
8.2	Results	107
8.3	Motivation	107
8.4	Future Work	108
	References	111
	A Data Set of Panoramic Dental Radiographs for Stomatologic Image Processing Purposes	117

List of Figures

1.1	The Outline of our developed work.	6
2.1	All cases of the fitting curve, image extracted from [10].	21
3.1	Picture of the Orthoralix 9200 DDE X-ray camera.	24
3.2	Examples of images of the Dental X-ray data set.	26
3.3	Number of images concerning the quantity of teeth per mouth.	27
3.4	Average sizes and corresponding confidence intervals, regarding the type of teeth.	28
4.1	Four lengths extracted of our images data-set.	32
4.2	Histogram of the R_i values.	32
4.3	Some results for the Region of Interest (ROI) definition.	33
4.4	Horizontal Projection of the intensities, for our starting point.	34
4.5	Results of the polynomial least squares fitting.	35
4.6	Results of the polynomial least squares fitting in images with accentuated missing teeth area.	36
5.1	Example of the minimum intensities for one input image.	38
5.2	Example of our method 1 results.	40
5.3	Example of failure in the teeth gap valley detection of our method 1.	41
5.4	Example one of the top and bottom hat transform.	43
5.5	Example two of the top and bottom hat transform.	44
5.6	Example of the Jaws division.	45
5.7	The resulting polar coordinates transformation for each separated jaw.	47

5.8	The example mask of our initialization for the active contour method.	48
5.9	Output example of the active contour without edges method, applied in each jaw.	49
5.10	The distance to the center of the lower jaw and the corresponding $0.10 * h$ cut.	52
5.11	The distance to the center of the upper jaw and the correspond $0.10 * h$ cut.	53
5.12	The gaussian filter for the convolution with the distance vector. . . .	54
5.13	Convolution applied to the upper and lower jaw distance vector, based on the filter in 5.12.	55
5.14	Output of the minimums extraction step, applied to the upper a lower jaw, respectively.	56
5.15	An example of our method 2 for the teeth gap valley detection. . . .	57
6.1	Example of our input images for the tooth segmentation stage.	62
6.2	Example of our input images for the tooth segmentation stage after the Top and Bottom Hat transformation.	62
6.3	Example images of the top and bottom hat transform.	63
6.4	Some mask examples of our initialization for the active contour method.	64
6.5	Examples of the active contours method applied to our input images.	65
6.6	Examples of the active contours method applied to our input images with overlap teeth.	65
6.7	Examples of the active contours method applied to our input images with extra noise present in the output image.	65
6.8	The main types of connectivity analysis, image extracted from [35]. .	75
6.9	The order that our pixels are visited, this figure is extracted from [35].	76
7.1	Results containing extra noise from the ROI definition stage.	89
7.2	Example of cutting teeth from the ROI definition stage.	89
7.3	Results where some teeth are overlapped by the polynomial fitting in the Jaws Partition stage.	91

7.4	Examples of output images with the presence of teeth overlap segmentation.	95
7.5	Examples of output images with the presence of extra noise.	95
7.6	Receiver Operating Characteristic (ROC) for the Artificial Neural Network (ANN) classifier for the normalized test set between [0, 1] test and training set.	101
7.7	ROC for the ANN classifier for the normalized test set between [0, 1] after the appliance of the Principal Component Analysis (PCA) for a 99% variance coverage.	101

List of Tables

3.1	Properties of the Orthoralix 9200 imaging device.	25
3.2	Statistic date retrieve of our images data-set.	29
6.1	Results from appliance of the PCA in our training set.	84
7.1	Results of our method stages up to the Dental Caries Detection. . . .	88
7.2	Confusion matrix for the ANN classifier after the PCA appliance for 95% variance covering.	98
7.3	Confusion matrix for the Support Vector Machine (SVM) classifier after the PCA appliance for 95% variance covering.	99
7.4	Confusion matrix for the Naive Bayes classifier after the PCA appliance for 95% variance covering.	99
7.5	Confusion matrix for the ANN classifier for the normalized test set between [0, 1] test and training set.	100
7.6	Confusion matrix for the ANN classifier for the normalized test set between [0, 1] after the appliance of the PCA for a 99% variance coverage.	102
7.7	Confusion matrix for the SVM classifier for the test set normalized between [0, 1] after the appliance of the PCA for a 99% variance coverage.	102
7.8	Confusion matrix for the Naive Bayes classifier for the test set normalized between [0, 1] after the appliance of the PCA for a 99% variance coverage.	103
8.1	Resume table of our results accuracy for each method stage.	108

Acronym

ROI Region of Interest

Socia Lab. Soft Computing and Image Analysis Lab

AM Ante-Mortem

PM Post-Mortem

MRI Magnetic Resonance Imaging

CT Computerized tomography

GVF Gradient Vector Flow

BMD Bone Mineral Density

PCA Principal Component Analysis

SVM Support Vector Machine

VIImage'09 II Eccomas Thematic Conference on Computational Vision and
Medical Image Processing

FFT Fast Fourier Transform

MPP Minimum-Perimeter Polygons

IDM Inverse Difference Moment

KNN K Nearest Neighborhood

NN Neural Network

ANN Artificial Neural Network

ROC Receiver Operating Characteristic

FPR False Positive Rate

FNR False Negative Rate

TNR True Negative Rate

TPR True Positive Rate

MLP Multi Layer Perceptron

Chapter 1

Introduction

The detection of dental caries, in a preliminary stage are of most importance. There is a long history of dental caries. Over a million years ago, hominids such as Australopithecus suffered from cavities. Archaeological evidence shows that tooth decay is an ancient disease dating far into prehistory. Skulls dating from a million years ago through the Neolithic period show signs of caries [2]. The increase of caries during the Neolithic period may be attributed to the increase of plant foods containing carbohydrates [39]. The beginning of rice cultivation in South Asia is also believed to have caused an increase in caries.

Dental Caries, also known as dental decay or tooth decay, is defined as a disease of the hard tissues of the teeth caused by the action of microorganisms, found in plaque, on fermentable carbohydrates (principally sugars). At the individual level, dental caries is a preventable disease. Given its dynamic nature the dental caries disease, once established, can be treated or reversed prior to significant cavitation taking place. There three types of dental caries [59], the first type is the Enamel Caries, that is preceded by the formation of a microbial dental plaque. Secondly the Dentinal Caries which begins with the natural spread of the process along the natural spread of great numbers of the dentinal tubules. Thirdly the Pulpal Caries that corresponds to the root caries or root surface caries.

Primary diagnosis involves inspection of all visible tooth surfaces using a good light source, dental mirror and explorer. Dental radiographs (X-rays) may show dental caries before it is otherwise visible, particularly caries between the teeth.

Large dental caries are often apparent to the naked eye, but smaller lesions can be difficult to identify. Visual and tactile inspection along with radiographs are employed frequently among dentists. At times, caries may be difficult to detect. Bacteria can penetrate the enamel to reach dentin, but then the outer surface maybe at first site intact. These caries, sometimes referred to as "hidden caries", in the preliminary stage X-ray are the only way to detect them, despite of the visual examination of the tooth shown the enamel intact or minimally perforated. Without X-rays wouldn't be possible to detect these problems until they had become severe and caused serious damage.

There are three main types of dental X-rays to perform the dental caries detection and other diseases [34]:

- **The panoramic dental X-ray:** Panoramic X-rays show a broad view of the jaws, teeth, sinuses, nasal area, and temporomandibular (jaw) joints. These X-rays do show problems such as impacted teeth, bone abnormalities, cysts, solid growths (tumors), infections, fractures and dental caries.
- **The bitewing dental X-ray:** Bitewing X-rays show the upper and lower back teeth and how the teeth touch each other in a single view. These X-rays are used to check for decay between the teeth and to show how well the upper and lower teeth line up. They also show bone loss when severe gum disease or a dental infection is present.
- **Periapical X-rays:** Periapical X-rays show the entire tooth, from the exposed crown to the end of the root and the bones that support the tooth. These X-rays are used to find dental problems below the gum line or in the jaw, such as impacted teeth, abscesses, cysts, tumors, and bone changes linked to some diseases.

1.1 Main Goal

The main goal of the work presented is to detect dental caries in panoramic dental X-ray images. Based on the panoramic input X-rays, our proposal is to mark infected teeth present in the image.

1.2 Motivation

The main motivation for this research work is the fact that is included in an area of great interest to the scientific community. Where increasingly the computer vision is become a part of our daily life, whether for security reasons as for reasons of leisure. The other major motivation is that this is an innovative work because there is not a complete case study in the literature as described in this document. In literature there are only some of the steps we implemented in our work. This work will be possible to the scientific community to be a basis for other methods or to the improvement of our method. Which method concerns the detection of dental caries in the panoramic X-ray images.

1.3 Computer Vision

Computer vision is the science e and technology that correspond to the vision of machines. As a scientific discipline, computer vision is the process to obtain information from images by building a artificial system. This images data can be presented to the system in many forms, the most popular views from a single camera, video sequence, views from multiple cameras, or multi-dimensional data from a medical scanner.

The classical problem in computer vision is that of determining whether or not the image data contains some specific object, feature, or activity. This task can normally be solved robustly and without effort by a human, but is still not satisfactorily solved in computer vision for the general case: arbitrary objects in arbitrary situations. The existing methods for dealing with this problem can at best solve it only for specific objects, such as simple geometric objects, human faces, printed or hand-written characters, or vehicles, and in specific situations, typically described in terms of well-defined illumination, background, and pose of the object relative to the camera in [15][55]. There are a variety of recognition problem, such as:

- **Recognition:** This problem consists in the recognition of objects that were pre-specified or learned by the system.
- **Identification:** As the name suggests the main goal in this problem is to

identify individuals, this can be fulfilled by the use of the individual face, fingertips, teeth contour, etc.

- **Detection:** Our work fits in this recognition problem, because the main goal is to detect, in the image, regions containing dental caries. In this case the purpose of capturing the images, is effectively limited for detection of dental caries.

The organization of a computer vision system is highly application dependent. Some systems are stand-alone applications which solve a specific measurement or detection problem, while others constitute a sub-system of a larger design which, for example, also contains sub-systems for control of mechanical actuators, planning, information databases, man-machine interfaces, etc. The specific implementation of a computer vision system also depends on if its functionality is pre-specified or if some part of it can be learned or modified during operation. There are, however, typical functions which are found in many computer vision systems [13][55][18]:

- **Image Acquisition:** A digital image is produced by one or several image sensors, which, besides various types of light-sensitive cameras, include range sensors, tomography devices, radar, ultra-sonic cameras, etc. Depending on the type of sensor, the resulting image data is an ordinary 2D image, a 3D volume, or an image sequence. The pixel values typically correspond to light intensity in one or several spectral bands (gray images or color images), but can also be related to various physical measures, such as depth, absorption or reflectance of sonic or electromagnetic waves, or nuclear magnetic resonance.
- **Pre-Processing:** This stage is performed before a computer vision method can be applied to image data in order to extract some specific information, it is usually necessary to process the data in order to assure that it satisfies certain assumptions implied by the method. For example noise reduction in order to assure that it does not introduce false information to the system, contrast enhancement to assure that relevant information can be detected and scale-space representation to increase relevant information in appropriate scales.
- **Feature Extraction:** Image features at various levels of complexity are extracted from the image data. There are two types of feature extraction, the low-level that corresponds to the basic features that can be extracted automatically

from an image without any shape information. And the high-level feature extraction concerns finding shapes in computer images [35].

- **Detection/Segmentation:** In computer vision, segmentation refers to the process of partitioning a digital image into multiple segments (sets of pixels) . The goal of segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyze [57]. Image segmentation is typically used to locate objects and boundaries (lines, curves, etc.) in images. More precisely, image segmentation is the process of assigning a label to every pixel in an image such that pixels with the same label share certain visual characteristics.
- **High-level processing:** In this step the input is generally a data-set containing a specific object. This is based on the features that describe the region. For example, in a process to detect diseases, the input in this step are the features that best fit on the disease classification. Associated to this step is the pattern recognition [19][1] that consists in the classification of data, composed by patterns. The methods to perform pattern recognition are divided in two main groups, supervised learning where we know from the start what are the categories of the input samples. The other group is the opposite, corresponding to the unsupervised learning, in this case the main goal is to divide in one, or more categories the input data.

1.4 Block Diagram

Our work consists in three different main stages with three sub stages as show in figure 1.1. The first stage is based on a statistically evaluation of the images to define a preliminary ROI taking out the nasal and chin bones. In the sub stage we make the detection of the upper jaw and the lower jaw, based on the extraction of primitive points between jaws and in a polynomial fitting process. The second stage is the teeth gap valley detection where firstly we detect the gap valleys with the active contour without edges technique [10], this process is apply in each jaw in polar coordinates. Secondly in this stage we divide each teeth with the extraction of the minimums points that correspond to the minimums points in the extracted contour of the previous step. Finally the third stage that consists in a first step that

segment each tooth extracted of the previous stage. In this first stage is also used the active contours to segment the tooth border. The second step of this final stage is the extraction of dental features to the classification of dental caries. A great variety of features were used, statistical, image characteristics, region based features, texture based features and features based on the tooth border.

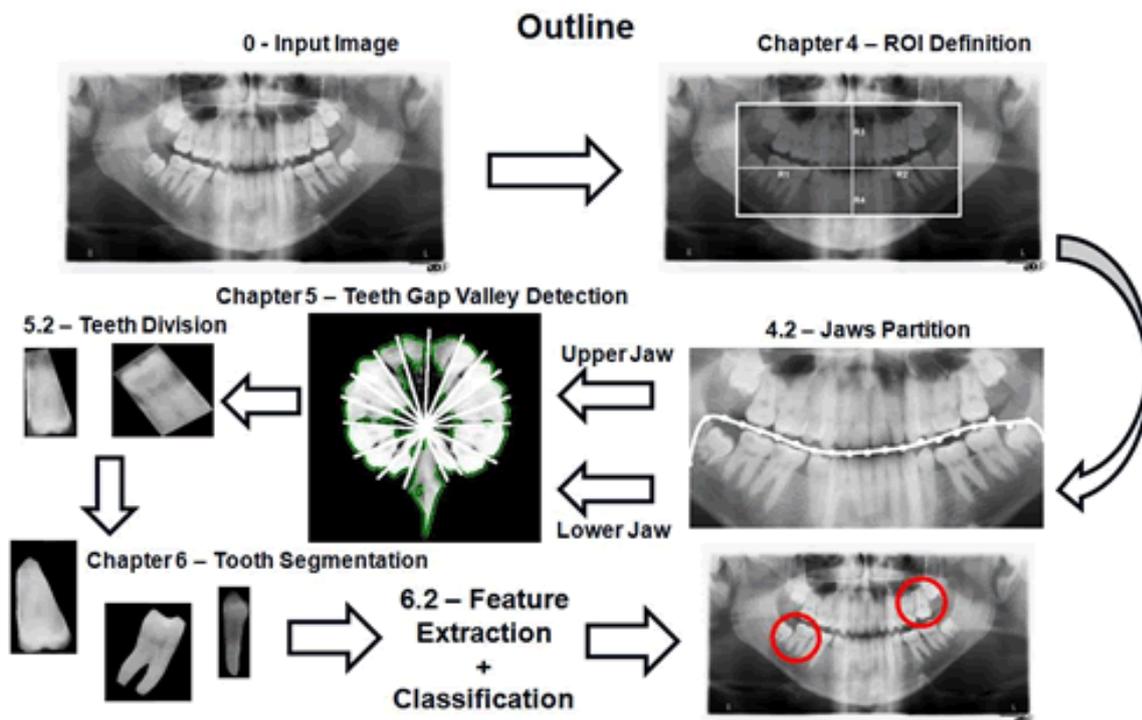


Figure 1.1: The Outline of our developed work.

1.5 Organization

This document is organized as follows.

Chapter 1 - Introduction Guidelines of the work developed for the dissertation.

Chapter 2 - State of the Art In a first stage of this chapter is introduced the fundamental concepts of Dental X-Rays and its applications. Secondly is presented part of the work developed in the literature that is important for this work.

Chapter 3 - Data-Set Images The presentation of our data-set images, and all the statistical analysis retrieved from the data-set images.

Chapter 4 - ROI Definition and Jaws Partition The first stage of our method, composed by the ROI definition where based on the extraction of four lengths starting in the center of the image to the four corners of the mouth. Finally the jaws partition process. This process consists in the retrieval of 21 primitive points between the jaws. Based on this extracted points we perform the connection between them with the polynomial fitting process.

Chapter 5 - Teeth Gap Valley Detection Second stage of our method, constitute by the teeth gap valley detection. In a first version and based on the output of the previous method we perform the extraction of the line sum intensities for each polynomial point. In a second version of our method, we use the active contours technique to perform the segmentation in the polar coordinates. Through this segmentation and based on the distances to the center of the image for each boundary tooth is extracted the border point that had the lowest distance value. Finally we perform the teeth division.

Chapter 6 - Dental Caries Detection Third and final stage of our method, representing the tooth segmentation and the features extraction plus the dental caries classification. For the tooth segmentation we use the active contour technique for better improve the extracted region of the previous method. For our final stage method is performed the dental caries feature extraction. Based in this extracted features and with the use of multiples classifiers we perform the dental caries detection in the input image containing only one tooth.

Chapter 7 - Results In this chapter is presented all the results for all the different stages of the method.

Chapter 8 - Conclusion Presentation of a few more remarks about the dissertation and the future work for the presented work.

Chapter 2

State of the Art

In this chapter is presented the fundamental concepts of Dental X-rays and its main applications. The second section of this chapter describe the segmentation methods that were developed using bitewing dental X-ray, in most cases for the individual identification and in other hand for the prediction of bone loss. Finally a small survey for the active contours technique, describing the active contours Snakes and the active contours Without Edges.

2.1 Dental X-ray

Dental X-rays are pictures of the teeth, bones, and soft tissues around them to help find problems within the teeth, mouth, and jaw. X-ray pictures can show cavities, hidden dental structures, and bone loss that cannot be seen during a visual examination. Dental X-rays may also be done as follow-up after dental treatments. As developed in the chapter 1 there are three main different types of dental X-ray. The bitewing, periapical and the panoramic. A full-mouth series of periapical X-rays are most often done during a person's first visit to the dentist. Bitewing X-rays are used during checkups to look for tooth decay. Panoramic X-rays may be used occasionally. Dental X-rays are scheduled when you need them based on your age, risk for disease, and signs of disease [34].

2.1.1 X-ray

In the 80s a German physicist (a person who studies the relationship between matter and energy) Wilhelm Conrad Roentgen (1845-1923) discovered X-rays in 1895 while he was experimenting with electricity. Because he did not really understand what these rays were, he called them X-rays because in mathematics X stands for the unknown. By 1900, however, doctors were using X-rays to take pictures (called radiographs) of bones, which helped them treat injuries more effectively. But scientists also discovered that overexposure to X-rays could cause burns, even death. In 1901, Roentgen was awarded the first Nobel Prize for physics for his discovery of this short-wave ray [52]. After this breakthrough grew a legitimate medical applications and practices, along with a new X-Ray industry. Hospitals, clinics, and private clinics became the center for one of the first functional, widespread medical advancements in years. The X-Ray set new precedents, especially in regard to training and certification of radiologists.

After the initial infatuation was over, the effects of the unknown ray were revealed. The early scientific pioneers were disfigured, cancer-ridden, and in constant pain. Yet explanations were still not available. The discovery of radium eventually answered the riddle of the X-Ray. Radioactivity was found to be the powerful side-effect causing the many ailments.

The X-Ray affected not only the medical industry, its influence was felt throughout society, culture, and the military. Culturally, the X-Ray was embraced in literature, art, and the emerging television and motion picture industries. However, it was the military's use of the invention in the World Wars, not through these pop culture venues, that ensured the American public's acceptance and trust in the X-Ray.

Today, the X-Ray has helped to created entire specialized fields within the medical industry. Almost everyone has been touched by the X-Ray in some way, whether it is the middle-aged woman having her yearly mammogram or the experienced business man having his luggage inspected before his flight to Europe, American culture has absorbed this invention into its everyday fabric. The frequency of today's scientific developments along with the increasingly pervasive world of computer technology creates an optimistic, almost limitless future for the use of the mistakenly discovered little ray.

2.1.2 Stomatologic

Is the branch of medicine that relates to the mouth and its diseases, originally practiced by physicians, it was a standard medical specialty through the early 20th century. The specialty is defined within Europe under the Medical Directive 2001/19/EC.¹ Stomatology studies diagnostics, treatment and prevention of diseases affecting teeth, oral cavity and tissues and organs which are topographically associated with it. Stomatology services are provided mainly in the form of an outpatient care, just a small part of the care is provided by inpatient stomatological facilities. Today's stomatology is a field that employs exclusively university-educated professionals, i.e. doctors after graduation from five to six years long studies at a university. A doctor-stomatologist's coworkers are health services staff: a nurse, a dental technician, an X-ray technician, and a dental hygienist.

Among the basic stomatological fields there are therapeutic stomatology, orthopedic stomatology, and surgical stomatology. Therapeutic stomatology (protective, conserving stomatology) deals with the diagnostics, treatment and prevention of a dental decay and its complications. Associated with this basic stomatology branch there are: children's stomatology (pedodontics) that deals with the care of the milk dentition or the developing permanent dentition of youngsters.

Periodontics deals with diseases of the periodontium tissues and the oral cavity mucous membrane diseases. This type of diseases are mainly detect in periapical dental X-ray.

Orthopedic stomatology (dental prosthetics) deals with the replacement of parts of crowns, individual teeth losses or provides for the total replacement of lost teeth by the production and application of various dental prostheses (crowns, bridges, removable dentures). An individual specialty is orthodontics (orthopedics of jaws) that deals with the diagnostics, treatment and prevention of irregularities of the individual teeth, groups of teeth and anomalies of jaws.

Surgical stomatology deals with the surgical treatment of the oral cavity diseases (dentoalveolar surgery) or as a specialty (maxillofacial surgery) provides for surgical treatment of larger oral diseases, mainly in the form of the inpatient care.

A graduate of the stomatology studies is prepared both theoretically and prac-

¹Medical Directive 2001/19/EC <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2001:206:0001:0050:EN:PDF>

tically for the praxis in prevention and cure in the basic stomatology fields. He or she receives only general knowledge in the specialized disciplines that allow for a responsible decision of a consequent treatment at highly specialized dental offices. As a graduate student he or she can receive a higher degree of qualification after passing necessary examinations and continue to work as a specialist in orthodontics or maxillofacial surgery fields [3].

2.1.3 Main Applications

There are several application in nowadays for the X-ray, for our work and for the most important area of application, we will only describe the importance it has in the medical field. In this area X-ray interpret a important role for the prevention and detection of innumeracy diseases. Since Röntgen's discovery that X-rays can identify bony structures, X-rays have been developed for their use in medical imaging. Radiology is a specialized field of medicine. Radiologists employ radiography and other techniques for diagnostic imaging.

X-rays are specially useful in the detection of pathology of the skeletal system, but are also useful for detecting some disease processes in soft tissue. Some notable examples are the very common chest X-ray which can be used to identify lung diseases such as pneumonia, lung cancer or pulmonary edema, and the abdominal X-ray which can detect intestinal obstruction. X-rays may also be used to detect pathology such as gallstones (which are rarely radiopaque) or kidney stones (which are often visible, but not always). Traditional plain X-rays are less useful in the imaging of soft tissues such as the brain or muscle. Imaging alternatives for soft tissues are computed axial tomography (Computerized tomography (CT)), Magnetic Resonance Imaging (MRI) or ultrasound. The use of X-rays as a treatment is known as radiotherapy and is largely used for the management (including palliation) of cancer, it requires higher radiation energies than for imaging alone.

X-rays are relatively safe investigation and the radiation exposure is low. But in pregnant patients, the benefits of the exposure process should be balanced with the potential hazards to the unborn fetus.

In stomatology X-rays also have an important role, the use of this technique provide the early detection of dental diseases, that prevent in some cases more several futures complications. The detection of dental caries is most common use

of the dental X-ray.

In Biometrics, the use of dental X-ray is also of great importance to identify deceased individuals. Here, having access to the Ante-Mortem (AM) dental X-ray and to the Post-Mortem (PM), the matching of the teeth border can positively identify these deceased individuals, later in this chapter will demonstrate the techniques used by authors to develop this task. When referring to natural disasters, and in other cases of the individual totally nonrecognition, dental X-ray is the only way to identify those individuals.

2.2 Applications of Dental X-ray in Computer Science

Automated image analysis processes have been achieving higher relevance for many purposes and the results can be considered satisfactory (e.g., biometrics and multiple types of medical image diagnosis). In the specific area of medical image processing these automated systems constitute a valuable tool as the earliest detection of many diseases and in the preprocessing of huge amounts of data.

2.2.1 Clinical Environments

Dental X-ray images also are used in clinical environments, in the detection/prediction of Bone Mineral Density (BMD) for the individuals identification that suffers of osteoporosis. In [58][30] this process is based on really medical data, there is no presence of more sophisticated methods or algorithms to be applied in the input images. The measures that the authors use to perform the prediction are based on the trabecular pattern on dental X-ray that can be used to predict BMD.

In a more sophisticated process, in [29], the authors propose a framework of support for dental clinics, through the segmentation of dental X-ray images. The segmentation contains two stages, the stage of training in which they manually select images representative of the whole. These images are segmented based on a hierarchy of regions of detectable "level set" [47][48][31]. It is then extracted using the characteristics of PCA [26][5] and those results are used to train SVM [41][14]. The segmentation of the new images are classified, first with the SVM. The classifier provides initial contours close to correct, that for diseases in which the article focuses

(Periodontics, Chronicle periapical and Periodontal Bone Loss). Finally the maps formed by the previous step is applied a scheme of analysis to the images.

2.2.2 Biometrics

Dental X-ray are used in biometrics for the identification of deceased individuals [24]. This process is based on the matching of the dental X-ray AM with the dental X-ray PM. In [12] the tooth contour is extracted for the matching of dental radiographs. This process is done by the directional Snake [22]. Three main stages that are performed, the initialization, the converge to the gradient and a fine adjustment in the end of the process. In the initialization the authors initialize the snake, for that is perform the gumline detection. The gumline is the "visual line" that separates the root area from the crow area. The method uses the property that there is an intensity increase at the gum lines from the crow area to the root area. The Gradient Vector Flow (GVF) field of the edges detected by the Canny operator is used as the external energy for the converge gradient stage. In the final stage the fine adjustments consists in the propriety that true pixels boundaries always had in their neighbor pixels with lower intensities. The external energy is defined on the 2.1, where $E_{ext,1}$ is the external energy of the teeth boundaries, $E_{ext,2}$ corresponds to the intensity image and finally ω that controls the trade-off between the gradient and the intensity.

$$E_{ext} = E_{ext,1} + \omega E_{ext,2} \quad (2.1)$$

The results of this paper are quite interesting, but it is to highlight the fact that the authors simply apply the method to bitewing dental X-ray images. So therefore the presence and position of the teeth is more accurate.

The matching of dental X-ray images for human identification, in [23], is based on the gap valley detection, tooth isolation, the contour extraction (crown and root contour extraction) and finally the matching of the contours. This method is initialize with a user interaction, where is ask to the user to mark a point between jaws. After that the horizontal projection of the intensities is calculated, where it's hope that near the point that the user introduced will be a gap intensity valley. Based on this point and using the probability function defined by 2.2.

$$p_{v_i}(D_i, y_i) = p_{v_i}(D_i)p_{v_i}(y_i), \quad (2.2)$$

Where

$$p_{v_i}(D_i) = c(1 - \frac{D_i}{\max_k D_k}), \quad (2.3)$$

and

$$p_{v_i}(y_i) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(y_i - \mu)^2 / \sigma^2}. \quad (2.4)$$

A set of points were extracted and later connected with a spline function [51] into a smooth curve. After that the teeth gap valley detection is based on the sum of the intensities done by the perpendicular lines starting in the spline function points extracted. The tooth isolation is obtained thorough the previous step, i.e., two extracted consecutive perpendicular lines form a single tooth between them. The contour extraction is based on the probability of each pixel in the ROI, in this case corresponds to the tooth isolated, belonging to the ROI, starting from the center of the region targeted in the previous steps. Finally is perform the contour matching for the human identification. This article was one of the articles in which we rely to a first method for the teeth segmentation. The input images were also the bitewing dental X-ray, and despite the good results obtained by the authors, this method is very dependent on the user point initialization, a wrong initialization has the consequence of not segmented any tooth.

Many authors have developed work in the identification of individuals using the dental X-ray images [25][45][43][44][69]. However all the work input images, are the bitewing and periapical dental X-ray. For our work this literature served us for a better way to get to know the properties that the X-ray images have and turn immediately exclude some methods for the realization of our ultimate goal.

2.2.3 Teeth Segmentation

The image segmentation problem is one of the most difficult tasks in image processing and it plays a important role in most subsequent image analysis, especially in pattern recognition and image matching. Segmentation means partitioning an image into its constituent regions and extracting the objects of interest. However there is hardly any image segmentation technique that performs well in all problems. In addition, the performance of a segmentation technique is greatly affected by noise embedded in images.

In [53], authors perform the teeth segmentation in digitalized dental X-ray films

using mathematical morphology. This article is a new approach to the problem of identification of PM through dental X-ray images. For that this article proposes a mathematical morphological method for segmentation of teeth and a decrease in processing of contrast in grayscale to improve the segmentation problem. The basis is the use of a mathematical model capable of reducing the errors inherent in segmentation, errors that come from the background very noisy and complex. The use of this method is due in large part because of its ability to draw various shapes and structures of the images. Morphological filters [33][42] used in large part to all sorts of problems for the extraction of features, representation of edges, detection of edges and recognition of them. In this article the ROI is a rectangular area in the original image containing a tooth. The extraction of the ROI in this article is divided as follows:

1. Initial extraction of the largest possible number of teeth of the image.
2. Operating in bitewing and periapical views.
3. In the worst case can at least extract one tooth of each image.

To perform the segmentation authors divide the grayscale range in three distinct regions:

1. Brighter Areas - which defines most teeth found in X-ray images.
2. Middle Areas - which defines the areas of most bones present in dental X-ray images.
3. Darker Areas - which defines the area corresponding to the background.

The segmentation process is divided in four principal stages, firstly the internal noise filtering that consists in the teeth gap valley detection and in the jaws partition. This step uses both horizontal and vertical projection of the sum intensities for the detection of the gap valley between the upper jaw and the lower jaw and for the detection of the bones and gap valleys between teeth, respectively. Based on the previous three distinct range regions of the grayscale input image. One important operation by the authors is the pre process perform in the images. They apply the closing top-hat transform, which is defined by subtracting the image from its morphological closing, this step reveal itself as very important for the emphasize

of darker pixels that are on the surrounding bright areas. The second step is the threshold operation, where the main goal is to divide the teeth of the background and the dental bones. Because of the shading effect present in the dental X-ray images, the extraction of only one threshold is not advised. Therefore the extraction of three thresholds is performed, based on the cumulative histogram of the filtered image. Thirdly after applying the threshold technique, the connected components labeling [18][15] is performed, in this case there are four different types of connected components, the teeth that are considered as ROI, more than one teeth, due to the overlap, connected components corresponding to the background or bones and finally corresponding to parts of the teeth, such as root or crowns. The final step is the refinement, in this step is carry out the selection of the best candidates of the connected components result. In this case the authors define two main rules:

- If two or more qualified ROI are generated from the three different thresholds for the same tooth, is return the union of the three regions.
- In the case of only two are generated is return the union of the two ROI if both centroid regions of each one are presented in each region.

In terms of results for the teeth segmentation, this article proof to have better results. Note that some of the pre-processing operations performed by the authors were used by us, despite of the readjust necessary applied on the structuring element for our input images.

Even within the segmentation of teeth images in dental X-ray, other authors [66][32] performed work in this area, and we emphasize the most important.

