Graduate Theses, Dissertations, and Problem Reports

2008

High performance image processing techniques in automated identification systems

Ayman Abaza
West Virginia University

Follow this and additional works at: https://researchrepository.wvu.edu/etd

High Performance Image Processing Techniques

In

Automated Identification Systems

Ayman Abaza

Dissertation submitted to the
College of Engineering and Mineral Resources
at
West Virginia University
in partial fulfillment of the requirements
for the degree of

Doctor of Philosophy
in
Electrical Engineering

Hany Ammar, Ph.D., Chair
Xin Li, Ph.D., Co-Chair
Arun Ross, Ph.D.
Tim McGraw, Ph.D.
Sam Mukdadi, Ph.D.

Lane Department of Computer Science and Electrical Engineering

Morgantown, West Virginia
2008

Keywords: Dental Identification, ADIS, Film Segmentation, Teeth Contour
Extraction, Teeth Classification, Dental Chart, Potential Match Search, Teeth
Registration, Ear Identification

ABSTRACT

High Performance Image Processing Techniques in Automated
Identification Systems

Ayman Abaza

This dissertation addresses different image processing problems faced during the development of two different identification systems (i) an automated system for postmortem identification using dental records (dental radiographs), (ii) an automated ear identification system. Automating the postmortem identification of deceased individuals based on dental characteristics is receiving increased attention especially with the large number of victims encountered in mass disasters, as 9/11 attack, and Tsunami. The Automated Dental Identification System (ADIS) can be used by law enforcement agencies to locate missing persons using databases of dental x-rays of human remains and dental scans of missing or unidentified persons. ADIS provides functionality for users to upload the reference records, and submit identification queries using submitted records. ADIS then produces a short matching list of possible matches for the dental experts to verify. The ear identification system can be used at access point of restricted areas; this system helps identify a person from surveillance videotapes.

This dissertation introduces new high performance approaches for three image-processing problems of the ADIS record preprocessing stage. For the first, we introduce an automatic hierarchical approach to the problem of cropping dental image records into films. Our approach is heavily based on concepts of mathematical morphology and shape analysis. Testing reflects an overall error of $\sim 3\%$. For the second, we address the problem of teeth contour extraction using active contour without edges. This technique is based on the intensity properties of the overall region of the tooth image. It extracts a very smooth and accurate tooth contour. For the third, we enhance the existing techniques for automatic classification of teeth into four classes (molars, premolars, canines, and incisors); as well as the

construction of a dental chart, which is a data structure that guides tooth-to-tooth matching. We tackle this composite problem using appearance-based features (low computational-cost) for assigning an initial class, followed by applying a string matching with don't care technique based on teeth neighborhood rules. Adding the don't care character allows the technique to work in the presence of missing tooth, which represents 21% of the database. Our approach achieves 82% teeth labeling accuracy based on a large test dataset of films.

For ADIS, also we introduce new techniques for the problem of fast dental image retrieval. We use Eigen images to reduce the dimensionality of each tooth, as well as other teeth contour descriptors. The main features of this search engine are that it completes the search in order of seconds and it reaches a reasonable accuracy with a relatively short candidate list.

For ear identification, we develop different components of a viable automated method for ear identification system. We automate the Iannarelli ear identification system, which had been used manually for years. We extract the ear external and internal curves, and use these curves to calculate the different Iannarelli distance measurements. We evaluated the system performance based on statistical analysis of a large dataset of thousands ear images, where the identification rate is 90% for rank 1 image.

Dedication

To my parents soul,

To my wife and kids with love,

To all my teachers, and

To all those who cooperated in this work.

Acknowledgements

Supporting Publications and Funding Sources

[1] D. Nassar, E. HajSaid, **A. Abaza**, S. Chekuri, J. Zhou, M. Mahoor, O. Nomir, H. Chen, G. Fahmy, H. Ammar, M. Abdel-Mottaled and A. Jain, "Automated Dental Identification System (ADIS)" Proc. of the 6th National Conference on Digital Government Research (dg.o.), Atlanta - GA, May 2005.

[2] S. Chekuri, D. Nassar, **A. Abaza**, A. Bahu, H. Ammar, and G. Fahmy, "A web-based Automated Dental Identification System (webADIS)" Proc. of the 5th IBIMA International Conference on Internet & Information Technology in Modern Organizations, Egypt, Dec 2005.

[3] X. Li, **A. Abaza**, D. Nassar, and H. Ammar, "Fast and Accurate Segmentation of Dental X-ray Records" Proc. of the 2nd International Conference on Biometrics (ICBA), Hongkong, Jan 2006.

[4] S. Chekuri, D. Nassar, **A. Abaza**, E. HajSaid, A. Bahu, U. Qurashi, G. Fahmy, and H. Ammar, "webADIS: A Flexible web-based Environment for the Automated Dental Identification System" Proc. of the 7th Annual International Conference on Digital Government Research (dg.o.), San Deigo - CA, May 2006.

[5] S. Shah, **A. Abaza**, A. Ross, and H. Ammar, "Automatic Teeth Segmentation Using Active Contour Without Edges" Proc. of the Biometric Consortium Conference (BCC), Baltimore, Sep 2006.

[6] S. Chekuri, D. Nassar, **A. Abaza**, A. Bahu, U. Qureshi, and H. Ammar, "webADIS: A Flexible web-based Environment for the Automated Dental Identification System" Proc. of the 2nd International Computer Engineering Conference (ICENCO), Cairo - Egypt, Dec 2006.

[7] **A. Abaza**, U. Qureshi, A. Bahu, E. HajSaid, D. Nassar and H. Ammar, "Automated Dental Identification System (ADIS) in Testing Mode", Proc. of the 8th Annual International Conference on Digital Government Research (dg.o.), Philadelphia - PA, May 2007.

[8] A. Mahmoud, D. Cortes, **A. Abaza**, H. Ammar, M. Hazey, P. Ngan, R. Crout, and O. Mukdadi, "Non-Invasive Assessment of Human Jawbone Using Ultrasonic Guided Waves", Proc. of the IEEE International Ultrasonics Symposium, New York - NY, Oct 2007.

[9] A. Mahmoud, D. Cortes, **A. Abaza**, H. Ammar, M. Hazey, P. Ngan, R. Crout, and O. Mukdadi, "Non-Invasive Assessment of Human Jawbone Using Ultrasonic Guided Waves", IEEE Transactions on Ultrasonics, Ferroelectric and Frequency Control (to appear in Special issue on Ultrasound in Bone, June 2008).

[10] D. Nassar, **A. Abaza**, X. Li, and H. Ammar, "Automatic Construction of Dental Charts for Post-Mortem Identification", IEEE Transactions on Information Forensics and Security, to appear June 2008.

[11] H. Ammar, D. Nassar, **A. Abaza**, and E. HajSaid, "Automated Dental Identification System", provisional US patent from (Jan 2006 – Jan 2007), filled Jan 2007.

[12] **A. Abaza,** A. Ross, and H. Ammar, "Retrieving Dental X-Ray Radiographs for Post-Mortem Identification", to be submitted to the 6th Biometrics Symposium (BSYM 2008)

[13] E. HajSaid, **A. Abaza,** A. Ross, and H. Ammar, "Ear Detection in Color Facial Images Using Mathematical Morphology", to be submitted to the 6th Biometrics Symposium (BSYM 2008)

[14] **A. Abaza,** A. Ross, and H. Ammar, "Retrieving Dental X-Ray Radiographs for Post-Mortem Identification", to be submitted to IEEE Transactions on Information Forensics and Security

[15] **A. Abaza,** A. Ross, and H. Ammar, "Automated Iannarelli Ear Identification System", to be submitted to IEEE Transactions on Information Forensics and Security

Table of Contents

1   Introduction

1.1   Overview

Humans have used body characteristics such as face, and voice for thousands of years to recognize each other. Reliable recognition is important for many applications, such as law enforcement, border control, or access control. Law enforcement agencies charged with the responsibility of investigating the evidence from cases involving violent crime, missing persons and mass disaster scenarios, have exploited biometrics for decades as key forensic recognition tools. With the evolution in information technology and the huge volume of cases that need to be investigated by forensic specialists, automation of forensic identification became inevitable.

1.1.1   Automated Dental Identification System (ADIS)

Post-Mortem (PM) identification i.e. identification after death, is a more difficult problem than Ante-Mortem (AM) identification, since few biometric characteristics can resist early decay of body tissues as well as withstand severe conditions usually encountered in mass disasters. Dental features are the best candidates for PM identification.

Recent disasters have brought the significance of automated dental identification systems. For example in 9/11 attack, many victims were identifiable only from pieces of their jaw bones. Forensic Dental experts manually using dental records were able to identify about 20% of the 973 victims identified in the first year after 9/11 attack [80]. Identifying the 2,749 victims of 9/11 disaster [81] took around 40 months [80]. The number of Asian tsunami victims identified during the first 9 months was only 2,200 (out of an estimated total of 190,000) [82]. Most of those identified Tsunami Victims were based on manual dental information, about 75% of the tsunami victims in Thailand were identified using dental records, compared to just 0.5% using DNA [82]. As can be concluded that manual dental identification is very powerful postmortem identification but it is also very time consuming process.

In 1997, the Criminal Justice Information Services Division (CJIS) of the Federal Bureau of Investigation (FBI) created a dental task force (DTF) whose goal is to improve the utilization and effectiveness of the National Crime Information Center's (NCIC) Missing and Unidentified Persons (MUP) files. The DTF recommended the creation of a Digital Image Repository (DIR) and an Automated Dental Identification System (ADIS). The proclaimed ADIS is an automated biometric system that when fed with raw subject dental record or records, will find a minimum set of candidate (or reference) records that have high similarities to the subject(s). Then, a forensic expert can examine the radiographs of the few candidates instead of the entire DIR.

1. Introduction

We view ADIS as a collection of the following components (as shown in Figure 1):

(i)   Record Preprocessing component which handles dental records cropping into dental films, grayscale contrast enhancement of films, classification of films into bitewing, periapical, or panoramic views, segmentation of teeth region from films, teeth contour extraction and finally annotating teeth with labels according to their location.
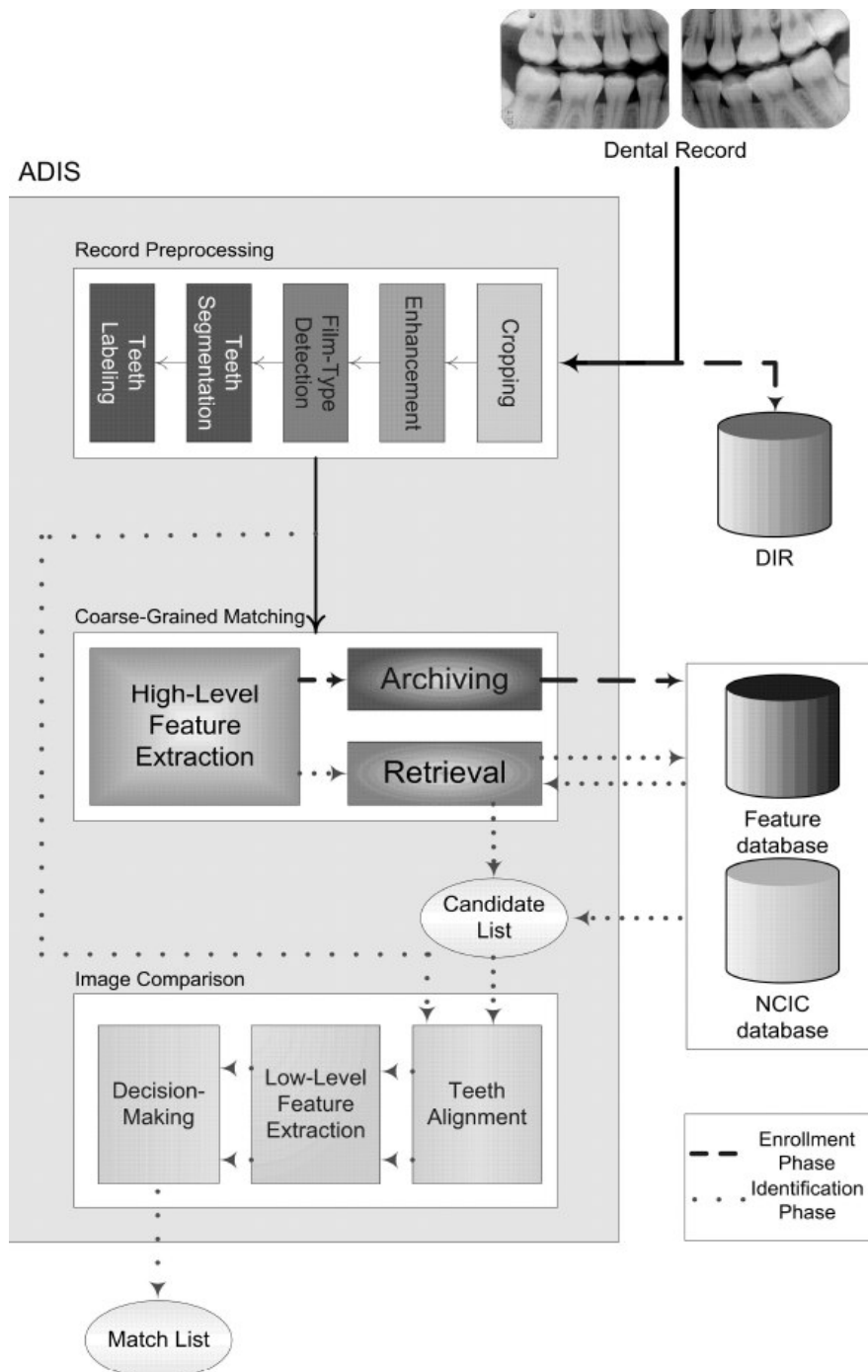
Figure 1: ADIS block diagram

(ii)   Potential Matches Search component which manages archiving and retrieval of dental records based on high-level dental features (e.g. number of teeth and their shape properties) or dimensionality reduction and produces a candidate list.

(iii)   Image Comparison component which mounts for low-level tooth-to-tooth comparison between subject teeth -after alignment- and the corresponding teeth of each candidate, thus producing a short match list.

1.1.2 Automated Ear Identification System (AEIS)

With the recent popularity in the biometric field, there has been a push to develop and study other possible characteristics that can be used for AM identification. Although we most closely identify with methods that we as humans would use to identify each other, automated computer based systems are able to utilize many other physical characteristics in the same manner. Ear structure is one such possible characteristic because the human ear has many variations between individuals, including the various curves and dimensions of the ear [99]. This was the motive behind building an Automated Ear Identification System (AEIS). We view AEIS as a collection of the following components (as shown in Figure 2):
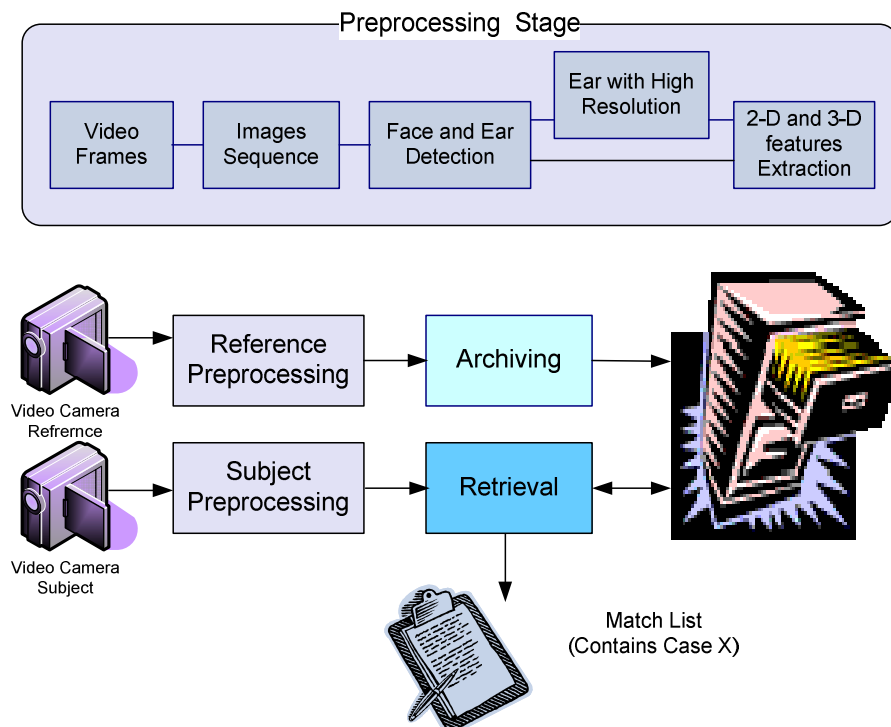
Figure 2: AEIS block diagram

- Face and Ear localization and tracking among the different input frames,

- Enhancement step of super resolution, also this enhancement step includes contrast and brightness adjustment, size (or dimension) normalization, and background elimination.

- Feature extraction which starts by the origin point detection followed by different feature point's measurements.

- According to the operational mode either these measurements are stored as a feature representation of the case; or the distance between the subject features and each reference case is used to retrieve the top matched cases from the database.

1.2 What is New in This Dissertation?

This dissertation introduces new approaches for different stages of ADIS, and AEIS (as discussed later in details):

- For ADIS, we propose an automated hierarchical approach to the problem of cropping dental image records into films. We propose a three-stage approach for cropping. In the first stage, we extract the background layer of the image record, extract connected components and classify them as either round-corner or right-corner connected components. In the second stage, we perform arch detection and dimension analysis, realization of this stage differs according to the corner type. In the third stage, we perform topological assessment of the cropping results in order to eliminate spurious objects.

- For ADIS, we address the problem of teeth contour extraction using active contour without edges. This technique is independent of the gradient of images for extracting the contour of teeth. It is based on the intensity properties of the overall region of the tooth image. It extracts a very smooth and accurate tooth contour. Also, the starting contour can be in any part of the image and does not have to be around the boundary of the tooth.

- For ADIS, an important problem in automated dental identification is automatic classification of teeth into four classes (molars, premolars, canines, and incisors). An equally important and related problem is the construction of a dental chart, which is a data structure that guides tooth-to-tooth matching. Dental chart is a key to avoid illogical comparisons that inefficiently consume the limited computational resources and may mislead decision-making. In other words, constructing dental chart improves the identification performance. We tackle this composite problem using a two-stage approach. The first stage utilizes low computational-cost, appearance-based features for assigning an initial class. The second stage applies a string matching technique, based on teeth neighborhood rules, to validate initial teeth-classes and hence to assign each tooth a number corresponding to its location in the dental chart, even in the presence of missing tooth.

- For ADIS, Potential Matches Search component is a fast search engine for a large database, which produces a longer candidate list as an input to the Image Comparison component. First, we use principal component analysis to reduce the dimensionality of each labeled segment (feature-based approach). This approach needs teeth registration, this task is accomplished using a novel projection based approach to estimate the rotation difference between teeth images. Second, we use shape-based features descriptors approaches. We fuse the Euclidian distances between corresponding teeth features, and hence calculate the match score between the subject record and the reference records.

- For AEIS, we develop different essential components for an automated ear identification system. The proposed system is an automation of the Iannarelli system, which has been manually ear used for more than 40 years. We present different image processing techniques, which help the image enhancement and hence ear curves extraction. We extract curves that represent different anatomical structure of the ear; that we use to calculate the different Iannarelli measurements, and hence identify or verify the person identity.

The contributions of this thesis are summarized as follows:

- Automate the methodology to segment dental films from records using morphological tool, and shape analysis.

- Develop a new teeth contour extraction method using active contour without edges.

- Generalize the teeth labeling technique to work even in the presence of a missing segment using string matching with don't care character.

- Boost the teeth labeling performance by considering the film class.

- Develop a new image registration method that can estimate the rotational difference based on circular projection. We used this method to align the different teeth segments to vertical models.

- Develop a methodology for the Potential Match Search Component based on appearance based, as well as different shape based. We fuse different tooth-to-tooth classifiers scores, to reach a record-to-record matching score; hence to retrieve reference records of high resemblance.

- Automate Iannarelli ear identification system that has been used manually for years.

1.3 Problem Statement

In pursue of our research effort on developing a research prototype of an Automated Dental Identification System (ADIS) and Ear Identification System (AEIS), we identify the following challenging problems:

- **Problem One: Record Cropping:** is an automatic global segmentation of dental films from their corresponding records. On one hand, we want segmentation results to be as accurate as possible since inaccuracy in cropping of dental records is likely to hinder the performance of subsequent processing steps and accordingly the overall performance of the entire identification system. On the other hand, we want the computational cost to be reasonably low, especially with the large volume of records that need to be processed.

    - Given:

    - o Different dental records have different scanned resolutions.

    - o A dental record D, with f radiographic films, a heterogeneous back ground.

    - o The background can have different color, texture, and sometimes contain text.

    - Problem One:

    - o To decompose dental image records into their constituent radiographic films.

- **Problem Two: Tooth Contour Extraction** is another level of segmentation, to extract the contour of the teeth. Tooth contour is one of the most important features for dental recognition. This is a very challenging task due to the fact that tooth can be defined by fuzzy contour. In other words, due to the smooth gray-level transient between the tooth and background, there is no sharp cutoff edge.

    - Given:

    - o A dental record D, with f cropped radiographic films, and each film contains Q dental ROIs.

    - o Each of the Q dental ROIs corresponds to a tooth (t_q), and each tooth has a contour.

    - o Tooth usually has smooth edges, and presented in a noisy background.

    - Problem Two:

    - o To segment a smooth and accurate tooth contour, in the presence of noise, and the absence of sharp contour gradient.

    - o To use a fast method, so as not to affect the overall system timeline performance.

- **Problem Three: Teeth Labeling** is an automatic classification of teeth into incisors, canines, premolars and molars as part of creating a dental chart that guides tooth-to-tooth matching. We seek to label teeth as such in order to avoid illogical comparisons between teeth, which inefficiently consume computational resources and may also mislead match decision-making. In other words, constructing dental chart improves the identification performance.

  - Given:

  - A training database of labeled dental Regions of Interest (ROIs) each assigned a class label $\omega_k \in \{$ 'I', 'C', 'P', 'M' $\}$, where 'I' stands for Incisor, 'C' stands for Canine, 'P' stands for Premolar, and 'M' stands for Molar.

  - Adult Dental Atlas with specific teeth arrangement.

  - A dental record D, with f radiographic films, and each film contains Q dental ROIs,

  - Each of the Q dental ROIs corresponds to an input tooth (t_q),

  - Cases of missing tooth.

  - Problem Three:

  - To assign a tooth-class label ω_k to each t_q (Initial teeth classification).

  - To assign a label to each tooth based on teeth neighborhood rules, even in the absence of a segment (missing tooth)

  - To validate the initial teeth-classes and hence to assign each tooth a number corresponding to the tooth location in the dental chart.

- **Problem Four: Potential Matches Search** is to search the dental database in a very fast manner to find a candidate list. This candidate list must have a very high probability of containing the target record, and have a suitable length compared to the original database size. Alignment is a key step for appearance-based features, same as extracting tooth contour for shape-based. Finally classifier fusion between multiple subject teeth and their corresponding reference teeth to reach matching score between the subject and each of the reference records.

  - Given:

  - A Subject dental record D_s, with f_s radiographic films, and each film contains Q_s dental ROIs,

- Each of the Q_s dental ROIs corresponds to a tooth view (t_s), and each tooth view is assigned a label corresponding to the tooth location in the dental chart.

- A data base of references, each reference record D_r with f_r radiographic films, and each film contains Q_r dental ROIs,

- Each of the Q_r dental ROIs corresponds to a tooth view (t_r), and each tooth view is assigned a label corresponding to the tooth location in the dental chart.

- There exist multiple views for some teeth either in the subject or the reference records.

- Problem Four:

- To determine the probability of match p_M (t_s, t_r) between subject tooth (t_s) and reference tooth (t_r).

- To determine the probability of match $p_M(S, R)$ between the subject record S and each of the reference records R.

- To form a candidate list that contains those reference records of high matching probability.

- **Problem Five: Automated Ear Identification System (AEIS)** is to search the ear database to find a reference image, and hence to identify, or verify the identity of a person using ear biometric.

- Given:

- A Subject video sequence. This sequence contains face profile images at different angles.

- A data base of references, each reference consists of video sequence, which in turn contains face profile images at different angles.

- Problem Five:

- Accurately calculating different Iannarelli distances; this problem can be further divided into detection of the origin calculation point, plus detection of the outer and inner ear curves.

- To determine the probability of match $p_M(E_S, E_R)$ between the subject segmented ear image E_S and each of the reference segmented ears E_R.

- To form a match list that contains those reference records of high matching scores.

1.4   Research Objectives

This research is broadly oriented towards developing: (i) Automated Dental Identification System (ADIS) that facilitates the task of the dental expert and cut down their searching time for missing, wanted and unidentified persons. (ii) Automated Ear Identification System (AEIS) that researched the ability to identify a person from surveillance cameras using ear biometrics. ADIS research problems, we studied, are part of the integral research framework of ADIS supported under the US National Science Foundation (NSF) and the US National Institute of Justice (NIJ) awards numbers EIA-0131079 and 2001-RC-CX-K013 respectively. AEIS research problems, we studied, are part of the integral research framework of ear identification system supported under the Technical Support Work Group (TSWG) award number W91CRB-05-C-0028.

The objectives of this research are:

- To automate the process segmenting films from records, using a fast and accurate morphological segmentation method with accuracy greater than 95%.

- To segment a smooth representative tooth contour for more than 85% and to reduce this process time to faction of second.

- To achieve better teeth classification accuracy greater than 85% by considering film type (periapical and bitewing), as well as considering other appearance based classifiers. To achieve correct teeth labeling, even in absence of one segment (missing tooth).

- To enhance the scalability of ADIS, by accomplishing high hitting rate and relatively short candidate list. Finally as we are working with large databases which involves hundred records (or thousands teeth), this searching engine needs to achieve record-to-record decision in a fraction of a second.

- To automatically measure distinctive ear features with a view to identify or confirm the identity of the owner.

1.5   Dissertation Organization

The dissertation consists of six chapters and two appendices organized as follows:

- Chapter 1 is titled "Introduction," gives a general introduction to ADIS, and AEIS and illustrates the different problems that we will present in this dissertation.

- Chapter 2 is titled "Literature review," presents the history of using dental identification and the development of the first automated dental identification system (ADIS). Then it presents a literature review about ear biometric system. Finally it presents a literature review about image based retrieval and multimodal biometric system.

- Chapter 3 is titled "Different Preprocessing Techniques for ADIS," explains three methods related to ADIS record preprocessing namely record cropping, teeth contour extraction and teeth labeling.

- Chapter 4 is titled "ADIS Potential Match Search," presents new techniques to achieve fast dental record retrieving which directly address the scalability problem of ADIS.

- Chapter 5 is titled "Automated Ear Identification System (AEIS)," presents novel approaches for automation of Iannarelli ear identification system.

- Chapter 6 is titled "Conclusions and Future Work," concludes the dissertation and gives some suggested future work as extension to our presented work.

- Appendix I is titled "Dental Glossary", gives definitions for the most common terminology used in dental identification.

2   Literature Review

Biometrics System is an automated method of recognizing an individual based on physiological or behavioral characteristics. Biometrics is being used in many locations to enhance the security and convenience of the society. Example deployments within the United States Government include the FBI's, the US-VISIT program, etc. Law enforcement agencies have exploited biometric identifiers, for decades, as a means for forensic identification.

Forensic identification is the application of forensic science and technology to identify specific objects (peoples) from the traces they leave. Forensic science is a scientific method of gathering and examining evidence. Crimes are solved with the use of pathological examinations that gather fingerprints, palm prints, tooth bite prints, blood, etc. Some people return this science to the time of Romans, where a criminal charge meant presenting the case. Both persons representing the two sides of the case gave speeches, and the one with best argument (best forensics skills) would win [1]. One of the first recorded applications of medical knowledge to the solution of crime, in 1248 Chinese book "Hsi DuanYu" or the Washing Away of Wrongs, described ways to distinguish between death by drowning and death by strangulation [2].

Forensic identification can be classified as Post-Mortem (PM) identification, i.e. identification after death, or Ante-Mortem (AM) identification. AM identification techniques includes but not limited to fingerprints, face, ear, or from other traces using other biometric techniques, while PM identification is carried out using [3]:

(i) Presumptive identification which include visual: May not be possible due to disfigurement or fragmentation due to disaster trauma, and it is the least reliable due to emotional stress and subjective assessment required of family member.

(ii) Finger/foot prints: can not survive in case of severe fires with disaster, or late discovery of the body.

(iii) Radiological: is useful only if the AM medical radiographs are available.

(iv) DNA: is becoming standard in military; however the civilian sector is still lagging behind. Still the development of DNA is providing forensic investigators a very promising PM identification tool.

(v) Dental: is widely used in criminal investigations and mass disasters, based on PM versus AM dental feature comparison.

Under severe circumstances encountered in mass disasters, such as those (e.g. aircraft accidents, Tsunami, hurricanes, etc.) or if identification is being attempted more than a couple of weeks after death (e.g. late

corpse discovery), soft tissues of the human body would have decayed to unidentifiable status. Because of their survivability, diversity and availability, the best candidates for biometric PM identification under such severe conditions are dental features.

This chapter is organized as follows, section 2.1 introduce literature review about dental identification, while section 2.2 introduce the first automated dental identification system (ADIS). Section 2.3 presents literature review about ear biometrics. In section 2.3, and 2.4 we present a review about image retrieval from database, and multimodal biometric system which we are going to use for dental image retrieval (chapter 4).

2.1   Dental Identification

The most important role of the forensic dentist is the identification of deceased individuals. The dental expert compares the PM and AM dental records and establishes (to a high degree of certainty) that whether they belong to the same person or not. A range of conclusions can be reached when reporting a dental identification. The American Board of Forensic Odontology recommends that these be limited to the following four conclusions [5]:

- Positive identification: The antemortem and postmortem data match in sufficient detail, with no unexplainable discrepancies, to establish that they are from the same individual.

- Possible identification: the antemortem and postmortem data have consistent features but, because of the quality of either the postmortem remains or the antemortem evidence, it is not possible to establish identity positively.

- Insufficient evidence: The available information is insufficient to form the basis for a conclusion.

- Exclusion: the antemortem and postmortem data are clearly inconsistent.

2.1.1   History of Dental Identification

One of the first reported cases of postmortem dental identification is the recognition of Dr. Warren body in 1776. This patriot was killed during the Battle of Breeds, and was buried in an unmarked grave by the British forces. Ten months later, his family wants to recover his remains. The grave was having two bodies, but Paul Revere recognized Warren's via artificial teeth special shapes. Warren was buried later with honors 1776. Dental identification was first used in the United States court to prove a murder crime of Dr. Parkman on 1850. Dr. Parkman disappeared after meeting with Dr. Webster that owed him a fortune. Portions of a porcelain denture were recovered from Dr. Webster's laboratory furnace. Dr. Keep who was the dentist for Dr. Parkman, identified those fragments of porcelain denture, and integrated them to Dr. Parkman's jaw plastic model. Dr. Webster was convicted of murder, and was executed 1880 [4].

Ndiokwelu et al [6] reported that in 1887, Gordon recommended to keep AM dental charts as a mean to help identify missing persons. In other words, this was the first step to have established dental forensic identification systems.

2.1.2 Dental Identification Process

Keiser-Nielsen in [7] presented a six-digit code for dental description of each tooth. This code is intended to simplify the establishment of AM/PM data files in a computer and consists of two digits for designating the individual tooth, two digits for describing the surface or surface combination involved (if any), and two digits for describing treatment type and material. Then a data comparison can be carried manually or by computer of one postmortem versus antemortem set.

Jones [8] introduced another dental identification process that is based on the fact that changes in the characteristics of teeth are unidirectional; that is, once any change (decay and dental work) happened to any tooth, it will never return to original condition, however this dental work can change. For his proposed system, he considered 5 surfaces of each tooth as either intact (no decay, no restoration) or changed in some way (filling, crown, missing, and so forth). He formalized a well known heuristic rule that if a postmortem examination indicates that a surface is intact then any antemortem records with that surface altered cannot possibly be a match. The same obviously applies for tooth present postmortem, tooth missing antemortem. Those rules are still used as exclusion rules by forensic experts.

Examples of the used antemortem and postmortem forms are presented in Figure 3, where "M" mesial; "O" occlusal or incisal; "D" distal; "F" facial, labial, or buccal; and "L" lingual. A sixth box, "X," indicate whether the tooth is present or absent [8].

Name: _____   Agency: _____

DENTAL MISSING PERSON REPORT — Overlay System

M	O	D	Tooth #	Description	F	L	X
O	O	O	1	_____	O	O	O
O	O	O	2	_____	O	O	O
O	O	O	3	_____	O	O	O
O	O	O	4	_____	O	O	O
O	O	O	5	_____	O	O	O
O	O	O	6	_____	O	O	O
O	O	O	7	_____	O	O	O
O	O	O	8	_____	O	O	O
O	O	O	9	_____	O	O	O
O	O	O	10	_____	O	O	O
O	O	O	11	_____	O	O	O
O	O	O	12	_____	O	O	O
O	O	O	13	_____	O	O	O
O	O	O	14	_____	O	O	O
O	O	O	15	_____	O	O	O
O	O	O	16	_____	O	O	O
O	O	O	17	_____	O	O	O
O	O	O	18	_____	O	O	O
O	O	O	19	_____	O	O	O
O	O	O	20	_____	O	O	O
O	O	O	21	_____	O	O	O
O	O	O	22	_____	O	O	O
O	O	O	23	_____	O	O	O
O	O	O	24	_____	O	O	O
O	O	O	25	_____	O	O	O
O	O	O	26	_____	O	O	O
O	O	O	27	_____	O	O	O
O	O	O	28	_____	O	O	O
O	O	O	29	_____	O	O	O
O	O	O	30	_____	O	O	O
O	O	O	31	_____	O	O	O
O	O	O	32	_____	O	O	O

Unusual Dental Characteristics: _____

Individual's Characteristics:

Age_____ Sex_____ Race_____ Ht._____ Wt._____ Date Missing: _____

Personal Dentist: _____

Reporting Agency: _____

_____ Phone No.: (   )_____

Copyright 1986 by Gerald Jones D.D.S

Figure 3: Dental antemortem (AM) form

2.1.3 Computer Aided Dental Identification System

Mass disasters which produce multiple fatalities consume the resources of forensic science teams. Even under ideal circumstances, the positive identification of a large number of victims is extremely stressful,

time-consuming, and costly. This increasing requirement for dental based identification motivated the development of assisting computer software. Lorton et al in [9] described the Computer-Assisted Postmortem Identification (CAPMI) system that was developed at the U.S. Army Institute of Dental Research to facilitate rapid identification of human remains. CAPMI System was originally developed by the U.S. Army Institute of Dental Research (USAIDR) in 1983, to automate the process of repetitive chart-by-chart comparison of traditional identification. Both non-dental and dental characteristics are compared to produce lists of probable antemortem records for the submitted postmortem one, thereby permitting forensic science examiners to concentrate his efforts on making positive identifications.

The CAPMI system has been effectively used in actual disaster identification problems (for example, the Gander, Newfoundland, crash of an Arrow Airlines' aircraft in December 1985 [9]). Figure 4 shows an example of an Optical Mark Read (OMR) form which is used for CAPMI. It includes fields of non-dental information as height, weight, eye and hair color, age, sex, and blood type. The CAPMI system uses a combination of the following primary descriptors may be entered: mesial (M), occlusal (O), distal (D), facial (F), lingual (L), missing (X), crown (C), or unerupted (U), as well as additional modifiers, such as pontic (PON), root canal therapy (RCT), or a dental restorative material code, may be added. If a tooth has had no restoration, no entry is made in the data base. CAPMI interprets this "no entry" as a virgin tooth. A list of used codes is given in Table 1.

X	missing	AU	gold
U	unerupted and/or impacted	CO	acrylic or other non-metal
CAR	carious	PG	porcelain or acrylic to metal
ANOM	anomaly	PJ	porcelain jacket
M	mesial	SS	Stainless steel
O	occlusal or incisal	TE	temporary
D	distal	3/4	three-quarter crown
F	facial	PON	pontic
L	lingual	RCT	root canal
C	crown	DEC	deciduous
AM	amalgam	PTD	partial denture

Table 1-Codes for use with OMR form depicted in Figure 4

Figure 4: OMR form for use with the CAPMI system at a disaster scene [9]

The limitations of the CAPMI system are similar to the limitations of any forensic science identification technique. An antemortem database for comparison purposes must both exist and be available. A high degree of accuracy must be assured by dental experts at a step before using the CAPMI process, in filling the required OMR form.

2.1.4 WinID

WinID [10] is representing an extension to the CAPMI system. WinID is also a dental computer system that matches the data of missing persons (filled in special forms) to that of unidentified human remains. WinID makes use of dental and anthropometric characteristics to rank possible matches. Information about restored dental surfaces, physical descriptors, and pathological and anthropologic findings can be entered into the WinID database. WinID stores data in a Microsoft Access Database, and provides extensive data filtering and data sorting capabilities.

WinID codes are an extension of CAPMI codes. A primary codes and up to five secondary codes may be specified. A dash (-) is placed between the primary and secondary codes. Figure 5 shows an example of AM record uded by WinID program, where the dental portion of the data entry form should be filled out by a Board Certified Forensic Odontologist. A list of used codes is given in Table 2.

Primary Codes	Secondary Codes
M - Mesial	A – Annotation
O - Occlusal	B – Deciduous
D - Distal	C – Crown
F - Facial	E – Resin
L - Lingual	G – Gold
I – Incisal	H – Porcelain
U – Unerupted	N - Non-Precious
V – Virgin	P – Pontic
X - Missing	R - Root Canal
J – MissCr MPM	S - Silver Amalgam
/ - No Data	T - Denture Tooth
	Z – Temporary

Table 2-Codes used by WinID

Antemortem Dental Record ID#: _____

Last: _____ First: _____ MI: _____
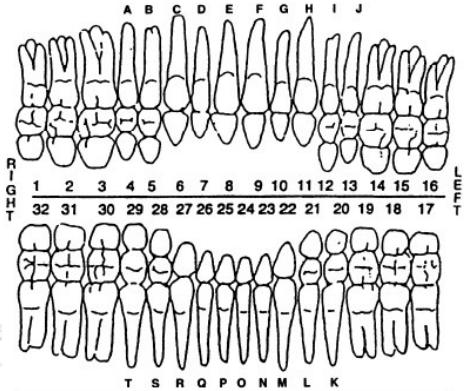
Date: _____ Sex: _____ Race: _____ Age/DOB : _____

Height: _____ Weight: ____ Eye: _____ Hair: _____ Blood Type: _____

Team Member: _____

  Confirm by: _____

Type, Date and Number of X-Rays _____

_____

					Description	Code
1	18					
2	17					
3	16					
4	15	A	55			
5	14	B	54			
6	13	C	53			
7	12	D	52			
8	11	E	51			
9	21	F	61			
10	22	G	62			
11	23	H	63			
12	24	I	64			
13	25	J	65			
14	26					
15	27					
16	28					
17	38					
18	37					
19	36					
20	35	K	75			
21	34	L	74			
22	33	M	73			
23	32	N	72			
24	31	O	71			
25	41	P	81			
26	42	Q	82			
27	43	R	83			
28	44	S	84			
29	45	T	85			
30	46					
31	47					
32	48					

Codes

Primary Codes	Secondary Codes
M – Mesial	A - Annotation
O - Occlusal	B - Deciduous
D – Distal	C - Crown
F – Facial	E - Resin
L – Lingual	G - Gold
I – Incisal	H - Porcelain
U – Unerupted	N - Non-Precious
V – Virgin	P - Pontic
X – Missing	R - Root Canal
J – Missing Cr	S - Silver Amalgam
/ - No Data	T - Denture Tooth
	Z - Temporary

A: _____
B: _____
C: _____

ID As: _____

Comments: _____
_____
_____

Copyright © 2001 James McGivney, DMD

Figure 5: Antemortem dental record used by WinID

With a given pair of antemortem and postmortem records, WinID and CAPMI will each return the same number of dental hits, the same number of dental mismatches and the same number of possible hits. The WinID comparison algorithm was finalized only when it returned the exact values that CAPMI yielded in a given situation. A significant difference between the two systems is that WinID can rank comparisons between records in two more ways than CAPMI. WinID will return a ranked list of *identifier matches*

(defined as matches of non-dental data as race, sex etc.). Lewis in [11] conducted a comparison study between those two principal computer programs are CAPMI4 and WinID2 (version 3 of this program is now available). She ran them on a sample of 100 simulated victims and 105 simulated postmortem fragments. CAPMI4 provided 48 correct matches and WinID2 provided 71 correct matches. In addition, comparisons were made within WinID2 to determine which of its three dental data sets was the most successful for suggesting correct matches.

However, none of these systems provide the desired level of automation, as they require a significant amount of human intervention. For example, in both *CAPMI* and *WinID*® feature extraction, coding, and image comparison are carried-out manually. Moreover, the dental codes used in these systems are entirely based on dental work. Hence, *CAPMI* and *WinID*® are more like sorting tools that help to cut down the time of forensic experts, but not identification systems. While forensic odontologists rely on teeth orientation, type of restorative materials, and radiographic appearance as basis for positive identification. These properties are neither incorporated in *CAPMI* nor in *WinID*®. Thus, the amount of automation offered by these dental identification systems resembles that of an automated fingerprint identification system, whereby a forensic expert is required to identify and classify the minutiae points of fingerprints before the system can produce a list of candidate matches to the subject.

2.2 Automated Dental Identification System (ADIS)

ADIS is required to speedup the postmortem identification process. It is estimated that at any point in time there are over 100,000 unsolved MUP cases in the National Crime Information Center (NCIC), 60 % of which have remained unsolved for 90 days or longer. Technically speaking, this large number of unsolved cases hinders the capabilities of search techniques currently employed. The benefit of ADIS to the society will surpass saving millions of dollars to also include psychological benefits of families of missing persons when knowing the disposition of their loved ones without having to wait extended period of time [19].

In 1997, the Criminal Justice Information Services (CJIS) division of the Federal Bureau of Investigation (FBI) formed a Dental Task Force (DTF), which recommended the creation of:

- Digital Image Repository (DIR): later in May of 2005, (CJIS) Management approved the creation of the National Dental Image/Information Repository (NDIR) to facilitate the identification of Missing, Unidentified, and Wanted persons. The NDIR became available in February 15, 2006 and is housed on Law Enforcement Online (LEO) and is permitting law enforcement agencies to store, access and supplement dental records which were housed in the Missing, Unidentified, and Wanted Persons (MUP) files in the National Crime Information Center (NCIC) system. The NCIC

system is a nationwide computerized database of documented criminal justice information maintained and supported by CJIS. Data in NCIC is exchanged in a shared-management partnership with criminal justice officials of local, state and federal governments in the United States, its possessions, and Canada [12].

- Automated Dental Identification System (ADIS) with goals and objectives similar to the Automated Fingerprint Identification System (AFIS). AFIS later on turned to Integrated AFIS, more commonly known as IAFIS, which is a national fingerprint and criminal history system maintained by the (CJIS) Division. The IAFIS provides automated fingerprint search capabilities, latent searching capability, electronic image storage, and electronic exchange of fingerprints and responses. As a result of submitting fingerprints electronically, agencies receive electronic responses to criminal ten-print fingerprint submissions within two hours and within 24 hours for civil fingerprint submissions [13].

2.2.1   ADIS Overview

The proclaimed dental identification system is fully automated system that when fed with raw subject dental records will find a minimum set of match reference records, ideally one, for each of the subject records. The forensic dental expert can examine the dental records of those few candidates to make a final decision on the identity of the missing or unidentified person [24].

Automated Dental Identification System (ADIS) is comprised of two stages of (1) search and retrieval stage based on potential similarities to produce a candidate list and (2) verification stage for matching based on low level comparison of dental images to produce a short match list (as shown in Figure 6). Some of the challenges behind developing ADIS were described in [23][26].
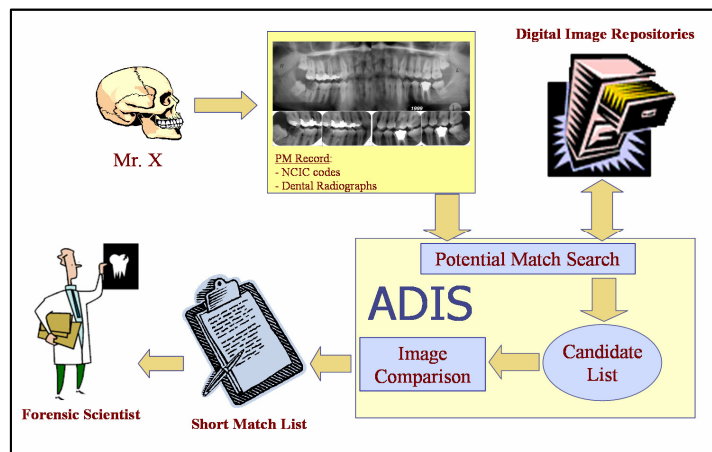
Figure 6: ADIS overview

2.2.2 ADIS Architecture

Nassar and Ammar [22] proposed a prelude research prototype of ADIS, which we later extended and refined - as depicted in Figure 1- to become the conceptual framework of the ADIS prototype [19]. At a high level of abstraction, we view ADIS as a collection of the following components: (i) the Record Preprocessing component, (ii) the Potential Matches Search component, (iii) the Image Comparison component [29][31][32][37].

2.2.2.1 Record Preprocessing

The record pre-processing component handles the following tasks:

(a) Records Cropping [39], which is the first segmentation level, where the dental films are segmented from the input records. This is our first problem and more details are given in section 3.1.

(b) Film Enhancement: A tooth maps to an area with mostly "bright" gray scales (except for the pulp tissue) while the gum maps to areas with "mid-range" gray scales, and air maps to "dark" gray scales. Thus, a "significant" contrast in gray scales within a "small" area of a dental radiograph indicates a transition from one object to another. In order to assist the coming preprocessing steps, Hajsaid et al [20] [28][40] introduces a gray scale contrast stretching to transform poor quality dental film in a way that insures an appreciable degree of contrast between the dominant gray scales used in capturing the different classes of objects; in other words to compensate for possible poor contrast.

(c) Film Type Classification: In the context of ADIS database, there are three types of dental images according to the way they capture the dental features. When taking radiographs of the teeth, the dental arches are divided into different regions to get a clear view of individual teeth. Each film placement is corresponding to a film type, and is designed to record specific teeth. Some commonly used film types are [51]:

Bitewing Film: records two-thirds (the crown portion) of opposing molar and premolar teeth, as shown in Figure 7-a.