2.2.4 Active Contours

Active contours [8] technique main goal is to delineating an object in a noisy input image. For that the goal is to minimize the boot external and internal energy functions, the external energy function is the energy perform by the exterior region of the contour, and the internal energy function is the energy generated by the region in the interior of the contour. The stop criteria is when the minimal energy is detected. In this case we should be in a presence of the object boundaries, because is in this region where is the supposed minimal value of the energy, ie, the contour that best approximates the perimeter of an object. It can be compared to a rubber

band of arbitrary shape that is capable of deforming during time, trying to get as close as possible to the object boundaries.

There are a variety of techniques that stem from the fundamental idea of the active contours models, we here list the most important and explain which are the main differences between them.

2.2.4.1 Snakes

For starter the Active Contour Snake [22][37][22], this technique is one of the pioneers with respect to active contours. In the case of snakes the initialization is of most importance, the initialize contour should be place near the object in the image. This contour must contain the entire object that we want to segment. During the iterative process, the snake is attracted towards the target contour by various forces that control the shape and location of the snake within the image. The energy function used in generally is a sum of several terms, each corresponding to some force acting on the contour, the suitable energy function is the sum of the following three terms, as shown in the equation 2.5.

$$E = \int (\alpha(s)E_{cont} + \beta(s)E_{curv} + \gamma(s)E_{image})ds. \quad (2.5)$$

Where the parameters α , β and γ control the relative influence that each energy has. Each energy terms corresponds to a different proposal:

- E_{cont} : it forces the contour to be continuous.
- E_{curv} : the term that controls the smooth of the contour.
- E_{image} : it attracts the contour toward the closest image edge.

The E_{cont} and the E_{curv} are denominated as the internal energy terms and E_{image} is the external energy term.

In conclusion the Snake does not guarantee convergence to the global minimum of the energy function and it works fine is the initial contour contains the object that we want to segment.

2.2.4.2 Geometric

Geometric active contours [28][50] were introduced as the solution to the problem of required topological changes during curve evolution. Modifications and enhancements have been added to change their behavior or improve their performance in a variety of applications. The theory that is on the base of this technique, is the curve evolution and the level set method. In this technique, curves evolve using only geometric measures, resulting in a contour evolution that is independent of the curve's parameterizations. This avoids the need to repeatedly parameterize again the curve or to explicitly handle topological changes. The parametric representations of the curves themselves are computed only after the evolution of the level set function is complete [7][67].

2.2.4.3 Level Set Methods

The level set methods [61][54] was devised by Osher and Sethian [46], as a simple and versatile method for computing and analyzing the motion of an interface. The advantage of the level set method is that one can perform numerical computations involving curves and surfaces on a fixed cartesian grid without having to parameterize these objects. This technique is the generalization of the previous two active contours technique with regard to the non-dependence on parametrization.

2.2.4.4 Geodesic

Although we have a level-set description of the function, we still do not have a meaningful analytical solution. Classical active contours including snakes and level sets have shown promise in a range of medical image segmentation problems. The geodesic active contour [60][49][7] technique has been shown to be simple, efficient and relatively accurate. These methods typically use a variational framework to obtain locally minimal contours by gradient descent of an energy functional. As a result the final segmentations are dependent upon their initialization, requiring a simpler and less reliable segmentation as input. They also have a tendency to leak through gaps in object edges due to noise or indistinct boundaries and may become caught in irrelevant local minima. In essence we are going to join the above 2 approaches into a single concept.

2.2.4.5 Without Edges

Based on techniques of curve evolution, Mumford-Shah [40] functional for segmentation and level sets. The active contour without edges model [10] can detect objects whose boundaries are not necessarily defined by the gradient. Through the minimization of the energy which can be seen as a particular case of the minimal partition problem. In the level set formulation, the problem becomes a "mean-curvature flow"-like evolving the active contour, which will stop on the desired boundary. However, the stopping term does not depend on the gradient of the image, as in the classical active contour models, but is instead related to a particular segmentation of the image. The authors give a numerical algorithm using finite differences.

The active contours without edges is a method of the minimization of an energy based-segmentation. Assuming that a image ω_0 is composed by two related regions of approximated constant intensities, of distinct values, ω_0^i and ω_0^j . The object that we want to segment is represented by the values of ω_0^i . Its boundary is denoted by C_0 . Therefore when we are inside the object we have $\omega_0 \approx \omega_0^i$ or inside the object boundary (C_0), and $\omega_0 \approx \omega_0^j$ outside the object boundary. Defining the fitting term, given by the equation 2.6 extracted from the [10].

$$F_1(C) + F_2(C) = \int_{inside(C)} |\omega_0(x, y) - c_1|^2 dx dy + \int_{outside(C)} |\omega_0(x, y) - c_2|^2 dx dy \quad (2.6)$$

Where C is the variable curve, and the constants c_1, c_2 , depending on C , are the intensities averages of ω_0 inside the C and respectively outside C . In this case the fitting term is given by the boundary of the object C_0 , as shown in equation 2.7 extracted from the [10].

$$\inf_C \{F_1(C) + F_2(C)\} \approx 0 \approx F_1(C_0) + F_2(C_0). \quad (2.7)$$

This can be observed for example, if the curve C is outside the object, then $F_1(C) > 0$ and $F_2(C) \approx 0$. If the curve C is inside the object, then $F_1(C) \approx 0$ but $F_2(C) > 0$. If the curve C is both inside and outside the object, then $F_1(C) > 0$ and $F_2(C) > 0$. Finally, the fitting energy is minimized if $C = C_0$, i.e., if the curve C is on the boundary of the object, this restrains are shown in the example figure 2.1 extracted from the [10].

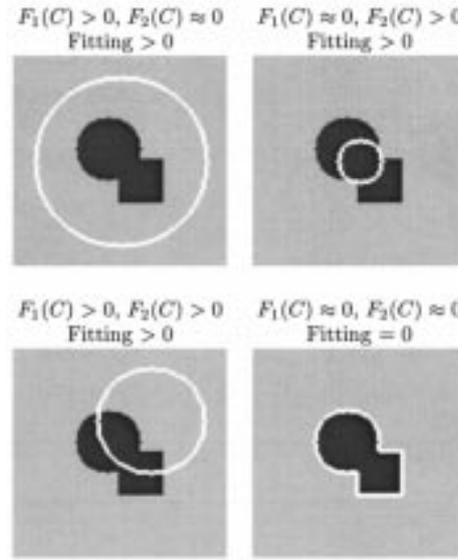


Figure 2.1: All cases of the fitting curve, image extracted from [10].

The authors [10] minimize the fitting term and they also add some regularizing terms, like the length of the curve C and the area of the region inside C , therefore the energy functional is defined by $F(c_1, c_2, C)$, as shown in the equation 2.8, extracted from the [10].

$$\begin{aligned}
 F(c_1, c_2, C) = & \mu * Length(C) + \nu * Area(inside(C)) \\
 & + \lambda_1 \int_{inside(C)} |\omega_0(x, y) - c_1|^2 dx dy \\
 & + \lambda_2 \int_{outside(C)} |\omega_0(x, y) - c_2|^2 dx dy
 \end{aligned} \tag{2.8}$$

Where $\mu, \nu \geq 0, \lambda_1, \lambda_2 > 0$ are fixed parameters.

This technique of active contours is used in our work in two different stages. In [56], the authors uses this method for the tooth segmentation in dental X-ray images, where the input image contains only a single tooth. The results proof why in our work we choose the active contours without edges for the extraction of the points between the teeth gap valleys, and in a later stage to perform the tooth segmentation.

With the active contour without edges the initialization of the contour cannot include totally or partially the object that we want to segment, due to is properties active contour without edges aren't depending on the initial contour.

2.3 Conclusion

In this chapter is presented the fundamental concepts of Dental X-rays and its main applications. In the second section we describe the importance of the work in the literature for our work. In this case there is not present in the literature a work that we could base our entire work, therefore is presented some work that partially can be used in our investigation work. Finally a small survey of the active contours methods that exists in the literature, and are in a certain way the most relevant in this topic. Next chapter is presented the properties of our data set images, and all relevant statistical data that we extract from the data-set.

Chapter 3

Data-Set Images

In this chapter we present our data-set images, the dental X-ray camera that captured our data-set images, some notes relative to morphology of the images, and their main properties, both statistical and structural.

3.1 Dental X-ray camera

The data set images were all captured by an Orthoralix 9200 DDE X-ray camera 3.1, which properties are described on the table 3.1.

3.2 Properties

There are a total of 1392 grayscale images in the data set, with varying types of dental structure, sizes of the mouth and number of teeth per image, as can be seen in figure 3.2. The gray levels of each image were stretched to the $[0, 255]$ scale, although both normalized and raw images will be soon available for download in the site of our research group Socia Lab.¹

Another major point of interest of this data set is the set of maps that enable the manual detection and localization of dental cavities - for the moment - and other diseases in a near future. This is achieved by corresponding a binary image to each dental X-ray image thereby showing the regions with dental cavities. This will turn

¹Socia Lab - Soft Computing and Image Analysis Laboratory, <http://socia-lab.di.ubi.pt/>



Figure 3.1: Picture of the Orthoralix 9200 DDE X-ray camera.

the data set into a preferable tool in the evaluation of method to perform automatic detection and localization of the respective diseases. Also, we believe that this set of images is useful to evaluate the current teeth segmentation methods as shown in 2.2.3.

3.2.1 Morphologic Properties

When compared with other types of stomatology images, radiographic images are highly challenging, due to several reasons that increase their heterogeneity:

1. Different levels of noise, due to the moving imaging device that captures a global perspective of the patient's mouth.
2. Low contrast, either global or on local regions of the images, the topology and morphologic properties are very complex.
3. The blurring that difficult the straightforward detection of edges.

Properties	Description
Power Supply	115-250 VAC pm 10%
Frequency	50/60 Hz pm 2 Hz
Maximum Line Power Rating	10A at 250V, 20A at 115V
Anode Voltage	60-84 kV, in 2kV steps
Anode Current	3-15mA, in 1mA steps
Exposure Time	12s for standard pan, 0.16-2.5 s for Cephalometric
Duty Cycle	1:20 at full power operation
Focal Spot	0.5 mm, IEC 336 (1993)
Vertical Reach	39" to 71" (100 to 180 cm) from floor to occlusal plane
Weight	410lbs (115 kg), 467lbs (212 kg) with Ceph arm
Active Area CCD Sensor	147x6m (Standard Pan), 220x6m (Ceph)
Image Size	1536x2725 pixels (Standard Pan), 2304x2529pixels (Ceph-Maximum)

Table 3.1: Properties of the Orthoralix 9200 imaging device.

4. The noise originated by the spinal-column that covers the frontal teeth in some images, as shown in figure 3.2.

3.3 Images Statistics

The statistical study of the input images, consists in the extraction of some statistical features that are important and relevant to our work for our images data-set. With this study we can assume that some rules in the future steps of our method, and in some cases this type of empirical data could reveal themselves to be very important to some steps in our work.

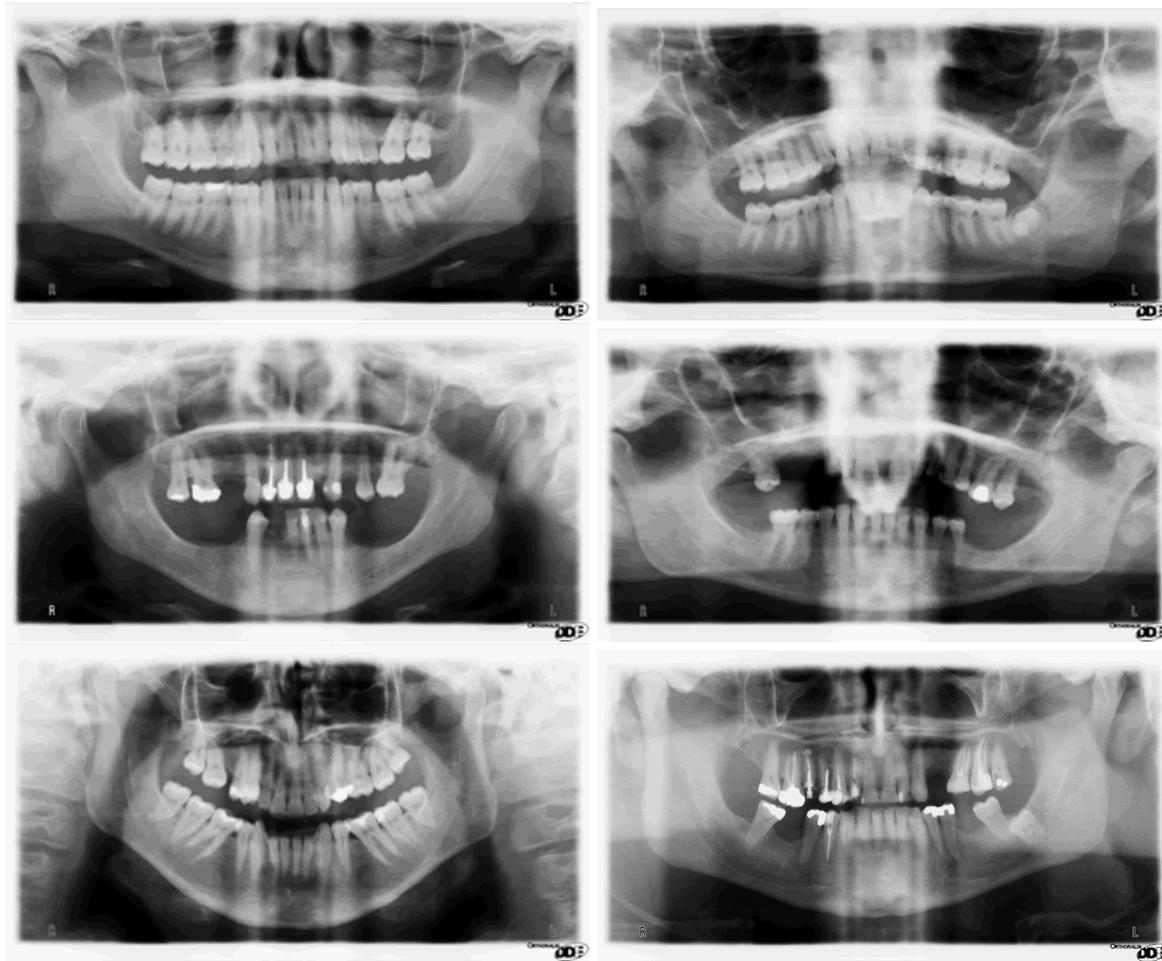


Figure 3.2: Examples of images of the Dental X-ray data set.

3.3.1 Number of Teeth per Image

We considered that images contains all teeth in two different circumstances:

1. When the images include only the first and second molar, superior or inferior and left or right along with all other teeth.
2. When the images contain all three molars superior or inferior and left or right in addition to all other teeth.

In order to more effectively use the images we named them according to their characteristics. With this arrangement we classify the images based on the number of teeth and in the existence of dental cavities. The naming takes the form *It1_t2_t3.tiff*, where *t1* corresponds to the image number, *t2* represents the quantity of teeth per

mouth (0 refers to no teeth, 1 to some teeth and 2 to all teeth) and $t3$ denominates the presence of dental cavities in the image (0 corresponds to no dental cavities and 1 to the presence of dental cavities). As an example, the name *I1200_2_0.tiff* corresponds to the image with id equal to 1200, which has all teeth and no dental cavities. In figure 3.3 is shown the count of the images data-set that fulfill the requirement of having all, some or none teeth.

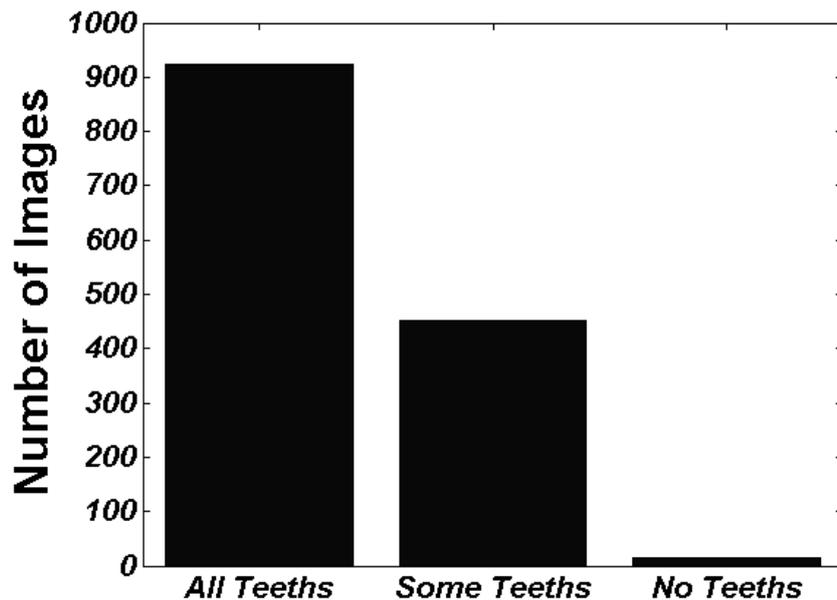


Figure 3.3: Number of images concerning the quantity of teeth per mouth.

3.3.2 Average Size

An adult individual has 32 teeth in is permanent teeth, the teeth are divided in two jaws, the upper a lower jaw, each one is constitute by 16 teeth. The composition of the mouth, concerning the teeth is on the left side constitute by three molars and two pre-molars, the front side contains four incisors two canine and finally the right side has three molars and two pre-molars. As reported in the beginning of this chapter, the data set contains many different dentition shapes and number of teeth per images. In figure 3.4 is given the interval for the average size of the teeth in 1000 sample size, with a 95% certainty. Note that for this study we only took into account the molars and pre molars, because they are in most cases the teeth

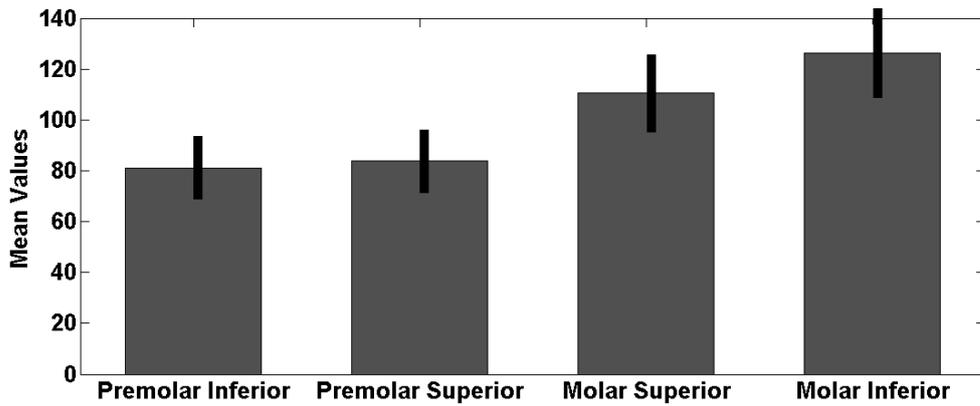


Figure 3.4: Average sizes and corresponding confidence intervals, regarding the type of teeth.

that missing in our input images. Based on our human nature, that missing teeth neighbors occupy the vacancy left by the missing teeth. This phenomenon is more visible in pre molars and molars.

3.3.3 Data Summary

Our main purpose of this section is to present the most important statistical characteristics of our data-set of images, consisting of 1392 grayscale images. In the table 3.2 are the statistics of all our images.

Note that the values are average values corresponding to all images, i.e., the mean value of intensities represented in the table 3.2, corresponds to the average value of intensities of all images.

As we can observe the mean intensities of all images present in our data-set images is a low mean value, this means that generally the intensities of each image is tend to be darker. This is also corroborated by the median value presented in the table 3.2. The maximum and minimum values show that exists a variety presence of almost all values belonging to the grayscale possible values. Finally the entropy that proves the fact that our images containing a large amount of information in the image, corresponding to the variety of types and properties of each image. The value of the entropy tend to decrease as we decrease the neighbor window. This proves that in small windows the changes in our images are inferior when comparing with larger neighbor windows.

Statistic Data	Values
Minimum intensity value	0
Maximum intensity value	252.1335
Mean intensity value	135.0296
Mode	240.7918
Variance	5227.9080
Standard Deviation	72.2699
Median	138.7947
Entropy	7.7920
Entropy (3x3)	2.0551
Entropy (9x9)	3.2412
Entropy (15x15)	3.6786

Table 3.2: Statistic date retrieve of our images data-set.

3.4 Conclusion

In this chapter we present our images data-set that have been essential to the development of our research. We present some relevant properties of our data-set images, such as the size of the teeth, the accounting number of images depending on the number of teeth present per image. To emphasize the importance of some data, boot statistic and empirical, in this chapter that will be used further on our method stages. In the next chapter is presented the firsts steps of our method, which are the ROI definition and the jaws partition.

Chapter 4

ROI Definition and Jaws Partition

This chapter will begin by describing the first steps of our method. In the first stage is elaborated the analysis of the distances between the center of the image and the four corners of the mouth. This with the aim of creating an area of interest, our ROI, an area which would contain only the mouth with their teeth. If so the removal of nasal and chin bones. In a second stage and having the first area of interest defined, we start to separate the jaws. Firstly by extracting points between the Jaws and then connect those points by the polynomial least squares fitting [63], which corresponds to derivation of the least squares fitting [38][62][17].

4.1 ROI Definition

The earliest stage is based in the statistical analysis of the sizes and positions of each component in each image, in order to define an initial region of interest. This eliminates non-useful information originated by the nasal and chin bones. Our purpose is to crop a region that contains the entire mouth and eliminate the maximum amount of noise possible. For each data set image we measured four distances (R_1, R_2, R_3, R_4), starting from the image center $(x_c, y_c) = (w/2 = 1408, h/2 = 770)$, as shown in figure 4.1.

Having these values of all the data set images, we obtained the four histograms as illustrated in figure 4.2. Here, the line series correspond to the approximated normal distribution obtained by a line fitting procedure, defined by the (μ, σ) parameters.

Based on this distribution we can approximate the minimum value for each R_i

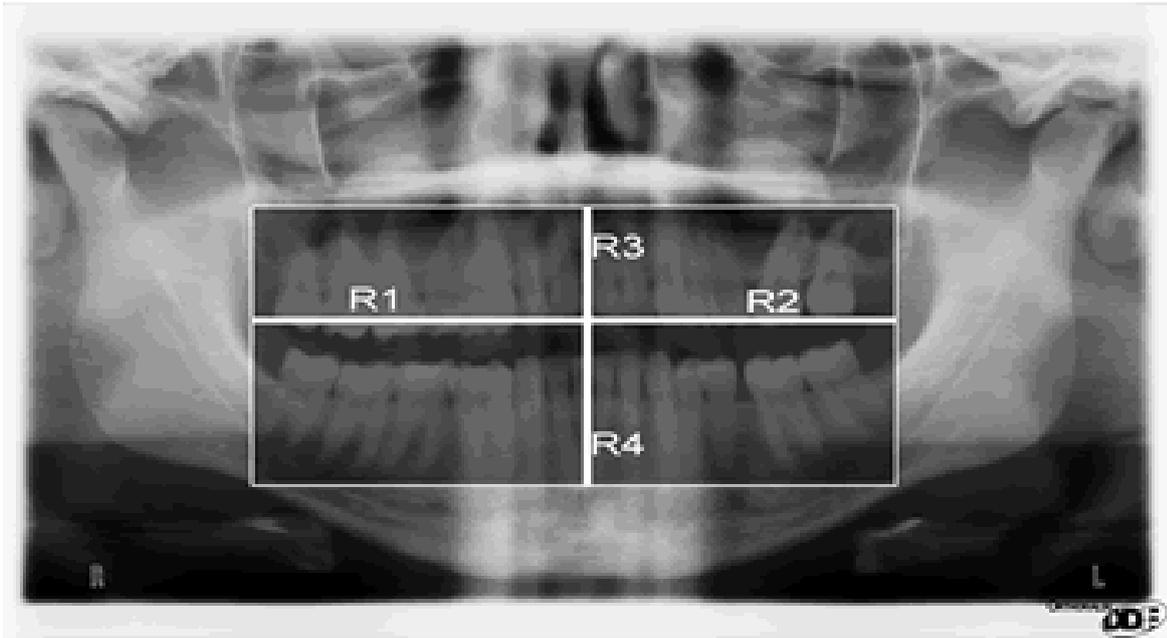


Figure 4.1: Four lengths extracted of our images data-set.

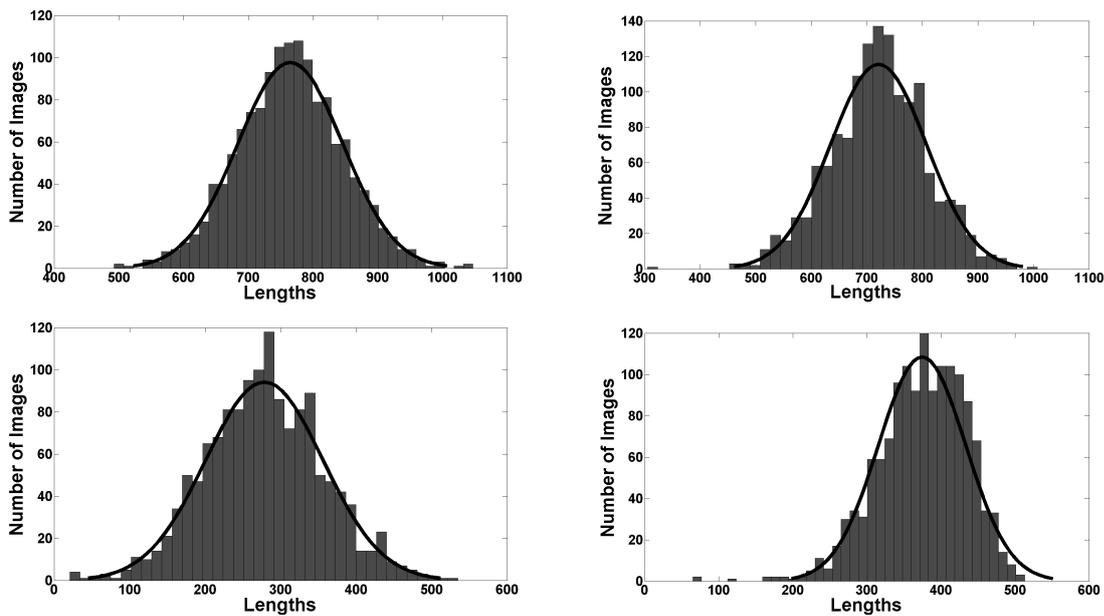


Figure 4.2: Histogram of the R_i values.

thereby appropriately cropping the images with a 95% certainty. Furthermore, a defined margin guarantees that slightly different images will be appropriately cropped, even if a small increment of non-useful regions is included. The values obtained were $R_1 \approx 897.77$, $R_2 \approx 863.36$, $R_3 \approx 406.31$, and $R_4 \approx 471.27$. Figure 4.3

illustrate the final results of this stage.

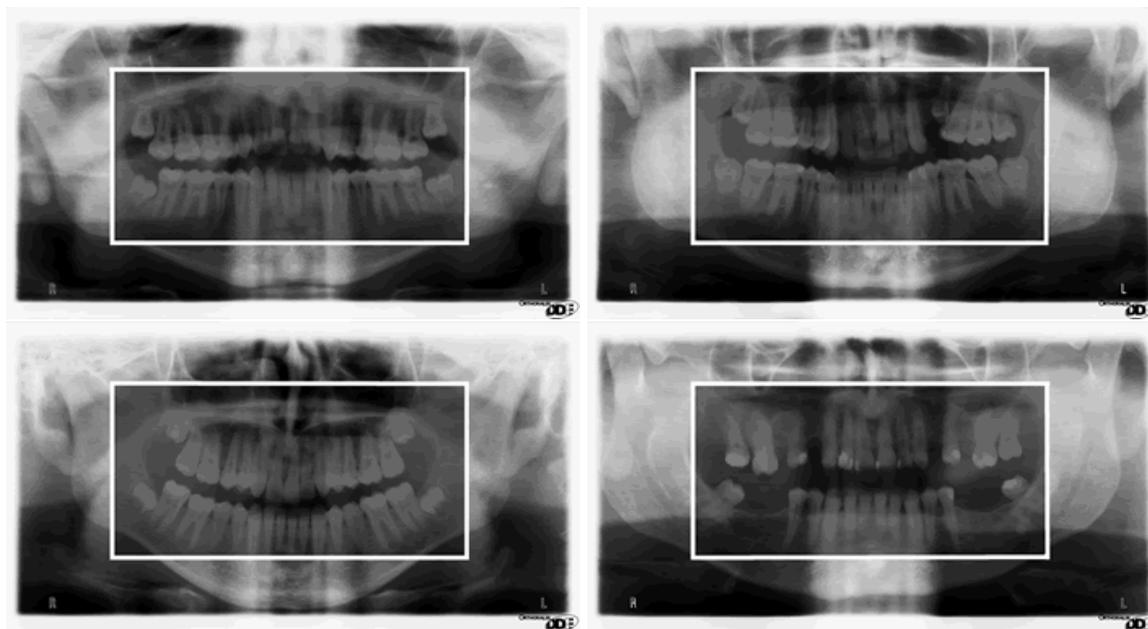


Figure 4.3: Some results for the ROI definition.

As shown, the images contain only the mouth with their teeth. The fact that we choose to cut with a confidence level of 95% leads to the images that emerge out of that range will have areas of teeth cut, which would originate a non detection of dental caries in the corresponding teeth. Furthermore the images are within this range can have over presence of noise, because of the wide variety of size and structure of the mouth becomes more complicated to define the ideal size for each image.

4.2 Jaws Partition

In this stage we separate the upper and the lower jaws, which is done by applying a polynomial fitting process to a set of primitive pixels located between jaws. This set of pixels is based on the horizontal projection, $v(u)$, of the images, given by 4.1, where $I(x, i)$ denotes the intensity value at line x and column i .

$$v(u) = \sum_{i=0}^w I(x, i) \quad (4.1)$$

The initial point is defined at the right extreme of the image and at the line that has the minimum $v(u)$ value, given by 4.2, where w is the image width and shown in figure 4.4.

$$p_0(x_0, w - 1) = \arg \min_x (v(u)) \quad (4.2)$$



Figure 4.4: Horizontal Projection of the intensities, for our starting point.