Periapical Film: either records teeth of the lower or upper jaw. It records the whole tooth (crown and root). It has different placement to record all teeth. Figure 7-b shows an upper periapical for the incisors and canine region. Periapical films are further classified into upper and lower.

Panoramic Film: records all the teeth at the same time, as shown in Figure 7-c.

Zhou and Abdel-Mottaleb [17][27] classify those types based on the number, location and relationship between extracted vertical and horizontal edges.
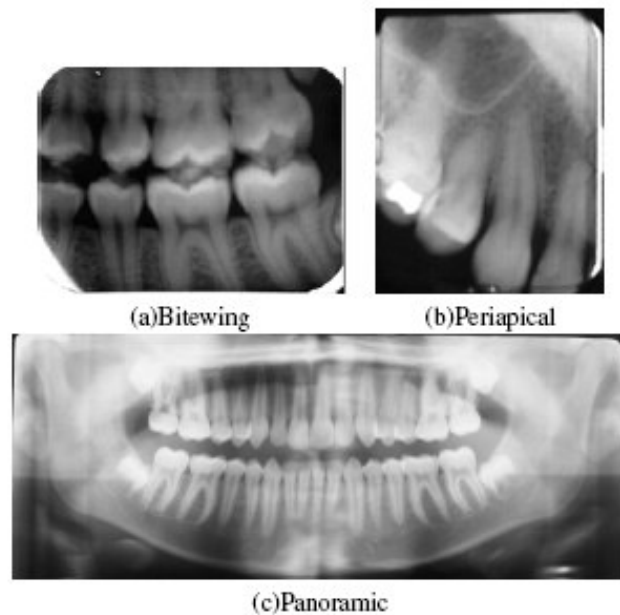
Figure 7: Different dental radiographic films

(d) Detecting whether the film is vertically flipped or not, Quereshi [46] usng a classifier based on molar root shape and jaw texture analysis.

(e) Segmentation of teeth regions from films, which is a second level of segmentation (as shown in Figure 8). This is a critical step for the remaining pre-processing steps that is why we had many realizations for it. Jain and Chen [14] proposed a semi-automated method based on integral projection and can be used with bitewing and panoramic views. Nomir and Abdel-Mottaleb [16] proposed an automated method based on iterative adaptive threshold and then integral projection, however it can be used with bitewing views only. Zhou and Abdel-Mottaleb [17] used automated method based on mathematical morphology followed by integral projection with the same limitation of handling bitewing views only. Hajsaid et al [20] proposed automated method based on mathematical morphology, which can handle both bitewing and periapical views. Nassar et al. [34] introduced performance evaluation criteria for different segmentation methods.

Figure 8: Tooth region segmentation

(f) Teeth contour extraction: the third segmentation level, where the target is to segment each tooth from the isolated region, this is recognized as teeth contour extraction [43]. This is our second problem and more details are given in section 3.2.

(g) Annotating teeth with labels that is corresponding their location. This is our third problem and more details are given in section 3.3.

2.2.2.2 Potential Match Search

This component manages archiving and retrieval of dental records based on high-level dental features (e.g. teeth intensity and shape descriptors) and produces a candidate list. The discriminative feature of this search engine is being very fast so as to be able to search a large database, with optimum accuracy. The forensic experts used to use dental and non-dental feature to reduce their search domain; in other words, they are implementing a similar technique to potential match search component, by using these features to reduce their search space. They used non-dental features as (sex, race, age, etc.) and dental features as presence/ absence of dental work and missing teeth.

In ADIS implementation, we allow the dental experts to use the same set of non-dental feature as (sex, race, age, etc.) to reduce the database, and then we use the potential match search component to search the reduced database to produce a relatively short candidate list. The Image Comparison component then handles another level of dental image comparison against this candidate list to produce a match list. The current in use image comparison component takes about 10 minutes to make record to record comparison. To search a database of 50 records directly using the image comparison, the system will run for about 8 hours. That is why we usually need this kind of fast image retrieving component that can reduce the database to a short candidate list. Implementing a fast and accurate potential match search is our fourth problem and more details are given in chapter 4.

2.2.2.3 Image Comparison

The Image Comparison component mounts for teeth alignment and comparison between subject and the corresponding teeth of each candidate, thus producing a short match list. Figure 9 illustrates the main steps of this component. In ROI pair selection step, guided by the output of the teeth-labeling step of the record-preprocessing component, we select corresponding segmented teeth pair (regions of interest "ROI") for subject and candidate records. In the alignment step, we conduct pair-wise region of interest (ROI) alignment. More details about alignment in section 4.2. We start with a hypothesis that the two objects (ROI) are matched, and accordingly we apply appropriate transformations that restore major geometric discrepancies between them. In the low-level feature extraction step, we use a set of nonlinear filters to map an the ROI in the corresponding feature spaces.

In the micro-decision step [21], we use Bayesian classification layer that computes the posterior probability of match between a pair of ROI's using the differences between spatially corresponding features of the ROI-pair. As a dental record usually comprises multiple films that may show more than a single view of a given tooth, we refer to the process of determining the match status of a subject/reference tooth-pair based on comparison of multiple views of this tooth-pair as micro-decision making.

Finally in the macro-decision step, with many micro-decisions, we combine these micro-decisions into a macro-decision that determines the match status of the subject/candidate record-pair and accordingly whether the candidate record should be placed on the match list. We also provide a ranking score to sort the match list.
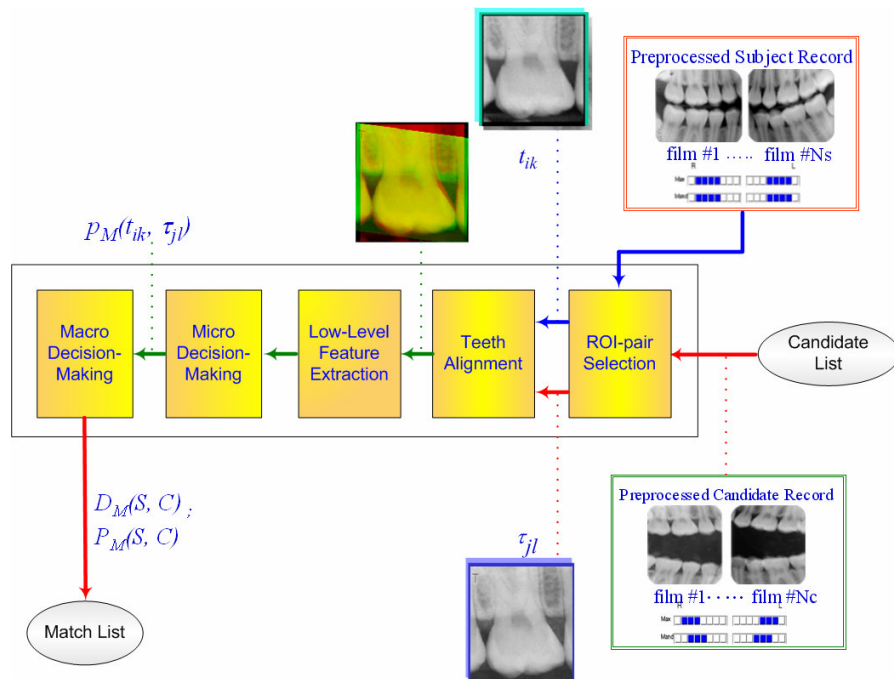
Figure 9: A block diagram of the image comparison component

2.2.3   webADIS Architecture

In developing an environment that integrates the various ADIS components, Chekuri et al. [38][42][45] adopted the 3-tier architecture. In the 3-tier architecture, volumes of data are maintained at a database layer; business rules (and logic) are processed (and validated) at the server tier, which also handles transactions with the client layer (the front end). Figure 10 illustrates the 3-tier architecture of webADIS, where a web browser component at the client tier communicates with a web server component at the middle tier, a central component (ADIS Server) handles transaction dispatching to the respective server components as well as to the database tier as needed.
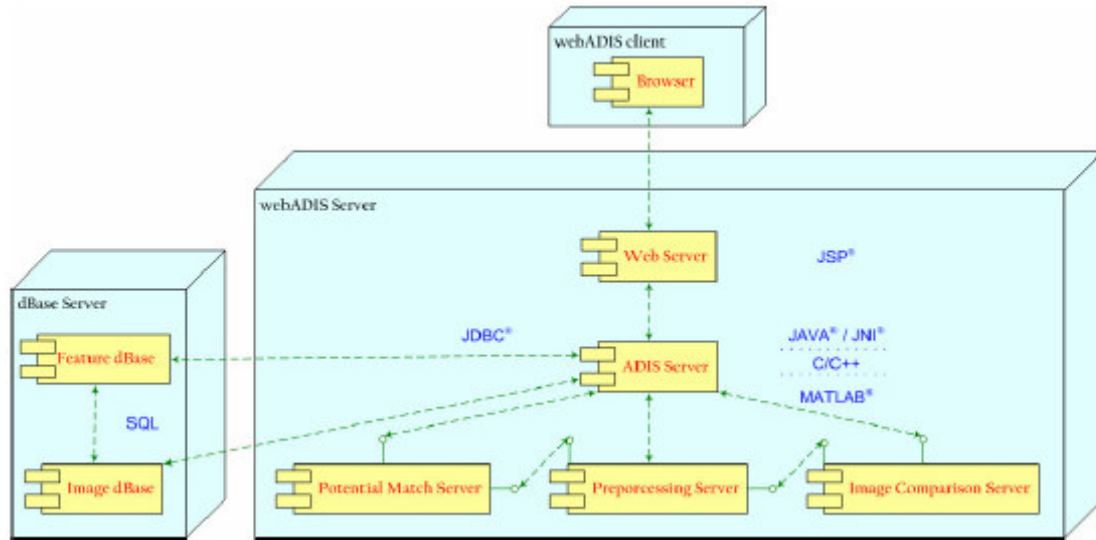
Figure 10: Architecture of webADIS; the web accessible ADIS environment

There are several benefits for using the 3-tier architecture in developing an environment like webADIS:

- First, the middle tier provides a layer of abstraction and hence client applications do not depend on the type of database server that actually stores the data. So any changes made to the database-tier would only require changes in middle tier.

- Second, a component in the business layer can be accessed by any number of components in presentation layer (clients), therefore any component implementation changes made at the business-tier will not mandate changes in components of other two tiers.

- Third, this decoupling of tiers -according to roles facilitates dynamic load balancing of servers, thus a server shares its load with other servers at times of bottlenecks.

Finally, 3-tier implementations can be designed to make business objects and data storage so close (in a network sense) that the network load is eliminated, thus reducing inefficient utilization of network bandwidth.

2.3 Ear Identification System

The Iannarelli System has been in use for more than 40 years [116], and is based upon the 12 measurements illustrated in Figure 11. Each photograph is aligned during development so that the lower tip of a standardized vertical guide on the development easel touches the upper flesh line of the cocha area whilst the upper tip touches the outline of the antitragus.

This method has the capability of identifying about 17 million different persons. Assuming an average standard deviation in the population of four units, and then the measurements provide for a space of (4^{12} =~17 million) distinct points. With the current automation, we can reach better measurement accuracy, so we can assume the average standard deviation in the population of ten units, then the measurements will provide a space of (10^{12} =~ Tera) distinct points.
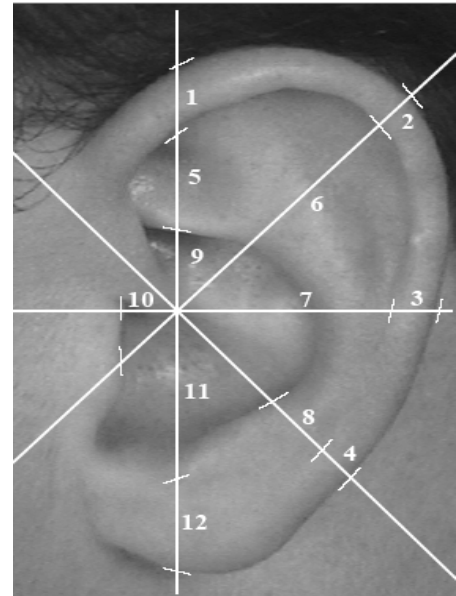
Figure 11: Iannarelli measurement [116]

Chang et al. have built multi-modal recognition using both the ear and face. Their research was based on the concept of Eigen faces and Eigen ears and used a principal component analysis (PCA) evaluation protocol. They reported a performance of 69.3% for face versus 72.7% for ear, in one experiment, compared to 90.9% for the multi-modal system [110].

Burge and Burger have researched automating ear biometrics; they used a 300 by 500 grayscale image of the subject's head in profile using a CCD camera. They located the ear by using deformable contours on a Gaussian pyramid representation of the image gradient. Then they made Edge extraction, Curve extraction and finally they build a graph model. They used a novel graph matching based algorithm for authentication, which takes into account the possible error curves [108]

Hurley et al. have used force field transformations for ear recognition. The image is treated as an array of Gaussian attractors that act as the source of the force field. The directional properties of that force field are exploited to automatically locate a small number of potential energy wells and channels. This feature extraction technique is used to form an ear biometric system, which was found to be robust and reliable [110][111] the reported a very high recognition rate of 98.4%.

Zhang et al. built a hybrid system for ear recognition purpose [112]. This system combines Independent Component Analysis (ICA) and Radial Basis Function (RBF) network. The original ear image database is decomposed into linear combinations of several basic images. Then the corresponding coefficients of

these combinations are fed up into RBF network instead of an original feature vector comprised of pixel values of grayscale images. The local features extraction of ICA and the adaptability of RBF neural network are then combined.

Abdel-Mottaleb and Zhou introduced a novel system for ear identification based on the face profile images [113]. The system consists of two major steps: 1) Ear detection from face profile based on template matching, where they used to module the ear by its external curve; 2) 2-D feature extraction based on force field transformation, then recognition based on these features.

Chen and Bhanu also used model-based (template matching) based technique for Ear-detection [114]. The model template is represented by an averaged histogram of shape index. The detection is a four-step process: step edge detection and thresholding, image dilation, connect component labeling and template matching.

Chen and Bhanu then developed another shape model-based technique for locating human ears in side face range images [115]; where the ear shape model is represented by a set of discrete 3D vertices corresponding to ear helix and anti-helix parts. They started by locating the edge segments and grouped them into different clusters which are potential regions containing ears. For each cluster, they register the ear shape model with the edges. The region with the minimum mean registration error is declared as the detected ear region.

The ear can be pressed against some material, e.g. glass, and the 'earmark' can be used as a biometric. This has been used in crime solving. In England four delinquents have been judged between the year1996 and 1998, by using only the earmarks. However in the Netherlands the court decided that the earmarks are not reliable enough for judging [109].

Cases of the ear being only partially occluded by hair, it is possible to recognize the hair and segment it out of the image. This can be done using thermogram images. A thermogram image is one in which the surface heat of the subject is used to form an image. Figure 12 is a thermogram of the external ear. The subject's hair in this case has an ambient temperature between 27.2°c and 29.7°c, while the external anatomy of the ear ranges from 30.0°c to 37.2°c. Removing hair is done by segmenting out the low temperature areas [108].

Figure 12: Ear thermogram[117]

In chapter 5, we present novel approaches for automation of Iannarelli ear identification system. This is our fifth and last problem in this dissertation.

2.4   Content-based image retrieval

To build ADIS potential match search component, we are using Eigen image (color, or appearance-based feature) of segmented tooth, as well as Edge direction histogram and Moment Invariant (shape-based features) of tooth contour for dental image retrieval. This section is reviewing the literature of image content retrieval in general, while more details on the features used for ADIS potential match search component are given in chapter 4.

Content-based image retrieval has been an active research area since the early 1990's. Many image retrieval systems, both commercial and research, had been built. The image features used by these systems can be generally classified as: color, texture, shape, spatial layout, and faces (or other specific object) [83].

2.4.1   Color Features

Color features are often easily obtained or calculated directly from the pixel intensities. Different color spaces as RGB (Red Green Blue) and HSV (Hue Saturation and Value) can be used. Example of color features are:

A) Global color histogram over the whole image:

An image histogram refers to the probability mass function of the image intensities. This is extended for color images to capture the joint probabilities of the intensities of the three color channels. Computationally, the color histogram is formed by discretizing the colors within an image and counting the number of pixels of each color.

B) Fixed subimages color histogram over fixed parts of the image:

Each image is divided in cells (sub-images) and for each cell a color histogram is calculated.

Ex.: FOCUS (Fast Object Color-based qUery System), where each image is divided in cells of 100x100 pixels and for each cell a color histogram in the HSV space, coarsely quantized along the saturation and value axes, is computed. The peaks of all local histograms are determined and combined in a list of unique peaks for the whole image by merging multiple copies of the same peak.

C) Region color histogram over a segmented region:

Image is divided into regions with associated color and texture. In other word specific regions are segmented and color histograms are calculated for these regions.

Ex. blobworld system is based on finding coherent image regions which roughly correspond to objects. Each image is segmented into regions by fitting a mixture of Gaussians to the pixel distribution in a joint color-texture-position feature space. Each region ("blob") is then associated with color and texture

descriptors. To match the color of two regions, the quadratic distance between their histogram is computed in such a way to give high score to two regions with similar colors, even if the colors fall in different histogram bins.

D) Spatial chromaticity histogram:

Bin i gives the ratio h(i) of pixels having color i, the relative coordinates of the baricenter b(i) of these pixels, and the standard deviation σ(i) of these pixels' distance to the baricenter.

E) Correlation histogram:

In the context of MIR (Multimodal Information Retrieval System), for the scenery pictures (where no face is detected) a color correlation histogram is computed. If {i, j, k} is a color triple in the 2 x 2 x 2 quantized HSV space, then $h_{ijk}`$ denotes the number of occurrences of three pixels of these colors as the vertices of a isosceles right triangle with the smaller sides of length $l$. The color correlation histogram $f_{ijk}`$ is $h_{ijk}`/$ $(4*h_i)$, where h_i is the ith bin of the traditional color histogram.

F) Farthest neighbor histogram:

The rationale behind this feature is to capture how abrupt the color changes are in an image. In photographs they tend to be smoother than in drawings. Therefore, drawings tend to score higher than photographs in this test.

Ex. WebSeer system picks a threshold T between 0 and 255 and counters C and S are initialized to 0. For each inner pixel P of the image, top, bottom, left, and right neighbor. For each neighbor, the absolute difference of its R value is calculated from the R value of P, and the maximum M of those four absolute differences is chosen. If M > 0, increase S. If, in addition, M > T, increase C. The score of the image is C/S.

G) Eigen image:

This representation offers a description relative to an average of a few prototypes by using the eigenvectors of a covariance matrix as an orthogonal coordinate system of the image space. First a pre-processing step is needed in order to normalize the input image for position, scale and orientation. In representing a new image region, only eigenvectors corresponding to the largest eigenvalues are used, in other word the image is projected into a lower dimensional space.

Eigen image is the color (appearance-based) feature that we use in implementation of the Potential Match Search Component; more details are given in chapter 4.

H) Color moments:

The image is divided into subimages (or regions) and within each subimage, the first moment (average color) is computed. Higher order moments can also be used.

Ex. ASSERT system uses five grey level features with each pathological bearing region (PBR) in images. These features are the grey level mean and standard deviation, and bins 6 , 9, and 16 from a 16-bin histogram.

I) Dominant colors:

This is represented by the most frequent color (MFC) vector.

Ex. C-Bird system uses a feature descriptor and a layout descriptor. A feature descriptor is a set of four vectors: a color vector, a most frequent color (MFC) vector, a most frequent orientation (MFO) vector, and a chromaticity vector. A 512-bin RGB histogram is stored in the color vector. The centroides of the regions associated with the 5 most frequent colors form the MFC vector and the centroides of regions of the 5 most frequent edge orientations form the MFO vector. The 36-dimensional chromaticity vector is then computed.

J) Color coherence vector:

Color-coherence vector (CCV) is describing the color content of an image. CCV is made of coherence pairs, each pair giving the number of coherent and incoherent pixels of a particular discretized color in the color space.

Ex. PicHunter system incorporates some redundancy information of color spatial distribution along with conventional color histogram. Besides a 64-bin HSV histogram, two other vectors – a 256-length HSV color auto-correllogram (CORR) and a 128-length RGB color-coherence vector (CCV) – are describing the color content of an image. CCV is made of 64 coherence pairs, each pair giving the number of coherent and incoherent pixels of a particular discretized color in the RGB space. In order to classify pixels in one of the two categories, the image is first blurred slightly by replacing each pixel value with the average value of the pixels in its 3 x 3 neighborhood. Then, pixels are grouped in connected components. A pixel will be classified as coherent if the connected component it belongs to is larger than a fixed threshold.

K) Number of regions with that color:

The colors of each image are quantized into number of colors bins. Values are associated with each color bin: the percentage of the image with colors in that bin, and the number of dots of that color found.

Ex. BDLP system quantizes the colors of each image into 13 colors bins. Six values are associated with each color bin: the percentage of the image with colors in that bin, and the number of 'very small', 'small', 'medium', 'large', and 'very large' dots of that color found.

L) Greyness and Background proportion:

Greyness is the percentage of gray valued pixels to the total number of pixels.

The background proportion is the percentage of the background in the total number of pixels, where the supposed background is the bin with the highest frequency in the hue or the gray value histogram.

Ex. ImageRETRO system clusters the images into 15-bin hue histogram, and then calculates many color features including greyness and background proportions.

M) Color channel variance:

This feature represents variances in each of the color channels of segmented regions.

Ex. Shoebox system uses extracted features as the average color in HSV color space, and the variances in each of the color channels of regions. The regions are either resulting from color segmentation, or a fixed partioning of the image, or the overlay of both.

2.4.2 Shape Features

There is no universal definition of what shape is either. Impressions of shape can be conveyed by color or intensity patterns, or texture, from which a geometrical representation can be derived. This is shown already in Plato's work Meno [84]. Here, we are considering shape as something geometrical. Therefore, I am considering for example edge orientation over all pixels as texture, but edge orientation at only region contours as shape information. Shape descriptors are diverse [85]. Examples of shape are (Template matching, Elastic models, Radius vector, Elementary descriptors, Curvature scale space, Bounding box/ ellipse, Generalized Hough transform, Number of regions, number of holes, Angles between edges, and cross ratios of them, Convex parts), as well as:

A) Edge direction histogram

Edge direction histogram is captures the general shape information in the image. Because the feature is local, it is robust to partial occlusion and local disturbance in the image. The major disadvantage is that two perceptually very different images may have similar edge histogram. A histogram of the edge directions is used to represent the shape attribute. The edge information contained in the database images is generated in the preprocessing stage using edge operator (as Canny [64] edge Sobel edge operators). A histogram intersection technique is used for shape-based retrieval; more details are given in chapter 4.

B) Fourier descriptors

Fourier descriptors (FD) method is one of the most widely used shape description methods. FDs are obtained by applying Fourier transform on shape boundary, the Fourier transformed coefficients are called the Fourier descriptors of the shape. By considering the contour of a 2D object as a closed sequence of successive boundary pixels (xs, ys), where $0 \leq s \leq N-1$ and N is the total number of pixels on the boundary. Then three types of contour representations, i.e., curvature, centroid distance, and complex coordinate function, can be defined. The Fourier transforms of these three types of contour representations generate three sets of complex coefficients, representing the shape of an object in the frequency domain [86]. FD is backed by the well-developed and well-understood Fourier theory. The advantages of FD over many other shape descriptors are (i) simple to compute; (ii) each descriptor has specific physical meaning; (iii) simple to do normalization, making shape matching a simple task [87]. Other advantages of FD method include easy normalization and information preserving [86]. With FD, global shape features are captured by the first few low frequency terms, while higher frequency terms capture finer features of the shape. The main feature is that this method is in variant to translation, rotation, scale, and starting point. The disadvantage of Fourier descriptors is local features cannot be located, because in Fourier transform only the magnitudes of the frequencies are known, and not the location [87]. Rui et al. [88] proposed a distance measurement to classify similarity transformed characters using Fourier transformed coefficients. This distance measurement is the weighting sum of the variance of magnitude ratios and the variance of phase difference between two sets of Fourier coefficients. The Fourier coefficients are derived from Fourier reconstructed shape boundary rather than from original boundary. This is not different from FD derived from a smoothed boundary.

C) Moment Invariants

The first significant work considering moments for pattern recognition was performed by Hu [89]. Based on these moments, Hu derived a set of seven moment invariants to translation, rotation, and scale. These moments are combinations of second and third order normalized central moments. We are using Hu's seven moments as tooth contour (shape-based) descriptors for dental image retrieval; more details are given in chapter 4.

D) Tensor Scale Descriptor

Tensor Scale Descriptor (TSD) exploits the notion of local structure thickness, orientation, and anisotropy as represented by the largest ellipse centered at each image pixel and within the same homogeneous region. Miranda et al. [90] proposed a method for TSD that uses the normalized histogram of the local orientation (the angle of the ellipse) at regions of high anisotropy and thickness within a certain interval.

They showed by experiment that TSD is invariant to translational, rotation and to some reasonable level of scale changes. TSDs are compared by using a correlation-based distance function [91].

2.4.3 Texture Features

Although a precise definition of texture is untraceable, the notion of texture generally refers to the presence of a spatial pattern that has some properties of homogeneity. In particular, the homogeneity cannot result from the presence of only a single color in the regions, but requires interaction of various colors. Examples of texture features are (Atomic texture features, Random field decomposition, Local binary patterns (LBP), Edge statistics, Histogram of polynomial of local intensities, Wavelet, Gabor, (Multi-scale) Intensity surface derivatives.

2.4.4 Layout Features and Faces (or other specific object)

Layout features is about the absolute or relative position of color, texture, or shape information. Higher level features are increasingly more specific, and thus less widely used.

Faces are frequently present in photo pictures and relatively often used as a feature [83]. This can be extended to the detection of object that can specify the picture nature. Automatic object recognition is needed, but most successful computer vision object recognition systems can only handle particular objects, such as industrial parts, that can be represented by precise geometric models. Content-based retrieval requires the recognition of generic classes of objects and concepts. A limited amount of work has been done in this respect, but no general methodology has yet emerged [92].

2.5 Multimodal Biometrics

ADIS can be described as a multi-biometric system that may use one or more instances of a single identifier, as well as more than one biometric identifier (as dental radiograph includes several teeth and sometimes the same tooth presences in more than one film of the same record). The future of the AEIS system is to utilize more instances of multiple biometric identifiers (e.g. face and ear image). In implementing the potential match search (Chapter 4), we need to move from tooth-to-tooth level matching score to record-to-record matching score. This section is reviewing the literature of multimodal biometrics specifically on how to fuse multiple classifiers in order to reach an overall score.

Based on the inputs, a biometric system can be classified as [100]:

- Uni-biometric: is a system that uses only single biometric identifier i.e., it relies on the evidence of a single source of information for authentication (e.g., single fingerprint, face, or single ear)

- Uni-modal biometric system: is a subset of Uni-biometric system that uses only single biometric identifier, single representation, single classifier and single matcher.

- Multi-biometric: is a system that uses more than one biometric identifier i.e., it relies on the evidence from independent or weakly correlated identifiers (e.g., fingerprint from different fingers, the two ears, or the different teeth images)

- Multi-modal biometric system: is a subset of Multi-biometric system that uses more than one correlated biometric measurement (e.g., multiple impression of the same finger, video sequence for the same ear, or different images of the same tooth), multiple classifiers, multiple matchers, or any combination thereof.

Most biometric systems deployed in real-world applications are Uni-modal, These systems have to contend with a variety of problems such as [93]:

(a) Noise in sensed data: A fingerprint image with a scar, or a voice sample altered by cold are examples of noisy data. Noisy data could also result from defective or improperly maintained sensors (e.g., accumulation of dirt on a fingerprint sensor), unfavorable ambient conditions (e.g., poor illumination of a user's face in a face recognition system), or poor quality of dental radio-graphic film.

(b) Intra-class variations: These variations are typically caused by a user who is incorrectly interacting with the sensor (e.g., incorrect facial pose), or different radio-graphic film orientation (e.g. the postmortem film is taken at a different angle from the antemortem one).

(c) Inter-class similarities: In a biometric system comprising of a large number of users, there may be inter-class similarities (overlap) in the feature space of multiple users. This is a typical case in dental feature space.

(d) Non-universality: The biometric system may not be able to acquire meaningful biometric data from a subset of users. A fingerprint biometric system, for example, may extract incorrect minutiae features from the fingerprints of certain individuals, due to the poor quality of the ridges. For dental identification system, we are facing the same problem for cases missing segments.

Some of the limitations imposed by uni-modal biometric systems can be overcome by including multiple sources of information for establishing identity; as incase of multi-biometric systems, which are expected to be more reliable due to the presence of multiple, (fairly) independent pieces of evidence [95]. These systems are able to meet the stringent performance requirements imposed by various applications. Maltoni et al. [100] presented an example of how the system performance enhanced by considering multi-classifier system:

Considering a two class classification problem (in the context of ADIS molar/ premolar), and an N classifier system (Eigen Images, teeth contour descriptors, and teeth texture descriptors); and majority vote between the classifier of whether the teeth belongs to molar class or premolar class. If p is the probability that a single classifier correctly classify the input tooth, then the probability of correct classification for the multi-classifier system is:

$$P_c(N) = \sum_{m=k}^{N} \binom{N}{m} p^m (1-p)^{N-m}$$

2.5.1 Integration Strategies

For a generic biometric system, it generally has four important modules [99]:

- o Sensor module which captures the *trait* in the form of raw biometric *data*;

- o Feature extraction module which processes the data to extract a *feature set* that is a compact representation of the trait;

- o Matching module which employs a classifier to compare the extracted feature set with the templates residing in the database to generate matching scores;

- o Decision module which uses the matching scores to either determine an identity or validate a claimed identity.

In multi-biometric system information integration/fusion can occur at any of the following levels [97] (as shown in **Figure 13**)

- (a) At the data (sensor level) or feature level: Either the data itself or the feature sets originating from multiple sensors/sources are fused. As an example, for ADIS, fusion at the feature level can occur if we fuse the different Eigen teeth weights to come with one general feature vector to represent the whole record. Method like Weighted Summation and Concatenation can be used for this fusion [98].

- (b) At the matches score level: The scores generated by multiple classifiers pertaining to different modalities are combined. As an example, the match scores generated by the face and ear modalities of a user may be combined via the simple min rule in order to obtain a new match score which is then used to make the final decision. The previous example architecturally is considered as parallel fusion. Other architecture combinations are serial fusion (cascading), or hieratical (treelike) [100].

(c) At the rank level: This type of fusion is relevant in identification systems where each classifier associates a rank with every enrolled identity (a higher rank indicating a good match). Thus, fusion entails consolidating the multiple ranks associated with an identity and determining a new rank that would aid in establishing the final decision. As an example, for ADIS, we can let each tooth give an identity rank, and then fuse the resulted ranks to come up with the whole record rank [97].

(d) Fusion at the decision level: The final output of multiple classifiers are consolidated via techniques such as majority voting, behavior knowledge space, etc…[96].
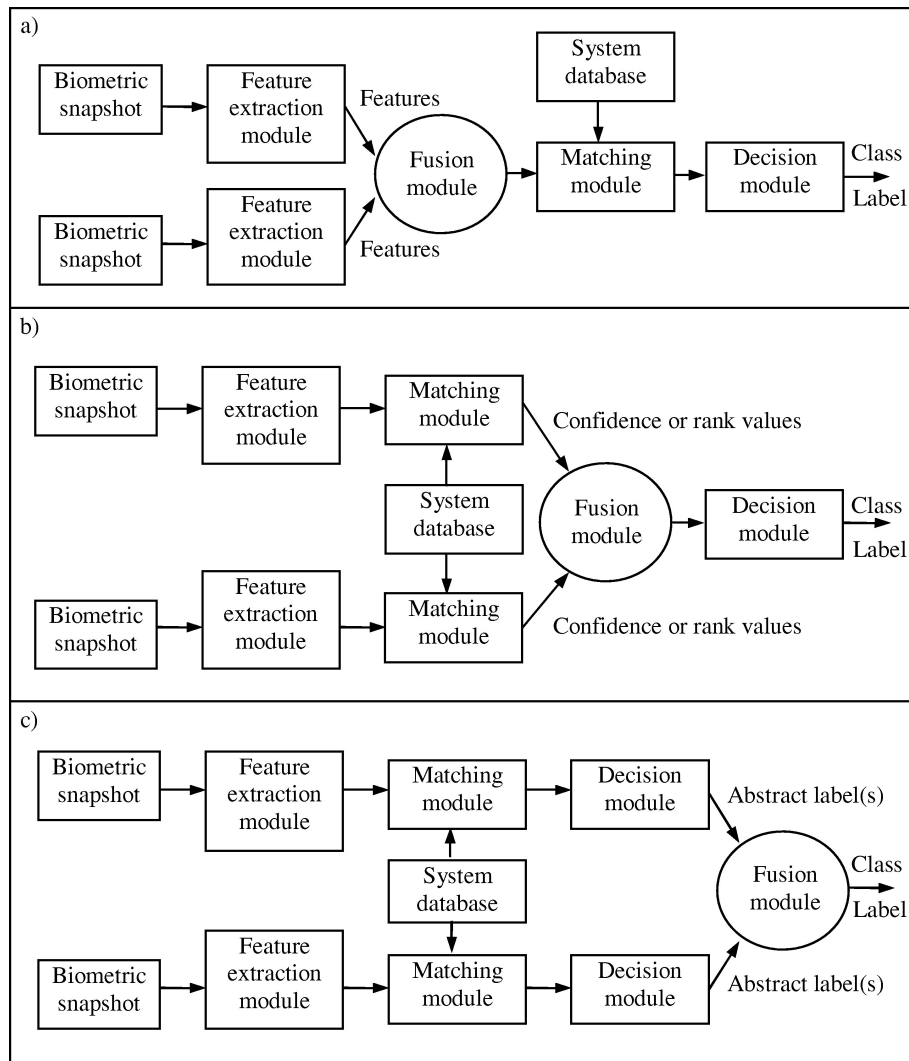
Figure 13: Different levels of fusion: (a) fusion at the feature level, and (b) fusion at matching score (confidence or rank) level, and (c) fusion at decision (abstract label) level, taken from [99].

Biometric systems that integrate information at an early stage of processing are believed to be more effective than those systems which perform integration at a later stage. Since the feature set contains richer

information about the input biometric data than the matching score or the output decision of a matcher, fusion at the feature level is expected to provide better recognition results. However, fusion at this level is difficult to achieve in practice because the feature sets of the various modalities may not be compatible [94] (e.g., eigen-coefficients of teeth and contour descriptor coefficients). Fusion at the decision level is considered to be rigid due to the availability of limited information. Thus, fusion at the match score level is usually preferred, as it is relatively easy to access and combine the scores presented by the different modalities.

2.5.2 Combination Strategies

Kittler et al. [101] developed a common theoretical framework for classifier combination and show that many existing schemes can be considered as special cases of compound classification. They demonstrated the commonly used classifier combination schemes such as the product rule, sum rule, min rule, max rule, median rule, and majority voting. They formulated the problem as follows:

Consider a pattern recognition problem where:

- pattern Z is to be assigned a class

- m possible classes $(w_1,..,w_m)$

- R classifiers each representing the given pattern by a distinct measurement vector

- measurement vector used by the ith classifier is x_i, where i=1…R

- each class w_k is modeled by the probability density function $P(x_i \mid \omega_k)$

- each class w_k *has* a priori probability of occurrence is denoted $P(\omega_k)$

Using *Bayes* decision rule, the class of novel pattern Z whose measurements are: $\{x_i , i = 1, .., R\}$ should be assigned to class w_j provided the a posteriori probability of that interpretation is maximum if:

$$Z \rightarrow \omega_j \mid p(\omega_j \mid x_1, .., x_R) = \max_{<k>} \ p(\omega_k \mid x_1, .., x_R) \tag{2.1}$$

The Product rule: assuming the representations used are conditionally statistically independent, then

$$p(x_1,...,x_R|\omega_k)\prod_{i=1}^{R} p(\mathbf{x}_i|\omega_k),$$

$$p(\omega_k \mid \mathbf{X}_1 .. \mathbf{X}_R) = \frac{p(\omega_k)\prod_{i=1}^{R} p(\mathbf{x}_i \mid \omega_k)}{\sum_{j=1}^{m} p(\omega_j)\prod_{i=1}^{R} p(\mathbf{x}_R \mid \omega_j)}$$

thus accordingly *Bayes* decision rule:

$$Z \longrightarrow \omega_j \Big| p(\omega_j) \cdot \prod_{i=1}^{R} p(\mathbf{x}_i | \omega_j) = \max_{<k>} p(\omega_k) \cdot \prod_{i=1}^{R} p(\mathbf{x}_i | \omega_k) \tag{2.2}$$

The Sum rule: assuming that the a posteriori probabilities computed by the respective classifiers will not deviate dramatically from the prior probabilities. Thus, $p(\omega_k | x_i) = p(\omega_k)(1 + \delta_{ki})$, substituting $p(\omega_k | x_i)$ in the product rule (2) and expanding the product then neglecting any terms of second and higher order:

$$Z_{\Sigma} \rightarrow \omega_j \Big| (1\text{-}R) p(\omega_j) + \sum_{i=1}^{R} p(\omega_j | \mathbf{x}_i) = \max_{<k>} \{ (1\text{-}R) p(\omega_k) + \sum_{i=1}^{R} p(\omega_k | \mathbf{x}_i) \} \tag{2.3}$$

The decision rules, product rule (2) and sum rule (3), constitute the basic schemes for classifier combination.

The Max rule: starting with the sum rule (3), if approximating the sum of posteriors using their maximum:

$$Z_{max} \rightarrow \omega_j \Big| (1\text{-}R) p(\omega_j) + R \max_{<i>} p(\omega_j | \mathbf{x}_i) = \max_{<k>} \{ (1\text{-}R) p(\omega_k) + R \max_{<i>} p(\omega_k | \mathbf{x}_i) \} \tag{2.4}$$

Under the assumption of equal priors:

$$Z_{max} \rightarrow \omega_j \Big| \max_{<i>} p(\omega_j | \mathbf{x}_i) = \max_{<k>} \{ \max_{<i>} p(\omega_j | \mathbf{x}_i) \}$$

The Min rule: starting with the product rule (2), and bounding the product of posterior probabilities:

$$Z_{min} \rightarrow \omega_j \Big| (p(\omega_j))^{-(R\text{-}1)} + R \min_{<i>} p(\omega_j | \mathbf{x}_i) = \max_{<k>} \{ (p(\omega_k))^{-(R\text{-}1)} + R \min_{<i>} p(\omega_k | \mathbf{x}_i) \} \tag{2.5}$$

under the assumption of equal priors:

$$Z_{min} \rightarrow \omega_j \Big| \min_{<i>} p(\omega_j | \mathbf{x}_i) = \max_{<k>} \{ \min_{<i>} p(\omega_j | \mathbf{x}_i) \}$$

The Majority vote rule: starting by the sum rule (3) and by hardening the posteriors according to

$$p(\omega_k | x_i) = \Delta_{ki} = \begin{cases} 1 & \text{if } p(\omega_k | \mathbf{x}_i) = \max_{<j>} p(\omega_j | \mathbf{x}_i) \\ 0 & \text{otherwise} \end{cases} :$$

$$Z_{Majority} \rightarrow \omega_j \Big| \sum_{i=1}^{R} \Delta_{ji} = \max_{<k>} \{ \sum_{i=1}^{R} \Delta_{ki} \} \tag{2.6}$$

All the above combination strategies and their relationships are represented in Figure 14.
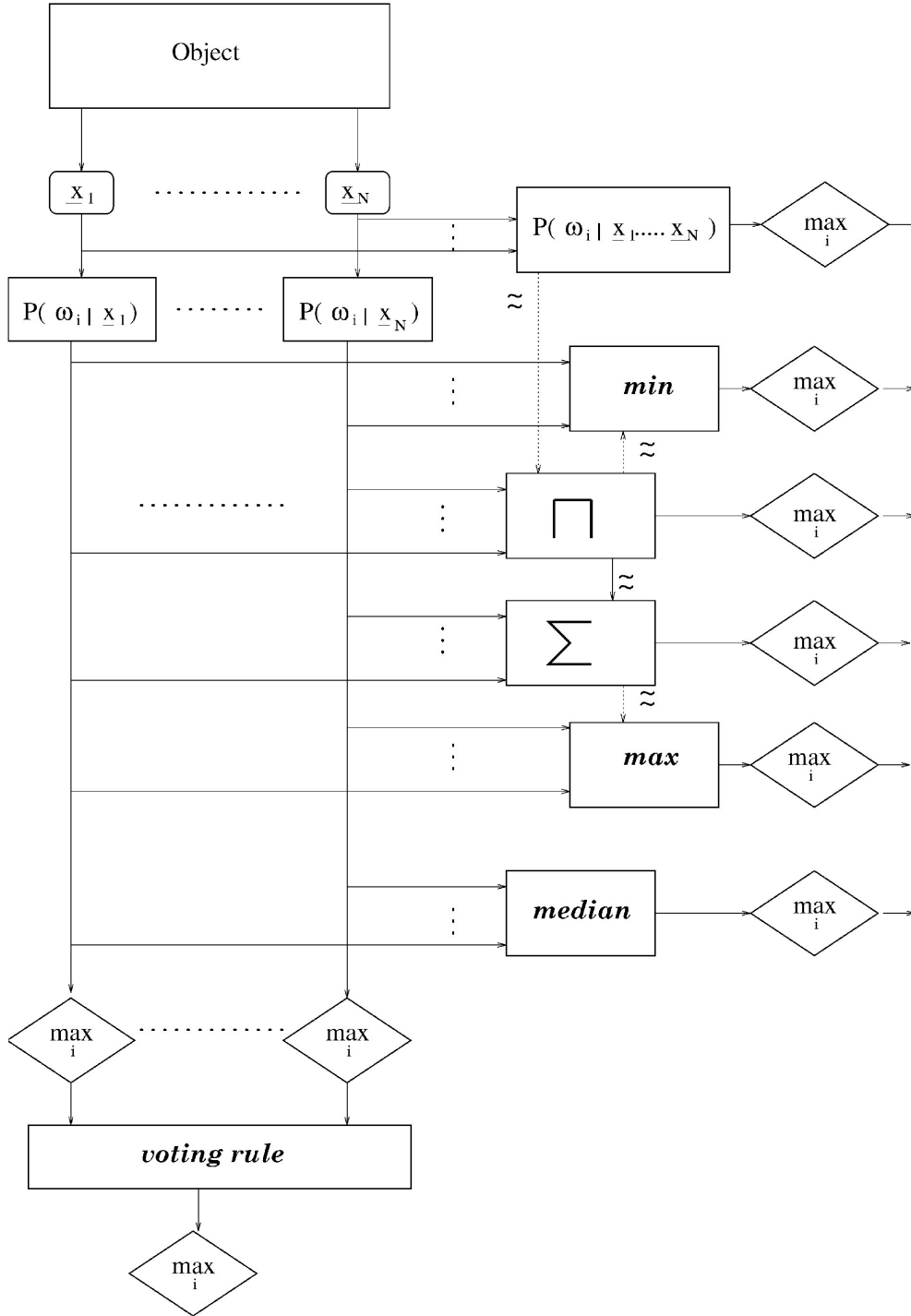
Figure 14: Classifier combination schemes, taken from [101].

In conclusion, a large number of combination strategies are available. According to Maltino et al.[100], it is not well understood which one of them will perform better, and if so under what circumstances. Traditionally, selection of a particular combination is based on experimental assessment of a set of combination strategies and schemes.

3 Different Preprocessing Techniques for ADIS

3.1 Cropping (Global Segmentation)

Dental radiographic films of an individual are usually composed into a digital image record. In order to achieve high level of automation in postmortem identification it is necessary to decompose dental image records into their constituent radiographic films, which are in turn segmented to localize dental regions of interest. In this section an automatic hierarchical treatment to the problem of cropping dental image records into films is proposed. This proposed approach is heavily based on concepts of mathematical morphology and shape analysis. Among the many challenges that face automatic cropping are non-standard assortments of films into records, variability in record digitization as well as randomness of record background both in intensity and texture.

3.1.1 Introduction

Though performing the film cropping task may seem trivial for a human observer, it is desirable to automate this process and to integrate it with the prototype ADIS system. In this section, we focus on the global segmentation (cropping) problem of dental X-ray records and seek a solution to achieve a good tradeoff between accuracy and complexity. On one hand, we want segmentation results to be as accurate as possible since inaccuracy in cropping of dental records is likely to hinder the performance of subsequent processing steps and accordingly the overall performance of the entire identification system. On the other hand, we want the computational cost to be reasonably low, especially with large volume of records that need to be processed.
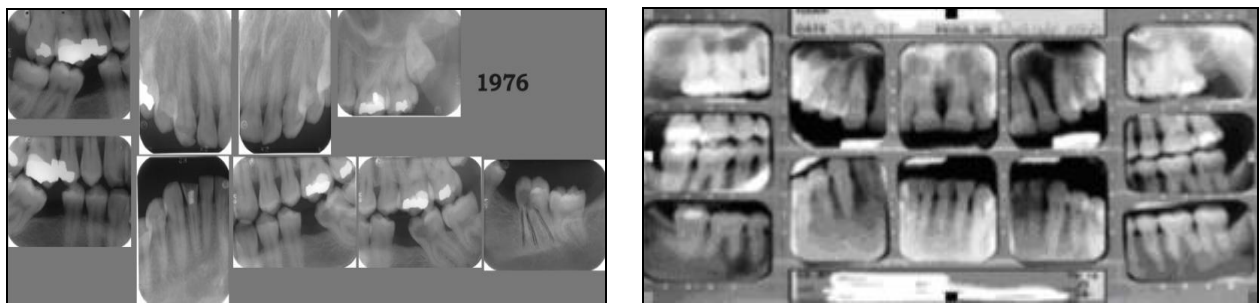
Figure 15: Sample dental records that show the heterogeneity among records of a typical dental record database. Note the varying background intensity, orientation and connectedness from record to record.

Fast and accurate cropping of dental X-ray records is a nontrivial challenge due to the heterogeneity of dental records. Traditionally, dental X-ray records are digitized by different facilities where human intervention is inevitable during the digitization process. Therefore, characteristics of dental X-ray records

vary not only from database to database but also from image to image. Figure 15 illustrates the heterogeneity among records of a typical dental record database [47][48].In order to develop an abstraction for varying characteristics of dental X-ray records, we classify the ambiguity sources of digitization process into the following three categories:

A) Intensity-related: Dental X-ray films are usually placed on the top of a uniform background; but we occasionally observe textured background. Dental fillings, which exist in a moderate portion of X-ray films, could have identical intensity values to the background color.