The remaining set of points p_i are regularly spaced, starting from $p_0 : p_i(x_i, (w - 1) - W/21)$, where x_i is obtained similarly to x_0 . To avoid too high vertical distances between consecutive p_i , we added the following constraints 4.3.

$$p_i(x_i, y_i) = \begin{cases} p_i(x_{i+1} + T, y_i), & |p_i(x_i, y_i) - p_{i+1}(x_{i+1}, y_{i+1})| > T \\ p_i(x_i, y_i), & \text{otherwise} \end{cases} \quad (4.3)$$

Empirically T is the threshold defined by us to avoid the high vertical distances, with the value of $T = 20$. We observed that this step plays a major role in dealing with missing teeth. This method to the extraction of the initial point is based on several proposals to perform the separation of jaws ([53][23][25]). Having the set of $p_i(x_i, y_i)$ primitive points, the division of the jaws is given by the 10th order polynomial, given by 4.4, obtained by the polynomial least squares fitting algorithm, based on the Vandermond matrix [65]. This technique is able to achieve impressive accuracy on the data set images, as illustrated in figure 4.5 and summarized in table 7.1.

$$p(x) = a_0 + \dots + a_{10}x^{10} \quad (4.4)$$

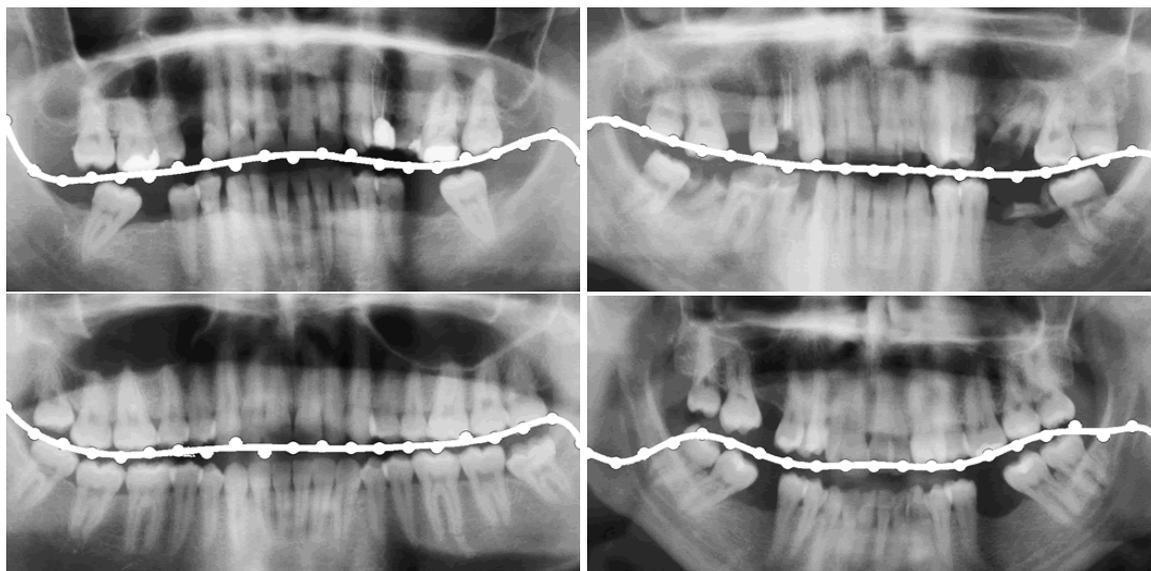


Figure 4.5: Results of the polynomial least squares fitting.

The process defined by us, to perform the jaws partition, is dependent on the size of the missing teeth present in the input image. If the teeth are missing and at the same time neighbors makes our method vulnerable to error, because the points extracted by us will tend to follow the darker areas of the image, and sometimes the areas in which missing teeth are darker when comparing to the gap between the jaws, after an considerable missing teeth area our method may be in some occasions capable to overcome the presence of a tooth, as shown in the figure 4.6.

A solution that can be implemented to improve this situation would be to divide the image into smaller stripes, when we are extracting the points, which in turn could make the method more sensitive to small changes of intensity in the image, making it even more difficult to separate the jaws.

4.3 Conclusion

In this chapter we present the first stage of our method, which consisted in defining a first area of interest, our ROI which with this step has been possible to eliminate much of the noise present in our input images. The second step was the jaws partition, by the extraction of 21 points between the jaws and the connection of this same points through the polynomial least squares fitting process, we have been able to separate the jaws for our data-set training images with a large percentage

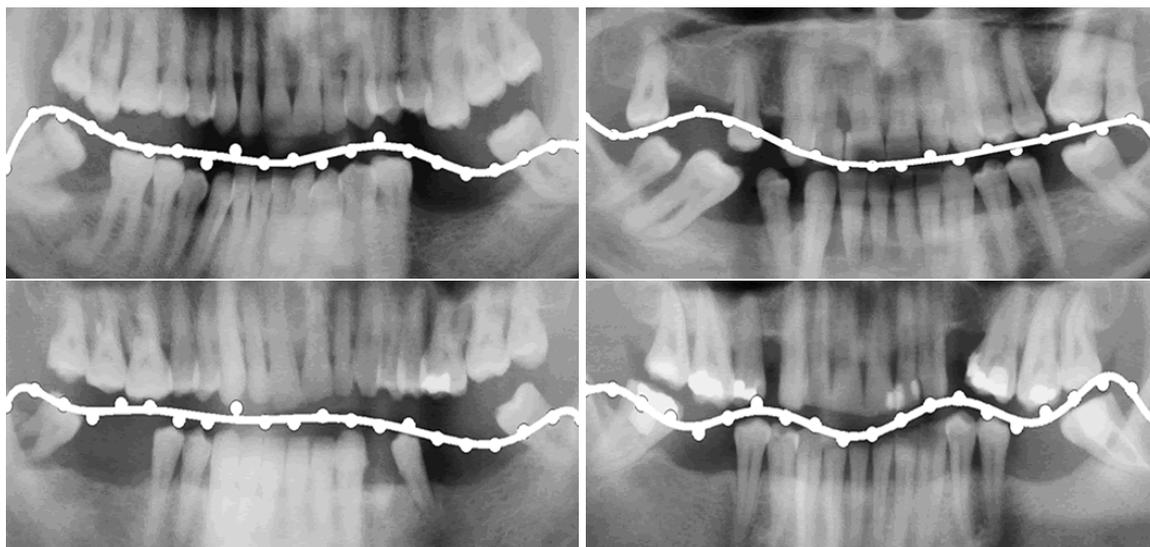


Figure 4.6: Results of the polynomial least squares fitting in images with accentuated missing teeth area.

of success. In the next chapter is staging the next step of our method, the teeth gap valley detection. This step is divide in two methods, the first one is our primary version of the teeth gap valleys detection, which culminates with a acceptable article submission in the II Ecomas Thematic Conference on Computational Vision and Medical Image Processing (VIPImage'09). The second is our better improvement of the first version, through the use of the Active Contours Algorithm.

Chapter 5

Teeth Gap Valley Detection

The next step in our method it is in the detection of the teeth gap valley where the aim is to get a zone containing only one tooth. The perfect output is when the output as many pictures then teeth in the image. To perform the detection of teeth gap valleys and based on the output generated by the polynomial obtained in the jaws partition. Next, we draw perpendicular lines to all points belonging to the polynomial. For each point it will be selected the line which the sum of intensities of the pixels belonging to the line is minimal. Finally the points are selected where the intensity is minimal, as can be seen in the figure 5.1. What we developed at this point as resulted in an article submitted to VIPImage'09.

5.1 Line Sum Intensities - Method 1

Having both jaws divided, our next goal was to localize the regions corresponding to each teeth, which is based on the detection of the gap valleys between them, because we are in presence of a darker zone, therefore the intensities between them is substantial inferior when compared with the teeth intensities. This rule is not applicable when referring to teeth overlap, in this case we have the opposite effect, an increase of intensity. This stage can itself be divided into three sub-stages:

1. For each pixel x between 0 and $w - 1$, where w is the image width, we obtain the equation of the perpendicular line to $r(x)$ at the point $(x, r(x))$. Then, the average intensity a of the pixels that fall in that line is obtained. Due to the image properties, the shape of the teeth and the shape of the polynomial, we

found convenient to vary the angle between the line and the polynomial in the $[-15^\circ; +15^\circ]$ interval.

2. Later, for each x we selected the line that minimizes the average intensities a , obtaining a set of $\{a_1, \dots, a_x\}$ values, which is illustrated in figure 4.5. The key insight is that the a_i values with lowest values should correspond to the partition between consecutive teeth. In order to compensate for abrupt variations in the values we smoothed these values through the use of a Gaussian kernel.
3. Finally, we extracted the local minimums of the smoothed signal, as shown in figure 5.1, hoping that they correspond to the desired partitions. However, we observed that generally the number of local minimums is higher than the gap valleys (false positives), but the key false negatives are almost inexistent and, for this reason, the subsequent use of an expert-system based approach will easily perform the detection/partition of each tooth.

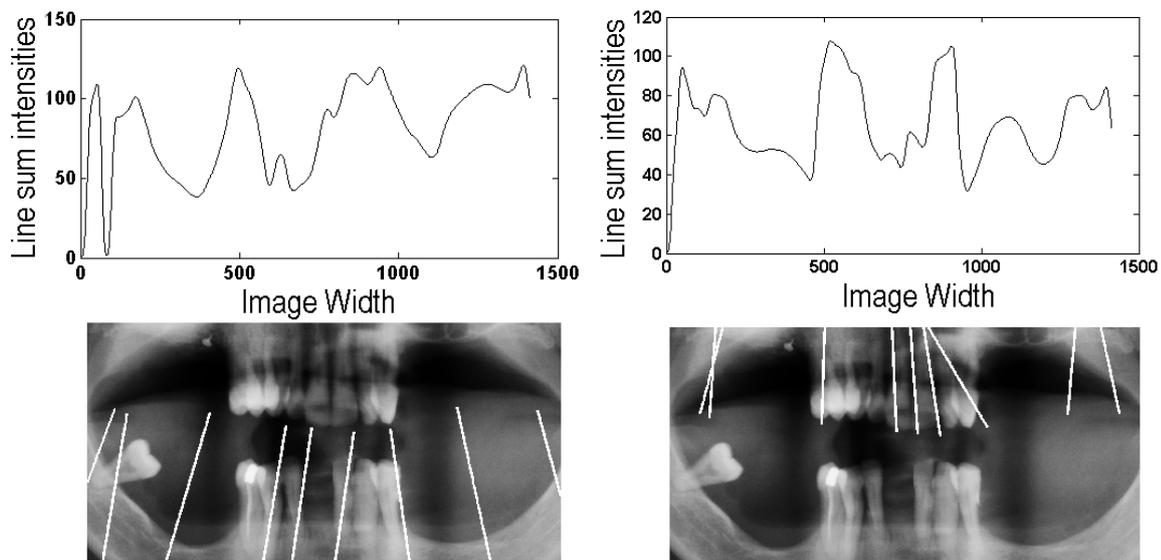


Figure 5.1: Example of the minimum intensities for one input image.

Based on the counting of the detected teeth, the final step crops the upper and lower jaw images into three regions. The left side is composed by three molars and two pre-molars, the front side contains the incisors and a canine and, finally, the right side has three molars and two pre-molars. As reported in the chapter 3, the

data set contains many different dentition shapes and number of teeth per images, which led us to crop the images based on two rules. In the first rule, let S_M and S_{PM} the average size of the molar and premolar, respectively, we define a cut limit given by 5.1.

$$L_{cut} = 3 * S_M + 2 * S_{PM} \quad (5.1)$$

Where L_{cut} corresponds to the maximum crop limit allowed. In the lower jaw it is similar, but in such case the teeth that matter are the inferior. Finally, the second rule measures the existence of five teeth partition lines after the imposed limit L_{cut} . If there are five lines with large differences between the middle point of each line, we suppose that all teeth exist and that the last division line gives the crop coordinates. Figure 3.4 gives the interval for the average size of the teeth in 1000 sample size, with a 95% certainty.

In figure 5.2 is illustrated one example of our results for the above described method. The results for stage 1 and stage 2 are very optimistic. For the stage 3 the result showed that improvements must be consider, due to the presence of multiple lines for the same gap valley, and in other cases the absence of lines. The result is positive due to the presence of distinguish darker zone between the teeth, which facilitates the straight lines detection that minimizes the intensity.

An example of a error case is showed in figure 5.3. The results for stage 1 and stage 2 demonstrates that the results are very optimistic. For stage three becomes essential to use pre processing in the input images, only then we will emphasize the teeth presence in the image. There are several reasons for the increased error in figure 5.3, among them stand out, regions where the overlap of the teeth is intense (brighter regions), the straight lines detection that minimize the intensity is difficult, because it is a region with little intensities variation. The same happens in the dark areas, because there isn't a contrast between the teeth and the background. For these reasons the straight lines detection that minimize the intensity, are dependent of a pre-processing stage to enhancement the image contrast, consequently obtain the best partition between background and teeth.

The proposed method is dependent on the image local intensity variations. Images where these variation is not so evidently becomes more complicated to detect the straight lines detection that minimize the intensity.

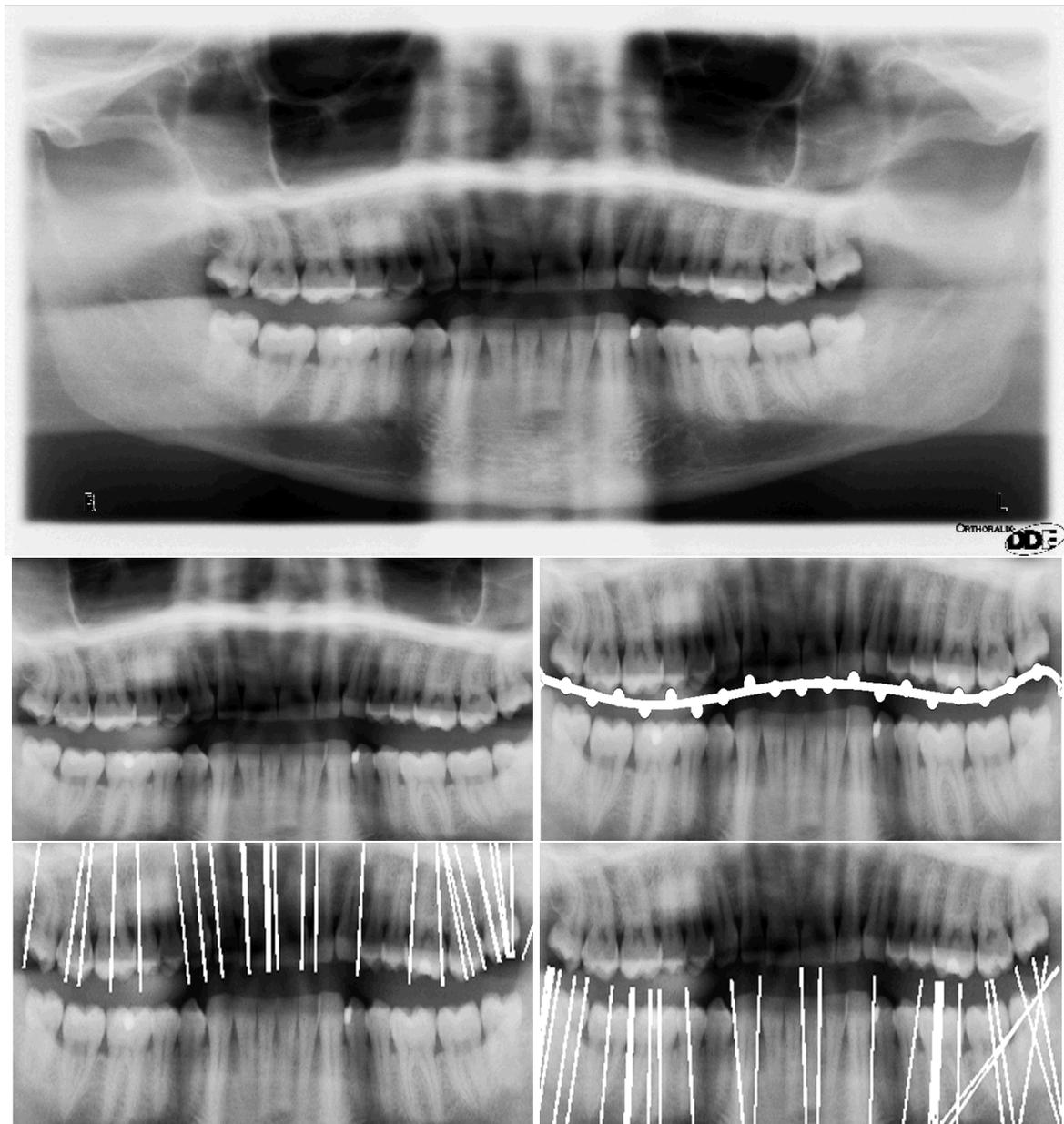


Figure 5.2: Example of our method 1 results.

Our approach at this point has led to an accepted article submission in the *VIPIImage'09*, the article is in the appendix 8.4, to call attention the conference is international and only serves to demonstrate that the work we have done so far is of most scientific importance, giving us strength and motivation to continue not only in the improvement of our method but also thinking in the implementation of new tasks related to computer vision and image processing.

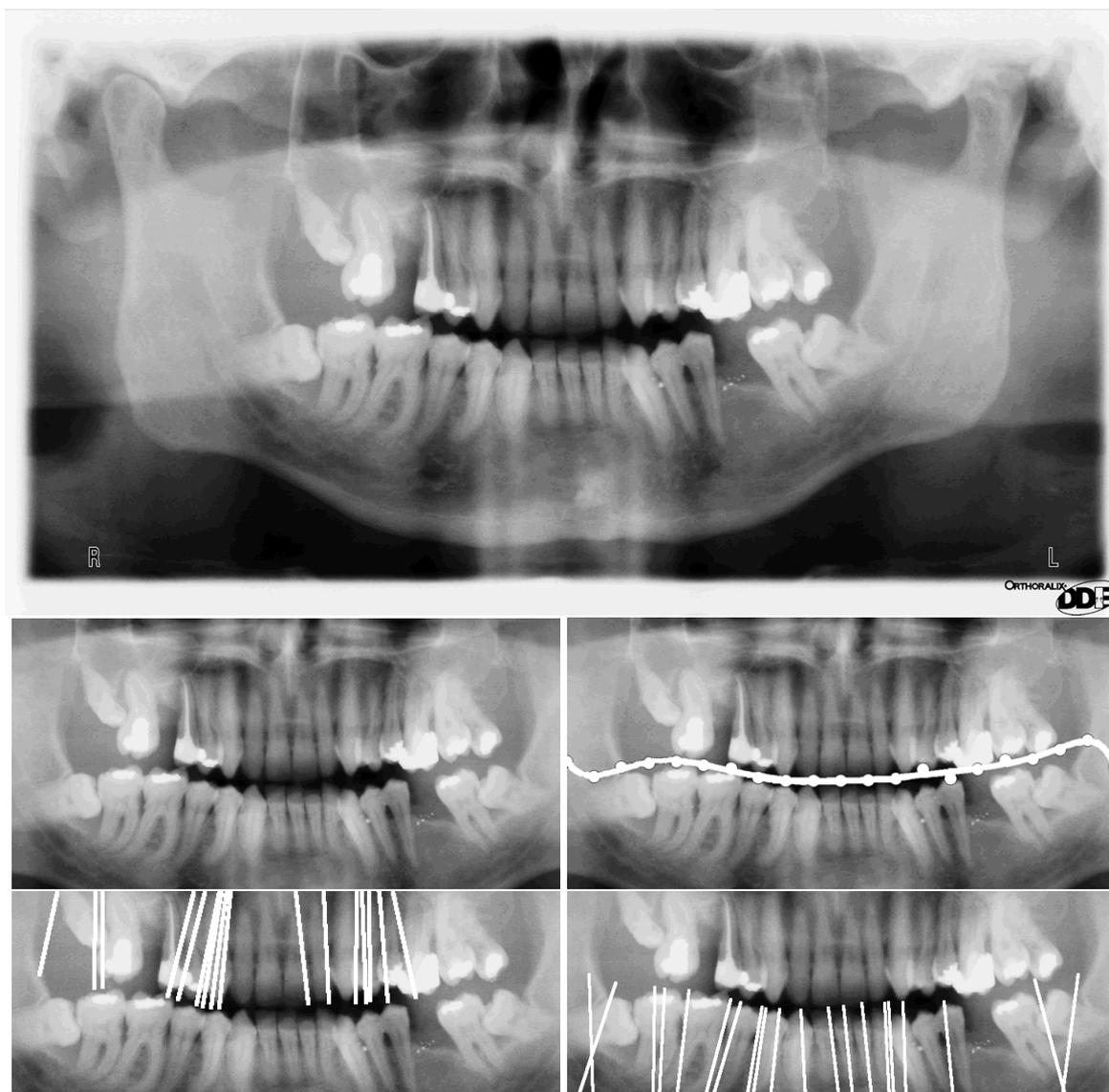


Figure 5.3: Example of failure in the teeth gap valley detection of our method 1.

5.2 Center Distance with the Active Contours - Method Two

In our second version of teeth gap valleys detection, above all the rest, our main focus was to try the selection improvement of region between the teeth. This is because in the previous method our most difficult task resides in the choice of the minimum points, obtained by the straight lines with lower intensity. To improve that process we decided to divide this second version in five distinct stages:

1. Initially is applied the pre-processing for the enhanced of the contrast in the input image.
2. Based on the previous step we divide our jaws in two different images.
3. After that we transform our images cartesian coordinates to the polar coordinates.
4. Having booth jaws in polar coordinates, our goal is to transform this in a typical active contour problem, where the object of interest is in the center of the image, therefore the contour can be easily initial as the image border.
5. With the output segmentation contour, we calculate for each contour point the distance to the center of the image. Having the distances vector we extract our minimum points. Which in this case corresponds exactly to the teeth gap valleys.
6. The final step is to perform the cut of the output regions.

5.2.1 Morphological Operators

To increase the contrast of the input images we apply a pre processing filter the top and bottom hat transform. As define in the equation 5.2.

$$\begin{aligned}
 I_{output} &= (I_{Original} + TH_{Original}) - BH_{Original} \\
 I_{outputFinal} &= (I_{output} + TH_{output}) - BH_{output}
 \end{aligned}
 \tag{5.2}$$

Where the $I_{Original}$ corresponds to our input image in grayscale [0;255], the $TH_{Original}$ is the result of applying the operation of the top hat transform to our original image and finally the $BH_{Original}$ that concerns to the applying of the bottom hat transform in our original image. The $I_{outputFinal}$ is the resulting image of this operation of pre processing. In figure 5.4 and in figure 5.5 is an example of the pre processing output applied in our images. We use a rectangular structuring element with dimensions $[w/4, h/2]$, where w and h are the width and height of the image, respectively. Our experimental choice of these structuring element is based on an experimental study on a set of 500 dental panoramic X-ray of our data-set images.

As we can see in the figure 5.4, the application of pre processing increased the contrast of the image, highlighting the dark areas and the bright areas of the original

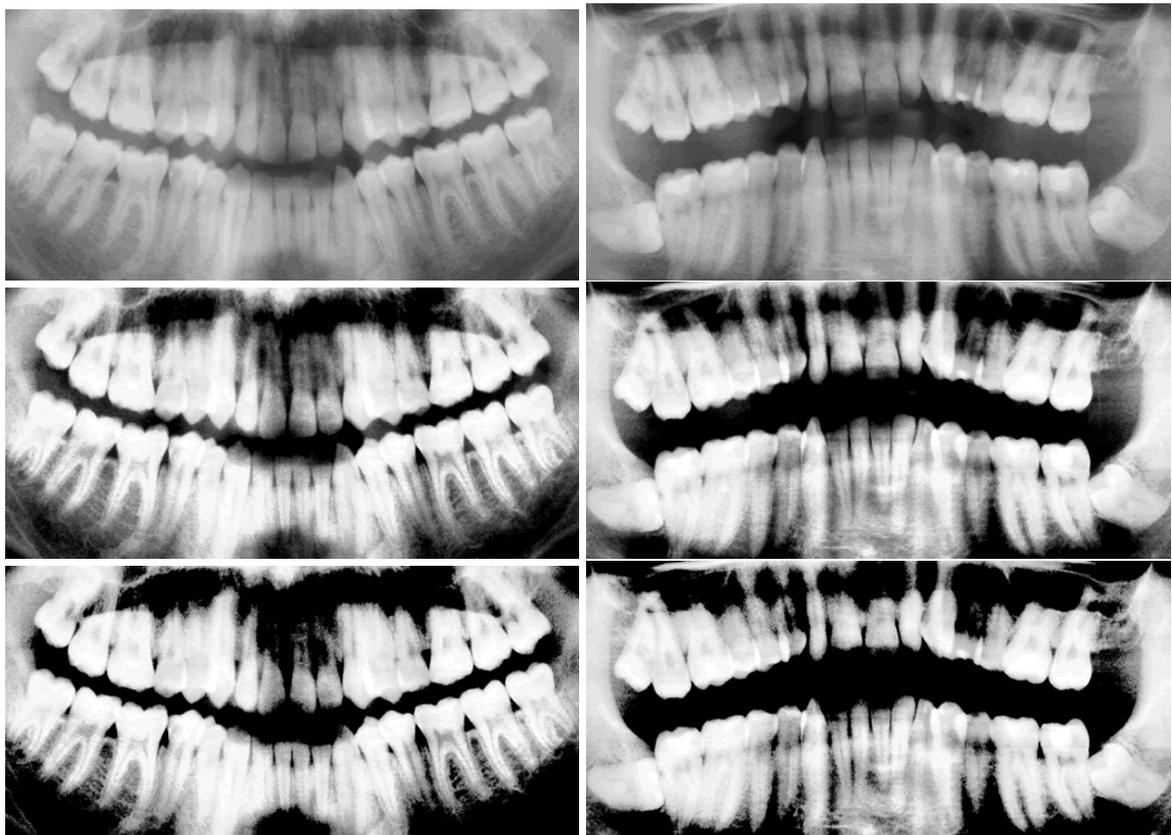


Figure 5.4: Example one of the top and bottom hat transform.

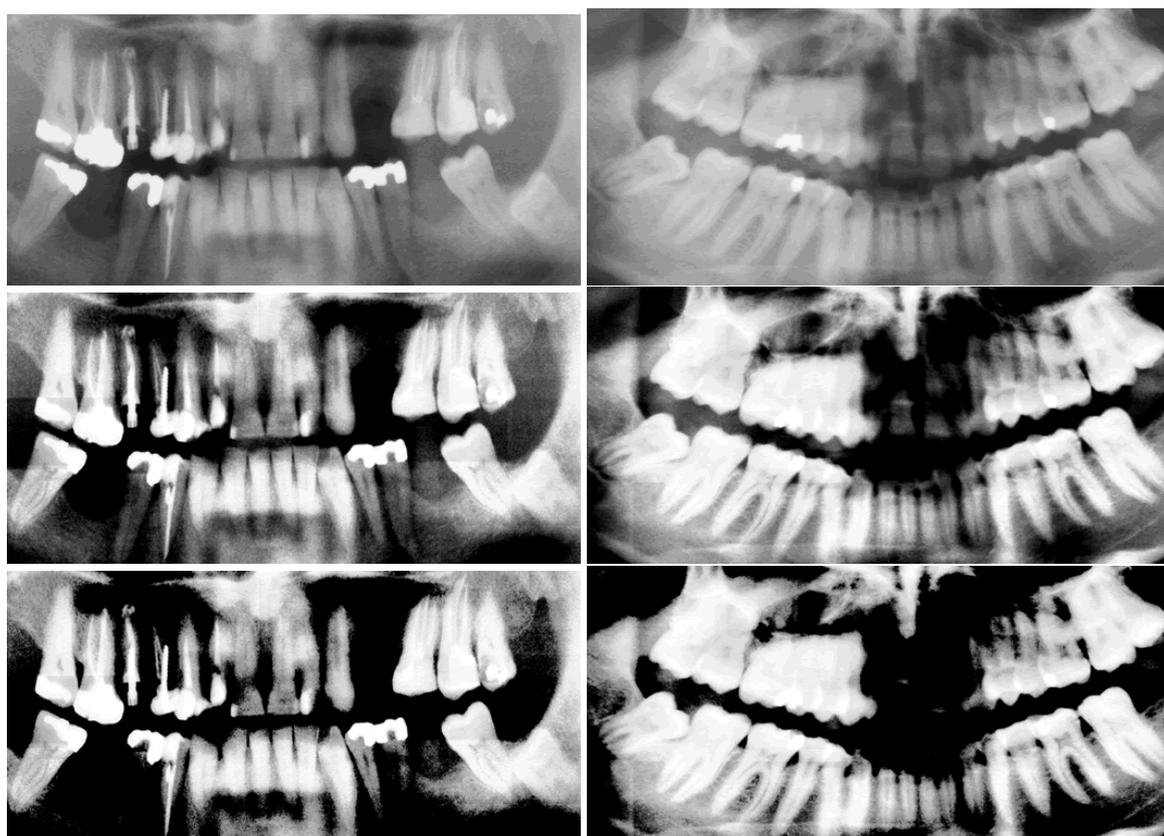


Figure 5.5: Example two of the top and bottom hat transform.

image. In other case we can see in the figure 5.5 the pre processing in teeth that were present in the original image that already do not contrast with the rest of the image, tend to disappear with the implementation of pre processing. This occurs only in the front teeth because they are thinner and can't reflect as much radiation as compared to the pre molars and molars. In other cases is due to the poor quality of the image that disables the good effectiveness of the pre processing.

5.2.2 Divide the jaws

In this step of the method the goal is to divide the jaws. To do this we use the information of the previous step by which we separate the jaws through the polynomial fitting process. In this step we apply a horizontally cut to the image. If we are in the upper jaw is the minimum value which the polynomial takes, if we are on the lower jaw the corresponding cut value is the maximum value that the polynomial takes. Thus we get two images containing only one of the jaws, the upper jaw and the lower jaw, as we see in the figure 5.6.

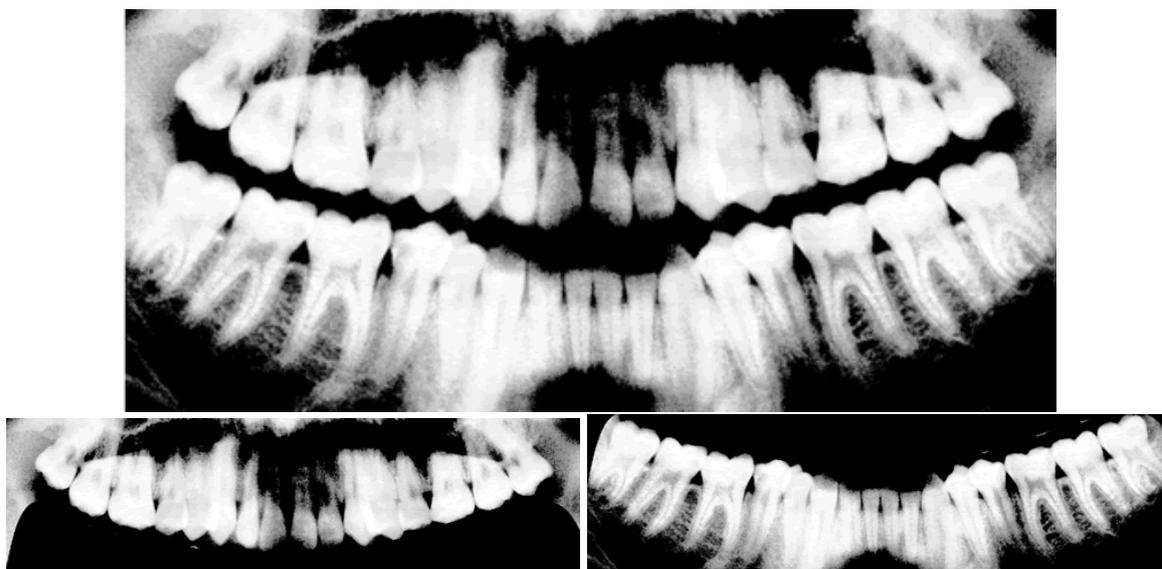


Figure 5.6: Example of the Jaws division.

5.2.3 Polar Coordinates

In mathematics, the polar coordinate system is a two-dimensional coordinate system, where each point is defined by an distance to a fixed point and the angle from

a fixed direction. The polar coordinates [6] r (the radial coordinate) and θ (the angular coordinate, often called the polar angle) are defined in terms of cartesian coordinates by the equation 5.3.

$$\begin{aligned}x &= r \cos(\theta) \\y &= r \sin(\theta)\end{aligned}\tag{5.3}$$

Where r is the radial distance from the origin, and θ is the counterclockwise angle from the x – axis. In terms of x and y . In the opposite in equation 5.4 is the method to transform an (x, y) point in the polar coordinates.

$$\begin{aligned}r &= \sqrt{x^2 + y^2} \\ \theta &= \tan^{-1} \frac{y}{x}\end{aligned}\tag{5.4}$$

Where $\tan^{-1} \frac{y}{x}$ corresponds to a two argument inverse tangent which takes the signs of x and y in to account to determine in which quadrant θ lies. The equation of a curve expressed in polar coordinates is known as a polar equation, and a plot of a curve in polar coordinates is known as a polar plot [64]. In figure 5.7 is the result of this step for each separated jaw.

The resulting images in polar coordinates, had the follow dimensions [$h * 2; h * 2$], where h corresponds to the height of the original input image. With this transformation of coordinates is thus possible to transform our problem, in a problem typical of active contours. One of the main reasons of the use of the polar coordinates is in fact referred to the radial division that is possible to perform in this image representation. Which allow us to better improve the result of the active contours therefore the teeth division become an easier task.

Although as we check in the following section, the transformation back to Cartesian coordinates can lead to loss of information, therefore transforming our process of teeth division not so accurate as we wish.

5.2.4 Active Contours

As explained in the chapter 2, the active contours technique is one of the most important and used on the literature. The main motivation for the use of this type of algorithms is their excellent ability for segment objects present in the images. For

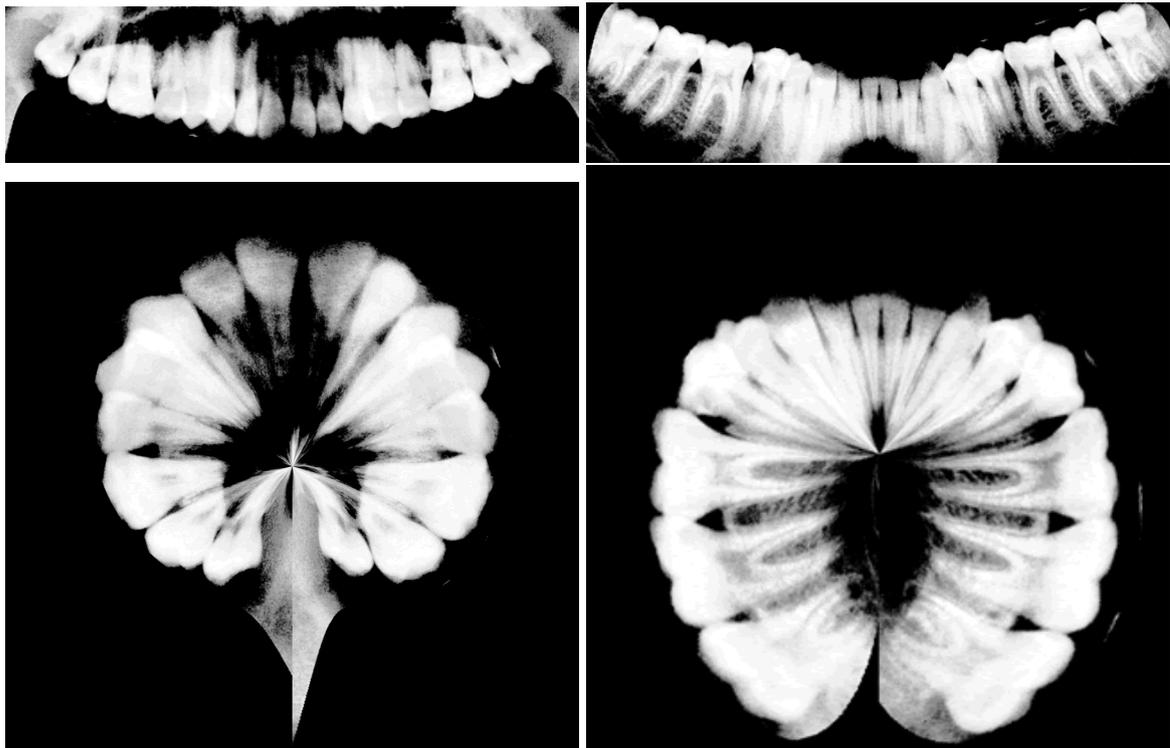


Figure 5.7: The resulting polar coordinates transformation for each separated jaw.

that fact and based on all work we exploit in the literature we choose for this step the use of the active contours without edges [56][10].

The reason that led us to choose this type of active contours, was the fact that this technique has a greater tolerance to areas of the image which do not correspond to the area we want to segment, which in the majority of the cases this area is originate from the noise of the previous steps. Making our process of the teeth gap valleys detection very precise and accurate.

The non use of the deformable active contours Snake, is mainly because it does not guarantee convergence to the global minimum of the functional, that is for example when we are in presence of a large amount of noise. And the active contours Snake success is based on the initialization. Which in this case is a problem because our initial contour is far from the desired solution, making a non desirable technique for usage.

5.2.4.1 Without Edges

The active contours without edges is a method of the minimization of an energy based-segmentation. Based on the exposed in chapter 2.2.4.5 and in the equation 2.8 that corresponds to the functional minimization of the energy.

After a study applied in 200 images of our data-set images, we define our original mask corresponding to a four independent elements in each image quadrant, as illustrated in the equation 5.5.

$$I_{mask}(x, y) = \begin{cases} ([w * 0.10...(w/2) * 0.90], [h * 0.10...(h/2) * 0.90]) = 255 \\ ([w/2 + (w/2) * 0.10...w * 0.90], [h * 0.10...(h/2) * 0.90]) = 255 \\ ([w/2 + (w/2) * 0.10...w * 0.90], [(h/2) + (h/2) * 0.10...h * 0.90]) = 255 \\ ([w * 0.10...(w/2) * 0.90], [(h/2) + (h/2) * 0.10...h * 0.90]) = 255 \end{cases} \quad (5.5)$$

Where w and h corresponds to the image width and height, respectively. An example of the mask is given in figure 5.8.

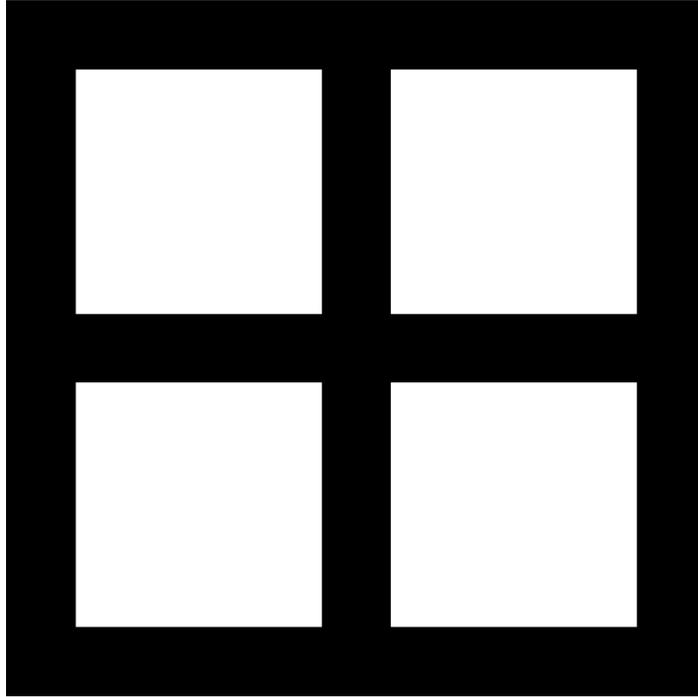


Figure 5.8: The example mask of our initialization for the active contour method.

The use of this mask is based on an experimental study, where through this initial mask we accomplished better results for the tested images. The test set corresponds

to a 200 set of images of our dental panoramic X-ray data-set images.

After the initialization mask definition, we are in conditions to perform the active contours without edges method in our images. In figure 5.9 is the output of the active contour without edges method applied in our images.

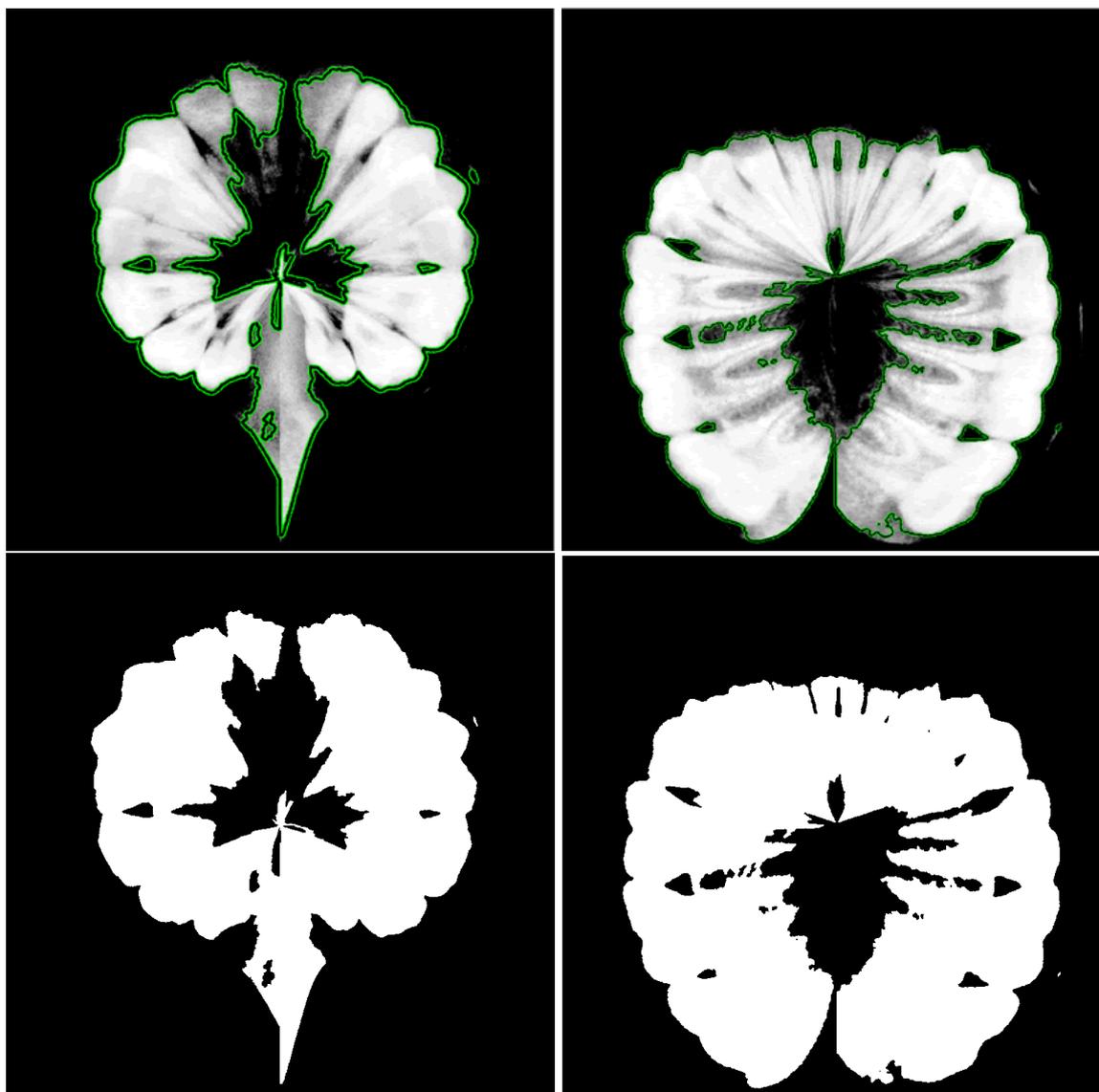


Figure 5.9: Output example of the active contour without edges method, applied in each jaw.

As we can observe the method of extracting the contour has a great performance for the segmentation of the teeth, which in turn improves the step of detecting the teeth gap valleys. Because the extraction points give us the guarantee that we are in a pixel corresponding to the teeth gap valley. For a faster computation

we apply a prior resize to the input image and the image corresponding to the mask. The resizing factor we chose was 4, i.e., the final size of the images where we apply the active contours had the following dimensions, $[w/4; h/4]$, where w corresponds to the width of the image in polar coordinates, and h is the value of the height input image in polar coordinates. In this case w and h , could be replaced by $h_{Original} * 2$, where $h_{Original}$ is the corresponding height of the original image in cartesian coordinates.

5.2.5 Teeth Division

The final step of our second method corresponds to the teeth division, this division is based on the minimums extracted from the previous step, the steps for the teeth division are:

1. Through the previous contour is calculated the distances from the center of the image to all points belonging to the contour.
2. Having the vector of distances are eliminated updated all distances that are less than $0.10 * h$, where h corresponds to the height of the input image. In this step also are update all distances between two points that had distances smaller than $0.10 * h$.
3. After that we apply a Gaussian filter [57][35] for the smooth of the output values of distances.
4. The extraction of the local minimums are the next step of our method.
5. Then is extracted another neighbor point whose intensity is the more approximated to the minimum point extracted. In this step is also take in consideration line that the first point perform with the center point of the image.
6. Finally is perform the regions cut, based on the lines that were extracted.

The distances to the center were calculated based on the Euclidean distance, as shown in the equation 5.6.

$$r = \sqrt{(c_x - b_x)^2 + (c_y - b_y)^2} \quad (5.6)$$

Where (c_x, c_y) corresponds to the center coordinates and (b_x, b_y) are the coordinates of the border points. After the extraction of the center distances, we apply a cut in to the distances values, based on $0.10 * h$, where h corresponds to the height of the input image in polar coordinates. This distance normalization is performed by updating all the points distances values between two distinct points that are smaller than $0.10 * h$, where h is the correspond height of the input image in polar coordinates. In figure 5.10 and 5.11 are the results of this step, concerning the lower and upper jaw, respectively.

Subsequent to the extraction of the distance vectors and the cut applied to the distance values points, we apply the convolution algorithm [20] to the distance vector and the central line of a Gaussian filter as shown in figure 5.12.

The gaussian filter corresponds to rotationally symmetric lowpass filter of size $hsize = 100$ with standard deviation $\sigma = 10$. The convolution algorithm consists in the multiplying the polynomials whose coefficients are the elements of input. In this case the distance vector and the central line of the gaussian filter. The convolution theorem also says that convolving two sequences is the same as multiplying their Fourier transforms [36]. In figure 5.13 is the result of the convolution applied to the distance vectors with the cut of the $0.10 * h$, where h corresponds to the image height, and using the gaussian filter as shown in figure 5.12.

Following the convolution step, we extract the minimums values from the distance vector convolve. The extraction of the minimums are based on the following rules:

- A window based local minimums extraction is used, the length of the window is $size = 5$.
- When all the first three elements has decreasing values and the fifth value is higher than the fourth element, we consider the fourth element a local minimum.
- The fourth element must be smaller than the first three elements and smaller than the fifth element, to be consider a local minimum.

The rules are expressed in the equation 5.7.

$$P(x) = (P(x - 3) > P(x - 2) > P(x - 1)) > P(x) \&\& P(x) < P(x + 1) \quad (5.7)$$

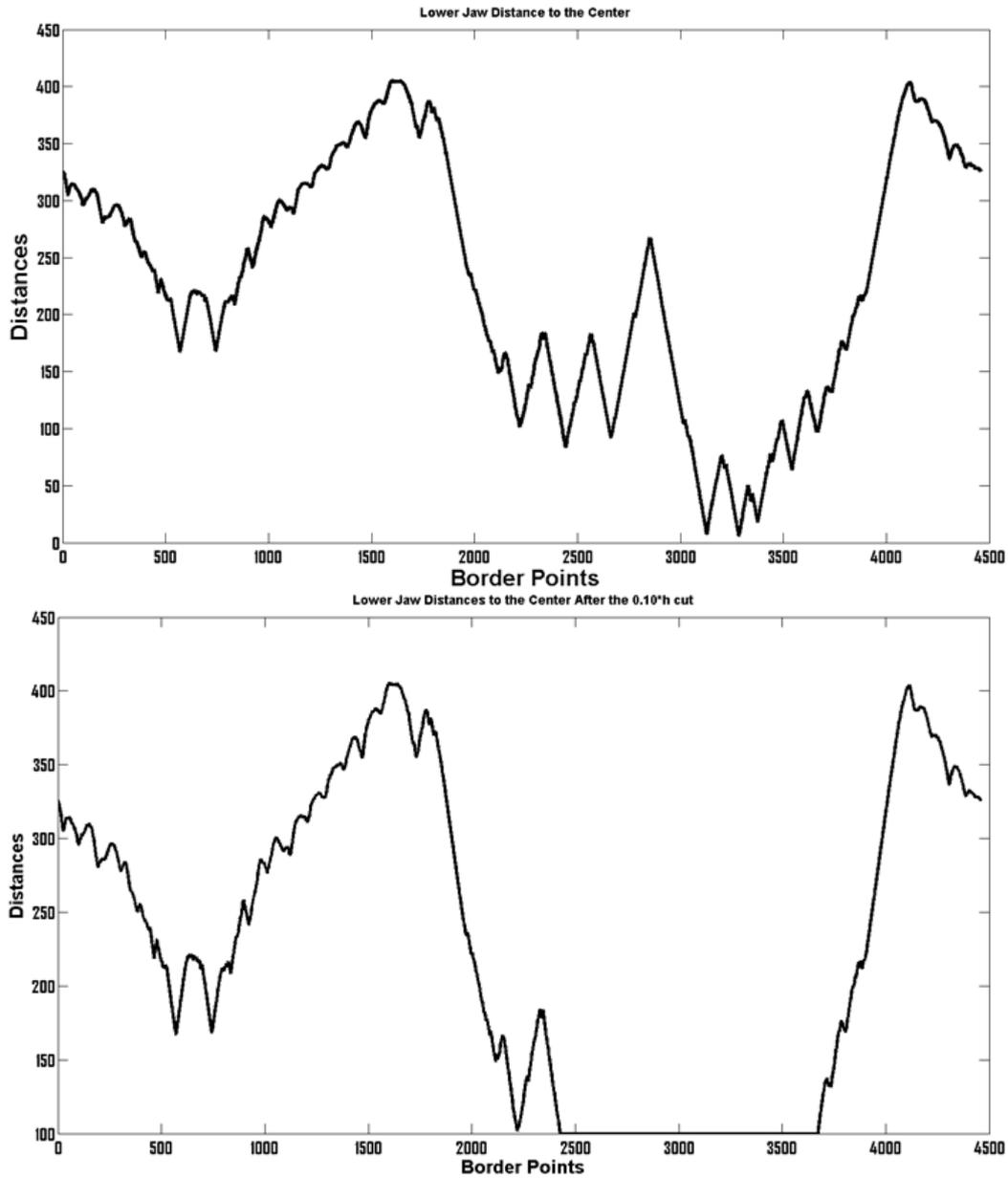


Figure 5.10: The distance to the center of the lower jaw and the corresponding $0.10 * h$ cut.

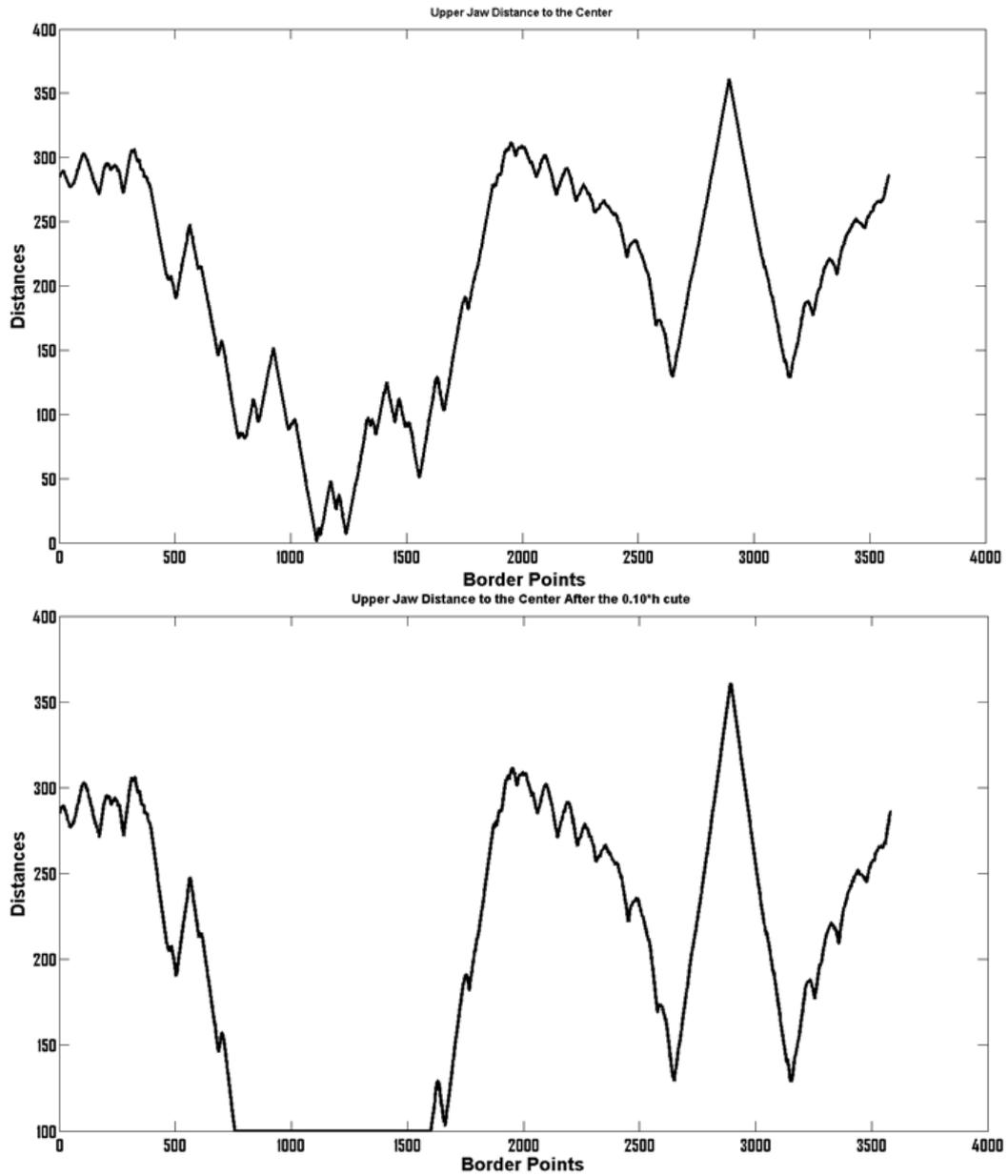


Figure 5.11: The distance to the center of the upper jaw and the correspond $0.10 * h$ cut.

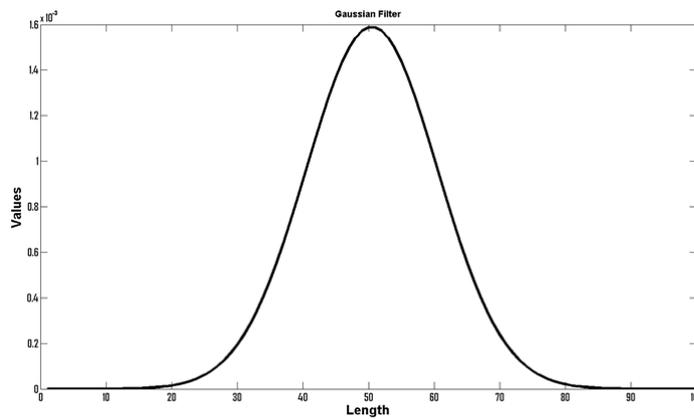


Figure 5.12: The gaussian filter for the convolution with the distance vector.

Where $P(x)$ corresponds to the fourth value of the window based local minimums extraction and x to the vector indices in analysis. In figure 5.14 is the result of the minimums extraction step applied to the upper and lower jaw, respectively.

As we can observe there are yet a small presence of false positives in the minimums extracted, for that extra minimums extracted we decide to eliminate some of these minimums based on two simple rules, the first corresponds to the presence of two approximated extracted minimums, where in this case the minimum with greater value is eliminated. The second rule is based on all the minimums extracted from small variations inside the previous region of cut that we perform, the cut based on the $0.10 * h$. All minimums extracted that are within this region are removed.

The final step consists in the extraction of the second point, this extraction is based on the line that is performed by the minimums points and the point corresponding to the image center. In this process is extracted the second point in the neighbor of the minimum point that belongs to the line (perform by the center point of the image and the minimum point).

After the extraction of this second point, we calculate the correspondent coordinates in the cartesian coordinates of the two points extracted. With these points coordinates we can now draw the respective line in the cartesian coordinates to perform the region cut of each teeth extracted.

As we can see in figure 5.15 there's an example of the final result of our method 2, for the teeth gap valley detection.

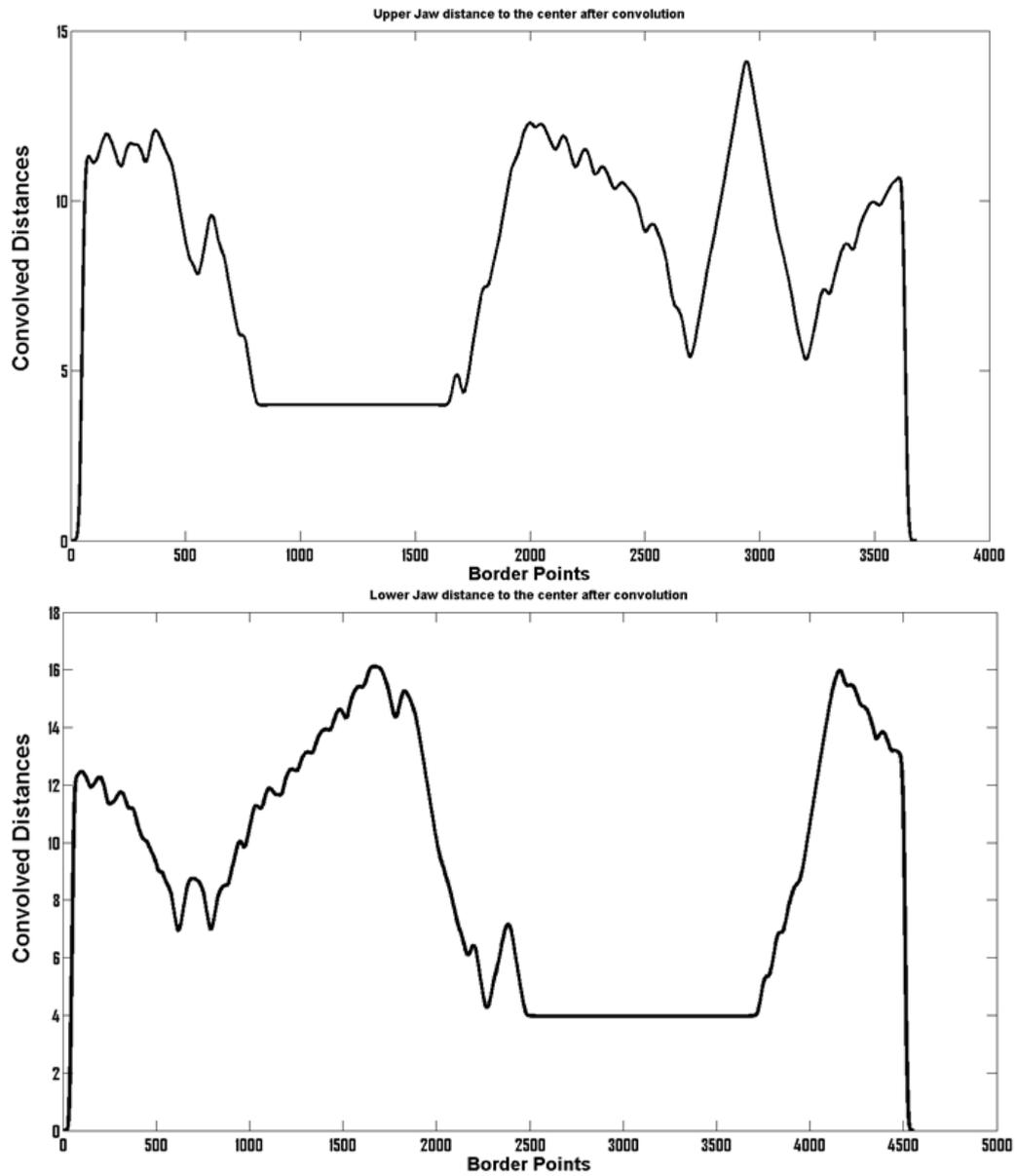


Figure 5.13: Convolution applied to the upper and lower jaw distance vector, based on the filter in 5.12.

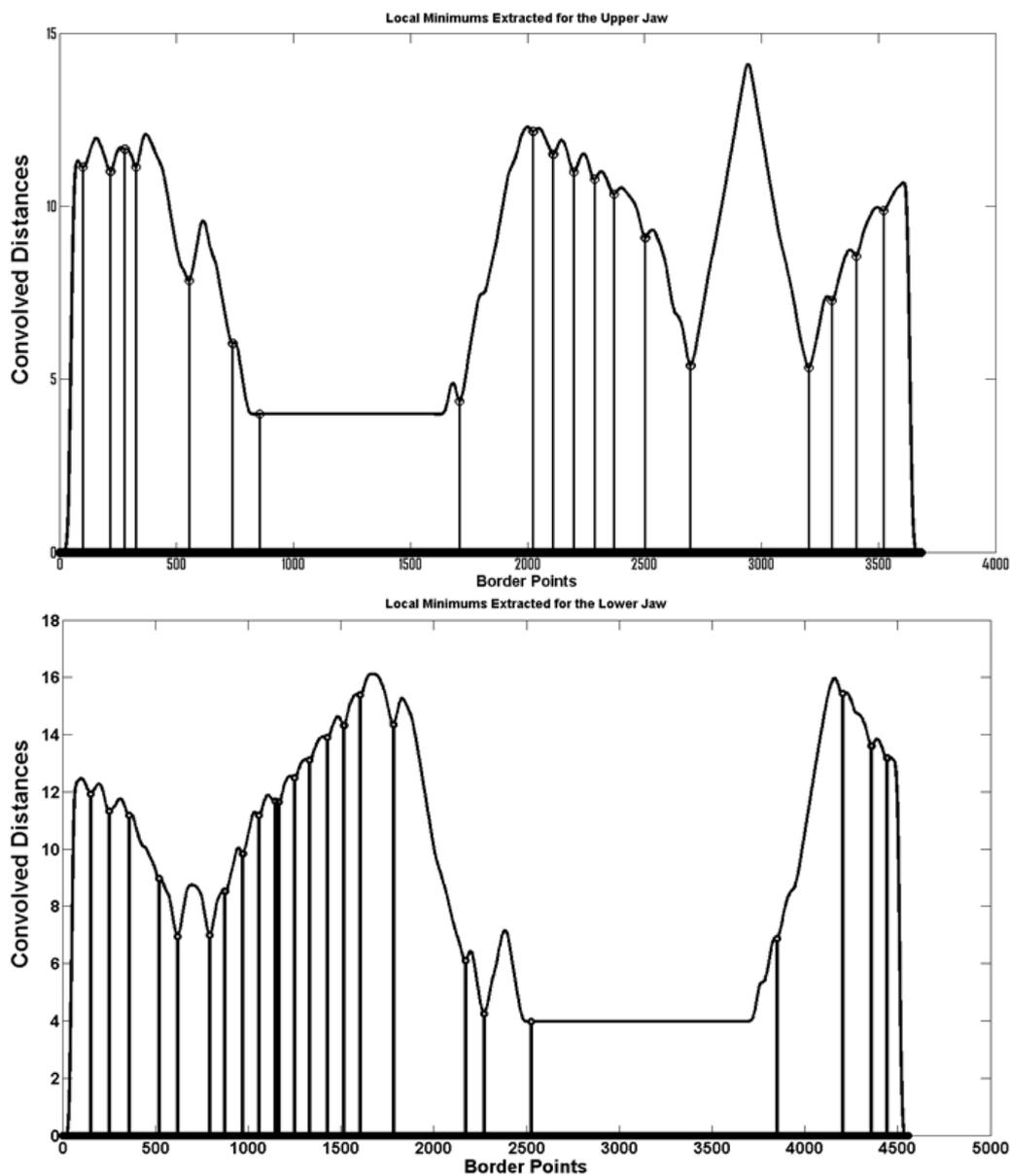


Figure 5.14: Output of the minimums extraction step, applied to the upper a lower jaw, respectively.