B) Geometry-related: Dental films have rectangular shape and two types of corners: round or right. However, the width, height and corner type of dental films change from record to record and sometimes vary within the same record. In some records, some portion of the border of dental films could get smeared out and become indistinguishable from the background.

C) Topology-related: Ideally we would like all dental films to be separated from each other during the digitization phase (more convenient from segmentation perspective). However, in typical databases used in forensic *odontology*, significant portion of the records contain films that are connected to each other.

These three classes of ambiguities jointly contribute to the difficulty of cropping dental X-ray records. Two aspects motivate our approach: First, *archness* of round-corner films provides valuable clues for segmentation. Even if several round-corner films are placed side-by-side in parallel, the arc segments still remain distinguishable. Second, the straight orientation of dental X-ray films significantly reduces the solution space - i.e., it is sufficient to consider only horizontal and vertical lines for cropping purpose. By fully exploiting geometric a priori information of dental films, we show that cropping can be achieved in a robust yet computationally efficient fashion.
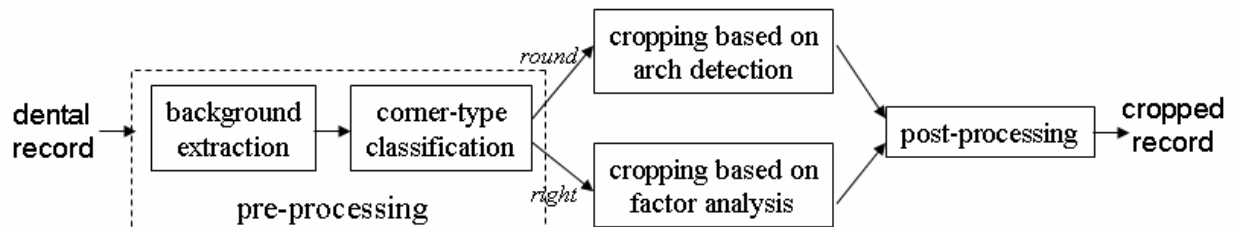
Figure 16: A block diagram illustrating the three-stage approach for dental record cropping

We use a three-stage approach for cropping as depicted in Figure 16: First a preprocessing stage whereby we extract the background layer of the image record, extract connected components and classify them as

either round-corner or right-corner connected components. The second stage performs arch detection and dimension analysis, realization of this stage differs according to the outcome of the preprocessing stage. The third stage is a post-processing stage that performs topological assessment of the cropping results in order to eliminate spurious objects.

In sub-section 3.1.2 we introduce the notations and terminology that we will use through the remainder of the cropping section. In sub-sections 3.1.3, 3.1.3, and 3.1.5 we elaborate on the preprocessing stage, the arch detection and dimension analysis stage, and the post-processing stage respectively. Finally in section 3.1.6 we present the experimental results and discussion on these results.

3.1.2 Notation and Terminology

In this sub-section, we introduce some notations and definitions for later use on the cropping section.

The dental X-ray record is a gray-scale image (or we convert it to gray-scale) denoted by $X(i, j) \in [0, 255]$ where $(i, j) \in \Omega = [1, H] \times [1, W]$ (H, W are the height and width of the image). The dimension of individual dental films is denoted by h, w.

Level set and its size

Level set L_k is a binary image defined by $L_k(i, j) = \begin{cases} 1 & X(i, j) = k \\ 0 & else \end{cases}$. The size of a binary image L_k, denoted by $|L_k|$, is simply the total number of ones in the image.

Connected Film Set and Boundary film

Multiple films that are not completely separated form a connected film set. Within a connected film set, a film is called boundary film if removing it does not affect the connectivity of remaining in the set.

Morphological area-open and area-close operators

Area open operator is the extension of morphological opening - it consists of three consecutive filters, namely erosion, small objects removal and dilation [49][50].

Area close operator is the extension of morphological closing - it consists of three consecutive filters, namely dilation, small objects removal and erosion.

90° V-corner and 180° V-corner

90° V-corner refers to the corner formed by a straight line and an arc segment;

180° V-corner refers to the corner formed by two adjacent arc segments; (as shown in Figure 17).

Inner L-corner and Outer L-corner

Inner L-corner refers to a right-corner with one quadrangle of white and three quadrangles of black; outer L-corner refers to a right-corner with one quadrangle of black and three quadrangles of white (as shown in Figure 17).

Note: Inner and outer L-corners can be detected by using morphological hit-or-miss operators (a specific case with window size of 3 is shown in Section 3.1.3.2.).
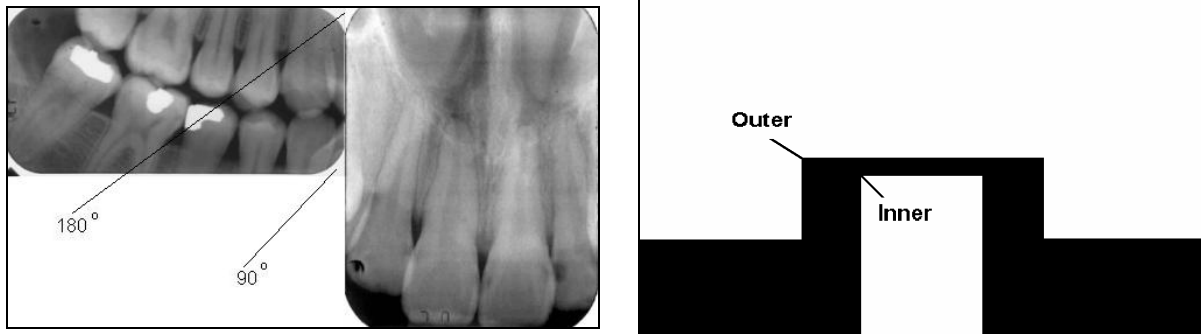
(a) 90° V-corner and 180° V-corner   (b) Inner L-corner and Outer L-corner
Figure 17: a) 90° V-corner and 180° V-corner; b) inner L-corner and outer L-corner

3.1.3 Record Cropping Preprocessing

3.1.3.1 Background Extraction

Although background typically consists of uniform color such as white, gray or black, intensity alone is not sufficient to distinguish the background from dental films. For example, the empty (no tooth) areas in a film often appear as dark regions and could become confused with black background. Similarly, dental filling often appear as bright regions in a film and could cause problem when the background is white as well. A more robust approach of extracting the background color is to count on geometric clues such as the shape of dental films. Since any dental film can be bounded by a rectangle, the boundary of background largely consists of vertical and horizontal lines.

Suppose the histogram of input image X (i, j) and the largest three peaks are n_1, n_2, n_3. We consider their corresponding level sets ∂L_k, k = n_1–n_3 and apply morphological filtering to extract the boundary of those three sets L_k. For dental films whose boundary is largely rectangular, the fitting ratio of ∂L_k by vertical or horizontal lines reflects its likelihood of being the true background. Specifically, we propose to extract vertical and horizontal lines from ∂L_k by direct run-length counting and define the fitting ratio by:

$$r_k = \frac{|R_k|}{\partial L_k}, k = n_1, n_2, n_3 \qquad (3.1.1)$$

where R_k is the binary image recording the extracted vertical and horizontal lines. The set with the largest fitting ratio among the three level sets is declared to be the background L_b. As soon as background is detected, we do not need intensity information but only the geometry of L_b for further processing (as shown in Figure 18).
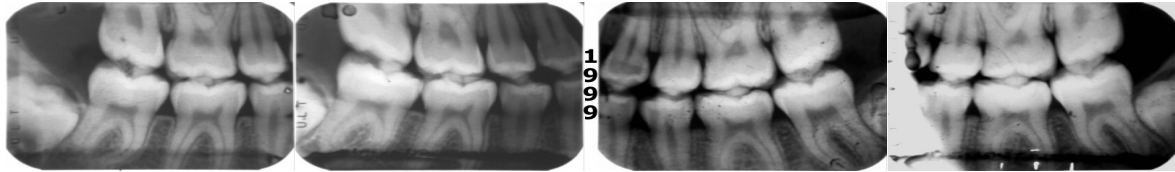
(a) original dental record

(b) level set L_1, fitting ratio $r_1=0.09$

(c) level set L_2, fitting ratio $r_2=0.05$

(d) level set L_3, fitting ratio $r_3=0.36$

Figure 18: Background extraction example

3.1.3.2 Arc Detection

The complement of detected background $\overline{L_b}$ consists of non-cropped dental films as well as various noises. The noise could locate in the background (e.g., textual information such as the date) or within

dental films (e.g., dental fillings that have similar color to the background). To eliminate those noises, we propose to apply morphological area-open operator to $\overline{L_b}$ and L_b sequentially. Suppose the

(a) Original dental record	(b) Extracted background

(c) filtered background	(d) Layered representation
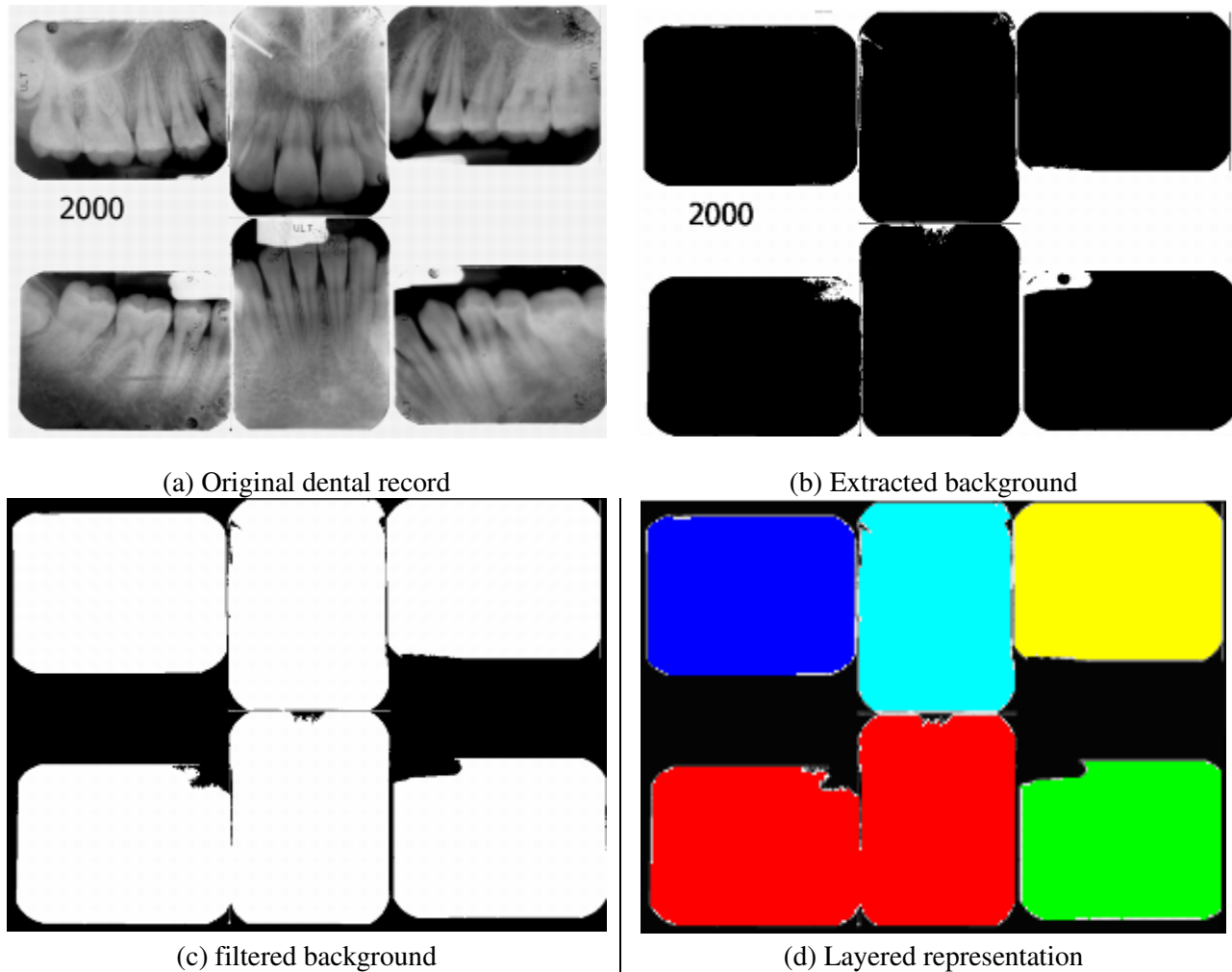
Figure 19: a) original dental record; b) extracted background; c) filtered background; d) layered representation

Euler number of filtered $\overline{L_b}$ is N, then we can label the N connected components in $\overline{L_b}$ by integers $1 - N$ (as shown in Figure 19). For each connected component (a binary map), we need to classify its corner type since a record could contain the mixture of round-corner and right-corner films.

The striking feature of a round-corner film is the arc segments around the four corners. In the continuous space, those arc segments are essentially 90°-turning curves (they link a vertical line to a horizontal one). In the discrete space, we propose to use the following Hit-or-Miss operator:

$$b_0 = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix}, b_1 = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}. \tag{3.1.2}$$

to detect corner pixels first and then morphological area-close operator to locate arc segments. The area-close operator is suitable here because it connects the adjacent corner pixels around a round corner to makes them stand out as an arc segment. By contrast, the corner pixels around a right corner will be suppressed by the area-close operator (as shown in Figure 20).
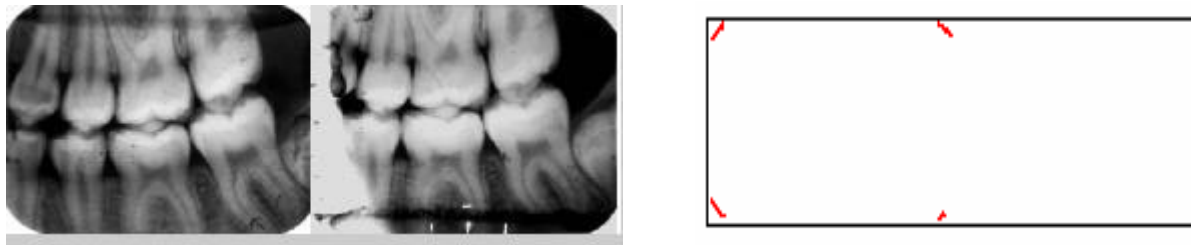
(a) Dental record with both round-corner and right-corner films   (b) arc detection result

Figure 20: Arc detection for corner-type classification

3.1.3.3 Dimension Analysis

For round-corner films, detected arc segments will be shown sufficient for cropping purpose in the next section. However, the location uncertainty of right-corner film boundary is more difficult to resolve because the inner boundary could disappear in the case of parallel superposition. We use purely geometric approach for estimating the dimension (h, w) of right-corner films. This dimension analysis approach is based on the heuristic observation that concatenation of two right-corner films could only give rise to inner L-corners. Therefore, the distance between any two outer L-corners must be:

$$d_i = m_i*h + n_i*w, \ i = 1 - k_1 \tag{3.1.3}$$

where (m_i, n_i) are a pair of nonnegative integers. Moreover, since the borders of right-corner films would not all align to each other in the case of non-parallel concatenation, it is reasonable to assume that min $\{d_i\} \in \{h, w\}$. In Figure 17, if the two outer corners corresponding to min $\{d_i\}$ are marked by A, B, then the closest corner to A, B must form a rectangular area of p*h*w, where p is an unknown positive integer.

To further resolve the uncertainty of p, we note the constraint on film ratio $\left(i.e., h\varepsilon\left[\frac{2}{3}w, \frac{3}{2}w\right] \right)$, and we had conducted an experiment, calculating the aspect ratio of 494 manually cropped dental films, and it proves that this constraint is valid (more details are shown in the experiment sub-section 3.1.6). Such constraint often reduces p to be at most two viable possibilities (combinations of h and w) when A or B is linked to an inner L-corner. If we denote the distance between any two inner L-corners by:

$$d'_i = m'_i * h + n'_i * w, \; i = 1 - k_2 > k_1 \tag{3.1.4}$$

then the weighting coefficients m'_i, n'_i could be arbitrary integers. By comparing the solutions (m'_i, n'_i) given by different combinations of (h, w), we pick out the one whose solution is closest to integers to be the most likely dimension.

3.1.4 Cropping Methodology

Preprocessing classifies each connected component in L_b to be either round-corner or right-corner. In this section, we present tailored cropping techniques for round-corner and right-corner components respectively. For round-corner component, we demonstrate that two types of V-corners associated with arc segments are sufficient for cropping. While for right-corner component, we use recursively crop out films one-by-one based on the estimated film dimension (h, w).

3.1.4.1 Round-corner Component

When multiple round-corner films are placed side by side, they form two types of V-corners as defined above. For 90° V-corner, its straight edge indicates where the cropping should occur. For 180° V-corner, we note that it is symmetric with respect to the target cropping line Figure 17. Therefore, the cropping of round-corner films can be fully based on locating and classifying the two types of V-corners. A V-corner is characterized by the intersection of two segments where the curvature experiences a sharp change. Such geometric singularity can be identified by local analysis of digital arc segment. Specifically, we define "curvature index" at a location (i, j) to be the maximum length of consecutive white pixels in $\overline{L_b}$ as we traverse its eight nearest neighbors in a clockwise order. The position is declared to be a V-corner if its curvature index is above 5 and it is close to one of the detected arc segments in corner-type classification.

Further classification of V-corner into 90° and 180° can be done based on symmetry analysis. For a 90° V-corner, neither horizontal nor vertical lines passing the V-corner divide the corner symmetrically. While for 180° corner, there exists a symmetric axis which tells the place to cut. We note that unlike generic

symmetry detection problems, the direction of symmetric axis is known to be either horizontal or vertical. Therefore, symmetry analysis can be conveniently carried out by correlation or differential techniques.

3.1.4.2 Right-corner Component

The cropping of right-corner films is based on the following intuitive observation with the boundary films. Due to the special location of boundary films, they can be properly cropped out with a higher confidence than the rest. Moreover, cropping out boundary films could make other non-boundary films become boundary ones and therefore the whole process of cropping boundary films can be recursively performed until only one film is left.

To implement recursive cropping, we require reliable detection of boundary films. It follows from the definition of boundary films that they must satisfy: 1) any boundary film contains a pair of outer L-corners; 2) the distance between the L-corner pair is either h or w. Therefore, detected outer L-corners in dimension analysis give useful clue for locating boundary films. We note that since the area of connected film set A_{cfs} and the film dimension are both known, the number of films contained in the set it approximately known $\left(n = \left\lfloor \dfrac{A_{cfs}}{hw} \right\rfloor \right)$. The iteration number of recursive cropping is given by $n - 1$.

3.1.5 Post-processing

Cropping techniques presented in the previous section mainly target separating connected films. There are other factors that can not be handled by cropping and may affect final segmentation results. For example, some dental films are contaminated before digitization such that a portion of film becomes indistinguishable from the background (as shown in Figure 19). The accuracy of cropping itself could also degrade due to errors in dimension analysis. For example, in the case of right-corner films, there might be some leftovers after cropping out $\left\lfloor \dfrac{A_{cfs}}{hw} \right\rfloor$ films. Consequently, it is desirable to have a post-processing stage to consummate the segmentation results.

One of prior information about dental films is that they are all convex sets, regardless of the corner-type. Such knowledge implies that the hole or cracks of any segmented component be filled in by finding its convex hull. Therefore, the first step in post-processing is to enforce the convexity of all connected components after cropping. Secondly, we propose to check the size and shape of each convex component. If the size of a component is too small or its shape significantly deviates from rectangle, we detect its outer L-corners and check if they correspond to one border of the film. If yes, we conclude that the film

was contaminated and derive its boundary using dimension information. Otherwise, we decide it is likely to be a non-film object and put it back to the background layer.

3.1.6 Experimental Results

First: we started by an experiment studying one of the dental film properties, film size (aspect ratio). Since dental films have 5 different sizes, so the ratio of width to height (aspect ratio) will have 5 different values (0.64, 0.66, 0.77, 0.5, 0.75) [51], but the films can slightly rotate while added to the record. By studying the aspect ratio of 494 films (periapical and bitewing types), we found it is to range between 0.7233 and 2. Moreover calculating the minimum of (aspect ratio, and inverse aspect ratio) for these films, which we call it "minimum side ratio"; we found it to range between 0.49 and 0.91. By studying the probability density function of those cases; we found that 94.13% have aspect ratio between 0.6 and 0.8, while app. 3% have minimum side ratio less than 0.6, and app. 3% have minimum side ratio greater than 0.8.

A performance assessment experiment: We evaluate both the yield and timeliness aspects of our film cropping approach using a randomly selected test sample of 100 dental records (images) from the CJIS ADIS database [47][48], the total film count in the test set is 722. We verified that the test sample has variability in background and contains films with both corner types (48 round-corner, and 52 right-corner records). We marked the cropped segments using the following convention: a perfect segment contains exactly one film, an under segmented (under cropped) region contains several whole films (Figure 21), and an erroneous segment is that contains part of the film, or region from the background texture, or both. In Figure 22, we summarize the yield analysis: 73.7% of the films were perfectly cropped, 23.8% were under segmented, and only 2.5% films developed into erroneous segments. Further cropping-yield analysis show that in the right corner records perfect segmentation rate is 70.1%, under segmentation rate is 28.2%, and erroneous segmentation rate is 1.7%. While in round corner records these rates are 76.9%, 19.9%, and 3.2% respectively. We measured the record cropping time of our algorithm using MATLAB® implementation running on a 2.0 GHz 512 MB RAM Intel® Pentium IV PC platform. The average cropping time is 30 kpix/sec and it varies depending on the number of films in the record and the amount of separation between films.