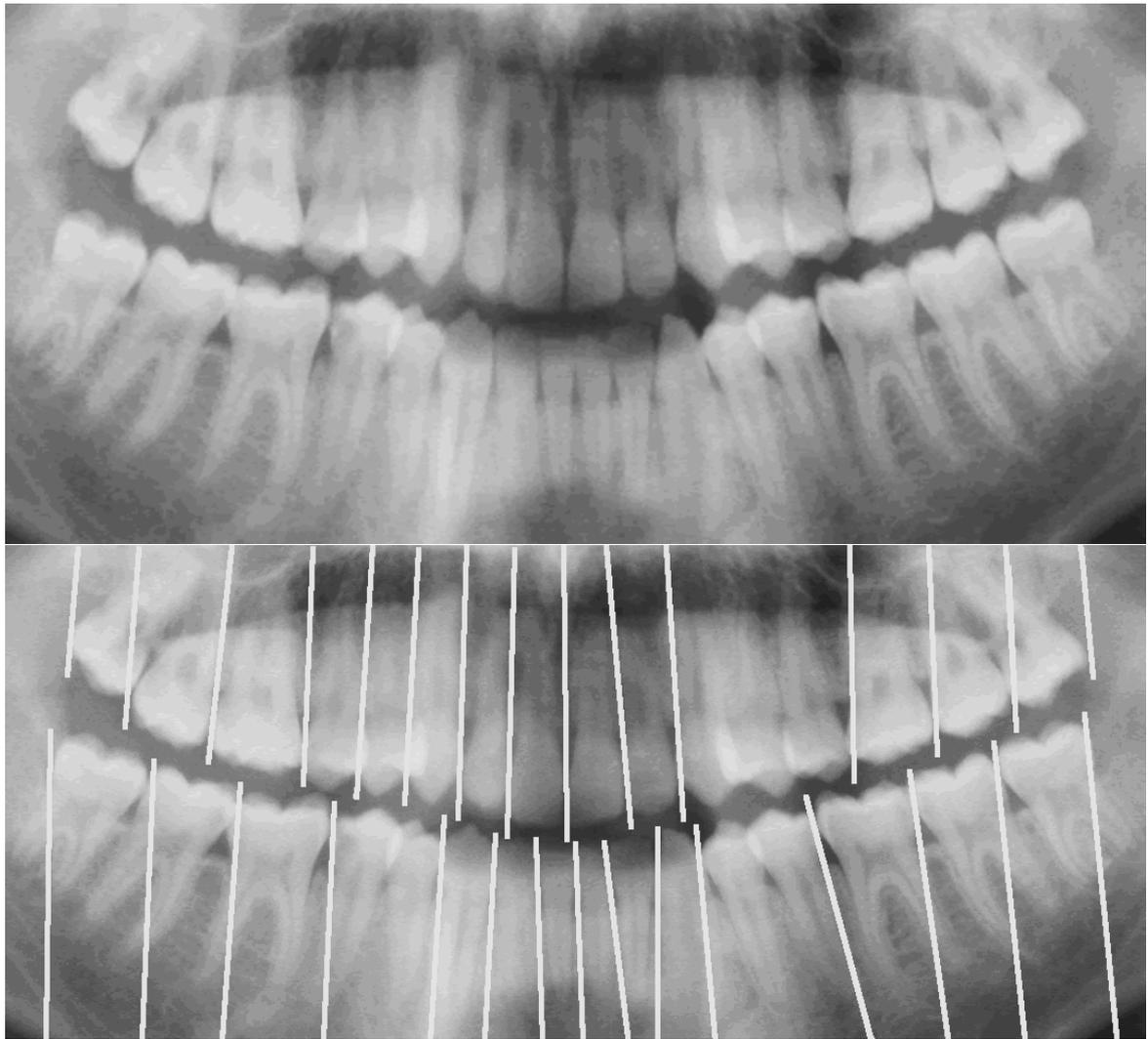


Figure 5.15: An example of our method 2 for the teeth gap valley detection.

The figure 5.15 shows that the extraction of the minimum is much better when compared with our original method, but as stated above, the transition back to cartesian coordinates corresponds in sometimes in the cut of some teeth when the lines are drawing in the cartesian coordinates system. Demonstrating that some improvements need to be carry out in this final step. When referring to the results obtained in the system of polar coordinates to Cartesian coordinate system. It is important also to highlight the non-detection when the teeth overlap is large, thus not enabling the detection of teeth gap valleys between them.

5.3 Conclusion

In conclusion in this chapter is presented two methods to perform the teeth gap valleys detection, booth methods are based on the previous step, jaws partition 4.

The first method of the line sum intensities consists in the extraction of the minimums. Minimums that are originated by the extraction of the lines with lower intensities. These lines are computed by the mining for each polynomial point of the previous step, of the sum of the intensities for each line. Where the values of the declivity vary in a $[-15^{\circ}; +15^{\circ}]$ interval. After the smooth of the curve, corresponding to the minimums values of each polynomial point, the extraction of the minimums is perform, as we can see in figure 5.3 and in figure 5.2. This method accuracy relies very much on the minimums selection, because the values between the points tend to had small variations. For the improvement of this method we implement the method 2, that improve substantially the accuracy of this step. The work that has been accomplished until this method, was recognized based on the publication of a paper to the VIPImage'09 conference. The article accepted and developed by us is in the appendix 8.4.

The method 2 that is the center distance with the active contours is divided in, applying the pre processing top and bottom hat method for the enhanced contrast image as shown in figure 5.4. After that is perform the jaws partition based on the polynomial extracted on the previous step, as demonstrated in figure 5.6. Then is performed the transformation of the image in polar coordinates in figure 5.7. Subsequent to this transformation we apply the active contour without edges technique for the segmentation of the teeth in each jaw, as shown in figure 2.1. Latter

we calculate for each border point the distance to the center of the image shown in figure 5.10 and in figure 5.11. Then through the distances vectors and applying the cut $0.10 * h$, where h corresponds to the image height for each point and in the end applying the convolution algorithm with a gaussian filter. Then we extract the minimums values of the distance vectors, corresponding to the figure 5.14. Finally and having the minimums points we perform the extraction of a second point in the neighbor of the minimum point and with this two points we execute the region cut of each teeth. This is based on the transformation of the extracted points back to the cartesian coordinates, the final method result is shown in the figure 5.15.

In the next chapter is introduced the tooth segmentation step and finally the dental caries detection, where for this step we present all features extracted and the features that best represent our training set, through the implementation of the PCA algorithm.

Chapter 6

Tooth Segmentation and Dental Caries Detection

The final steps of our method are described in this chapter. Firstly and based on the result of the previous step of the teeth gap valleys detection, we perform the tooth segmentation. As we can observe in the previous chapter the output results don't containing only the region corresponding to the segmented tooth, it contains a larger region with one single tooth. For this tooth segmentation we use the information of the images where we apply the pre processing top and bottom hat method. Therefore the actual tooth zone that we segment are applied on those images. After the segmentation is performed we extract the correspondent region in the original input image. Where through this region we perform the dental caries feature extraction. That corresponds to our final method step. The features extracted are based on the region, statistic, region boundaries, image region features, and more. In this chapter we also present the best features of all features extracted. This is execute by the use of the PCA algorithm, that corresponds to a reduction of the features space. This chapter is divided in two main sections one that refers to the tooth segmentation and the other to the dental caries detection based on the feature extraction. Therefore the evaluation of the various steps of this method are demonstrated in the results chapter 7.

6.1 Tooth Segmentation

In this step is execute the tooth segmentation, this process is quite similar to the method 2 of the teeth gap valleys detection in chapter 5. The input images of this stage are the images containing only one tooth. Where the problem of the input images is that they only contains one tooth but in the majority of the cases all images contains extra background corresponding to the gum line and the gum itself. In other cases the teeth overlap that corresponds to a presence of a neighbor tooth.

In figure 6.1 are some examples of the input images for this stage. That shows exactly the type of input images.

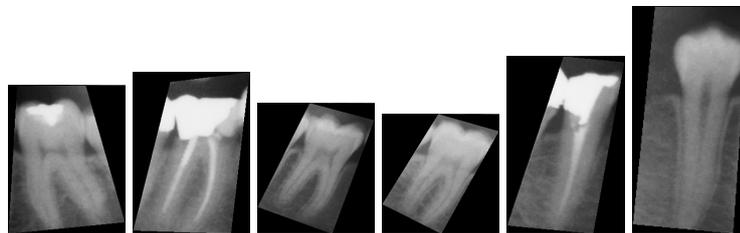


Figure 6.1: Example of our input images for the tooth segmentation stage.

For this step our main goal is to perform the tooth segmentation, where our desirable output it's the region that contains exactly the tooth present in the image. Unfortunately the output doesn't correspond exactly to this desirable output.

In the segmentation the images that we use to segment are the corresponding images of the figure 6.1 in the top and bottom hat transformation, of the previous step method 2 5.2. In figure 6.2 is shown the same images but after the pre processing transformation of the bottom and top hat method.



Figure 6.2: Example of our input images for the tooth segmentation stage after the Top and Bottom Hat transformation.

The segmentation in this step is perform by the use of the active contours without edges method. After the appliance of the active contours without edges we perform

the pre processing transaction, in effort to the removal of the maximum background that is present in our input images. The pre processing applied is the top and bottom hat transformation. As in the previous step the pre-processing is applied as shown in equation 6.1.

$$\begin{aligned} I_{output} &= (I_{Original} + TH_{Original}) - BH_{Original} \\ I_{outputFinal} &= (I_{output} + TH_{output}) - BH_{output} \end{aligned} \quad (6.1)$$

Where the $I_{Original}$ corresponds to our input image, in grayscale [0;255], the $TH_{Original}$ is the result of applying the operation of the top hat transform to our original image and finally the $BH_{Original}$ that concerns to the applying of the bottom hat transform in our original image. The $I_{outputFinal}$ is the resulting image of this operation of pre processing. In figure 6.3 is an example of the pre processing output applied in our images. We use a rectangular structuring element with dimensions $[w/4, h/2]$, where w and h are the width and height of the image, respectively. Our experimental choice of these structuring element is based on an experimental study on a set of 500 input tooth images, as shown in figure 6.1.

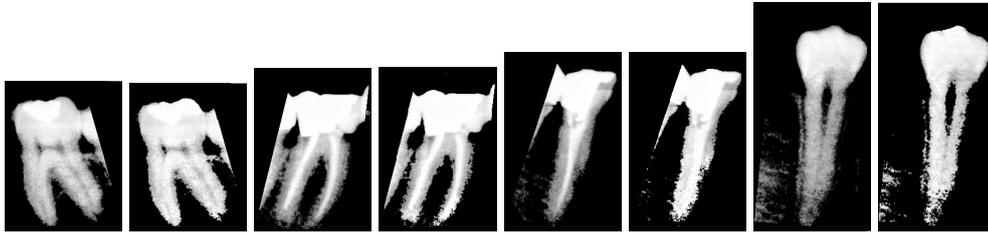


Figure 6.3: Example images of the top and bottom hat transform.

As we can see in the figure 6.3, the application of pre processing increased the contrast of the image, highlighting the dark areas and the bright areas of the original image. In other case the pre processing in teeth that were present in the original image that already do not contrast with the rest of the image tend to disappear with the implementation of pre processing. This occurs only in some due to the poor quality of the input image that disables the good effectiveness of the pre processing.

After the pre processing step, we perform the tooth segmentation through the active contours without edges that is a method of the minimization of an energy based-segmentation. Based on the exposed in chapter 2.2.4.5 and in the equation 2.8 that corresponds to the functional minimization of the energy.

After a study applied in 500 images of our input images, as shown in figure 6.1, we define our original mask corresponding to a four independent elements in each image quadrant, as illustrated in the equation 6.2.

$$I_{mask}(x, y) = \begin{cases} ([w * 0.20... (w/2) * 0.40], [h * 0.20... (h/2) * 0.40]) = 255 \\ ([(w/2) + (w/2) * 0.60... w * 0.80], [h * 0.20... (h/2) * 0.40]) = 255 \\ ([(w/2) + (w/2) * 0.60... w * 0.80], [(h/2) + (h/2) * 0.60... h * 0.80]) = 255 \\ ([w * 0.20... (w/2) * 0.40], [(h/2) + (h/2) * 0.60... h * 0.80]) = 255 \end{cases} \quad (6.2)$$

Where w and h corresponds to the image width and height, respectively. An example of the mask is given in figure 6.4.

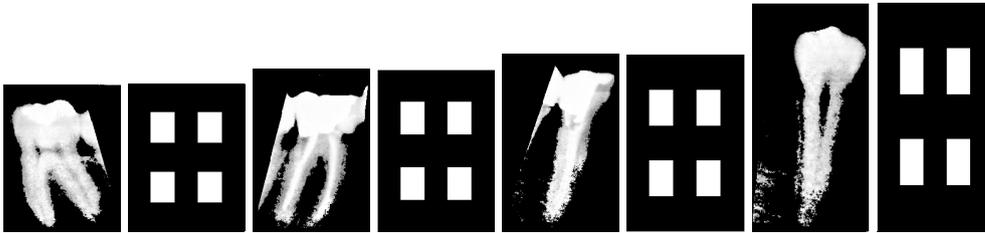


Figure 6.4: Some mask examples of our initialization for the active contour method.

The use of this mask is based on an experimental study where through this initial mask we accomplished better results for the tested images. The test set corresponds to a 500 set of our input images. Whereas observed in figure 6.1 the extra noise and the low contrast denotes in some cases of a ineffective pre processing, therefore some improvements in this step are required.

Taking our images already processed with the bottom and top hat transform our next goal is to perform the active contour without edges algorithm. Some results of this operation are presented in the figure 6.5 and in figure 6.6. In the figure 6.7 are some examples for this step where the presence of extra noise in the image make the output image not containing only one tooth but also the gum line and the gum itself.

As shown in the figure 6.5 the application of active contours appears to have been a very good option. It must be stressed that all the images segmented contain the tooth, i.e., the region corresponding to the tooth is always present in the resulting image, although sometimes due to the presence of extra noise as mentioned early the

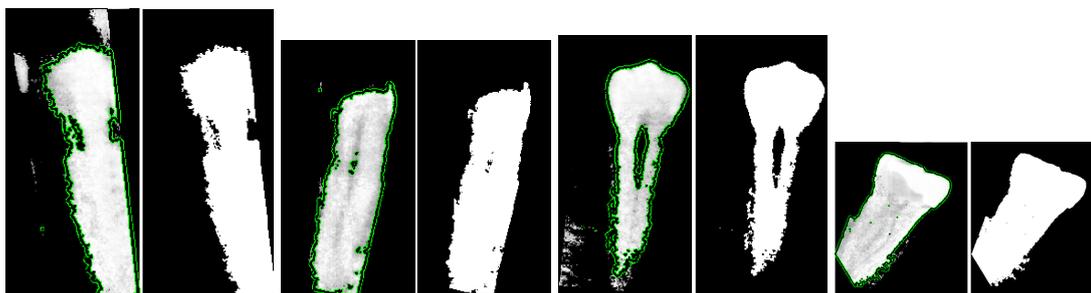


Figure 6.5: Examples of the active contours method applied to our input images.

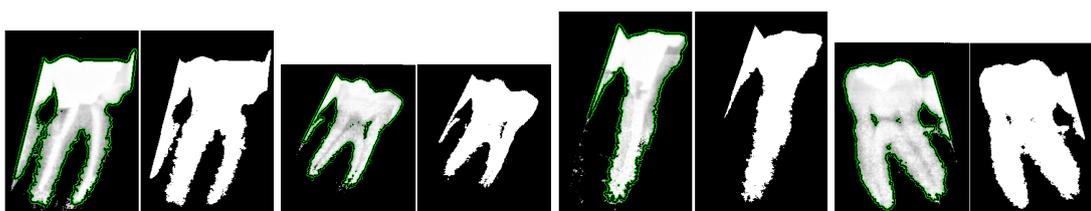


Figure 6.6: Examples of the active contours method applied to our input images with overlap teeth.

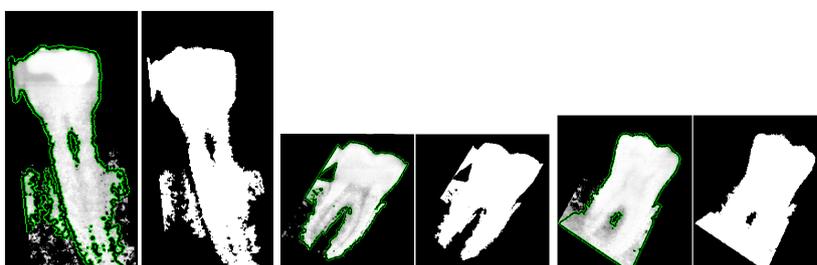


Figure 6.7: Examples of the active contours method applied to our input images with extra noise present in the output image.

segmentation result is not so accurate, this is shown in figure 6.7. In other cases the presence of the overlap teeth where these teeth in the segmentation result are also considerable as belonging to the tooth that we want to segment, as demonstrated in figure 6.6.

We also studied the possibility of applying the active contours technique directly into our original images and to the original images with the appliance of the pre-processing. But this test proved to be ineffective because the resulting images were not as satisfactory as those obtained in this step now.

In this step it is essential to recognize the type of tooth that is present in the image as for the pre molars and molars teeth the pre-processing operation can be more rigid. With a larger structuring element but unlike what happens to the incisors and canines teeth where this pre processing time ends up by eliminating much of the teeth, thus making it very difficult to detect cavities or dental caries. Hence the importance of future work as we realize the recognition of the type of teeth that are presented to the system. So we can define rules based on more specific characteristics, contrary to what happens now we have some more general rules for the proper functioning of the system and the correct classification of dental caries.

6.2 Dental Caries Detection

The next step is the detection of cavities or dental caries, for that we perform the extraction of features which in our view better characterized the problem in question. To do this we create a training set of 1098 images, which contain 549 images with dental caries and 549 images containing healthy teeth. In this set of training exists all sorts of teeth, but with more presence of pre molars and molars, because even in our original data-set of images all the cavities were largely in molars and pre molars. Doing so the reason for our choice to include more of this type of teeth, therefore we are consistent with the data that we have.

As will be explained later there are different types of features extracted, hence the importance of using algorithms for selecting the best features, which we choose to use the PCA algorithm. Ahead we will make a brief introduction to the PCA algorithm and its main features and applicability in the selection of the best features.

6.2.1 Dental Caries

As referred in previous chapters the dental caries, or cavities or also known as tooth decay is due to an Acid Base Chemistry that has excess amount of acid. The disease is associated with sugar bugs. Sugar bugs are adult bacteria that thrive on white refined sugar. Remains of sticky candy and other white sugar sweets attach in between the teeth and in the grooves of the teeth. The sugar bugs attack the remaining particles because they have the enzymes to digest the white sugar to use for energy. The sugar bug release byproducts of strong acids that eat through the enamel of the teeth. The sugar bug dominates the environment and influences other remaining bacteria to become allies in decaying the teeth. The strong acids are not the right environment for enteric bacteria and they seem to disappear. They either die or they go back to being spores. The key pathogen has been identified as "Streptococcus mutans".

6.2.2 Feature Extraction

In the process of extracting features is to emphasize that far contributed the literature reviewed by us, only then we could reach the knowledge of certain characteristics which have proved to be fairly representative of our data set training. Therefore it is essential to perform a study by characteristic, this study will be shown in the next pages of our work. For starter all the features that were extracted on our images and then what were the best or those that best represent our training set.

In this section we will present all features that were extracted from our input images. This features consists in five groups, features based on the image characteristics like pixel intensity, maximum pixel intensity, etc. Statistic features as the entropy, mean, variance and more. Region based features such as the Hu moments, the Zernike moments, area, perimeter, etc. Features based on the region segmented boundary, like chain code, Fourier descriptors, signature, angular function, among others. And finally features based on the image texture, like the energy, third moment, etc. There are several features that can be extracted from the image properties, for our work we extracted the following features:

- Maximum pixel value, as $max_{InputImage}$ corresponds to the maximum pixel value in the input image.

- Minimum pixel value, as $min_{InputImage}$ corresponds to the minimum pixel value in the input image.
- $max_{InputImage} - min_{InputImage}$ that gives the value of the maximum pixel value minus the minimum pixel value.

The others features extracted based on the image pixel values are shown in the next sub section.

6.2.2.1 Statistical Features based on the image properties

The statistical characteristics of our input images is to extract the information that exists which may be valuable for the detection of dental caries. For this is to highlight the fact that all the statistical features that are extracted have a great emphasis on literature as well as the first option in many of the problems of pattern recognition and the like.

Such features include a large number of applications, in our work we use several times this type of characteristics for different groups of features that we extracted from our images. For example for the characteristics based on the border, or the tooth boundary, when using this type of characteristics it enables the retrieval of some information that is present in the tooth contour or in the tooth boundary. Knowing that not always such characteristics are recommended because they may add more noise to the training set. Therefore its use is rationed according to the relevance of the value for the feature in question.

The statistical features extracted are the arithmetic mean, that corresponds to the equation 6.3.

$$\bar{x} = \frac{1}{h * w} * \sum_{x=1}^w \sum_{y=1}^h I(x, y) \quad (6.3)$$

Where h and w corresponds to the image height and image width, respectively. And $I(x, y)$ is the corresponding pixel value for the x and y coordinates. The median that corresponds to the medium value of all the pixel values present in the input image. Although, the median is less efficient than the mean, it is less sensitive to outliers than the mean. Other extracted features is the mode that corresponds to the most frequent pixel values in the input images because the input images could be

multimodal, i.e, exists various pixel values that equally occurs in the image. For that fact the distribution of the values are denominated as multimodal. The standard deviation calculated as shown in equation 6.4.

$$s_{(w*h)-1} = \sqrt{\frac{1}{(w * h) - 1} * \sum_{x=1}^w \sum_{y=1}^h (I(x, y) - \bar{x})^2} \quad (6.4)$$

The w corresponds to the image width, and h to the image height. The $I(x, y)$ is the corresponding pixel value for the x and y coordinates and the \bar{x} is the mean as in the equation 6.3. The result $s_{(w*h)-1}$ is the square root of an unbiased estimator of the variance of the image pixel values, from which $I(x, y)$ is drawn. The variance that corresponds to the square of the standard deviation, which in this case is represented as demonstrated in equation 6.5.

$$\begin{aligned} \sigma^2 &= \left(\sqrt{\frac{1}{(w*h)-1} * \sum_{x=1}^w \sum_{y=1}^h (I(x, y) - \bar{x})^2} \right)^2 \\ &= \frac{1}{(w*h)-1} * \sum_{x=1}^w \sum_{y=1}^h (I(x, y) - \bar{x})^2 \end{aligned} \quad (6.5)$$

In this equation w corresponds to the image width, and h to the image height. The $I(x, y)$ is the corresponding pixel value for the x and y coordinates and the \bar{x} is the mean as in the equation 6.3. Finally the statistic feature extracted is the entropy, entropy is the measure that gives us the "disorder" of a system in this case the use of the entropy in our images give us the measure of the variety of pixel values in the image. Entropy is given by the equation 6.6

$$H(I(X, Y)) = - \sum_x^w \sum_y^h P(I(x, y)) * \log_2 P(I(x, y)) \quad (6.6)$$

Where w and h corresponds to the image height and image width, respectively. The $I(x, y)$ is the corresponding pixel value for the x and y coordinates. The $H(I(X, Y))$ corresponds to the probability that $I(X, Y)$ is in the state $I(x, y)$.

These are the main statistical features extracted and they will be used more than once for the extraction of other statistical data. These data will be used in features based on the boundary or tooth border.

6.2.2.2 Region based features

In this section we will demonstrate all the features extracted based on the region of the previous step, the tooth segmentation. At This stage is important to stress that the region used is from the segmented area of the previous step, i.e., the binary image that corresponds to image mask of the output segmentation, as demonstrated in equation 6.7.

$$I_{region}(x, y) = \begin{cases} I_{region}(x, y) = 255, & \text{if } I_{region}(x, y) \subset B_{boundary} \\ I_{region}(x, y) = 0, & \text{otherwise} \end{cases} \quad (6.7)$$

One of the most popular region features are the moments. Moments describe the shape's layout of the object in this case the arrangement of the image pixels. M.K. Hu (1962) derived a transformation of the normalized central moments to make the resulting moments rotation invariant. These moments continue to be published in books on image processing. The first six of these are also invariant to reflection, while the last one changes sign.

In many applications such as shape recognition, it is useful to generate shape features which are independent of parameters which cannot be controlled in an image. Such features are called invariant features. There are several types of invariance. For example, if an object may occur in an arbitrary location in an image then one needs the moments to be invariant to location. For binary connected components this can be achieved simply by using the central moments, m_{pq} .

If an object is not at a fixed distance from a fixed focal length camera, then the sizes of objects will not be fixed. In this case size invariance is needed. This can be achieved by normalizing the moments. The third common type of invariance is rotation invariance. This is not always needed, for example if objects always have a known direction as in recognizing machine printed text in a document. The direction can be established by locating lines of text.

The non-orthogonal centralized moments are translation invariant and can be normalized with respect to changes in scale. Nevertheless, to enable invariance to rotation they require some reformulation. Hu moments described two different methods for producing rotation invariant moments. The first used a method called principal axes, however it was noted that this method can break down when images do not have unique principal axes. Such images are described as being rotationally

symmetric. The second method Hu described is the method of absolute moment invariants. Hu derivative these expressions from algebraic invariants applied to the moment generating function under a rotation transformation. They refers to a set of groups of nonlinear centralized moment expressions. The result is a set of absolute orthogonal (i.e. rotation) moment invariants, which can be used for scale, position, and rotation invariant pattern identification. In equation 6.8 are the seven Hu moments[18] [35].

$$\begin{aligned}
\phi_1 &= \eta_{20} + \eta_{02} \\
\phi_2 &= (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \\
\phi_3 &= (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \\
\phi_4 &= (\eta_{30} - \eta_{12})^2 + (\eta_{03} + \eta_{21})^2 \\
\phi_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\
&\quad (\eta_{03} - 3\eta_{21})(\eta_{03} + \eta_{21})[(\eta_{03} + \eta_{21})^2 - 3(\eta_{12} + \eta_{30})^2] \\
\phi_6 &= (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + \\
&\quad 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{03} + \eta_{21}) \\
\phi_7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + \\
&\quad (\eta_{30} - 3\eta_{12})(\eta_{03} + \eta_{21})[(\eta_{03} + \eta_{21})^2 - 3(\eta_{12} + \eta_{30})^2],
\end{aligned} \tag{6.8}$$

Where

$$\phi_{pq} = \frac{\mu_{pq}}{\mu_{00}^\gamma}, \tag{6.9}$$

And

$$\gamma = \frac{p+q}{2} + 1, \tag{6.10}$$

The properties of invariant to the rotation and translation, are obtained through the relation between the normalized moments and central moments, as defined in equation 6.11.

$$\mu_{pq} = \sum_x^w \sum_y^h (x - \bar{x})^p (y - \bar{y})^q I(x, y) \Delta A, \tag{6.11}$$

Where \bar{x} and \bar{y} corresponds to the center mass of the object, given by the equation 6.12, the w and h corresponds to the image respectively width and height and $I(x, y)$

concerns the pixel value in the position (x, y) .

$$x = \frac{m_{10}}{m_{00}} \quad y = \frac{m_{01}}{m_{00}} \quad (6.12)$$

These moments are of finite order, therefore unlike the centralized moments they do not comprise a complete set of image descriptors. However, higher order invariants can be derived, as the Zernike moments explained below. It should be noted that this method also breaks down, as with the method based on the principal axis for images which are rotationally symmetric as the seven invariant moments will be zero.

The Zernike polynomials were first proposed in 1934 by Zernike. Their moment formulation appears to be one of the most popular outperforming the alternatives (in terms of noise resilience, information redundancy and reconstruction capability). Complex Zernike moments are constructed using a set of complex polynomials which form a complete orthogonal basis set defined on the unit disc $(x^2 + y^2) \leq 1$. They are expressed as A_{pq} Two dimensional Zernike moment, as shown in equation 6.13.

$$A_{mn} = \frac{m+1}{\pi} \int_x \int_y f(x, y) [V_{mn}(x, y)]^* dx dy \text{ where } x^2 + y^2 \leq 1 \quad (6.13)$$

where $m = 0, 1, 2, \dots, \infty$ and defines the order, $f(x, y)$ is the function being described and $*$ denotes the complex conjugate. While n is an integer (that can be positive or negative) depicting the angular dependence, or rotation.

For our work the value of the Zernike [21][27] moments defined by us are $m = 0 \dots 3$, our experimental choice of these Zernike moments is based on an experimental study on a set of 100 dental images containing only one tooth, as shown in the figure 6.1.

The major disadvantage of moments in general is that they are global features rather than local. This makes them not suited for recognizing objects which are partially obstructed. Moments are inherently location dependent so some rules must be adopted to insure location invariance (like the centroid). In conclusion the note that we decide to use also Zernike moments because they are superior to the Hu moments concerning the rotation invariance.

More classical region features were extracted, as listed below:

- **Area:** that measures the number of pixels inside the region. Let $I(x, y) = 255$ for any $x \in [1; w]$ and for any $y \in [1; h]$, where h and w corresponds respectively to the image height and width, the pixel values corresponding to the pixels belonging to the image region so the area is calculated by the equation 6.14.

$$Area_{I(x,y)} = \sum_x^w \sum_y^h I(x, y) = 255 \quad (6.14)$$

- **Perimeter:** corresponds to the number of pixels that composed the boundary image, or the border image. Let $B(x, y) = 255$ for any $x \in [1; w]$ and for any $y \in [1; h]$, where h and w corresponds respectively to the image height and width, the pixel values corresponding to the pixels belonging to the image boundary, so the perimeter is calculated by the equation 6.15.

$$Perimeter_{I(x,y)} = \sum_x^w \sum_y^h B(x, y) = 255 \quad (6.15)$$

- **Diameter:** the diameter is the maximum distance between any two pixels in the corresponding region.
- **Major Axis:** Contains the coordinates for the endpoints of the major axis to the corresponding region.
- **Minor Axis:** Contains the coordinates for the endpoints of the minor axis to the corresponding region.
- **Basic Rectangle:** Defines the region enclosing rectangle defined by the major and minor axes, for this feature e save the distance of the diagonal of this rectangle.
- **Euler Number:** This feature specifies the number of objects in the figure minus the number of holes in those objects. This feature is very dependent on the active contour initialization, because if we initialize the contour with one single region, it will never segment the holes inside the tooth region therefore the importance of the active contour initialization. If there is a hole in the tooth image this could represent a immediate dental caries.
- **Centroid:** Corresponds to the coordinates of the center mass of the region, this feature for the future when calculation the distances of all boundary pixels to the centroid of the image.

- **Convex Hull:** Defines the smallest convex polygon that the region can contain. Associated to this feature is also the **Convex Area**, that corresponds to the area of the convex polygon.
- **Solidity:** This corresponds to the proportion of the number of pixels that are in the Convex Hull and in the same time in the region containing the tooth.
- **Eccentricity:** Specifies the eccentricity of the ellipse that has the same second-moments as the region. The eccentricity is the equivalent to ratio of the distance between the foci of the ellipse and its major axis length. The output value is between 0 and 1, in the cases that the output is 0, the ellipse is actually a circle, in cases where the eccentricity is 0 the ellipse becomes a line segment.
- **Orientation:** Based on the previous feature, we can extract the orientation of the image based on the ellipse that corresponds to the the angle between the x-axis and the major axis of the ellipse. Having the same second-moments as the region the same as the previous feature.

The characteristics based on region features are essential in the pattern recognition whose shape can vary, which is the case when referring to the teeth cavities that partly or completely destroy the tooth crown. The system becomes more inaccurate in matter of the detection of dental caries. Besides the use of region based features would be important the use of a system of rules that automatically classify these teeth, thus in part also to conduct in future work. In the following section we will discuss the features based on the border of the region extracted by us such features are also important in the detection of partial or total destruction of the tooth border or the tooth boundary.