An exploratory experiment to examine potential future yield-boosting opportunities: Some film geometric properties, like the minimum sides ratio 'γ', may provide clues to judge cropped segments. We found that by checking the rule $\gamma > 0.49$, we could mark under segments, containing ~14.8% of the films, as γ-violating. Furthermore, by observing that most records comprise films of about the same area (except for

panoramic films), we could also mark under segments, containing ~8.4% of the films, as area-violating. In the future we may exploit these properties for defining additional post-processing rules whose violations call for further subsequent processing and hence boost the yield performance. Figure 22 is showing an example of how the technique performs.
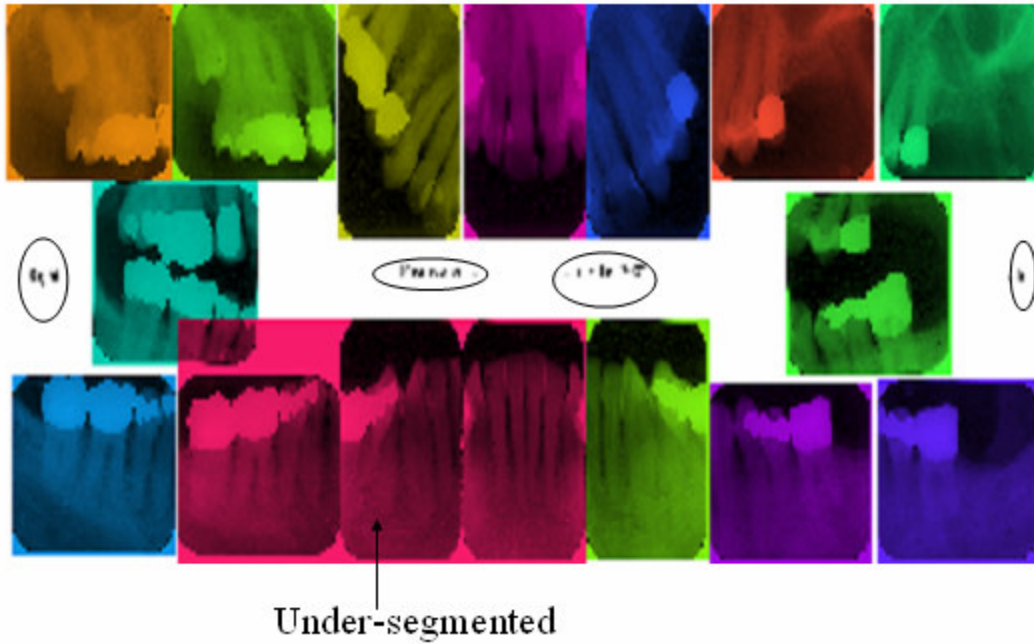
Figure 21 Record cropping example
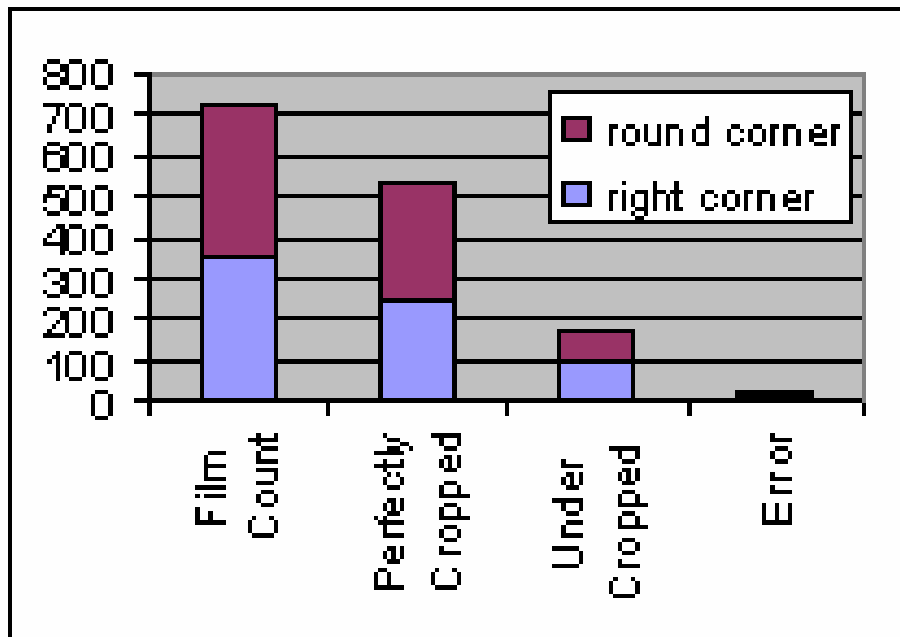
Figure 22: Experimental results

3.2 Teeth Contour Extraction

Dental features extraction is one of the major pre-processing steps. In this section, the problem of teeth contour extraction using level sets and active contour without edges are addressed. Preliminarily experimental results prove that the active contour without edges approach for teeth contour extraction achieves higher accuracy and requires significantly less time as compared to the other approaches currently being used within the context of ADIS.

3.2.1 Introduction

After the dental films are cropped from dental records, they passed through the enhancement step followed by a film type classification step. According to the film type (bitewing, upper periapical, or lower periapical), each cropped film is segmented to multiple teeth, as shown in Figure 23, which is recognized as a local segmentation problem of isolating each tooth in order to facilitate the extraction of features like crown contour (crown is part of the tooth above the gum line) and root contour (root is part of the tooth below the gum line "inside the bone") for identification use. Haj said et al. [20] use an automated approach for segmenting dental x-rays films into multiple teeth segments based on mathematical morphology.
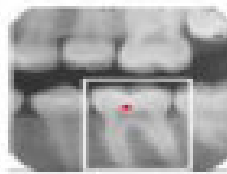
Figure 23: Local segmentation

Last, segmenting each tooth from the isolated region, this is recognized as teeth contour extraction (as shown in Figure 24). Jain and Chen [14][25] propose a semi-automatic contour extraction method. First, the crown center is marked inside the rectangle tooth segment and then edge detection algorithm is used to extract the contour of teeth. Later Chen and Jain [18] extract the teeth contour using the active contour models (snake). A new dynamic energy term was used for the directional snake [52] to discriminate boundaries of adjacent teeth. Zhou and Abdel-Mottaleb [17] also have used active contour models (snake) for extracting the contour of the teeth. Active contour models based on edges (snakes) are driven by the gradient of the images. In many teeth images, the gradient between the teeth and the background is not prominent, and hence the techniques like snakes will not be able to extract the exact tooth contour. Also, snakes are Lagrangian formulations where the contour is represented in a parametric form. Parametric
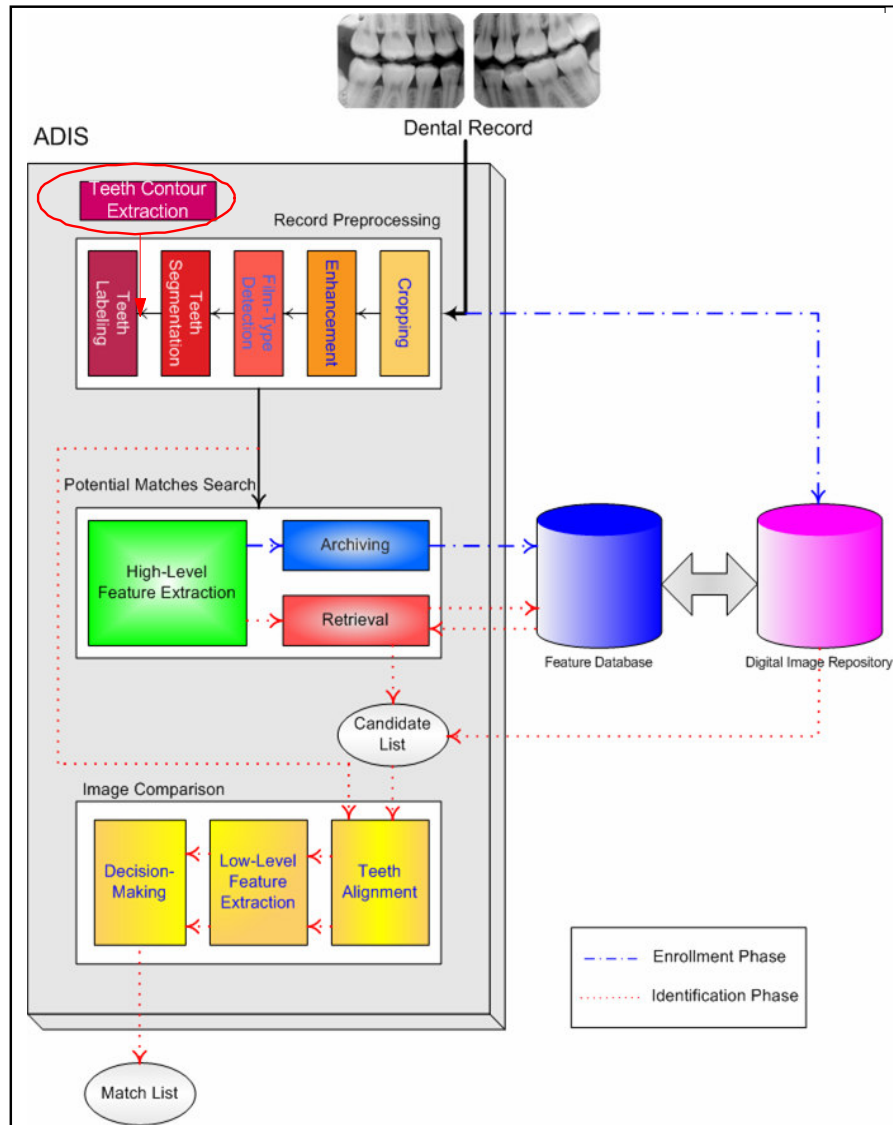
Figure 24: ADIS modified architecture

representations typically fail to evolve in noisy conditions [53]. Further, the parametric contours are unable to split and merge on encountering local minima in the image [54]-[57].

In this section, we show how we started using level sets then we moved to a novel technique using Active Contour without Edges for teeth contour extraction. The latter technique is independent of the gradient of images for extracting the contour of teeth. It is based on the intensity of the overall region of the tooth image. Also, the starting contour can be in any part of the image and does not have to be around the boundary of the tooth as in case of snakes. Preliminarily experimental results on the teeth images show that the proposed technique using active contour without edges performs extremely well (even compared to level sets) for extracting the teeth contour.

The remainder of the section is organized as follows: Sub-section 3.2.2 introduces some of the notations and terminology that will be used throughout the rest of this section. The technique for extracting the teeth contour using level sets and active contour without edges are explained in sub-section 3.2.3 and 3.3.3 respectively. The experimental results and a discussion on these results are reported in sub-section 3.3.4.

3.2.2   Notations and Terminology

Some notations and definitions are introduced in this section.

Teeth segmentation: Teeth segmentation means partitioning an image into its constituent regions and extracting the objects of interest. In the context of ADIS, segmentation is a step required to identify the extent of teeth comprised in a digital image of a dental radiographic film.

Tooth segment: A tooth segment represents a Region of Interest (ROI) that contains distinctive features used in the subsequent steps of identification. A qualified ROI is defined as a rectangular area in the image of a dental film that bounds one tooth as shown in Figure 23.

Tooth contour extraction: Tooth contour extraction follows the segmentation of teeth and extracts the tooth contour from its background.

3.2.3   Teeth Contour Extraction using Level Sets

Malladi et al. [53] proposed a shape modeling (A level set) approach. This approach to shape modeling retains some of the attractive features of active contour models (snake) and overcomes some of their limitations. Level set can split and merge freely to represent more than one object, can have self-intersection, and can represent sharp of other discontinuities. Level sets can be applied to model arbitrarily complex shapes, which include shapes with significant protrusions. The interface (front) is a closed, non-intersecting, hyper-surface flowing along its gradient field with constant speed or a speed that depends on the curvature $F(k)$. It is moved by solving a "Hamilton-Jacobi" type equation written for a function in which the interface is a particular level set:

$$\frac{\partial \psi}{\partial t} + F(\kappa)|\nabla \psi| = 0 \tag{3.2.1}$$

A speed term synthesized from the image is used to stop the curve evolution in the vicinity of object boundaries. The resulting equation of motion is solved by employing entropy-satisfying upwind finite difference schemes. The segmentation problem is now reduced to finding an appropriate function $F(k)$. We can factor $F(k) = K(F_0 + F_1)$, where F_0 is the advection term (inflation force) and is usually a constant, F_1 is the curvature of the curve provides the smoothing constraints on the level sets and reduces

the total curvature of level sets, K is a stopping term, to slow evolution near the object boundaries. Generally, the form is $F(k) = K (\alpha + \varepsilon\,k)$, , and the evolution equation for Ψ (x, y, t) is given by:

$$\psi_t = -K(\alpha + \varepsilon\kappa)|\nabla\psi| \tag{3.2.2}$$

where ε is a very small constant, the term α acts like a balloon force and it pushes the curve outward towards the tooth boundary. The goal of the stopping term K is to slow the evolution when it reaches the boundaries. The stopping function used for the evolution of level sets is derived by the Gaussian smoothed image gradients. k is the curvature of the curve and is calculated by:

$$\kappa = -\frac{\psi_{xx}\psi_y^2 - 2\psi_x\psi_y\psi_{xy} + \psi_{yy}\psi_x^2}{\left(\psi_x^2 + \psi_y^2\right)^{\frac{3}{2}}} \tag{3.2.3}$$

where Ψ_x is the gradient of the image in the x direction; Ψ_y is the gradient in the y direction; Ψ_{xx} is the 2nd order derivative in the x direction; Ψ_{yy} is the 2nd order derivative in the y direction; and Ψ_{xy} is the 2nd order derivative first in the x direction and then in the y direction.
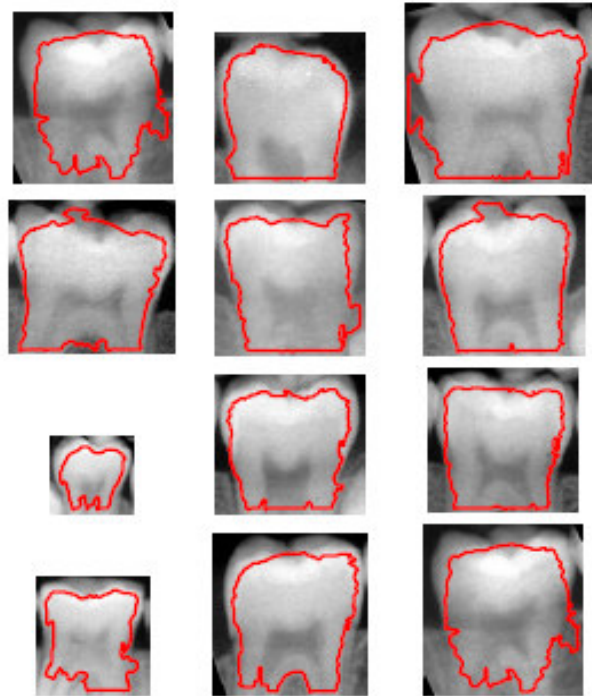
Figure 25: Teeth contour extraction using the level set approach

The previously mentioned technique based on the Level Set approach (active contour with edges) was used for teeth contour extraction. Here, the curve evolution is driven by the gradient of the image. Figure 25 shows the extracted contour on segmented teeth images. It is evident from the results that the contour

extracted using Level Sets is not smooth. Also, when edges are not prominent in the teeth images, the stopping function is not able to terminate the evolution of the level sets, and thus an incorrect tooth boundary is extracted. This was the motive behind investigating the performance of active contours without edges algorithm.

3.2.4 Teeth Contour Extraction using Active Contours without Edges

To advance the research in this area, of extracting the contour of a tooth, a novel method using active contours without edges [58] based on the Mumford-Shah functional for image segmentation [59] has been used. This is a region based segmentation technique where the evolution of the contour is guided by a model fit error term. Here an embedding function Ψ (x, y, t) evolves from inside the tooth image under the influence of a model fit error term and stops on the boundary of the tooth. The image is assumed to be made of two regions viz. the foreground and background of approximately piecewise constant distinct Intensities. The embedding function Ψ (x, y, t) is designed as a signed distance function, $|\Psi$ (x, y)$| =$ distance of (x, y) to the curve:

$$\psi(x, y) = \begin{cases} = & 0 \quad if \quad (x, y) \quad is \quad on \quad the \quad curve \\ < & 0 \quad if \quad (x, y) \quad is \quad inside \quad the \quad curve \, ; \\ > & 0 \quad if \quad (x, y) \quad is \quad outside \quad the \quad curve \end{cases} \tag{3.2.4}$$

Let C_0 be the contour separating the two regions. Let the energy of the fitting be defined as:

$$E(C) = \int_{inside(c)} |I_0(x, y) - c_1|^2 \, dxdy + \int_{outside(c)} |I_0(x, y) - c_2|^2 \, dxdy \tag{3.2.5}$$

where C is any contour in the image, and c_1 and c_2 are the mean intensities of the image, I_0, inside and outside of the contour C, respectively. The above equation will be minimized only when the contour $C = C_0$, i.e. when the contour C is segmenting the boundary of the object. It will be nonzero when the contour is either inside or outside the object. Thus, in this formulation, the embedding function, Ψ (x, y, t), is evolved from inside the tooth image under the influence of the above mentioned model fit error term. It stops on the boundary of the tooth.

The contour is initialized as a circle centered at the center of the image. Then the evolution equation is given by:

$$\frac{\partial \psi}{\partial t} = \left[\mu div \left(\frac{\Delta \psi}{|\Delta \psi|} \right) - \left(\lambda_1 (I_0 - c_1)^2 + \lambda_2 (I_0 - c_2)^2 \right) \right] \tag{3.2.6}$$

where λ_1, λ_2 and μ are constant (In our implementation λ_1, and λ_2 are both one and μ is 2). The first term (curvature of the curve $div \left(\frac{\Delta \psi}{|\Delta \psi|} \right)$), is calculated as:

$$div\left(\frac{\Delta\psi}{|\Delta\psi|}\right) = -\frac{\psi_{xx}\psi_y^2 - 2\psi_x\psi_y\psi_{xy} + \psi_{yy}\psi_x^2}{\left(\psi_x^2 + \psi_y^2\right)^{1.5}} \qquad (3.2.7)$$

where Ψ_x is the gradient of the image in the x direction; Ψ_y is the gradient in the y direction; Ψ_{xx} is the 2nd order derivative in the x direction; Ψ_{yy} is the 2nd order derivative in the y direction; and Ψ_{xy} is the 2nd order derivative first in the x direction and then in the y direction. The curvature term provides the smoothing constraints on the evolution of the contour by minimizing the total curvature of the contour. The second term, the model fit error term, drives the evolution of the contour. The model fit error is minimized when the contour divides the image in such a way that the mean intensities c_1 and c_2 are having the maximum difference.
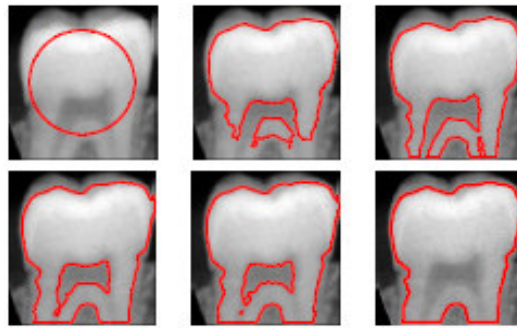
Figure 26: Evolution of the active contour without edges

Figure 26 shows an example of how the proposed technique evolves over time to segment different teeth contours. In some dental images a part of the neighboring tooth may also be erroneously included in the segmented result (note that local segmentation precedes the contour extraction process). To avoid this, a post processing step was used after contour extraction. The contour on the left and right side of the tooth is always convex in shape. Hence, whenever the contour was becoming concave due to neighboring tooth image, that part of the contour was deleted.

3.2.5 Experimental Results

We tested the proposed Teeth Contour Extraction using Active Contours without Edges with different types of teeth and the preliminary experiment proves the concept. Figure 27 shows the results that were obtained using this algorithm on molars, pre-molars, canines and incisors. As can be seen, a very smooth

and accurate teeth contour is obtained using this algorithm. Also, the contour extraction procedure is extremely fast (Converges in 0.16 second/tooth using a MATLAB environment).
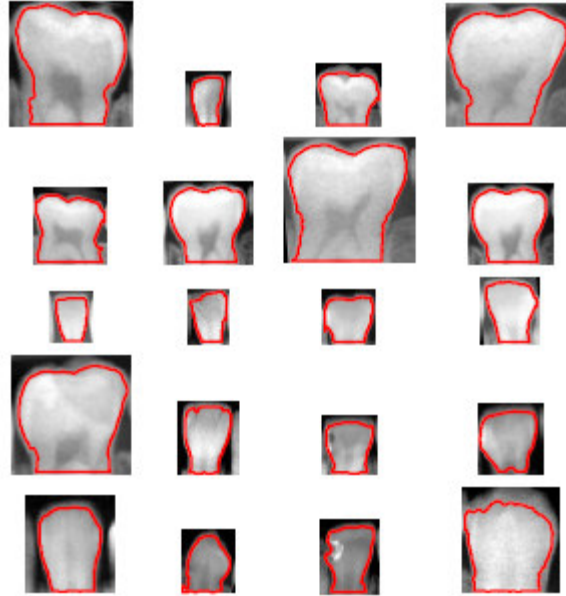
Figure 27: Teeth contour extraction using active contour without edges

The snakes algorithm as implemented by Chen and Jain [30] was also used to extract the contour of the teeth images in dental films. Figure 28 (a) and (c) shows the contour extracted from two dental films using their technique. As is evident from the results, the snake is not able to detect the contour of many tooth entities in the dental films (e.g. lower left tooth in Figure 28 (c)). The extracted contour does not represent the real teeth contour but, rather, is an approximate shape of the teeth. This is perhaps due to the previously mentioned limitations of the snake-based evolution techniques. The contour extracted for the same set of films using the proposed active contour without edges technique is shown in Figure 28 (b) and (d). It can be seen that the contour is extracted for most of the tooth images in the films and the real boundaries of the crown of the teeth images are perfectly segmented. The error in the extracted contour in Figure 28 (d) for the lower molar (red color) is due to incorrect teeth image segmentation in the previous stage.
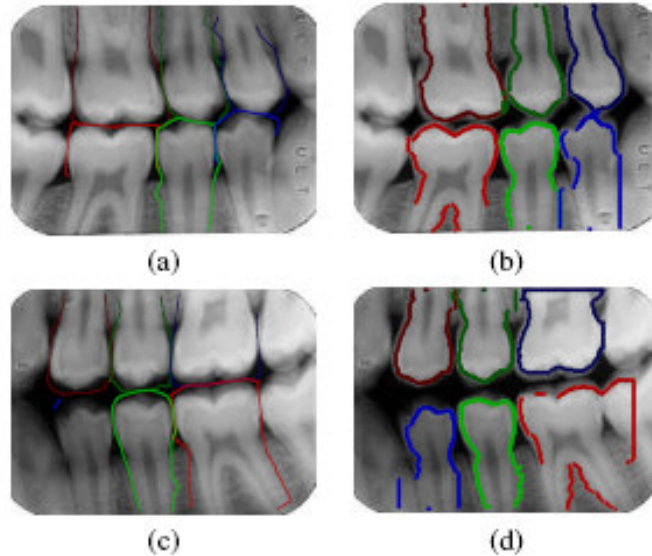
Figure 28: (a) and (c) contour extracted using the snake-based algorithm;
(b) and (d) contour extracted using active contour without edges

The aforementioned procedure was used to extract the teeth contour of images from CJIS Division's Digitized Radiographic Images databases [47][48]. These databases have two types of dental radiographs. The radiographs obtained when the person is alive are labeled as Ante-mortem (AM) radiographs and those obtained after the death of the person are labeled as Postmortem (PM) radiographs. From this database a total of 10 AM and 10 PM records were randomly selected. Each record contains approximately 3 films. There are a total of 340 teeth represented in these films. The teeth contour algorithm was executed on these images and the results were manually categorized into the following four types: (a) Perfectly segmented tooth (whole tooth) (P), (b) Perfectly segmented crown (PC), (c) Contour that can be corrected at a later stage (C), (d) Error in contour extraction (E). Table 3 summarizes the performance of the proposed contour extraction technique.

Records	Correct or partially correct contour extraction (%)			Errors in contour extraction (E)	Average time/tooth (sec)
	Perfect contour (P)	Perfect crown contour (PC)	Partially corrected (C)		
10 (AM)	56.2	14.05	16.75	13.0	0.15
10 (PM)	60.0	11.61	14.83	13.56	0.17
20 (Total)	58.10	12.83	15.79	13.28	0.16

Table 3: Performance of the proposed teeth contour extraction algorithm

3. Preprocessing Techniques for ADIS

Figure 29 shows an example for each of the above mentioned categories.
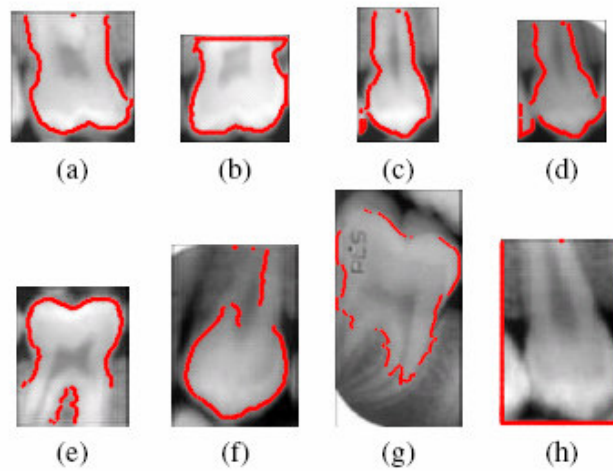
Figure 29: Contour extracted using active contour without edges.
(a, b) Perfectly segmented (P), (c, d) Perfect Crown (PC),
(e, f) Correctable error (C), (g, h) Error in contour (E)

Figure 30 shows several examples of the extracted teeth contour on other dental films.
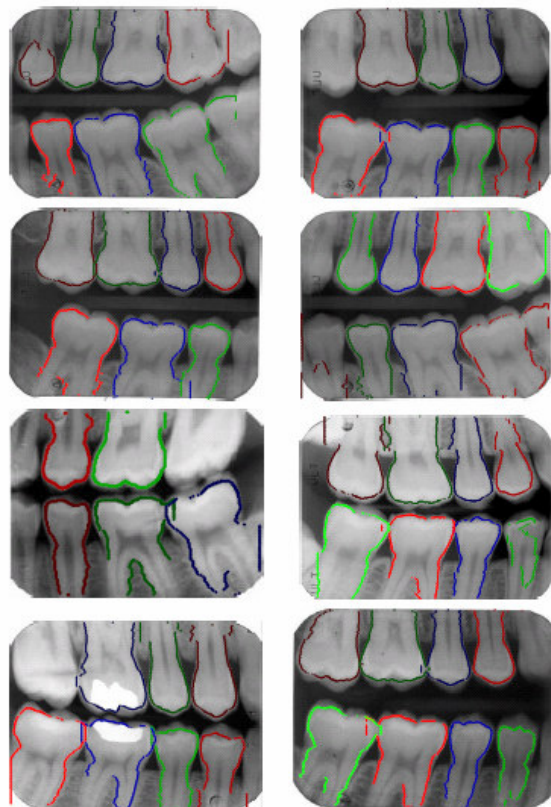
Figure 30 A few more examples of extracted teeth contour using active contour without edges.

3.3 Teeth Labeling

An important problem in automated dental identification, is automatic classification of teeth into four classes (molars, premolars, canines, and incisors) as part of creating a data structure that guides tooth-to-tooth matching, thus avoiding illogical comparisons that inefficiently consume the limited computational resources and may mislead decision-making. We tackle this problem by first using low computational-cost, appearance-based on features to assign an initial class, and second a string matching technique, based on teeth neighborhood rules, is applied validate initial teeth-classes and hence assign each tooth a number corresponding to its location in the dental chart. The proposed approach works very fast, as it does not need any complex feature extraction steps, and achieves accurate teeth labeling even in the absent of one tooth.

3.3.1 Introduction

The problem of automatic construction of dental charts was addressed in [15][33], where Mahoor and Abdelmottaleb deal only with bitewing films (2 class problem) to classify molar and premolars. They used the Fourier Descriptors of teeth contour and Bayesian classification. They assigned labels based on the location of premolar and molars. The teeth arrangement without missing tooth should match one of these patterns:

PM, PMM, PMMM, PPM, PPMM and PPMMM (left quadrant image)

MP, MMP, MMMP, MPP, MMPP and MMMPP (right quadrant image)

They discard the cases with only molar or premolar teeth such as M, MM, MMM, P and PP.

Jain and Chen [36] used a hidden Markov model (HMM) as an underlying representation of the dental atlas, where the states representing the available teeth have discrete observations, namely the class of each tooth, and the states representing the missing teeth have continuous observations (They can assign one or two missed teeth according to the distance between neighboring teeth).

All these techniques use teeth contour features (shape based); while automatic extraction of teeth contour is computationally intensive, and to the best of our knowledge, none of the techniques exploited in teeth contour extraction, so far, achieves this task reliably [43]

The proposed approach for constructing dental charts relies on low computational-cost, appearance-based features to initially independently classify each segmented tooth of a dental film, i.e. without considering neighboring teeth. The proposed technique is an extension to what Nassar presented [60], where we consider:

- The film type, where with periapical films, a tooth t_q is initially assigned a class label $\omega_k \in \{$ 'I', 'C', 'P', 'M' $\}$, where 'I' stands for an Incisor, 'C' stands for a Canine, 'P' stands for a Premolar, and 'M' stands for a Molar, while with bitewing film, the initial classes are 'P', 'M'.

- Testing the performance of Linear Discriminant Analysis (LDA) for the teeth classification of *FisherTeeth* is an interesting approach that is worth pursuing.

- Validating and possibly correct initial teeth class assignments, even in the presence of missed tooth, we use teeth class-neighborhood rules that govern the assortment of teeth in a human mouth.

This section is organized as follows, sub-section 3.3.2 introduce the notations and terminology that will be used through the remainder of this section. In sub-sections 3.3.3, 3.3.4 and 3.3.5 we elaborate on the input segment normalization steps and the initial class assignment using reconstruction based on PCA, followed by using the Linear Discriminant Analysis (LDA) for the initial class assignment. In sub-section 3.3.6 we show how we can make validation based on string matching techniques including 'don't care' character. In sub-section 3.3.7 we present the experimental results on the performance of the different class assignment methods and a discussion on these results.

3.3.2 Notation and Terminology

Dental Atlas: An adult has 32 permanent teeth (8 incisors, 4 canines, 8 premolars and 12 molars). Each tooth has a specific structure and position in the mouth. Figure 31 shows left half of the upper jaw.
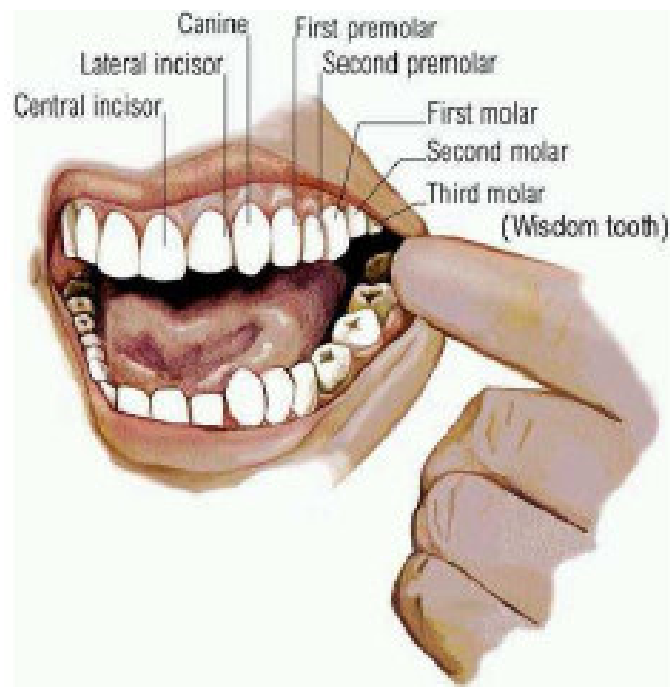
Figure 31: Dental atlas, taken from [61]

Periapical and Bitewing Different Placements [62] (as shown in Figure 32)

Central incisor periapical – records the crowns and roots of the central incisors and the mesial (side towards the midline) aspect of the lateral incisor

Lateral incisor periapical – records the crown and root of the lateral incisor

Canine periapical – records the crown and root of the canine tooth (The lateral and canine views are often combined into one projection on each side.)

Premolar periapical – records the crowns and roots of the distal (side away from midline) of the canine, 1st and 2nd premolar (bicuspid) teeth and 1st molar teeth

Molar periapical – records the crowns and roots of the 1st, 2nd and 3rd molar teeth

Premolar bitewing – records the crowns of the distal (side away from midline) of the canine, 1st and 2nd premolar (bicuspid) teeth and 1st molar teeth

Molar bitewing – records the crowns of the 1st, 2nd and 3rd molar teeth
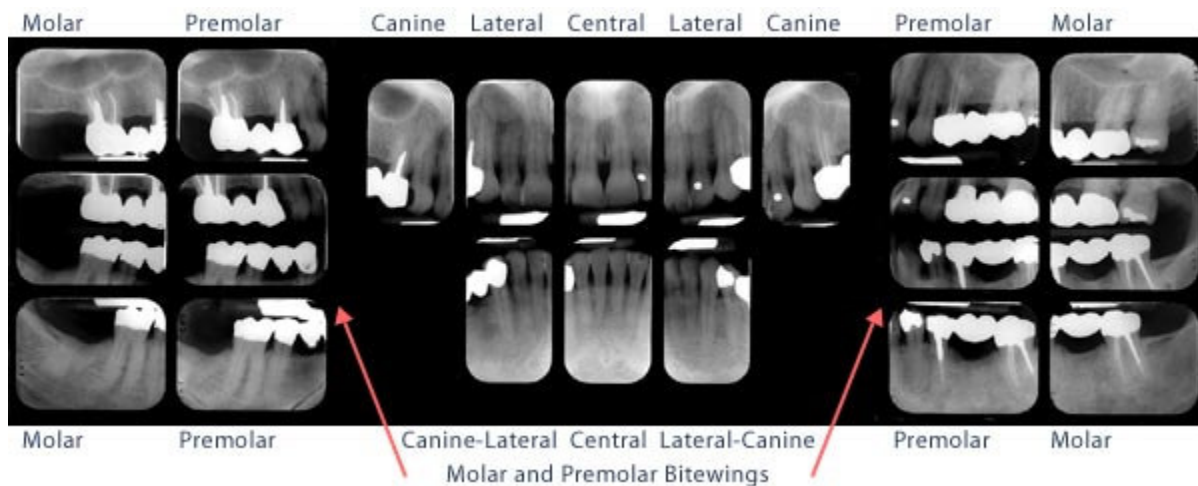
Figure 32: Periapical and bitewing different placements, taken from [62]

Numbering systems for Adult: have been developed in order to have a standard way of referring to particular teeth (there are more than 32 different systems). Two are commonly in use in the US today: 1) Universal Numbering System, which has been adopted by the American Dental Association, and 2) Palmer Notation Method, which is used by some orthodontists, pedodontists (child dental specialist), and oral surgeons. While internationally the two-digit FDI World Dental Federation notation is widely used [63]. Table 4 shows the difference between those systems.

Universal Numbering System															
Upper (Maxillary) Left								Upper (Maxillary) Right							
16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1
M	M	M	P	P	C	I	I	I	I	C	P	P	M	M	M
M	M	M	P	P	C	I	I	I	I	C	P	P	M	M	M
17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
Lower (Mandibular) Left								Lower (Mandibular) Right							

Palmer Notation															
8	7	6	5	4	3	2	1	1	2	3	4	5	6	7	8
8	7	6	5	4	3	2	1	1	2	3	4	5	6	7	8

FDI Two-Digit Notation															
28	27	26	25	24	23	22	21	11	12	13	14	15	16	17	18
38	37	36	35	34	33	32	31	41	42	43	44	45	46	47	48

Table 4: Different teeth numbering systems

3.3.3 View Normalization

In practice, dental radiographic films show teeth that need not comply with the standardizing assumptions on view that we made. On one hand, crowns need not appear on the topside of the segmented regions, as in the case with upright films showing teeth of the upper jaw (or flipped films showing teeth of the lower jaw). While some of the teeth segmentation algorithms annotate teeth as either upper or lower, and thus partly solve the first type of deviations from standard view [14][16][17], other teeth segmentation algorithms do not discriminate between upper and lower teeth [20]. To accommodate the different teeth segmentation algorithms, a simple test is conducted for bitewing films (For periapical films; the film type classification stage differentiates between upper and lower periapical films). This test advises based on whether the center of the segmented tooth is closer to the top or the bottom part of the bitewing film. On the other hand, the lateral surfaces of radiographed teeth do not necessarily appear as predominantly vertical lines. In view normalization, there are two types of transformations that we usually need to address. First is the photometric (or intensity) transformation.

3. Preprocessing Techniques for ADIS

The photometric transformation was addressed at earlier step of dental record preprocessing (film enhancement). Hajsaid et al. [20], proposed a grayscale contrast stretching transformation using a parametric sigmoid transform. Thus they mapped an input grayscale g to T_{Enh} (g, α, β, g_{th}) according to:

$$T_{Enh}(g,\alpha,\beta,g_{th}) = \begin{cases} g_{TH} \cdot \dfrac{e^{\alpha \cdot \frac{(g-g_{min})}{(g_{th}-g_{min})}} - 1}{e^{\alpha} - 1} & g_{min} \leq g \leq g_{th} \\[3em] g_{TH} + (255 - g_{TH}) \cdot \dfrac{e^{\beta \cdot \frac{(g_{th}-g)}{(g_{max}-g_{th})}} - 1}{1 - e^{\beta}} & g_{th} \leq g \leq g_{max} \end{cases}$$

(3.3.1)

where α, and β designate compression and expansion factors, respectively, g_{th} is a threshold grayscale between the compression and the expansion regions of T_{Enh}, g_{TH} is the image of g_{th} under T_{Enh}, g_{min} and g_{max} are the minimum and maximum grayscales of the input image, respectively. We choose g_{th} such that its proportionality to the new grayscale range [0; 255] resembles that of g_{TH} to the grayscale range of the input image. Thus $g_{TH} = 255 \cdot (g_{th}-g_{min})/(g_{max}-g_{min})$.

The next step towards normalizing the view of a segmented tooth is to ensure that its lateral surfaces appear predominantly vertical. Starting with a tooth with an upward oriented crown as an input image (as shown in Figure 33-a), the view normalization task proceeds in the following steps:

(a) Tooth Image   (b) Edge Image   (c) Histogram of Orientations

(d) Rotated Image   (e) Rotated Edge Image   (f) Tooth ROI
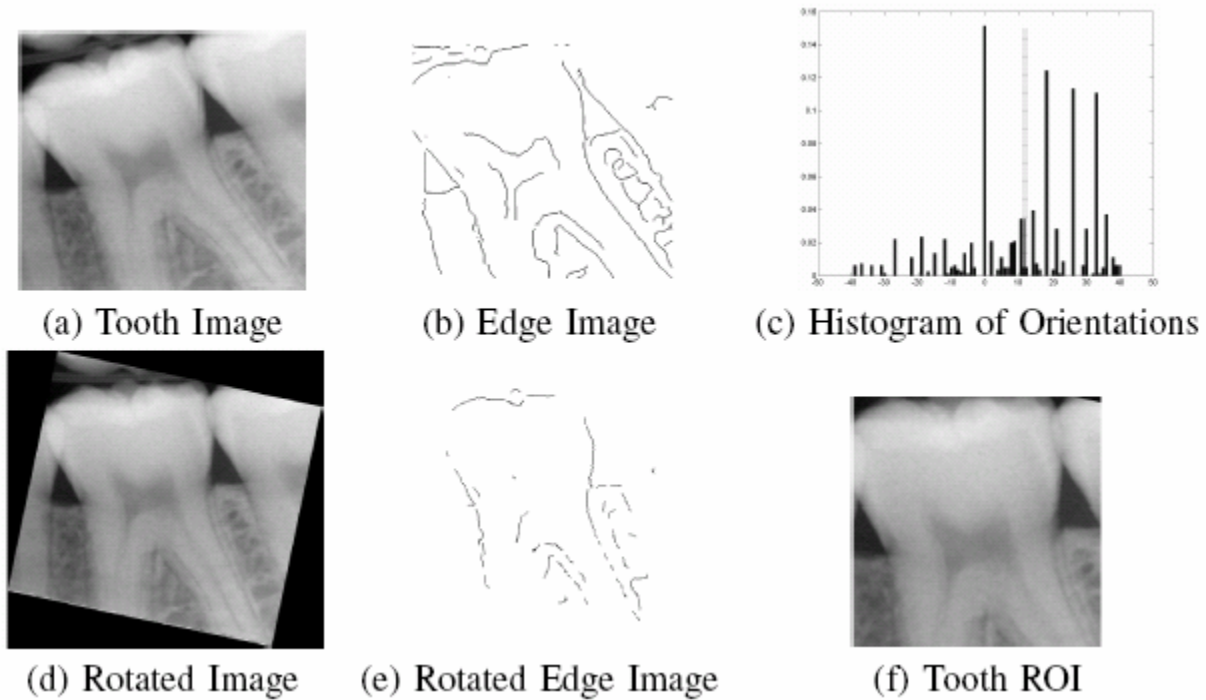
Figure 33: illustration of the orientation adjustment and Region of Interest (ROI) confinement steps of a segmented tooth

- An edge image is calculated from the input image (using Canny edge detection [64]), as shown in Figure 33-b. The choice of the width of the smoothing kernel is such that to avoid any excessive number of edge points due to intensity variations not related to the tooth surfaces.

- Using the gradients of the input image, the orientation at each of the remaining edge points is estimated, and accordingly classified as either horizontal or vertical. Hence a histogram of orientations of the vertical edge points is computed (as shown in Figure 33-c)

- The mean of the vertical orientations is computed, and its negative is used as the rotation angle for the input image such that the lateral surfaces of the tooth appear predominantly vertical (as shown in Figure 33-d).

- By analyzing the vertical and horizontal span of the tooth, the tooth ROI is confined to a square area, whose edge is just above the predominantly horizontal line segment belonging to the crown and such that the lateral sides of the tooth are also contained in the square area (as shown in Figure 33-e).

- The resulting square images are then resized to a fixed sized image of nxn pixels. Based on experimental study in section 3.3.7, we used n = 32.

3.3.4 Teeth Reconstruction and Classification

In designing a classification system, there are many factors that determine the choice of a set of features. The computational cost and reliability of automatic feature extraction are two key factors that determine our choice of practically feasible features. By visually inspecting images of radiographed teeth of the four different classes, one can easily find similarities between teeth of the same class and discrepancies between teeth of different classes. These similarities (and differences) are primarily broadly manifested in teeth peripheries. However, reliable automatic extraction of teeth contours is a very difficult problem and its current solutions are still inaccurate [43]. This motivates us to consider alternative reliable appearance-based, low computational-cost approaches.

The key idea behind the initial stage of classification in Nassar's approach [60] was to establish four image subspaces corresponding to the four teeth classes, then to use the projections of a novel tooth onto these subspaces as basis for classification. Because of some resemblance between his method for teeth classification and the eigenfaces (at least at the abstraction level) [65], he referred to the resulting basis vectors defining the teeth subspaces as the eigenteeth. With these image subspaces constructed, he achieved initial teeth classification as depicted in Figure 34, which can be summarized as follows:
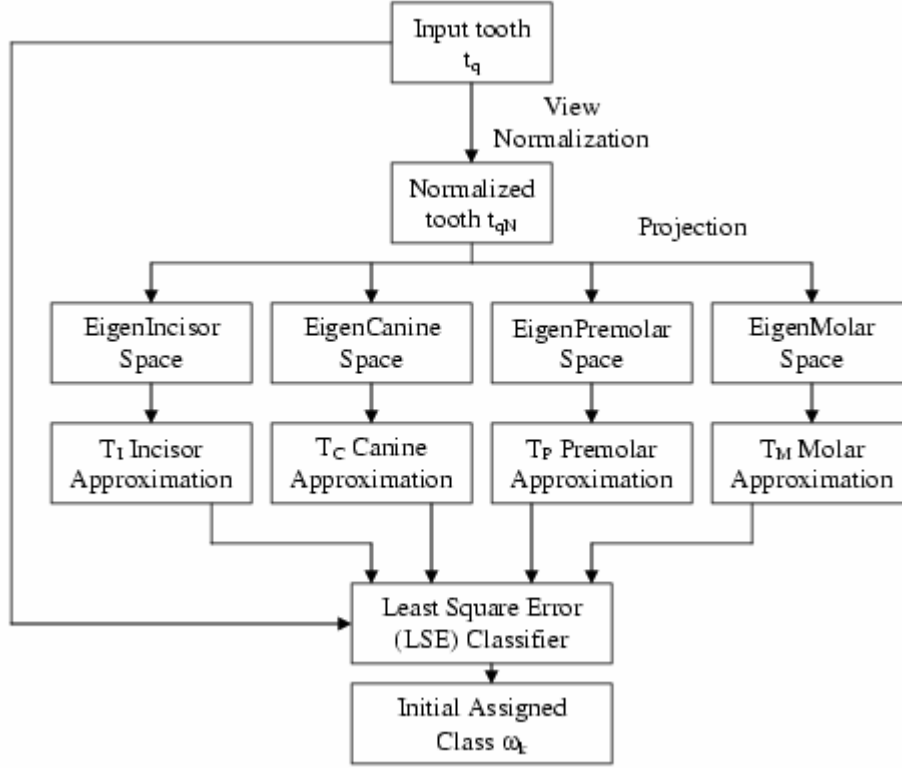
Figure 34: EigenTeeth reconstruction and initial classification

- An input tooth t_q is view-normalized (as described in subsection 3.3.3) to compensate for possible geometric variations that may cause significant differences between that tooth and the exemplar sets used for constructing the four subspaces.

- The view-normalized input tooth t_{qN} is projected onto the four image subspaces. Hence, we obtain four coefficient sets w_I, w_C, w_P, and w_M, corresponding respectively to the projections of t_{qN} onto the incisors subspace, the canines subspace, the premolars subspace, and the molars subspace. The weight sets are given by:

$$w_{\omega_k} = (\underline{t}_{qN} - \underline{\mu}_{\omega_k})^T \times [\underline{u}_{1\omega_k} \ \underline{u}_{2\omega_k} \ ..\underline{u}_{i\omega_k}.. \ \underline{u}_{K\omega_k}], \ \omega_k \in \{'I','C','P','M'\} \qquad (3.3.2)$$

where $\mu_{\omega k}$ is the mean vector, and $u_{i\omega k}$ is the ith eigenvector of class ω_k, $\omega_k \in \{I; C; P; M\}$.

- The obtained weight sets are used in conjunction with the sample mean of each of the four teeth classes to reconstruct the view-normalized tooth t_{qN} in the four image subspaces, thus obtaining the approximations $(\hat{T}_I, \hat{T}_C, \hat{T}_P, \hat{T}_M)$ according to:

$$\hat{\underline{T}}_{\omega_k} = \underline{\mu}_{\omega_k} + [\underline{u}_{1\omega_k} \ \underline{u}_{2\omega_k} \ ..\underline{u}_{i\omega_k}.. \ \underline{u}_{K\omega_k}].w_{\omega_k}^T, \ \omega_k \in \{'I','C','P','M'\} \qquad (3.3.3)$$

- Select K eigenvectors for each set whose corresponding eigenvalues are the largest. Based on experimental study in section 3.3.7, we used K = 15. The eigenvectors are then arranged in a descending order according to their corresponding eigenvalues.