6.2.2.3 Boundary or Border Tooth Features

As shown in the previous section a region usually describes contents, or interior points, which are surrounded by a boundary which is often called the region's contour, or the region's boundary. The contour form is in generally referred as the object shape. A point can be defined to be on the contour (boundary) if it is part of the region and in the same time there is at least one pixel in its neighborhood that is not part of the region. This means in the type of images that we use and based on the tooth segmentation that the pixel to be considered in contour must

have a pixel with intensity equal to zero. The boundary itself is usually found by iterative contour following, i.e., based on one first point on the boundary the iterative process is to extract the other boundary points based on the progress in the clockwise direction or anti-clockwise direction. To accomplish this task we must have in consideration the relationship between pixels in the neighborhood. The relationships are described by means of connectivity rules. There are two common ways of defining connectivity: 4-way (or 4-neighborhood) where only immediate neighbors are analyzed for connectivity or 8-way (or 8-neighborhood) where all the eight pixels surrounding a chosen pixel are analyzed for connectivity. These two types of connectivity are illustrated in figure 6.8. In this figure 6.8, the pixel is shown in light grey and its neighbors in dark grey. In 4-way connectivity a pixel has four neighbors in the directions north, east, south and west, its immediate neighbors. The four extra neighbors in 8-way connectivity, are those in the directions north east, south east, south west and north west, the points at the corners [35].

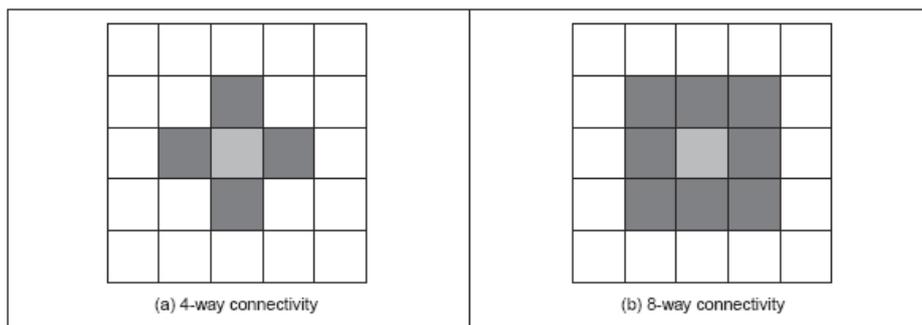


Figure 6.8: The main types of connectivity analysis, image extracted from [35].

There exists a variety of boundary representation algorithm, one of the oldest is the Chain Codes firstly introduced in 1961 by Freeman. The basics is the representation of the relative position between consecutive pixels. This is performed by the use of the different main types of connectivity as shown in figure 6.8. But in this case it is necessary to define extra rules, consisting in the order of the pixel visits. As shown in figure 6.9 and based on the [35] we define in that way our priority pixel access, in clockwise direction.

The chain code is formed by concatenating the number that designates the direction of the next pixel. That is, given a pixel the successive direction from one pixel to the next pixel becomes an element in the final code. This is repeated for each point until the start point is reached when the (closed) shape is completely

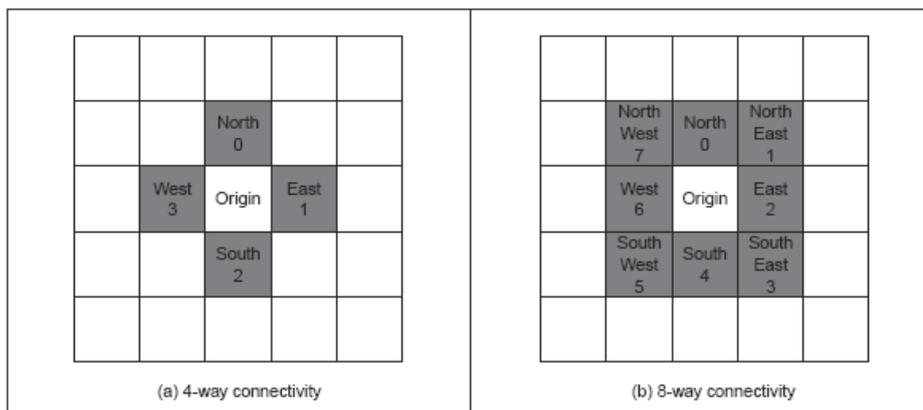


Figure 6.9: The order that our pixels are visited, this figure is extracted from [35].

analyzed. But this process is very dependent on the initial point because different starting points originates different chain codes, so the goal is to make this chain code invariant to the initial point. The process consists in actually look to the chain code as an integer, and make the shift position operation. The stopping criterion is when we are in the presence of the smallest integer represented by our chain code. Although our chain code with this step became invariant to the initial point, but it's not invariant to the rotation of the object. This is perform by the use of differential codes where the new chain code is the result of the $C'_{Image}(i) = C_{Image}(i) - C_{Image}(i-1)$, where C_{Image} is the original chain code and C'_{Image} is the resultant chain code after the appliace of the differential codes.

After having our chain codes the features extracted were based on the statistic features that we describe above and the representation of the chain code in the frequency space, where more statistic features were extracted and others like frequency, the difference of the real part of the signal and the imaginary part of the signal.

Another representation of the boundary of the object, or the object shape, is the use of the angular function. The angular function is represented by $\phi(s)$, and measures the direction of the tangent line to a object point taking in consideration the length of the arc. Given an initial point of the tooth contour is measure (in regularly intervals) the declination of the respective tangent lines. But the major problem with the angular function is that is based on the tangent and the tangent is a discontinuity function, closed in the interval $[0, 2\pi]$. For that is used the cumulative angular function that defines angular variation of the cumulative differential curvatures (k)

since the initial point ($k(0)$). Therefore the cumulative angular function in a point s of the object boundary is defined by 6.16.

$$\gamma(s) = \int_0^s K(r) dr - k(0) \quad (6.16)$$

The cumulative angular function avoids the discontinuity of the angular function but nevertheless continues to be discontinuous at the extremes, thus creating the representation of closed objects open. Also varies between the size of the object analyzed. These problems can be solved through a process of normalization set by the equation 6.17.

$$\gamma^*(t) = \gamma\left(\frac{L}{2\pi}t\right) + t \quad (6.17)$$

Where t vary in the interval of $[0, 2\pi]$. The normalizing factor $\frac{L}{2\pi}$, normalizes the function so the values not varies concerning the size of the object. The term t avoid discontinuities in the extreme of the function and guarantees that $\gamma^*(0) = \gamma^*(2\pi) = 0$.

The normalized cumulative angular function denotes a great performance for the boundary description. For this boundary representation we also extracted all the statistical features and the features in the frequency domain, through the Fast Fourier Transform (FFT). This algorithm implements the discrete Fourier Transform to transform data from time into the frequency domain.

Based on the Fourier theory exists a method to represent the boundary of a region as a periodic function which can be expanded in a Fourier series, this method is known as the Fourier Descriptors. These frequency-domain descriptions provide an increasingly accurate characterization of shape as more coefficients are included. The main idea is to characterize a contour by a set of numbers that represent the frequency content of a whole shape. Based on the frequency analysis is selected a small set of numbers (know as Fourier coefficients), first we have to define a representation of a curve in this case the cumulative angular function, and the expand it using the Fourier theory. In the end and after some algebraic simplifications, our Fourier coefficients are defined as shown in equation 6.18 for

the discrete image processing analysis, based on [35].

$$a_k^* = -\frac{1}{\pi k} \sum_{i=1}^m \kappa_i \tau_i \sin\left(\frac{2\pi k}{L} S_{i-1}\right) \quad (6.18)$$

$$b_k^* = -\frac{1}{\pi k} \sum_{i=1}^m \kappa_i \tau_i \cos\left(\frac{2\pi k}{L} S_{i-1}\right)$$

These descriptors a_k and b_k are invariant to the scale and translation, for the rotation invariance we must calculate the magnitude of the descriptors, as defined in equation 6.19.

$$|c_k^*| = \sqrt{(a_k^*)^2 + (b_k^*)^2} \quad (6.19)$$

The cumulative angular function transforms the two-dimensional description of a curve into a one-dimensional periodic function suitable for Fourier analysis. In a different point, the elliptic Fourier descriptors maintain the description of the curve in a two-dimensional space (Granlund, 1972). This is achieved by considering that the image space defines the complex plane. That is, each pixel is represented by a complex number. The first co-ordinate represents the real part whilst the second co-ordinate represents the imaginary part. The coordinates of the image are represented in complex plane as in the equation 6.21.

$$s(t) = x(t) + jy(t) \quad (6.20)$$

Where the parameter t is given by the arc-length parametrization. The fourier elliptic descriptors invariant to rotation, translation and scale are defined as shown in equation 6.21.

$$\frac{|A_k|}{|A_1|} + \frac{|B_k|}{|B_1|} \quad (6.21)$$

Where

$$\frac{|A'_k|}{|A'_1|} = \frac{\sqrt{a_{xk}^2 + a_{yk}^2}}{\sqrt{a_{x1}^2 + a_{y1}^2}} \quad \text{and} \quad \frac{|B'_k|}{|B'_1|} = \frac{\sqrt{b_{xk}^2 + b_{yk}^2}}{\sqrt{b_{x1}^2 + b_{y1}^2}} \quad (6.22)$$

For our work the value of the Fourier coefficients [35] defined by us are $m = 20$, our experimental choice of these number of coefficients is based on an experimental

study on a set of 100 dental images containing only one tooth, as shown in the figure 6.1. The coefficients defined were used for the fourier descriptors and the elliptic fourier descriptors.

To these different types of representation of the image boundary, based on the Fourier Theory, we extract the statistic features of these coefficients and the 20 coefficients originated by the types of tooth border representation.

The signature of a border of an object is the functional and dimensional representation of that border. The simplest way to extract a signature of the border is the calculation of distances from all points of the border to the center of the region, i.e., the centroid according to an angle. The signature alone is invariant to translation but it is not for rotation and changes of scale. To be invariant to the rotation will always extract those angles starting in the same point, when concerning the extraction of distances. For example as performed by us, always started at the border point farthest from the region centroid. To become the signature invariant to changes of scale will make a mandatory standardization of it, for example by dividing the function $\rho(\theta)$ by its standard deviation σ . Based on this representation we perform also the representation of the border by the distance to the centroid, but in this case, performing the distance in cartesian coordinates, not in polar coordinates as the signature process do. In this case we also start always from the farthest point in the image border to the centroid, and perform also the normalization of the function dividing it by its standard deviation.

In this particular case the representation based on the distance to the centroid of the region, is more accurate in absolute terms of the distance that a border point is from the object border. This is important when dealing with various types of input image, concerning the molars, pre molars, incisors and canine, because due to the teeth morphology is important to had some features that could perform the difference between the various types of input image.

A digital boundary can be approximated with arbitrary accuracy by a polygon. For closed boundary, the approximation becomes exact when the number of segments of the polygon is equal to the number of points in the boundary so that each pair of adjacent points defines a segment of the polygon. The goal of a polygonal approximation is to capture the essence of the shape in a given boundary using the fewest possible number of segments. This problem is not trivial in general and can turn into a time-consuming iterative search. However approximation techniques

of modest complexity are well suited for image processing tasks. Among these, one of the most powerful is a representation of the image boundary that is based on the polygonal approximations using Minimum-Perimeter Polygons (MPP). The theoretical underpinnings and an algorithm for finding MPP are discussed in the paper of Sklansky et al. [1972]. The method that we used is based on the [18], that is restricted to simple polygons, i.e., polygons with no self-intersections, and regions with peninsular protrusions. The finding algorithm of the MPP is based on the Sklansky's shrinking rubber band approach, as explained in [18].

The size of the MPP is then used as a feature for our tooth segmented region.

In conclusion of the region based features extracted it should be stressed that all the important statistic features of the above section were used for the extraction of features, concerning the chain codes in a 4 way-connectivity and in a 8 way-connectivity, regarding the fourier descriptors and the elliptic fourier descriptors, in the cumulative normalized angular function. The FFT were also used for the extraction of these features in the domain frequency.

In the next section is presented the region texture features that were used as features for our problem of the Dental Caries detection.

6.2.2.4 Region Texture Features

An important feature of the region is its texture, unfortunately there is no objective definition of texture, although there are measures steps to try to quantify. The texture can describe properties such as granularity, the smoothness and regularity of a region. The notion of texture admits to no rigid description, but a dictionary definition of texture as "something composed of closely interwoven elements". The description of interwoven elements is intimately tied to the idea of texture resolution, which one might think of as the average amount of pixels for each discernable texture element. If this number is large, we can attempt to describe the individual elements in some detail. However, as this number nears unity it becomes increasingly difficult to characterize these elements individually and they merge into less distinct spatial patterns. For a image in grayscale we define the co-occurrence C as a square matrix with size equal to the number of the different grayscale values, i.e, $N = 256$, in each position of $C(i, j)$ contains the probability of a pixel with intensity equal to i is the distance p of another with intensity j , in the

direction of θ . Therefore having two different matrix for each value of p and θ that we wish to consider. Through this matrix C is possible to extract various texture measures, for our work we computed the follows:

1. Energy:

$$E(p, \theta) = \sum_{i=0}^N \sum_{j=0}^N C(i, j)^2 \quad (6.23)$$

2. Entropy that measures the randomness:

$$H(p, \theta) = \sum_{i=0}^N \sum_{j=0}^N C(i, j) \log_2(C(i, j)) \quad (6.24)$$

3. Inertia:

$$\iota(p, \theta) = \sum_{i=0}^N \sum_{j=0}^N (i - j)^2 C(i, j) \quad (6.25)$$

4. Inverse Difference Moment (IDM):

$$IDM(p, \theta) = \sum_{i=0}^N \sum_{j=0}^N \frac{1}{1 + (i - j)^2} C(i, j) \quad (6.26)$$

5. Third Moment that measures the skewness of a histogram. This measure is 0 for the symmetric histograms, positive by histograms to the right and negative for histograms skewness to the left:

$$\mu_3(p, \theta) = \sum_{i=0}^N \sum_{j=0}^N (C(i, j) - m)^3 C(i, j) \quad (6.27)$$

As mentioned before, the texture allows, among other things, to characterize the regularity of the structures of the image. This regularity is in the form of periodic patterns. The Fourier transform spectrum is the ideal to identify periodic patterns or pseudo periodic patterns. One way to do this is based on the transformation of the spectrum representation, for it be expressed in the polar coordinates $S(\rho, \theta)$. In each direction of θ the spectrum can be considered as a unidimensional function, $S_\theta(\rho)$, which allows us to evaluate the spectrum along a given direction based on the radial component ρ . By analogy, we can fix a frequency ρ and analyze the behavior

of the circumference spectrum, using the function $S_{rho}(\theta)$. We can take advantage of properties that we have just seen and create the following functions:

$$S(\rho) = \sum_{\theta=0}^{\pi} S_{\theta}(\rho) \quad (6.28)$$

$$S(\theta) = \sum_{\rho=1}^{R_0} S_{\rho}(\theta) \quad (6.29)$$

Where θ varies on the interval $[0, \pi]$ because the spectrum is symmetric, and R_0 is the maximum value in consideration for ρ . The equation 6.28 gives us the sum of all the values from the spectrum in a semi-circumference that distance ρ of the origin. The equation 6.29 is the correspond to the sum of all the values from the spectrum along a given direction θ between $\rho = 1$ and R_0 . This two functions can be used for describing the region texture. Based on the feature extraction of the border of the region or the boundary of the region, in this step we also extract statistical features from the spectral texture analysis.

In this section we present all the features extracted of our training data-set, this features were based on the image properties, in the region based features that describes the output region of the tooth segmentation. Statistic features that corresponds to the image properties in the statistical analysis, these features were used not only directly in the image but also used in the feature extraction of other techniques used to represent our output region. The features based on the object boundary where in a first step the representation of the object boundary that in our case corresponds to the tooth region segmented output, this object boundaries descriptors were based on multiples techniques, like the chain codes, fourier descriptors, fourier elliptic descriptors and the cumulative normalized angular function, among others. Finally the features based on the region textures, performing a suitable region description in many problems of pattern recognition. In the next section is present the PCA that is used for the dimensionality reduction of our extracted features, by doing so it emphasizes the best features for our main goal, that corresponds to the dental caries detection.

6.2.3 PCA

When we are in presence of a great variety of features, becomes imperative the selection of the best ones, this selection could perform in many ways, empirically, visually but none taking the real advantage of the profound study of the features values. Fortunately the PCA exists. The main goal of the Principal Component Analysis is to perform a dimensionality reduction while preserving as much of the randomness (variance) in the high-dimensional space as possible. The PCA is the oldest technique in multivariate analysis, the PCA is also known as the Karhunen-Loève transform (communication theory), but the PCA was first introduced by Pearson in 1901, and it experienced multiple modifications until it was generalized by Loève in 1963.

After this historical introduction let's pass to the explanation of the PCA, for the extraction of the PCA in a set, we must effectuate the following steps:

1. Center data values, for each value we must subtract the mean value of the correspondent feature.
2. Retrieve the covariance matrix of the data, the covariance matrix corresponds to a square matrix, where size is equal to the number of features used for the data representation.
3. Perform the extraction of the eigenvectors and the eigenvalues of the covariance matrix.
4. Finally transform the data, i.e., the data projection in the coordinates formed by the eigenvectors.

The eigenvalues are proportional to the data variance associated to the corresponding eigenvector, i.e., that the eigenvector with the greatest eigenvalue corresponds to the eigenvector with more data variance. The second eigenvector with the greatest eigenvalue corresponds to the eigenvector with the second most data variance, and so on. The main components are exactly the directions for the largest variability in the data, i.e., the directions indicated by the eigenvectors associated with higher values of eigenvalues.

6.2.3.1 PCA applied to our Features

In this sub section will demonstrate the results of PCA in our training set where the main goal is to demonstrate, in all the features present in our training set which are the best that characterize our training set. As we can see in the table we present the application of PCA in our training set to cover 95% and 99% of the total variance of all training.

Features Based On	Variance	Number Features	Variance	Number Features
	95% (%)	65	99% (%)	128
Image properties	6.1	4	3.1	4
Region	97	63	93.75	120
Boundary	100	65	100	128
Texture	46.2	30	39	50
Statistical	93	60	92.2	118

Table 6.1: Results from appliance of the PCA in our training set.

Shown in the table 6.1 is the proportion of the features type after the appliance of the PCA. In numerical terms the percentage of each group, is based on the ten best features of each feature of the PCA, corresponding to each type of features. This happens because as explained above, a feature of the PCA is one or more features of our initial set. Thus we can see that the features based on the region boundary and the region itself, had a significant preponderance when compared with other types of features. In turn we can study from the table that the features based on the image properties do not add anything new to our features collection. Note also that the statistical features based on the boundary and the region, corresponds to a significant portion of the resulting set by the PCA algorithm.

6.3 Conclusion

In this chapter is introduced the finals steps of our implemented method, for the detection of dental caries. In this chapter we first start with the tooth segmentation, based on the output of the previous step, our main goal is to segment the tooth in the region that only contains one tooth. This is carry out by the appliance of the active

contours without edges technique. The results as happen in the method 2 of the Teeth detection Gap Valley were quite satisfactory, although in some images the pre processing stills need some improvements as referred early the recognition of the different types of teeth will absolutely make possible the appliance of directed pre processing, based on the type of tooth in the input image. In the second part of this section we present our features that were extracted based on the tooth segmented region. We implement several types of features divided in five major groups, features based on image properties, features based on the tooth segmented region, based on the statistical analysis, the image border or boundary also is one type of extracted features and finally the features based on the region texture. With this numerous extracted features, becomes imperative the use of a extraction of the best features, i.e, the use of algorithms that gives us in certain way the guarantee that the features we will use in the future are the features that better characterize the training set. This is perform by the use of PCA, with this transformation our training set will only contains the features that best covers the variance of the features in the training set. In the subsection 6.2.3.1 and in the table 6.1 is demonstrated the best features we extracted after the appliance of the PCA. Corresponding to the covered variance and the number of features extracted for the corresponding variance for each type of features.

Chapter 7

Results

In this section we discuss the results in all the stages of our method. First of all make a few observations regarding the mode of extraction of these results as well as all the sets we defined. For the ROI definition, jaws partition, teeth gap valley's detection and tooth segmentation stages the process for obtaining the results were extracted by visual inspection, with a 2-folder cross validation. It should be stressed that all the results obtained were based on the perfect input for each stage, i.e., that in our results we will not present the results for the entire process, this will be performed in future work, after our error rates in each stage are quitter satisfactory for testing the entire method as one. This perfect input does not mean that no noise were presented in our input images, this perfect input means that for each stage we do not take in consideration previous errors that could be driven by the previous stages. Therefore the results obtained really illustrates the precision in each stage.

In table 7.1 is shown the results obtained by us in the first three stages of our method, this results were also shared with the scientific community in our paper 8.4, because in this table are the results for the stages that we perform and demonstrated in our accepted paper in VIPImage'09. As explained above the results were obtained by visual inspection and the evaluation of our method is based on a completely independent data sets. It should be stressed that all the results were obtained through 2-folder cross validation.

As shown in the table 7.1 the results in the ROI definition and in the jaws partition are quite satisfactory, which correspond to an accuracy in a set of 700 images above 90%. For these two steps and concerning the visual inspection we

Stages	Results (% of correct)	Number of Images
ROI Definition	95.7	700
Jaws Partition	92.6	700
Teeth Gap Valley Method 1	74.6	700
Teeth Gap Valley Method 2	87.5	700
Tooth Segmentation	71.91	1098

Table 7.1: Results of our method stages up to the Dental Caries Detection.

consider correct, the presence of all teeth in the output image resulting of the ROI definition. Regarding the jaws partition we considered correct any polynomial returned by this step that divides each jaw properly, without cutting any tooth, regardless of the number of changes that may contain.

7.1 ROI Definition

The results for this step gives us the idea that empirical approach achieved good results in the ROI definition, however further work is important as the implementation of more comprehensive set rules to improve the outcome of this step. Because although the percentage of success is great, it should be stressed that in some output images the performed cut of the image brought some extra noise, that is, despite being correctly defined to the whole area that contains teeth, the noise produced by the chin and nasal bones still has a significant presence in the resulting images, as shown in the figure 7.1. In these cases it will be essential to improve the cut thus reduce noise in the image. For example applying the horizontal and vertical projection to try to detect the maximum top line of the teeth, thus approaching the original defined area to the cutting edge area belonging to the teeth.

In the case of images where the cut is outside of the limit teeth area, as shown in the figure 7.2, a possible solution would be to see if the cut area is in a possible area containing teeth, and thus require that the cutting is done in an area more darker, compared to the initial cutting area defined.

In conclusion for this stage improvements are required for the minimization of the noise present in the output images and to the minimization of the teeth cut. Although this last improvement had a small impact because there are a small



Figure 7.1: Results containing extra noise from the ROI definition stage.



Figure 7.2: Example of cutting teeth from the ROI definition stage.

portion of images that are cut. Despite this improvements the results obtained by us are optimistic for encouraging the improvements.

7.2 Jaws Partition

In this step the algorithm that we used for the connection of all points extracted comes from a careful choice because our extraction of items is subject to misinformation from the presence of other dark areas without necessarily being the areas between the jaws. For example the areas caused by the failure of teeth, or even the dark areas between the teeth. Those areas that influence the extracted points thus creating the presence of error in this step. Note that to overcome this sensitivity in the choice of the points we decided to use the polynomial interpolation, more precise the polynomial least squares fitting algorithm. Through this process our final result will be only slightly affected by an outlier point, despite what happens with successive outlier points, for example in the long absence of teeth, the result already implies the cutting of teeth. At this point the most important would be to consider these unwanted effects and decided to implement this way to the connection between the points extracted from the jaws gap valley, where the results obtained were excellent concerning the jaws partition. As shown in the figure 7.3 we have examples of images that were not well partitioned, although sometimes the affected corresponds only to one tooth. In other cases we can see that when there is a demarcated area to separate each jaw our method follows the darker areas which sometimes does not corresponds to a good division. In conclusion to this step the main improvements are related to the optimization of the extraction points applying new heuristics that can inform us with more confidence that the extracted point actually corresponds to points between the jaws.

In short at this stage of our method results lead to a correct interpretation of the problem, which led to a correct choice regarding to the process of the points connection, so to emphasize as future work will address the optimization of the points extraction for example increasing the number of stripes that divide the image while reducing the maximum possible value of the vertically step allowed for the points, this optimization could result in a compensation effect due to the absence of the teeth where in this case the effect would be much less noticed, thus the method may be able to positively react to the presence of the next tooth after the absence of

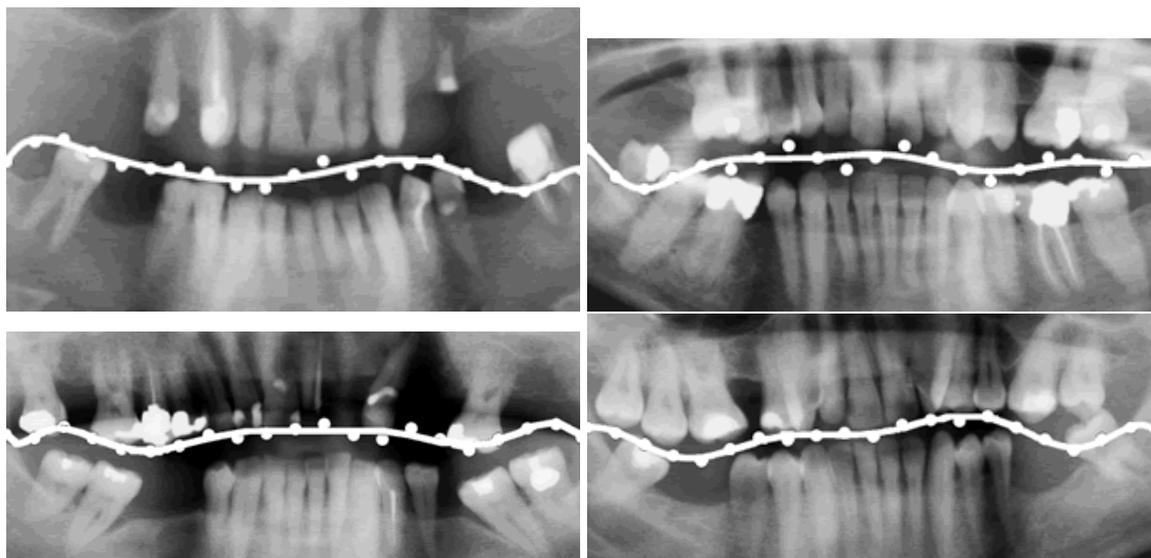


Figure 7.3: Results where some teeth are overlapped by the polynomial fitting in the Jaws Partition stage.

the teeth.

7.3 Teeth Gap Valley Detection

In this section the detection of the teeth gap valleys, as already noted, we implement two different methods, where the second it is already an optimization of the first method. Regarding the first method we have the results in the table 7.1, which evidence that improvements are needed for this important stage of our method.

7.3.1 Line Sum Intensities - Method 1

For this first method the main problem was formed by extracting the minimum points calculated by the minimum sum of intensities perform by the lines for each polynomial point. Because due to the fact that for each point was always chosen the lines with the lower intensity which exactly made our decision very permissive. In some cases corresponds to the removal of some good minimum after the curve smoothing. Thus it became essential to optimize the extraction of the minimum points corresponding to the gap valleys between the teeth, for this achievement contributed the use of active contours to the second method. In this step is to

emphasize that it was based as mentioned earlier in the footsteps of other authors for the detection of the teeth gap valleys. Since the authors in most cases for that only had into account the line with an angle of 90° to each point that was in the polynomial. This was also tested by us but could not reach good results, due to the variation of the polynomial. In most of the cases the lines with that angle never corresponds to the teeth gap valley which led us to opt for the solution of vary the angle of each line with each point of the polynomial. In this case the results were much better where the real difficult was based on the detection of the minimum which would correspond in the best case to the teeth gap valleys.

For this method one of the improvements we made is our second method, however we believe that with the use of the pre processing performed in the second method this method would achieve better results because with the darken of the zone of interest that we want to detect there's no doubt that this method will achieved better results. The same with the previous step of jaws partition, more darker the area between the jaws more performance will achieve our method.

7.3.2 Center Distance with the Active Contours - Method 2

Our second method turns out to be a method with some complicated steps. Both methods are based on the output of the previous step, where through the least squares polynomial fitting we obtain the separation of the jaws. Through this separation and throughout the entire process since the applicability of the pre processing to enhance the contrast between the teeth and to contrast the background with the teeth, until the transition to the polar coordinates for the implementation of active contours. After the detection of the minimum generated by the technique of active contours without edges, we achieve the teeth separation with a satisfactory rate of success. However this method sometimes face the same dilemma that we had in the previous method, which is related to the choice of the best minimal. For this we used the convolution technique for best accentuate the local minimum found in the distance vector. Another problem we had is the loss of information that occurs in the transition from cartesian coordinates to polar coordinates, we initially tried to connect with the corresponding line, the center of the image with the minimum point extracted and then pass this results again for the cartesian coordinates, but the results we obtained prove to be a bad because the straight cut overlaps almost

all teeth. In the actual process this flaw is less noted because the extraction of the point is always based on the polar coordinates point. Although for this method this decision turns out to be the best. Not excluding some improvements in the future work. One of the other problems associated with this method relates to the differences in each jaw when represented in polar coordinates, that happens because the lower jaw center of the image was not corresponding to the upper jaw, so we had to perform a symmetry in the image of the lower jaw giving the look of a "stretched" image.

One of the improvements for this step and more precisely to the lower jaw is to make the transition to polar coordinates where the center corresponds to the middle point in the last line under the picture. As in the upper jaw where the center corresponds to the first line above the image. Regarding the presence of unwanted minimum, only happens on the molars, because in this type of tooth exists a slight decrease in the middle of the tooth, and sometimes causing a minimum in that undesirable area. To solve this problem one of the implemented solutions is based on the threshold imposition and thus require the presence of another minimum in a more significant distance, when comparing to the middle of the teeth. This is based on the statistical data from our data-set of images more precisely related to the size of the molars and pre molars.

In conclusion to this second method the use of the active contour without edges technique proved to be a good choice, when comparing to the result of our previously method.

7.4 Tooth Segmentation

In the last step before the final step that corresponds to the dental caries detection, the primary aim with this step were to restrict the maximum area corresponding to the tooth in the output image of the previous step. For this the use of active contours without edges has proved to be an ideal choice for problems of segmentation where the contrast between the background and area of interest is elevated. However this contrast was not always obvious, which sometimes originate the presence of more noise in the resulting tooth, more than the expected. For this we use the techniques of pre processing of the previous step, the top and bottom hat transform. Whose

process increased significantly the contrast for the most of the images. The main problem associated with this step is when due to the nature of the input images, the quality is not the best and transforming the edge detection problem in a very complicated one. In this cases some examples results are shown in the figure 7.5, the presence of extra noise in the output image and in the figure 7.4 the presence of the overlapped teeth. But as mentioned earlier this occurs in some images of our test set and hence the reason why we obtain the results approximate to 72% of correctness. Note that for this step the results are obtained by take in to account the number of misclassify pixels over the number of the real pixels belonging to the tooth region. As demonstrated in the equation 7.1.

$$ER_{Total} = \frac{\sum_{k=1}^{T_{images}} ER_{I_k}}{T_{images}}, \quad (7.1)$$

Where $T_{images} = 1098$, corresponding to the total test set images. And the ER_{I_k} corresponds to the error for each image, given by 7.2.

$$ER_{I_k} = \frac{N}{T} \quad (7.2)$$

Where N denotes the total number of misclassified pixels in the image I_k and T corresponds to the total number of pixels that corresponds to the tooth region.

In conclusion this step corresponding to the tooth segmentation has the main goal to achieve the optimization of our next and final step. If the features that we extract from our output images of previous step contains only the limited area restrict to the tooth we have the guarantee that these features are relate only to the region of interest. Where in this case is the area corresponding to the tooth. If exists the presence of extra noise this could in fact influence the results obtained in our next step because this noise can prevent a good separation of the features values corresponding to the teeth containing cavities and teeth that do not contain any cavities.

7.5 Dental Caries Detection

In the last stage of our work, and as already stated and demonstrated above, the main goal is to extract features to our output images of the previous stage so

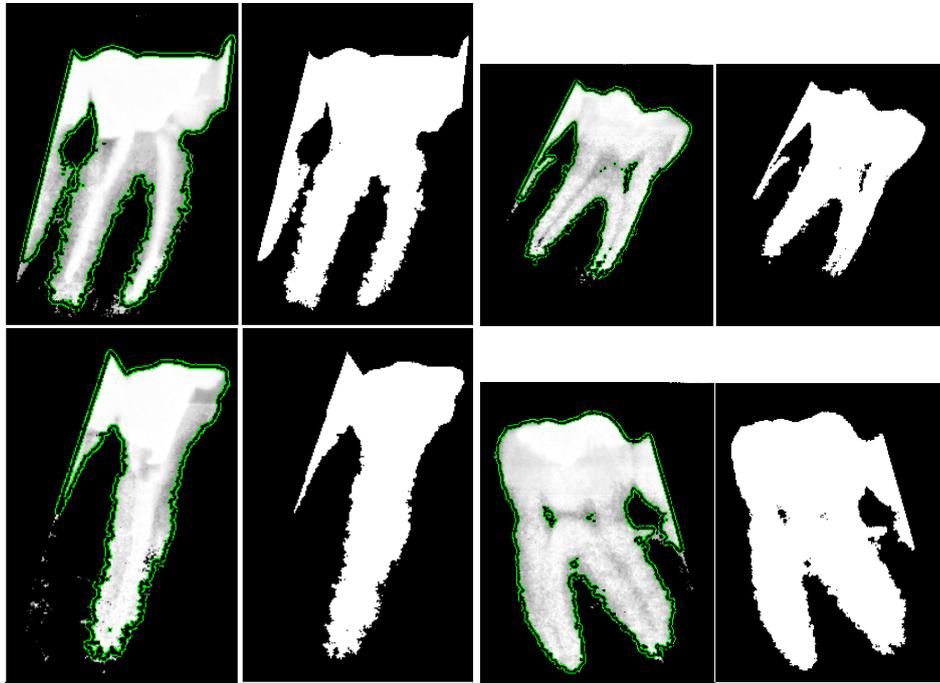


Figure 7.4: Examples of output images with the presence of teeth overlap segmentation.

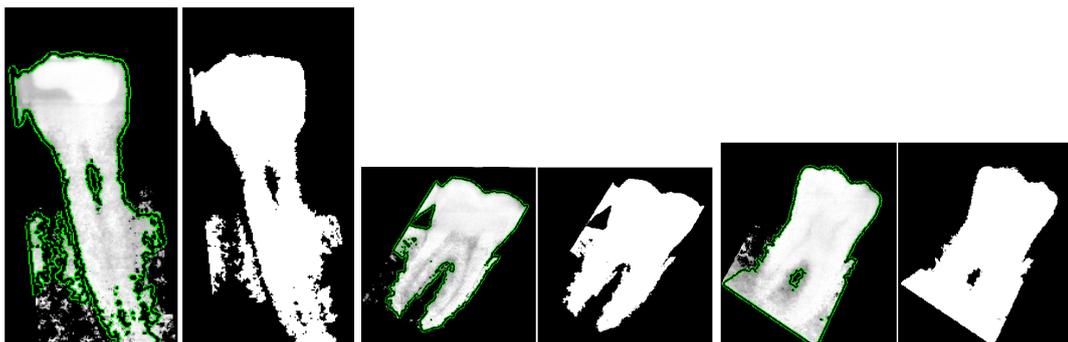


Figure 7.5: Examples of output images with the presence of extra noise.

we can pass them to a classifier for the dental caries detection in the teeth. This process dental caries classification is dependent on the classifier that in our case perform the classification based on the follows classifiers, the ANN, the K Nearest Neighborhood (KNN), the Naive Bayes and finally the SVM. To test our last method stage we created two sets of training and testing, in which one is the use of the tooth segmentation of previous stage described above. And the other corresponds to the real segmentation that was delimited by us, or where the segmentation of the tooth is perfect. With this we can compare the extent of our tooth segmentation stage in affect the results of the caries detection when the input segmented image is perfect.

For this chapter our main accuracy measures are the classification accuracy and the area under the ROC curve, which corresponds to the graphical representation of the true positives with the false positives. The classification accuracy is calculated by the confusion matrix. This give us the results of the:

- True Negative Rate (TNR): Or the rate of the true negative, i.e., the correct of no dental caries detected.
- True Positive Rate (TPR): Or the rate of the true positive, i.e., the correct of dental caries detected.
- False Positive Rate (FPR): Or the rate of the false positive, i.e., the tooth that are badly classified as dental caries.
- False Negative Rate (FNR): Or the rate of the false negative, i.e., the tooth that are badly classified as no dental caries.

Concerning the confusion matrix, in the horizontal lines is the real value expected for all the input images. The vertical lines corresponds to the value obtained for each image. Where this expected values are 0 for the presence of no dental caries in the tooth and 1 for the presence of dental caries in the input image containing a single tooth.

7.5.1 Data-Set Normalization

The features normalization is a necessary step because in many situations the features values lie within different dynamic ranges. Thus, features with large values may have larger influence in the cost function than features with small

values. Although this does not necessarily reflect their respective significance in the design of the classifier. The problem is overcome by the features normalization, so their values lie within similar ranges. A straightforward technique is normalization via the respective estimates of the mean and the variance. For N available data of the k th feature we have:

$$\bar{x}_k = \frac{1}{N} \sum_{i=1}^N x_{ik}, \quad k = 1, 2, \dots, l \quad (7.3)$$

$$\sigma_k^2 = \frac{1}{N-1} \sum_{i=1}^N (x_{ik} - \bar{x}_k)^2 \quad (7.4)$$

$$\hat{x}_{ik} = \frac{\bar{x}_k - x_{ik}}{\sigma_k} \quad (7.5)$$

Where the \bar{x} and σ are the sample mean and the sample standard deviation of the k th feature [4]. In other words this corresponds that all the resulting features will now have zero mean and the unit variance. Which corresponds to a linear method. Other linear techniques limit the features value in the range of [0, 1] by proper scaling. There are other normalization techniques that corresponds to nonlinear methods that can also be employed in cases in which the data are not evenly distributed around the mean. In such cases transformations based on nonlinear (i.e., logarithmic or sigmoid) functions can be used to map the data within specified intervals. For our problem we only use the two above linear methods of normalization. As future work it will be interesting to study the impact of all different types of normalization on the dental caries detection problem.

7.5.2 Results with our Tooth Segmentation

Here are the best results obtained for the test set which the tooth segmentation is performed by our method. As we noted the results obtained for this test set, demonstrated that the tooth segmentation stage influences in a large scale our results in the dental caries detection. When compared to the perfect segmentation of the tooth. Where the results are much better. This conclusion are always based on the tested classifiers.

7.5.2.1 Artificial Neural Network

An ANN, usually called Neural Network (NN), is a mathematical model or computational model that tries to simulate the structure and/or functional aspects of biological NN. It consists of an interconnected group of artificial neurons and processes information using a weight connection between each other. In most cases an ANN is an adaptive system that changes its structure based on external or internal information that flows through the network during the learning phase. Neural networks are non-linear statistical data modeling tools. They can be used to model complex relationships between inputs and outputs or to find patterns in data [11][9].

Class	No Dental Caries	Dental Caries	Total
No Dental Caries	317	232	549
Dental Caries	148	401	549
Total	465	633	718
Accuracy	65,39%		

Table 7.2: Confusion matrix for the ANN classifier after the PCA appliance for 95% variance covering.

7.5.2.2 Support Vector Machine (SVM)

A SVM performs classification by constructing an N -dimensional hyper plane that optimally separates the data into two categories. SVM models are closely related to ANN. In fact, a SVM model using a sigmoid kernel function is equivalent to a two-layer, perceptron ANN. SVM models are a close cousin to classical Multi Layer Perceptron (MLP) NN. Using a kernel function, SVM are an alternative training method for polynomial, radial basis function and MLP classifiers in which the weights of the network are found by solving a quadratic programming problem with linear constraints, rather than by solving a non-convex, unconstrained minimization problem as in standard NN training [14][41]. In table 7.3 are the results obtained by this classifier to our test set, with our tooth segmentation. This classifier corresponds to our second better classifier, with a accuracy of 62.02%.

Class	No Dental Caries	Dental Caries	Total
No Dental Caries	263	286	549
Dental Caries	131	418	549
Total	394	704	681
Accuracy	62.02%		
ROC Area	62%		

Table 7.3: Confusion matrix for the SVM classifier after the PCA appliance for 95% variance covering.

7.5.2.3 Naive Bayes

A naive Bayes classifier is a simple probabilistic classifier based on the appliance of the Bayes theorem (from Bayesian statistics) [16] [68], with strong (naive) independence assumptions. A more descriptive term for the underlying probability model would be "independent feature model". Depending on the precise nature of the probability model, naive Bayes classifiers can be trained very efficiently in a supervised learning setting. In many practical applications, parameter estimation for naive Bayes models uses the method of maximum likelihood, i.e., that turns this type of classifiers very intuitive for the user. In table 7.4 are the results obtained by this classifier for our input test set with the appliance of the PCA with the variance of 95% covered, that corresponds to a total of 104 features.

Class	No Dental Caries	Dental Caries	Total
No Dental Caries	255	294	549
Dental Caries	180	369	549
Total	435	663	624
Accuracy	56,83%		
ROC Area	59,2%		

Table 7.4: Confusion matrix for the Naive Bayes classifier after the PCA appliance for 95% variance covering.

For this test set the classifier that obtains better accuracy is the ANN, with a 65.39%. Next will demonstrate the results achieved for the perfect tooth segmentation. Later in this section will discussed and compared the results obtained for the

different types of test sets.

7.5.3 Results with the perfect Tooth Segmentation

At this stage we will demonstrate the results obtained by us for the perfect tooth segmentation. This tooth segmentation only contains the region corresponding to the tooth. This segmentation when compared to our tooth segmentation method, we have an 72% of pixels correctly segmented for our test set images. As we will discuss later on this section the tooth segmentation corresponds to an important pre processing stage for the improvement of the dental caries detection.

7.5.3.1 Artificial Neural Network

In this section is shown the results obtained for the test set for the ANN, composed by 1098 images, containing only one tooth per image. The results for the test training normalized between $[0, 1]$ are demonstrated in the table 7.5 and in the figure 7.6, containing the ROC for this test.

Class	No Dental Caries	Dental Caries	Total
No Dental Caries	351	198	549
Dental Caries	21	528	549
Total	372	726	879
Accuracy	80,10%		

Table 7.5: Confusion matrix for the ANN classifier for the normalized test set between $[0, 1]$ test and training set.

In the table 7.6 is the best result for the test set, after the appliance of the PCA with the covered variance of 99%, in the figure 7.7 is the correspondent ROC.

As we can observe the results obtained are very good, confirming that with a good tooth segmentation the results for the previous test set could increase very easily. There is an increase of the accuracy rating after the appliance of the PCA for a 99% variance coverage. Which denotes the importance of the attributes selection performed by the PCA.

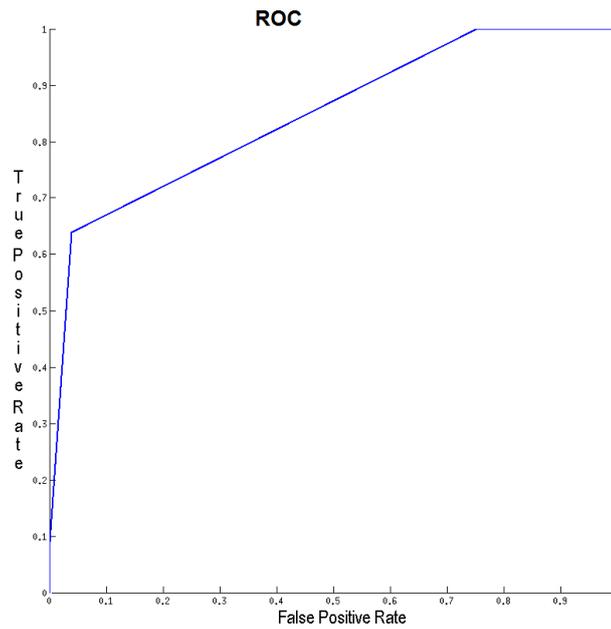


Figure 7.6: ROC for the ANN classifier for the normalized test set between [0, 1] test and training set.

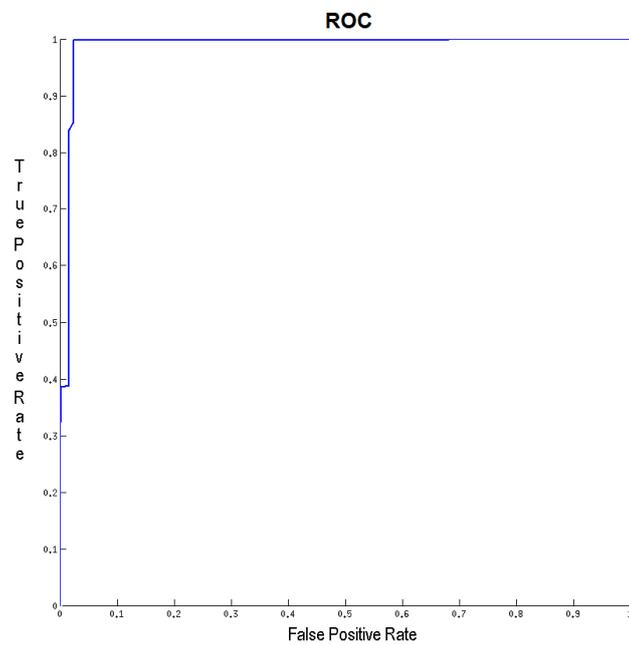


Figure 7.7: ROC for the ANN classifier for the normalized test set between [0, 1] after the appliance of the PCA for a 99% variance coverage.

Class	No Dental Caries	Dental Caries	Total
No Dental Caries	548	1	549
Dental Caries	13	536	549
Total	561	537	1084
Accuracy	98,70%		

Table 7.6: Confusion matrix for the ANN classifier for the normalized test set between [0, 1] after the appliance of the PCA for a 99% variance coverage.

7.5.3.2 Support Vector Machine (SVM)

For the SVM is shown in the table 7.7 the best result obtained by us for the SVM classifier, this results is based on the set teste normalize between [0, 1] and a covered variance of 99% extracted by the PCA.

Class	No Dental Caries	Dental Caries	Total
No Dental Caries	532	17	549
Dental Caries	10	539	549
Total	542	556	1071
Accuracy	97.54%		
ROC Area	97.66%		

Table 7.7: Confusion matrix for the SVM classifier for the test set normalized between [0, 1] after the appliance of the PCA for a 99% variance coverage.

7.5.3.3 Naive Bayes

The Naive Bayes classifier obtained the worst result among the the three classifiers, although the accuracy rating stills very high. Demonstrating the good results that we are obtaining by the use of the perfect segmentation. In table 7.8 are the results obtained by this classifier for our input test set with the appliance of the PCA with the variance of 99% covered.

Class	No Dental Caries	Dental Caries	Total
No Dental Caries	528	21	549
Dental Caries	26	523	549
Total	554	544	1051
Accuracy	95,72%		
ROC Area	95,50%		

Table 7.8: Confusion matrix for the Naive Bayes classifier for the test set normalized between [0, 1] after the appliance of the PCA for a 99% variance coverage.

7.5.4 Discussion

As we can observe exists a great discrepancy in the results with the perfect segmentation and the segmentation extracted from our method. And due to the error that is inherent to our tooth segmentation stage, an error of approximately 28%, the dental caries detection process with our initial segmentation has a very low income, when compared with the accuracy that we obtain with the perfect segmentation.

The results enforces the idea of the importance of tooth segmentation as future work, because in terms of features we can achieved a great success of positively correctly rate of dental caries or cavities. Proofing that the process of segmentation is essential for the successfully detection of dental caries.

Another of the conclusions drawn from the results obtained is that the features we extracted, when in the presence of the perfect tooth segmentation, can infer with a great degree of certainty if the tooth in the image is or not a tooth with dental caries. Demonstrating and reinforcing why we opted for the feature extraction for each tooth, i.e., looking for the tooth image as a whole. Rather the feature extraction based on the pixel approach.

The best results were achieved by the ANN for the two types of training and test sets. This demonstrates the high robustness and effectiveness that these type of classifiers can accomplish. Although the difference when comparing to the others classifiers is not to different. This results are mainly because the good features that we extracted, providing to each classifier good accuracy results.

Should also be noted that the PCA had a leading role in the improvement of the accuracy rating of all classifiers, it is not explicit in this work for the SVM and for

the Naive Bayes, because we only demonstrate the better results obtained. But it should be stressed that for all the results of the test set with the appliance of the PCA were substantially better. As proven in the ANN for the perfect segmentation test set.

For future work it will be important besides the improvement of the tooth segmentation, the test with a more larger group of test images. Where through that it will be possible to affirm the great results obtained by us in this stage of the method. Proving in fact the reliability of our method to caries detection.

7.6 Conclusion

In this section we present all the results obtained by the different type of tests performed in each stage of our method. As we can observe above the results are quite satisfactory where in some methods the results are more convincing of our method reliability, and in others the sign the improvements must be considered.

For all the stages the results were obtained considering the best inputs, not the inputs without noise but the input that actually corresponds to the main goal of each method stage. For example the perfect input for the dental caries detection is the corresponding to a region segmented with a single tooth. In terms of testing all the results tests presented were based always in two different sets, one corresponding to the training and the other corresponding to the test. In most of the stages the results accuracy were extracted by visual inspection. Which corresponds to a non-automatic process of extracting the results.

Chapter 8

Conclusion

8.1 Contribute

The direct contribute of this dissertation for the scientific community is the development of a complete case study for the dental caries detection in dental panoramic X-ray. The outline of the developed method is in the figure 1.1. The method consists in five main stages.

ROI Definition 4.1 - That consists in the statistical cut of our input images for the elimination of extra noise present in the image corresponding to the chin and nasal bones. In this step is defined four length's R_i where each one corresponds to the distance of the center of the image to the left, up, down and right sides of the mouth. Then the cut is performed by the approximation of the distances of each length to a normal distribution that enables the cut of all our images with a certain level of confidence. The training set for the evaluation of this length's corresponds to half of the total images present in our data-set.

Jaws Partition 4.2 - In this stage our main goal is to divide the mouth upper a lower jaw these process is performed by the extraction of a set of primitive points between the jaws by the horizontal projection of the image intensities. After having this set of points the next step is to perform the connection between them this is obtained by the implementation of the polynomial fitting process that connects all points between the jaws.

Teeth Gap Valley Detection 5.2 - After having both jaws divided our next goal is to detect the teeth gap valley. This gap valleys tends to be more darker than the teeth itself but not always present due to the teeth overlap. This is the second version of a possible solution to the problem. This optimization of our initial method consists in having both jaws divided represent each jaw in the polar coordinates and after applying some pre processing on the image the top and bottom hat transform. Through the representation of each jaw in polar coordinates turns the problem in a classical active contours problem. By doing so and based on the great variety of these methods in the literature we decided to choose the Active Contours Without Edges, to be more accurate and not so noise sensitive. After having the contour of the teeth the extraction of distances of each border point to the center of the image is calculated. With this distance vector and through the curve smooth using the convolution with a gaussian filter and then we extract the minimums points that corresponds in the majority of the tested cases to the teeth gap valleys present in each jaw.

Tooth Segmentation 6.1 - Based on the previous output that is an image contain only one tooth we perform the tooth segmentation. This step is perform by the use of the Active Contour Without Edges. After applying a pre processing process in the input images. As the same before the pre processing choose by us is the top and bottom hat transform.

Dental Caries Detection 6.2 - The process of the dental caries detection consists in the extraction of features in our input images this features were based on the image properties. In the region based features that describes the output region of the tooth segmentation. Statistic features that corresponds to the image properties in the statistical analysis these features were used not only directly in the image but also used in the feature extraction of other techniques used to represent our output region. The features based on the object boundary, that in our case corresponds to the tooth region segmented output. This object boundaries descriptors were based on multiples techniques like the chain codes, fourier descriptors, fourier elliptic descriptors and the cumulative normalized angular function among others. Finally the features based on the region textures. After the extraction of these features we perform the selection of the best ones using the PCA where the goal is to

perform the reduction of the dimensionality while preserving the variance in the high-dimensional space as possible. The final step consists into apply the learning classifiers to obtain the accuracy of our dental caries detection method.

The other main contribute of our work is the publication of an accepted article (in the appendix 8.4) in an international conference VIPImage'09. The paper focussed the two first stages of our method and the first version of our third stage method. The next goal is to publish more articles based on the final steps of our method.

The final contribute of our work is to announced the availability of a new data set of panoramic dental X-ray images, which can constitute a tool for the research community in the development of the stomatologic-related applications. This data set has varying morphologic properties that make it more valuable to the scientific community. These varying properties include the number of teeth per image, the shape of the mouth and teeth as well as the levels of noise.

8.2 Results

As shown in the table 8.1 concerning the first two stages of our method the results are quite enthusiastic, obtaining in the ROI definition a accuracy rating of 95.7%, and in the jaws partition a result of 92.6%. The other three stages had the accuracy rating of 87.5% concerning the teeth gap valleys detection for the tooth segmentation our accuracy is 71.91%. Finally the accuracy of the dental caries detection is 98.88%, these results are relative to each stage as an independent task which leave as future work the process of linking these tasks, and hopefully, with some improvements in the key stages achieve better results for the dental caries detection. Where it demonstrates in fact why the results obtained confirms that the problem that we want to solved had some excellent indicators of success.

8.3 Motivation

The motivation for the realization of this work, is based on the importance of the computer vision and image processing in the scientific community. Other motivation is that in the literature there is no presence of a method the we could base our entire work, the only literature relevant data were in the pre processing

Stages	Classification Accuracy (%)	Number of test images
ROI Definition	95.7	700
Jaws Partition	92.6	700
Teeth Gap Valley Detection	87.5	700
Tooth Segmentation	71.91	1098
Dental Caries Detection	98.70	1098

Table 8.1: Resume table of our results accuracy for each method stage.

applied in dental X-ray images and finally in the teeth segmentation and jaws partition but always in bitewing dental X-ray images. In the literature does not exist until know the presence of methods performed in dental panoramic X-ray.

8.4 Future Work

The optimization of the ROI definition based on the actual ROI applying a method for the removal of the extra noise that presents in some images. For example applying the pre processing top and bottom hat operation for the enhanced of the contrast. Therefore the detection of the zone where starts and ends the chin bones and the nasal bone, could be more accurate.

In the jaws partition as future work, it is very important to detect missing teeth because in this situations our algorithm tend to not compensate the missing teeth present in the input image. This process could be perform in two situations, after the jaws partition or during the process of jaws partition. After the process of jaws partition it will be important to detect if the polynomial extracted as in the various stripes equally mean intensities. In a strip where this propriety is not fulfill we are in the case where the polynomial is passing over a brighter area of the image, and possibly a teeth. During the process of jaws partition one possible optimization is when the extracted point is in a brighter area we move this same point up or down for a darker zone, therefore extracting always points in darker areas of the image. In the step of jaws partition the appliance of pre process of top and bottom hat may also be determinant for the performance improvement.

The method based on the center distance with the active contours, our method 2 of the teeth gap valleys detection is already a optimization to our first method. But

also in this second method some optimizations could be perform for future work, for example:

- Divide the image in three different zones (left, frontal and right) for the correct appliance of the top and bottom hat transform.
- Apply in practice the various techniques of active contours to evaluate each performance individually.
- The extraction of the line in polar coordinates for the cartesian coordinates.
- The improvement concerning the teeth overlap, applying local erosion in brighter areas, that in many cases corresponds to teeth overlap zones.

Concerning the dental caries detection the future improvements to implement are the creation of a machine learning system. Where before the real feature extraction the images containing a single tooth will be evaluate to detect the dental caries that destroy partially the tooth crown. To detect darker zones in the teeth where in this cases corresponds with a great probability to dental caries. Due to the different properties that each teeth contains, it will be important to classify each tooth image concerning the type (molars, pre molars, incisors and canine), because the dental caries in each type of teeth manifests in very different ways. For example concerning the size of the darker zones, in smaller teeth the darker zones are smaller when comparing with larger teeth.

For the improvement of the dental caries detection is important to improve the tooth segmentation. With the machine learning system the tooth segmentation stage could loose a little of his key stage for the correct dental caries detection. But is vital for the good dental caries detection the improvement of the tooth segmentation, which in our tests corresponds to a higher increase of the rating accuracy.

Finally as future work our main goal is to start the detection of other types of dental diseases, using the described dental X-ray data set. For now we are starting by the detection of dental cavities in the future hopefully move to other types of problems.

References

- [1] Iapr - international association for pattern recognition. <http://www.iapr.org/>.
- [2] In epidemiology of dental disease, 1 2007. Hosted on the University of Illinois at Chicago website.
- [3] Spemd - sociedade portuguesa de estomatologia e medecina dentária, 2008. <http://www.spemd.pt/>.
- [4] Selim Aksoy and Robert M. Haralick. Feature normalization and likelihood-based similarity measures for image retrieval. *Pattern Recognition Letters*, 22:563–582, 2000.
- [5] Howard Anton. *Elementary Linear Algebra*. John Wiley & Sons Inc.
- [6] Howard; Irl Bivens; Stephen Davis Anton. *Calculus*. Anton Textbooks, Inc., 7 edition, 2002.
- [7] Ben Appleton. Optimal geodesic active contours: Application to heart segmentation, 2003.
- [8] Andrew Blake and Michael Isard. *Active Contours*. Springer.
- [9] Maureen Caudill and Charles Butler. *Understanding Neural Networks; Computer Explorations*. MIT Press, Cambridge, MA, USA, 1992.
- [10] Vese L.A. Chan T.F. "active contours without edges". *Image Processing, IEEE Transactions*", 10:266–277, 2 2001.
- [11] C. Charalambous. Conjugate gradient algorithm for efficient training of artificialneural networks. *Circuits, Devices and Systems, IEEE Proceedings*, 139(3):301–310, 6 1992.

- [12] H. Chen and A.K. Jain. Tooth contour extraction for matching dental radiographs. *Proc. 17th Int'l Conf. Pattern Recognition*, III:522–525, August 2004.
- [13] Christopher M. Brown Dana H. Ballard. "*Computer Vision*". "Prentice-Hall, Inc.", 1982.
- [14] Friedrich Leisch David Meyer and Kurt Hornik. "the support vector machine under test". "*Image Processing, IEEE Transactions*", 8 2003.
- [15] E. Davies. *Machine Vision: Theory, Algorithms and Practicalities*. Academic Press, 1990.
- [16] Pedro Domingos and Michael J. Pazzani. On the optimality of the simple bayesian classifier under zero-one loss. *Machine Learning*, 29(2-3):103–130, 1997.
- [17] S.A. Dyer and Xin He. Least-squares fitting of data by polynomials. *IEEE Instrumentation and Measurement Magazine*, 3(2):46–51, December 2001.
- [18] R. C. Gonzales and R. E. Woods. *Digital Image Processing*. Prentice-Hall, 2 edition, 2002.
- [19] Ricardo Gutierrez-Ossuma. "*Introduction to Pattern Analysis*". "Texas A&M University".
- [20] Michiel Hazewinkel. *Encyclopaedia of Mathematics*, volume 1. Kluwer Academic Publishers, 2000. <http://eom.springer.de/>.
- [21] Patrick C. Hew. *Geometric and zernike moments*, 1996.
- [22] T. HyunWook Park, Schoepflin and Yongmin Kim. "active contour model with gradient directional information: directional snake". "*Circuits and Systems for Video Technology, IEEE Transactions*", 11:252–256, 2 2001.
- [23] A. K. Jain and H. Chen. Matching of dental x-ray images for human identification. *Pattern Recognition*, 37:1519–1532, 2004.
- [24] A. K. Jain, L. Hong, and S. Pankanti. Biometrics: promising frontiers for emerging identification market. *in Comm. ACM*, pages 91–98, February 2000.
- [25] A.K. Jain, H. Chen, and S. Minut. Dental biometrics: Human identification using dental radiographs. *AVBPA, UK*, pages 429–437, 2003.

- [26] I.T. Jolliffe. *Principal Component Analysis*. Springer Series in Statistics, 2 edition, 2002.
- [27] Alireza Khotanzad and Yaw Hua Hong. Invariant image recognition by zernike moments. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(5):489–497, May 1990.
- [28] Satyanad Kichenassamy, Arun Kumar, Peter Olver, Allen Tannenbaum, and Anthony Yezzi Jr. Gradient flows and geometric active contour models, 1994.
- [29] Shuo Li, Thomas Fevens, Adam Krzyzak, Chao Jin, and Song Li. Semi-automatic computer aided lesion detection in dental x-rays using variational level set. *Pattern Recognition*, 40:2861–2873, 2007.
- [30] Shuo Li, Thomas Fevens, Adam Krzyzak, and Song Li. An automatic variational level set segmentation framework for computer aided dental x-rays analysis in clinical environments. *Computerized Medical Imaging and Graphics*, 30:65–74, 2006.
- [31] Krzyzak A Li S, Fevens T. Level set segmentation for computer aided dental x-ray analysis. *SPIE Conference on Medical Imaging*, 5747:580–9, 2005.
- [32] J.S. Lin, K.S. Cheng, and C.W. Mao. A fuzzy hopfield neural network for medical image segmentation. *IEEE Trans. Nucl. Sci.*, 43(4):2389–2398, August 1996.
- [33] R.W. Maragos, P. Schafer. Morphological systems for multidimensional signal processing. *Proceedings of the IEEE*, 78(4):670–710, April 1990.
- [34] ELS Maria G. Essig, MS. "dental x-rays". "Susan Van Houten, RN, BSN, MBA", 8 2008.
- [35] Alberto S. Aguado Mark S. Nixon. "Feature Extraction and Image Processing". Newnes, 2002.
- [36] Hamish D. Meikle. *A New Twist to Fourier Transforms*. Wiley-VCH, 1 edition, 2004.

- [37] Andrew Witkin Michael Kass and Demetri Terzopoulos. "snakes: Active contour models". *International Journal of Computer Vision, Springer*, 1(4):321–331, 11 2004.
- [38] Cleve Moler. *Numerical Computing with MATLAB.*. SIAM - Society for Industrial and Applied Mathematics.
- [39] Richards MP. "a brief review of the archaeological evidence for palaeolithic and neolithic subsistence". *Eur J Clin Nutr*, pages 16–1262, 2002.
- [40] D. Mumford and J. Shah. "optimal approximations by piecewise smooth functions and associated variational problems.". *Communications on Pure and Applied Mathematics*, 42(5):577–685, 1989.
- [41] Jonh Shawe-Taylor Nello Cristianini. *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. Cambridge University Press, 1 edition, 2002.
- [42] J. Alison Noble. The effect of morphological filters on texture boundary localization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(5):554–561, May 1996.
- [43] O. Nomir and M. Abdel-Mottaleb. A system for human identification from x-ray dental radiographs. *Pattern Recognition*, 38:1295–1305, 2005.
- [44] O. Nomir and M. Abdel-Mottaleb. Human identification from dental x-ray images based on the shape and appearance of the teeth. *IEEE Transactions on Information Forensics and Security*, 2(2), 2007.
- [45] O. Nomir and M. Abdel-Mottaleb. Fusion of matching algorithms for human identification using dental x-ray radiographs. *IEEE Transactions on Information Forensics and Security*, 3(2), June 2008.
- [46] Stanley Osher and James A. Sethian. Fronts propagating with curvature dependent speed: Algorithms based on hamilton-jacobi formulations. *Journal of Computational Physics*, 79:12–49, 1988.
- [47] Fedkiw R. Osher S. Level set methods: An overview and some recent results. *J Comput Phys*, 169:463–502, 2001.

- [48] Fedkiw R. Osher S. Level set methods in image science. *International conference on image processing*, 169:631–4, 2003.
- [49] Nikos Paragios, Olivier Mellina-Gottardo, and Visvanathan Ramesh. Gradient vector flow fast geodesic active contours. In *In IEEE International Conference in Computer Vision*, pages 67–73, 2001.
- [50] Nikos Paragios, Olivier Mellina-Gottardo, and Visvanathan Ramesh. Gradient vector flow fast geometric active contours. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(3):402–407, 2004.
- [51] Christian H. Reinsch. Smoothing by spline functions. *Numerische Mathematik, Springer Berlin / Heidelberg*, 2005, 10:177–183, October 1967.
- [52] Peggy Saari. "medicine and disease - who discovered the x ray?.". *"History Fact Finder. Ed. Julie L. Carnagie. UXL-GALE"*, 2001.
- [53] Eyad Haj Said, Diao Eldin M. Nassar, Gamal Fahmy, and Hany H. Ammar. Teeth segmentation in digitized dental x-ray films using mathematical morphology. *IEEE Transactions on Information Forensics and Security*, June 2006.
- [54] Christophe Samson, Laure Blanc-Féraud, Gilles Aubert, Josiane Zerubia, and Josiane. A level set model for image classification, 1999.
- [55] Mubarak Shah. *"Fundamentals of Computer Vision"*, 12 1997. "Computer Science Department, UCF, FL, USA".
- [56] Samir Shah, Ayman Abaza, Arun Ross, and Hany Ammar. Automatic tooth segmentation using active contour without edges. *IEEE Biometrics Symposium*, 1(2), June 2006.
- [57] Linda G. Shapiro and George C. Stockman. *"Computer Vision"*. "Prentice-Hall", 2001.
- [58] S. Shiffman, G. D. Rubin, and S. Naple. Semi-automatic computer aided lesion detection in dental x-rays using variational level set. *IEEE Trans. Med. Imag.*, 19(11):1064–1074, November 2000.
- [59] Ajoy K. Ray Tinku Acharya. *"Image processing: principles and applications"*. "John Wiley and Sons", 2005.

- [60] R. Kimmel V. Caselles and G. Sapiro. "geodesic active contours". *International Journal of Computer Vision*, 22(1):61–79, 1997.
- [61] Luminita A. Vese, Tony F. Chan, Tony, and F. Chan. A multiphase level set framework for image segmentation using the mumford and shah model. *International Journal of Computer Vision*, 50:271–293, 2002.
- [62] Eric W. Weisstein. "least squares fitting.". <http://mathworld.wolfram.com/LeastSquaresFitting.html>.
- [63] Eric W. Weisstein. "least squares fitting–polynomial.". <http://mathworld.wolfram.com/LeastSquaresFittingPolynomial.html>.
- [64] Eric W. Weisstein. "polar coordinates.". <http://mathworld.wolfram.com/PolarCoordinates.html>.
- [65] Eric W. Weisstein. "vandermonde matrix.". <http://mathworld.wolfram.com/VandermondeMatrix.html>.
- [66] Dina Eldin Nassar Xin Li, Ayman Abaza and Hany Ammar. Fast and accurate segmentation of dental x-ray records. *Lecture Notes in Computer Science, Springer*, 3832:688–696, December 2005.
- [67] Chenyang Xu, Anthony Yezzi, and Jr. On the relationship between parametric and geometric active contours. pages 483–489, 2000.
- [68] Harry Zhang. The optimality of naive bayes. In Valerie Barr and Zdravko Markov, editors, *FLAIRS Conference*. AAAI Press, 2004.
- [69] J. Zhou and M. Abdel-Mottaleb. A content-based system for human identification based on bitewing dental x-ray images. *Pattern Recognition*, 38:2132–2142, 2005.

A Data Set of Panoramic Dental Radiographs for Stomatologic Image Processing Purposes

Dental X-Ray: A Data Set of Panoramic Dental Radiographs for Stomatologic Image Processing Purposes

João Oliveira & Hugo Proença

Department of Computer Science, IT - Instituto de Telecomunicações
University of Beira Interior, Covilhã, Portugal

This paper has two major purposes: at firstly, to announce the availability of a new data set of panoramic dental X-ray images. This data set contains 1392 images with varying types of noise, usually inherent to this kind of images. Furthermore, the number of teeth per image and their dental morphology were not constant. Secondly, we propose a method to approximate the panoramic images in bitewing images, which are the most common type of images used in the human identification and in the tooth segmentation for the diagnosis of dental diseases.

Keywords: Dental X-ray; Data Set; Medical Image; Image Processing; Computer Vision;

1 INTRODUCTION

Automated image analysis processes have been achieving higher relevance for many purposes and the results can be considered satisfactory (*e.g.*, biometrics and multiple types of medical image diagnosis). In the specific area of medical image processing these automated systems constitute a valuable tool as the earliest detection of many diseases and in the preprocessing of huge amounts of data.

This paper announces the availability for research purposes of a new data set of panoramic dental radiograph images. Nowadays, there are a variety of areas where this type of images can be helpful: biometrics identification (Jain, Chen, and Minut 2003)(Jain, Hong, and Pankanti 2000)(Chen and Jain 2004) (Jain and Chen 2004) (Nomir and Abdel-Mottaleb 2005) (Zhou and Abdel-Mottaleb 2005) (Nomir and Abdel-Mottaleb 2007) (Nomir and Abdel-Mottaleb 2008), detection of dental cavities, bone loss and periodontitis (Li, Fevens, Krzyzak, and Li 2006) (Li, Fevens, Krzyzak, Jin, and Li 2007) (Shiffman, Rubin, and Naple 2000). However, the types of images that are usually chosen for these purposes are the bitewing and the periapical, used to perform the tooth segmentation (Shah, Abaza, Ross, and Ammar 2006) (Said, Nassar, Fahmy, and Ammar 2006) (Lin, Cheng, and Mao 1996) and to identify deceased individuals (Chen and Jain 2004) (Jain and Chen 2004) (Nomir and Abdel-Mottaleb 2005) (Zhou and Abdel-Mottaleb 2005) (Shah, Abaza, Ross, and Ammar 2006) (Said, Nas-

sar, Fahmy, and Ammar 2006) (Nomir and Abdel-Mottaleb 2008). Here, having access to the AM (*ante-mortem*) and PM (*post-mortem*) dental X-ray, is in some cases the only way to identify such individuals. Finally, these types of images are also used in clinical environments, in the detection of dental diseases (Li, Fevens, Krzyzak, and Li 2006) (Li, Fevens, Krzyzak, Jin, and Li 2007).

According to the above discussion, in this paper we also propose a method that approximate the panoramic dental X-ray images into bitewing dental X-ray images used by many authors. This method comprises three stages, the first stage is based on a statistical evaluation of the images to define a preliminary region-of-interest (ROI), taking out the nasal and chin bones. In the second stage we make the detection of the upper and lower jaws, based on the extraction of primitive points between jaws and a polynomial fitting process. The third stage is the partition of each jaw into three parts. The left side contains the left molars and the left pre-molars, the front side includes the canines and the incisors and, finally, the right side that contains the right molars and the right pre-molars.

Another major point of interest of this data set is the set of maps that enable the manual detection and localization of dental cavities - for the moment - and other diseases in a near future. This is achieved by corresponding a binary image to each dental X-ray image thereby showing the regions with dental cavities. This will turn the data set into a preferable

tool in the evaluation of method to perform automatic detection and localization of the respective diseases. Also, we believe that this set of images is useful to evaluate the current teeth segmentation methods.

The remaining of this paper is organized as follows. In Section 2 we present the most relevant features of the data set. Section 3 describes the proposed method to transform the panoramic dental X-ray images into bitewing images. Section 4 reports our experiments and, finally, Section 5 gives the conclusions and future work.

2 Data Set

In this sub-section we describe the main points of the data set images which were all captured by an Orthoralix 9200 DDE X-ray camera. There are a total of 1392 grayscale images in the data set, with varying types of dental structure, sizes of the mouth and number of teeth per image, as can be seen in figure 1. The gray levels of each image were stretched to the $[0, 255]$ scale, although both normalized and raw images will be soon available for download in the site [SocialLab¹](http://socialab.di.ubi.pt/)

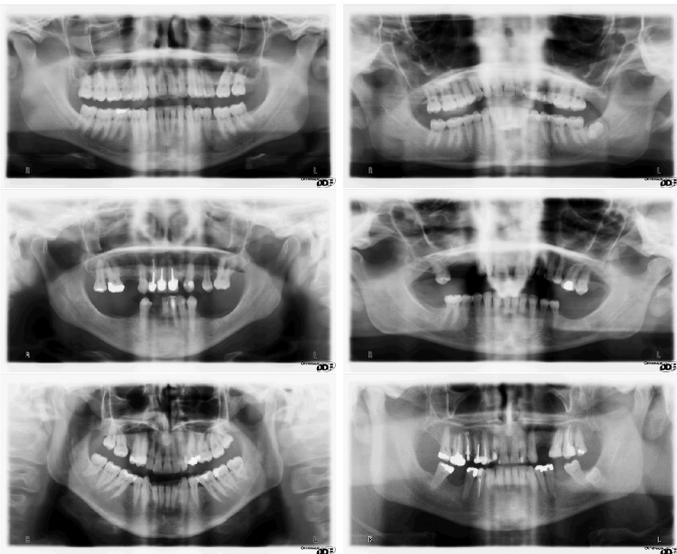


Figure 1: Examples of images of the Dental X-ray data set.

When compared with other types of stomatology images, radiographic images are highly challenging, due to several reasons that increase their heterogeneity. Firstly, different levels of noise, due to the moving imaging device that captures a global perspective of the patient's mouth. Secondly, low contrast, either global or on local regions of the images, the topology and morphologic properties are very complex. Thirdly the blurring that denies the straightforward detection of edges and finally the spinal-column that

¹Socia Lab - Soft Computing and Image Analysis Laboratory, <http://socialab.di.ubi.pt/>

covers the frontal teeth in some images, as shown in figure 1.

We considered that images contains all teeth in two different circumstances, first when the images include only the first and second molar, superior or inferior and left or right along with all other teeth. Second is when the images contain all three molars superior or inferior and left or right in addition to all other teeth. In order to more effectively use the images we named them according to their characteristics. With this arrangement we classify the images based on the number of teeth and in the existence of dental cavities. The naming takes the form *It1 t2 t3.tiff*, where *t1* corresponds to the image number, *t2* represents the quantity of teeth per mouth (0 refers to no teeth, 1 to some teeth and 2 to all teeth) and *t3* denominates the presence of dental cavities in the image (0 corresponds to no dental cavities and 1 to the presence of dental cavities). As an example, the name *I1200_2_0.tiff* corresponds to the image with id equal to 1200, which has all teeth and no dental cavities.

3 Proposed Method to Approximate Panoramic images Into Bitewing images

In this section we detail the proposed method to transform the original images into bitewing images. It consists in three different stages, as shown in the block diagram of figure 2:

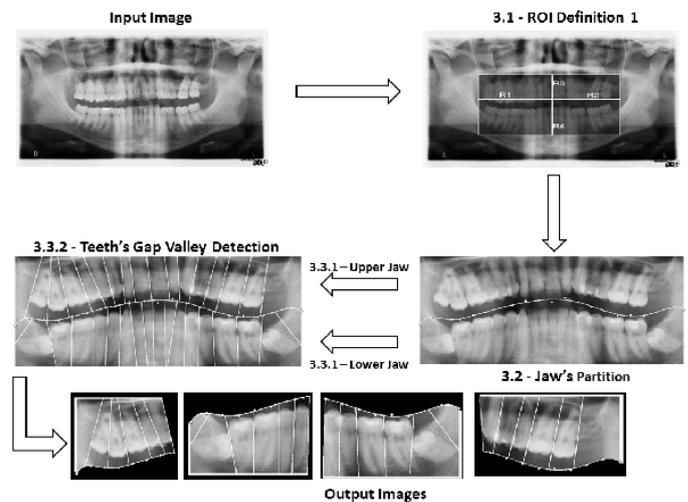


Figure 2: Block Diagram of the proposed method.

3.1 Statistical Analysis

The earliest stage is based in the statistical analysis of the sizes and positions of each component in each image, in order to define an initial region of interest. This eliminates non-useful information originated by the nasal and chin bones. Our purpose is to crop a region that contains the entire mouth and eliminate the maximum amount of noise possible. For each data set image we measured four distances

(R_1, R_2, R_3, R_4) , starting from the image center $(x_c, y_c)=(w/2 = 1408, h/2 = 770)$, as shown in figure 3.

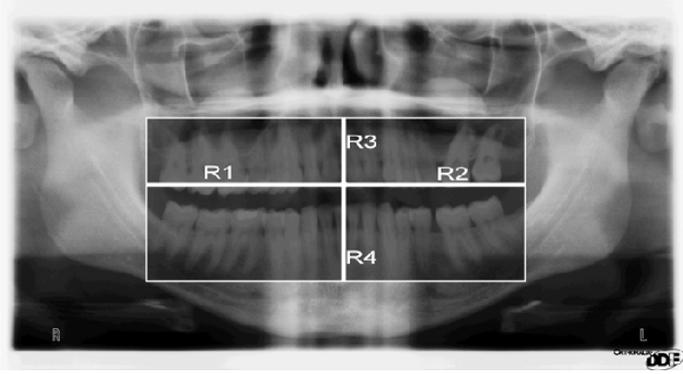


Figure 3: Four lengths extracted of our images database.

Having these values of all the data set images, we obtained the four histograms illustrated in figure 4. Here, the line series correspond to the approximated normal distribution obtained by a line fitting procedure, defined by the (μ, σ) parameters.

Based on this distribution we can approximate the minimum value for each R_i thereby appropriately cropping the images with a 95% certainty. Furthermore, a defined margin guarantees that slightly different images will be appropriately cropped, even if a small increment of non-useful regions is included. The values obtained were $R_1 \approx 897.77$, $R_2 \approx 863.36$, $R_3 \approx 406.31$, and $R_4 \approx 471.27$. Figure 5 illustrate the final results of this stage.

3.2 Jaws Partition

In this stage we separate the upper and the lower jaws, which is done by applying a polynomial fitting process to a set of primitive pixels located between jaws. This set of pixels is based on the horizontal projection, $v(u)$, of the images, given by 1, where $I(x, i)$ denotes the intensity value at line x and column i .

$$v(u) = \sum_{i=0}^w I(x, i) \quad (1)$$

The initial point is defined at the right extreme of the image and at the line that has the minimum $v(u)$ value, given by 2, where w is the image width.

$$p_0(x_0, w - 1) = \arg \min_x (v(u)) \quad (2)$$

The remaining set of points p_i are regularly spaced, starting from $p_0 : p_i(x_i, (w - 1) - W/21)$, where x_i is obtained similarly to x_0 . To avoid too high vertical distances between consecutive p_i , we added the fol-

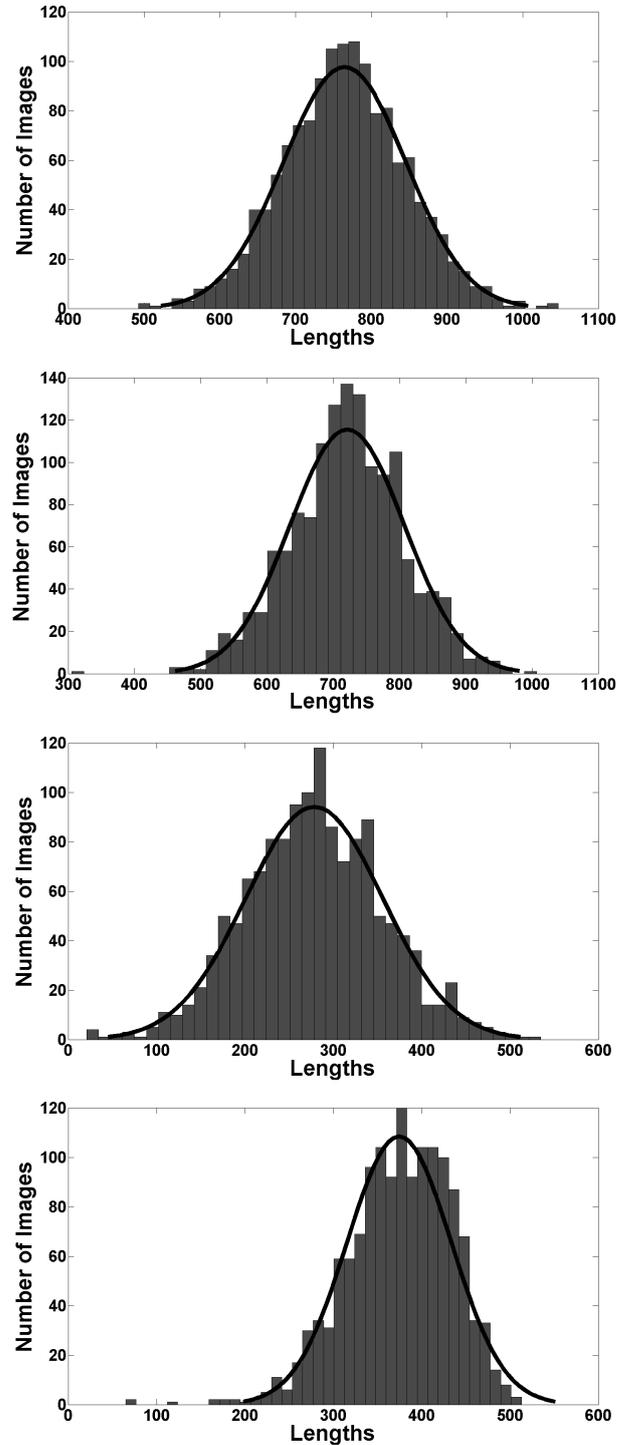


Figure 4: Histogram of the R_i values.

lowing constraints 3.

$$p_i(x_i, y_i) = \begin{cases} p_i(x_{i+1} + T, y_i), \\ |p_i(x_i, y_i) - p_{i+1}(x_{i+1}, y_{i+1})| > T \\ p_i(x_i, y_i), \text{ otherwise} \end{cases} \quad (3)$$

Empirically T is the threshold defined by us to avoid the high vertical distances, with the value of $T = 20$. We observed that this step plays a major role in dealing with missing teeth. This method is based on several proposals to perform the separation of jaws ((Chen and Jain 2004) (Jain and Chen 2004) (Shah,

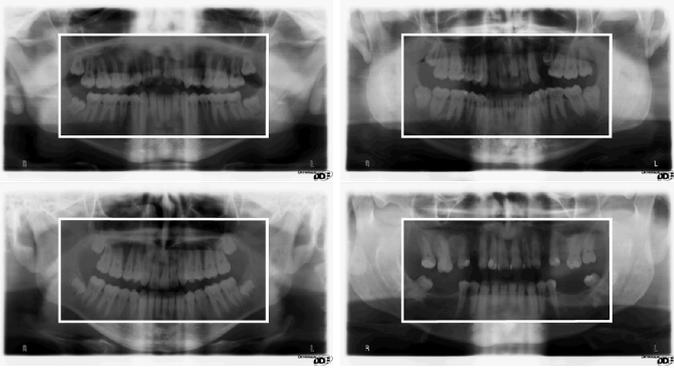


Figure 5: Some results for the ROI definition.

Abaza, Ross, and Ammar 2006)). Having the set of $p_i(x_i, y_i)$ primitive points, the division of the jaws is given by the 10th order polynomial, given by $p(x) = a_0 + \dots + a_{10}x^{10}$, obtained by the least squares fitting algorithm, based on the Vandermond matrix (Dyer and He 2001). This technique is able to achieve impressive accuracy on the data set images, as illustrated in figure 6 and summarized in table 1.

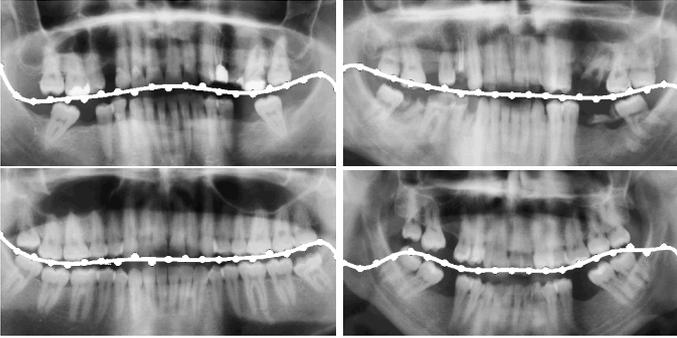


Figure 6: Results of the polynomial fitting by the least squares.

3.3 Teeth Segmentation

Having both jaws divided, our next goal was to localize the regions corresponding to each teeth, which was based in the detection of the gap of intensities between them. This stage can itself be divided into three sub-stages: For each pixel x between 0 and $w - 1$, where w is the image width, we obtain the equation of the perpendicular line to $r(x)$ at the point $(x, r(x))$. Then, the average intensity a of the pixels that fall in that line is obtained. Due to the image proprieties, the shape of the teeth and the shape of the polynomial, we found convenient to vary the angle between the line and the polynomial in the $[-15^\circ; +15^\circ]$ interval. Later, for each x we selected the line that minimizes the average intensities a , obtaining a set of $\{a_1, \dots, a_x\}$ values, which is illustrated in figure 6. The key insight is that the a_i values with lowest values should correspond to the partition between consecutive teeth. In order to compensate for abrupt variations in the values we smoothed these values through the

use of a Gaussian kernel. Later, we extracted the local minimums of the smoothed signal, hoping that they correspond to the desired partitions. However, we observed that generally the number of local minimums is higher than the gap valleys (false positives), but the key false negatives are almost inexistent and, for this reason, the subsequent use of an expert-system based approach will easily perform the detection / partition of each tooth. Based on the counting of the detected teeth, the final step crops the upper and lower jaw images into three regions. The left side is composed by three molars and two pre-molars, the front side contains the incisors and a canine and, finally, the right side has three molars and two pre-molars. As reported in section 2, the data set contains many different dentition shapes and number of teeth per images, which led us to crop the images based on two rules. In the first rule, let S_M and S_{PM} the average size of the molar and premolar, respectively, we define a cut limit given by $L_{cut} = 3 * S_M + 2 * S_{PM}$, where L_{cut} corresponds to the maximum crop limit allowed. In the lower jaw it is similar, but in such case the teeth that matter are the inferior. Finally, the second rule measures the existence of five teeth partition lines after the imposed limit L_{cut} . If there are five lines with large differences between the middle point of each line, we suppose that all teeth exist and that the last division line gives the crop coordinates. Figure 7 gives the interval for the average size of the teeth in 1000 sample size, with a 95% certainty.

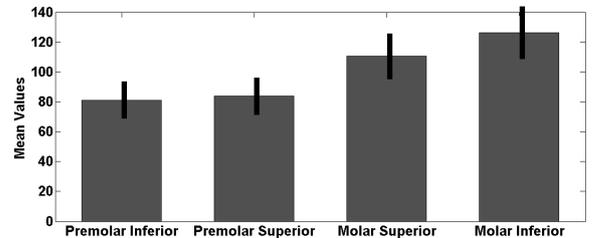


Figure 7: Confidence level for the average size of each type of teeth.

4 Results

In this section we illustrate the results of the above described method that were obtained by visual inspection as showed in table 1. The evaluation of our method is based on a completely independent data sets. It should be stressed that all the results were obtained through 2-folder cross validation.

Stages	Stage 1	Stage 2	Stage 3
Results (% of correct)	95.7	92.6	74.6

Table 1: Results of our method for the test data set.

In figure 8 is illustrated one example of our results for the above described method. The results for stage

1 and stage 2 are very optimistic. For the stage 3 the result showed that improvements must be consider, due to the presence of multiple lines for the same gap valley, and in other cases the absence of lines. The result is positive due to the presence of distinguish darker zone between the teeth, which facilitates the straight lines detection that minimizes the intensity.

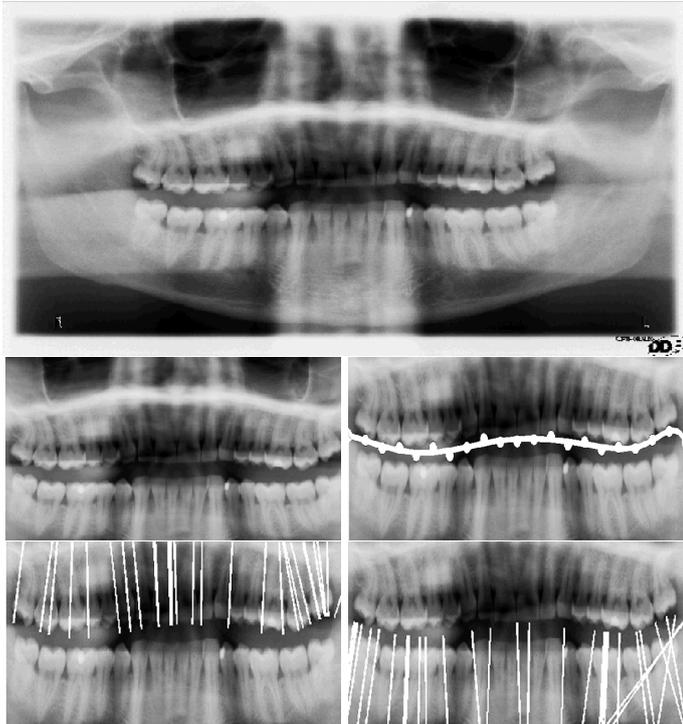


Figure 8: Example of our method results.

An example of a error case is showed in figure 9. The results for stage 1 and stage 2 demonstrates that the results are very optimistic. For stage three becomes essential to use pre processing in the input images, only then we will emphasize the teeth presence in the image. There are several reasons for the increased error in figure 9, among them stand out, regions where the overlap of the teeth is intense (brighter regions), the straight lines detection that minimize the intensity is difficult, because it is a region with little intensities variation. The same happens in the dark areas, because there isn't a contrast between the teeth and the background. For these reasons the straight lines detection that minimize the intensity, are dependent of a pre-processing stage to enhancement the image contrast, consequently obtain the best partition between background and teeth. The proposed method is dependent on the image local intensity variations. Images where these variation is not so evidently becomes more complicated to detect the straight lines detection that minimize the intensity.

5 Conclusions / Future Work

In this paper it is announced the availability of a new data set of panoramic dental X-ray images, which can

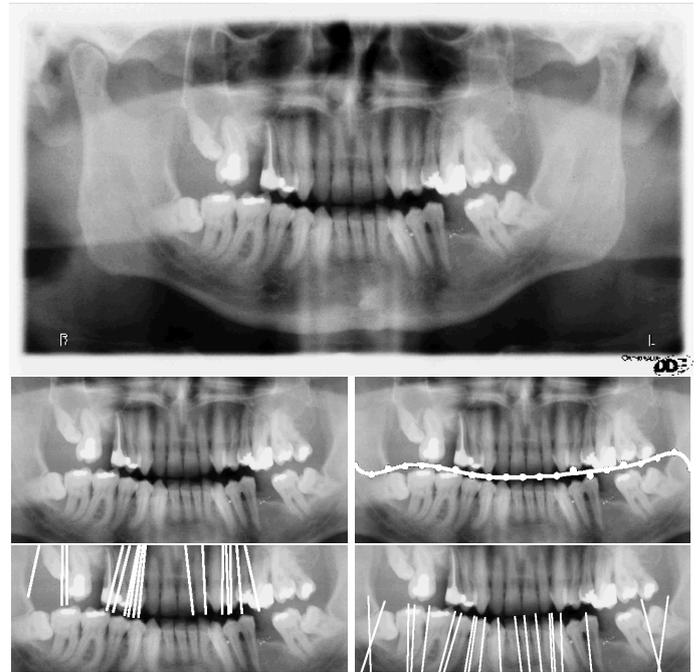


Figure 9: Example of failure in the teeth segmentation.

constitute a tool for the research community in the development of the stomatologic-related applications. This data set has varying morphologic properties that make it more valuable to the scientific community. These varying proprieties include the number of teeth per image, the shape of the mouth and teeth as well as the levels of noise. Also, we gave a method to approximate the panoramic dental X-ray images into bitewing dental X-ray images, which is the type of stomatologic images generally used in image processing methods. This method starts by the definition of a region of interest, followed by the partition of the data into both jaws (through a least squares polynomial fitting process) and, finally, the detection of the teeth in each image which are subsequently counted in order to make the partition into the final desired types of images. As future work our main goal is to start the detection of dental diseases, using the described dental X-ray data set. We plan to start by the detection of dental cavities and move to other types of problems. For that we must improve the accuracy of the teeth segmentation. We are currently studying the use of active contours in the teeth segmentation stage, with this technique and the increase of segmentation accuracy we look to obtain good results in our main goal.

REFERENCES

- Chen, H. and A. Jain (2004, August). Tooth contour extraction for matching dental radiographs. *Proc. 17th Int'l Conf. Pattern Recognition III*, 522–525.
- Dyer, S. and X. He (2001, December). Least-squares fitting of data by polynomials. *IEEE Instrumentation and Measurement Magazine* 3(2), 46–51.
- Jain, A., H. Chen, and S. Minut (2003). Dental biometrics: Human identification using dental radiographs. *AVBPA, UK*, 429–437.
- Jain, A. K. and H. Chen (2004). Matching of dental x-ray images for human identification. *Pattern Recognition* 37, 1519–1532.
- Jain, A. K., L. Hong, and S. Pankanti (2000, February). Biometrics: promising frontiers for emerging identification market. *in Comm. ACM*, 9 1–98.
- Li, S., T. Fevens, A. Krzyzak, C. Jin, and S. Li (2007). Semi-automatic computer aided lesion detection in dental x-rays using variational level set. *Pattern Recognition* 40, 2861–2873.
- Li, S., T. Fevens, A. Krzyzak, and S. Li (2006). An automatic variational level set segmentation framework for computer aided dental x-rays analysis in clinical environments. *Computerized Medical Imaging and Graphics* 30, 65–74.
- Lin, J., K. Cheng, and C. Mao (1996, August). A fuzzy hopfield neural network for medical image segmentation. *IEEE Trans. Nucl. Sci.* 43(4), 2389–2398.
- Nomir, O. and M. Abdel-Mottaleb (2005). A system for human identification from x-ray dental radiographs. *Pattern Recognition* 38, 1295–1305.
- Nomir, O. and M. Abdel-Mottaleb (2007). Human identification from dental x-ray images based on the shape and appearance of the teeth. *IEEE Transactions on Information Forensics and Security* 2(2).
- Nomir, O. and M. Abdel-Mottaleb (2008, June). Fusion of matching algorithms for human identification using dental x-ray radiographs. *IEEE Transactions on Information Forensics and Security* 3(2).
- Said, E. H., D. E. M. Nassar, G. Fahmy, and H. H. Ammar (2006, June). Teeth segmentation in digitized dental x-ray films using mathematical morphology. *IEEE Transactions on Information Forensics and Security*.
- Shah, S., A. Abaza, A. Ross, and H. Ammar (2006, June). Automatic tooth segmentation using active contour without edges. *IEEE Biometrics Symposium 1*(2).
- Shiffman, S., G. D. Rubin, and S. Naple (2000, November). Semi-automatic computer aided lesion detection in dental x-rays using variational level set. *IEEE Trans. Med. Imag.* 19(11), 1064–1074.
- Zhou, J. and M. Abdel-Mottaleb (2005). A content-based system for human identification based on bitewing dental x-ray images. *Pattern Recognition* 38, 2132–2142.