- The image of a view-normalized tooth t_{qN} and each of its four approximations $(\hat{T}_I, \hat{T}_C, \hat{T}_P, \hat{T}_M)$ are feed to a least square error classifier. $\varepsilon_{\omega k}$, the square of the representation error between t_{qN} and each of the four approximation, is defined as follows:

$$\varepsilon_{\omega_k} = \sum (\hat{\underline{T}}_{\omega_k} - \underline{t}_{qN})^2, \ \omega_k \in \{'I','C','P','M'\},$$

The least square error classifier assigns t_{qN} to ω_k based on the minimum representation error as follows:

$$t_q \rightarrow \omega_k \,|\, k = \arg \min_{<\omega_k>} \{\varepsilon_{\omega_k}\}, \omega_k \in \{'I','C','P','M'\}.$$

(a) EigenIncisors

(b) EigenCanines

(c) EigenPremolars

(d) EigenMolars
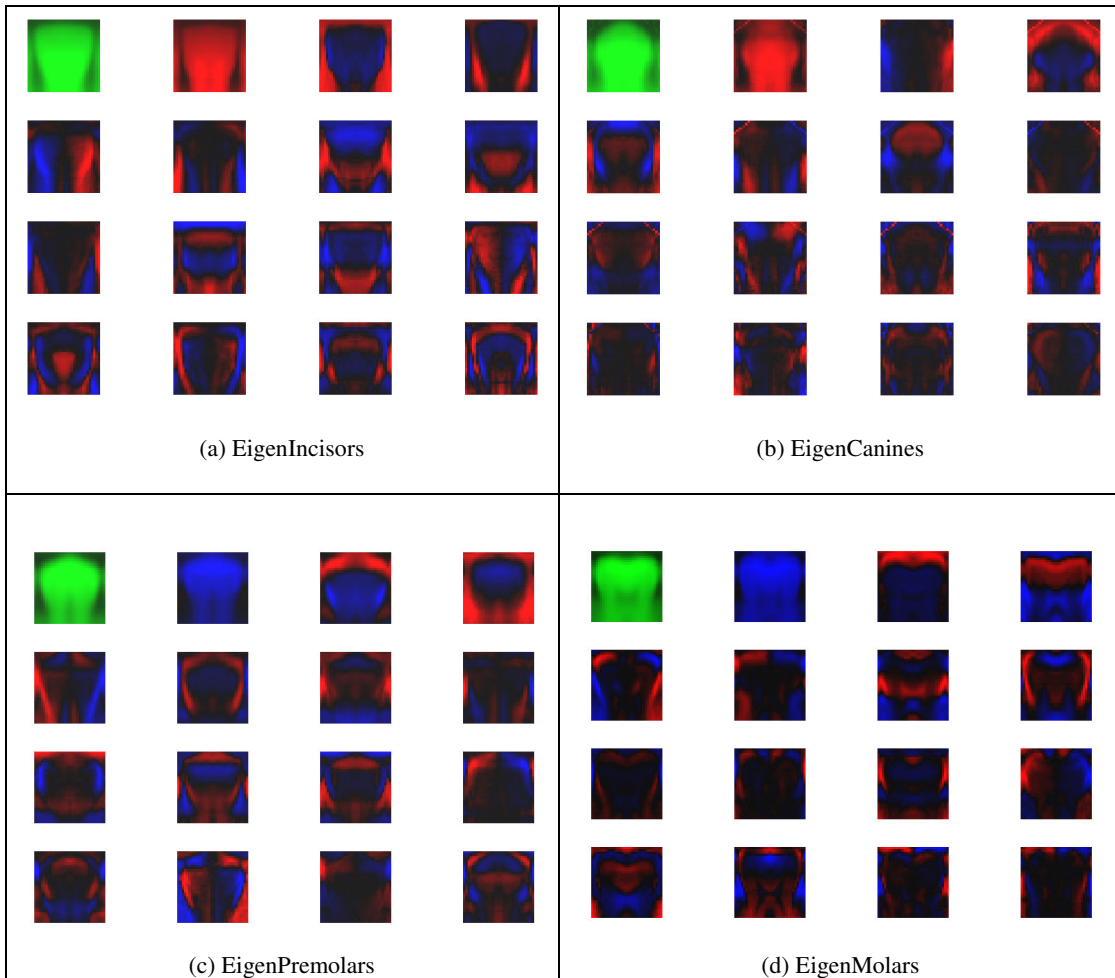
Figure 35: EigenTeeth first principal components representation.

Figure 35 shows image representations of the first 15 eigenvectors of each of the four subspaces resulting from the PCA of four sample sets. These image representations are originally gray images. We selected to use pseudo colors instead, to help presenting the presentation. The image (in green) on the upper left corner of each set represents the sample average (ω_k, $\omega_k \in \{I; C; P; M\}$). The eigenvectors arranged according to their corresponding eigenvalues such that the first EigenIncisor, for instance, is to the immediate right of the average incisor and so on.

3.3.5 Teeth Classification using Linear Discriminant Analysis

Because of some resemblance between the proposed method for teeth classification and the FisherFaces [66] (at least at the abstraction level), we refer to the resulting basis vectors defining the teeth images subspace as the FisherTeeth. While PCA techniques choose a dimensionality reducing linear projection that maximizes the scatter of all projected samples, LDA choose another projection, in such a way that maximize the ratio of between-class (inter-class) and the within (intra-class) class scatter. Figure 36 illustrates the proposed method.
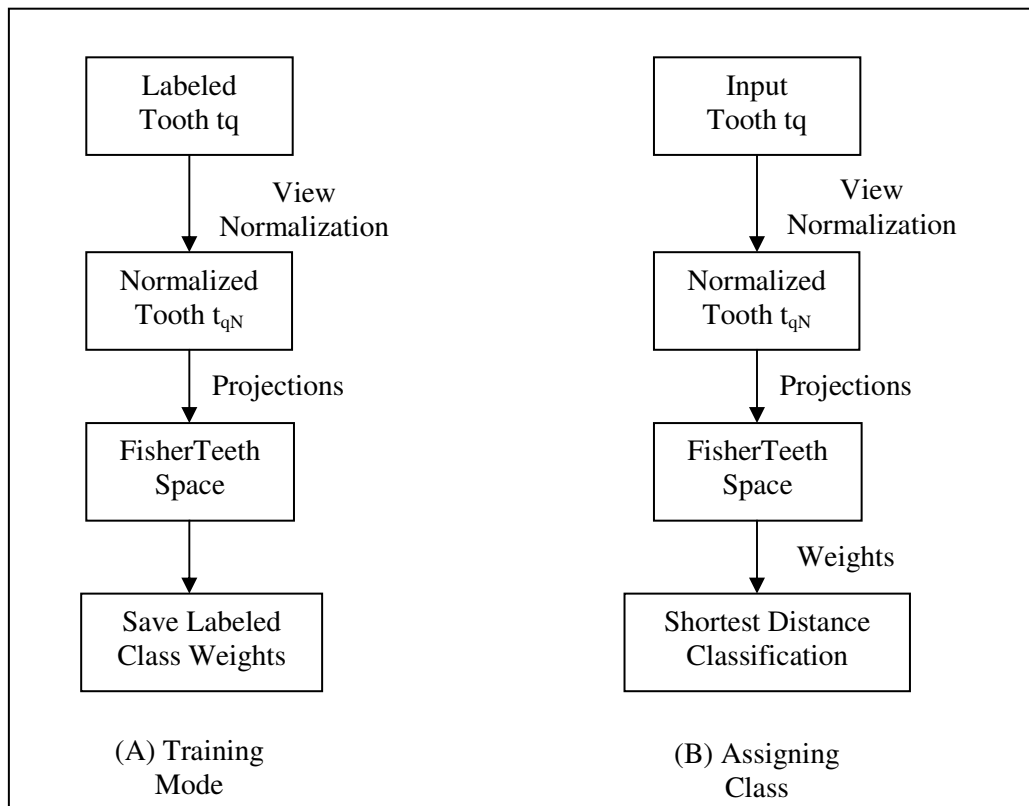
Figure 36: FisherTeeth and initial classification

3.3.5.1 Samplers Preparation

First representative teeth samples of each of the four teeth classes are selected. Figure 37 depicts a sample of exemplars from each teeth class, where different teeth segments sizes and intensity levels are noted. Second each selected tooth segment is view-normalized (as described in sub-section 3.3.3).

(a)Incisor   (b)Canine   (c)Premolar   (d)Molar
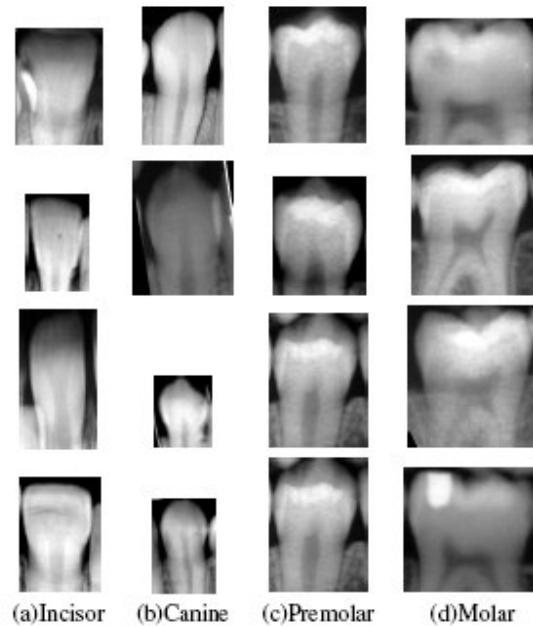
Figure 37: Sample of the four teeth classes

(a)   (b)
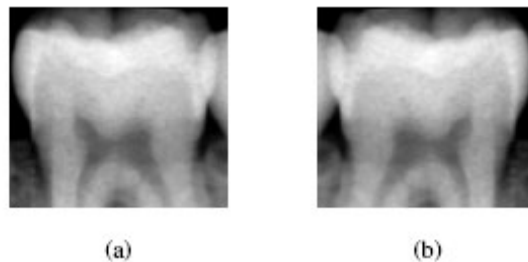
Figure 38: (a) Upward pointing crown; (b) conjugate of (a) under horizontal flipping

Third we augment the exemplar sets with conjugates of the exemplars produced by horizontal flipping (as depicted in Figure 38), i.e. around a vertical axis. The value of this is to substitute for orientation biases in the sample, thus in effect we end up with equal number of exemplars from the left and side quadrants.

3.3.5.2 Fisher Teeth

With the sets of exemplars prepared, we proceed as follows:

- Each training sampler is reshaped as a vector x_i, where the dimension of x_i is (n^2), and as we resize to window of n.n, Based on experimental study in section 3.3.7, we used $n = 32$.

3. Preprocessing Techniques for ADIS

- The N training sampler images are arranged as $\{x_1, x_2 \ldots x_N\}$, where each sample belong to one of the four classes.

- The average (Mean) of each of the four samples is computed:

$$\underline{\mu}_k = \frac{1}{N_k} \sum_{<i>} \underline{x}_{ik}, k \in \{'I','C','P','M'\} \tag{3.4.1}$$

where N_k is the number of exemplars of class k, and $\underline{x}_{ik}$ is the vector representation of the ith k exemplar. Also the general mean (μ) for all the N samples is computed.

- Let the inter-class scatter be defined as

$$S_B = \sum_{<k>} |N_k| (\mu_k - \mu)(\mu_k - \mu)^T, k \in \{'I','C','P','M'\} \tag{3.4.2}$$

and the intra-class scatter matrix defined as

$$S_W = \sum_{<k>} \sum_{x_i \in x_k} (x_i - \mu_k)(x_i - \mu_k)^T, k \in \{'I','C','P','M'\} \tag{3.4.3}$$

- If S_w is nonsingular, the optimal projection W_{opt} is defined as follows:

$$W_{opt} = \arg\max_W \frac{|W^T S_B W|}{|W^T S_W W|} = [w_1, w_2, \ldots w_m] \tag{3.4.4}$$

where $[w_1, w_2, \ldots w_m]$ is the set of generalized eigenvectors of (Opt=S_B/S_W), which is computed by first solving the eigenvalue problem $S_B w_i = \lambda_i S_w w_i, i = 1,2,\ldots m.$, where λ_i is the ith eigenvalue of (Opt=S_B/S_W).

- Select K eigenvectors for each set whose corresponding eigenvalues are the largest. Based on experimental study in section 3.3.7, we used K = 15. The eigenvectors are then arranged in a descending order according to their corresponding eigenvalues.

Figure 39 shows image representations of the first 16 eigenvectors resulting from the LDA of four sample sets. These image representations are originally gray images. The eigenvectors arranged according to their corresponding eigenvalues such that the first FisherTeeth, for instance, is the upper left incisor and so on.
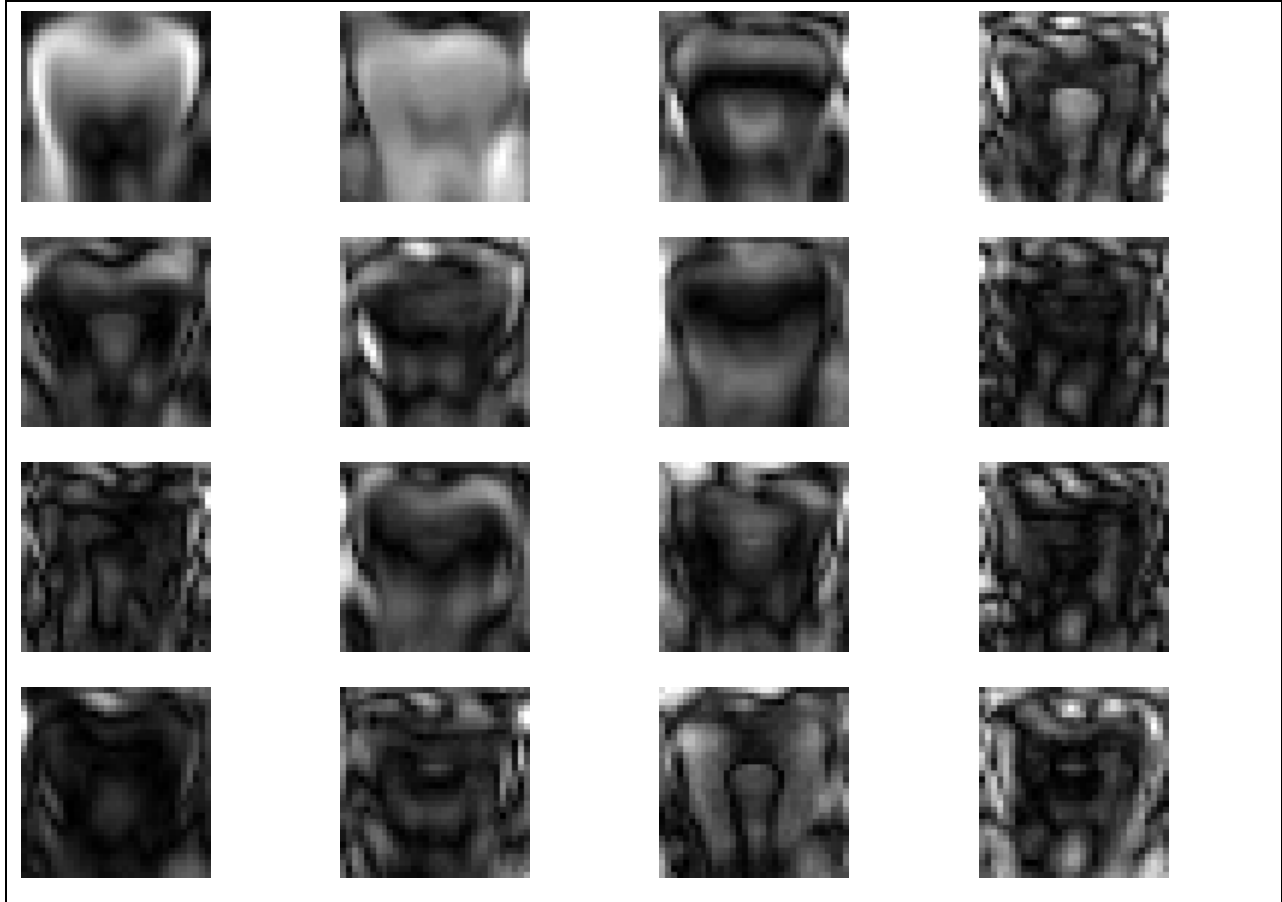
Figure 39: FisherTeeth first principal components representation.

3.3.6 Teeth Sequence Validation and Labeling

As in most of the classification problems, the initial class labels assigned to each tooth are prone to errors. However, a dental film usually shows a number of teeth, and because the assortment of teeth in a human mouth follows a specific pattern, we rely on teeth neighborhood rules to validate the detected sequence of teeth class labels. We also seek to correct, if possible, sequences that do not conform to the reference pattern of possible sequences. And finally if we find the validated/corrected sequence unique, we assign a number to each tooth corresponding to its position in its dental quadrant (as described in Figure 40). In this figure, the lower sequence was correctly classified, so in the teeth labeling stage it was validated and labeled as (4P - 5P - 6M - 7M). The upper sequence wasn't correct, in this case, so it was corrected to (CPPMM) and then labeled (3C - 4P - 5P - 6M - 7M).

Our approaches for class validation is based on string matching with don't care character. When validating bitewing sequences, the horizontal distance between teeth in the upper and lower jaws is taken into consideration. So let:

- X denote the 16 character reference string 'MMMPPCIIIICPPMMM'; and corresponds to labels 'L8-L7-L6-L5-L4-L3-L2-L1-R1-R2-R3-R4-R5-R6-R7-R8' according to Palmer Notation Teeth Numbering System.

- $S_F = s_1...s_j...s_n$ such that; $1 \le n \le 16$ and $s_j \in$ ('I', 'C', 'P', 'M'), denote the sequence of the initially assigned labels of the segmented teeth of the radiographic film F.

- Incase of detecting a missed tooth (segment); $S_F = s_1..s_j.s_i..s_n$ such that $1 \le n \le 16$, $s_i \in$ ('I', 'C', 'P', 'M'), and $s_j = d$, where 'd' is a don't care and can be any value ('I', 'C', 'P', or 'M'). S_F denotes the sequence of the initially assigned labels of the segmented teeth from upper or lower jaw of a radiographic film.
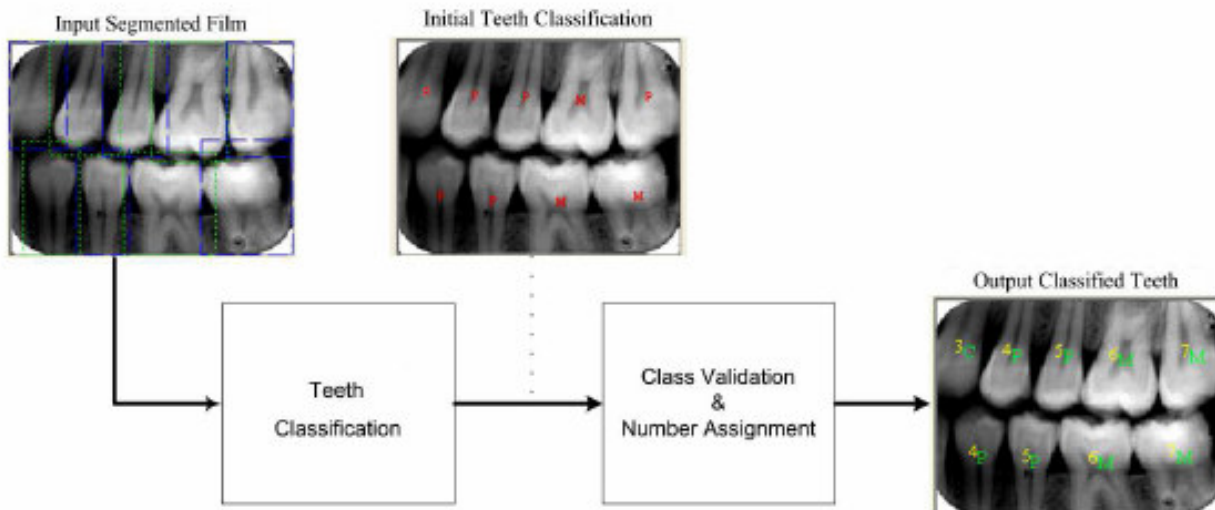
Figure 40: Block diagram of the two stages of teeth labeling

Based on our understanding and studying to the problem of handling teeth labeling for a sequence that contains a missed tooth. To decide whether we have a missed segment (tooth) or not, we have to have the following:

- Presence of 2 or more segments per sequence (one jaw).

- The missing tooth is in the middle of the sequence; otherwise it will not affect the later processing.

- The missing tooth is leaving a gap with width (d).

- So the proposed technique will detect a missing tooth if the gap width between two consecutive segments (d) > c.W_{avg}, where c is a constant that is experimentally defined to 0.39, and W_{avg} is the average width of the two neighbor segments (W_1+W_2)/2, as shown in Figure 41.
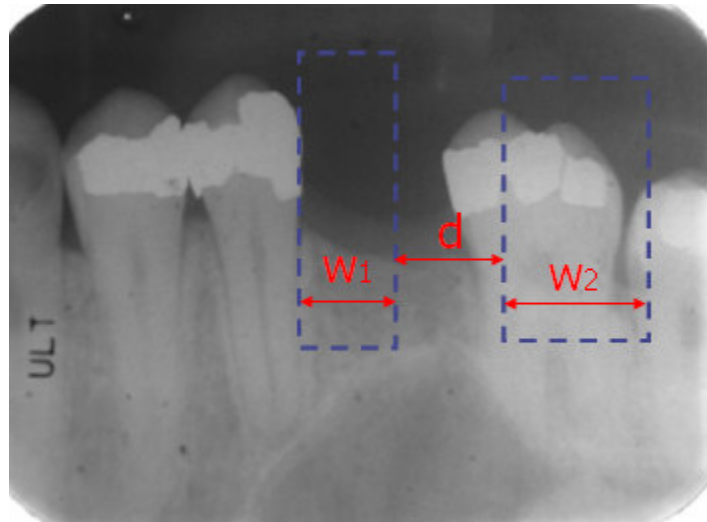
Figure 41: Missing tooth detection

We post the class validation problem as a string-matching problem with error [73] and it may include one don't care character [67][68]. We seek to match the pattern S_F to the text X with the possibility of error in the former. A number of string matching algorithms have been reported in the literature [69][70]. Generally, string matching can be thought of as the maximization/ minimization of a certain cost function such as the edit distance. The use of edit distance began in the 1970s [71]. A key paper by Wager and Fischer proposed the fundamental of edit distance algorithm for string matching with errors and they showed that it was optimal [68]. Here we are interested in a simplified version of their edit distance algorithm, which allows only for exchange character (replace a letter from *SF* by the corresponding character from X).

For the initially correct case, when S_F is a factor of text X, S_F will be assigned the corresponding label. For cases when S_F consists of only one label, or two labels that are not having unique location as ('MM' and 'II'); we assign class only ('M', 'P', 'C', 'I') and not number. Still assigning class only can guide the image comparison stage and reduce the number of required tooth-to-tooth registration.

For cases, when S_F is not a factor of X (due to an error in initial classification), we are interested in finding the closest match, and to change S_F to $S_{F'}$, where $S_{F'}$ will be a factor of X. To measure the similarity (or difference) between the two strings, we need to define the cost of changing from one label

to another. One simple solution was to assign equal cost of say one (as suggested in [73]) for each label change. This solution means assigning equal cost for changing the initial class to any new class. As we have the initial labeled assigned using least square error classifier, we used this prior information to defined the cost of changing the class label of the ith tooth from its initially assigned class k_o to the label of class k (where k $\in$ ('I', 'C', 'P', 'M')) as follows:

$$c_{k_0 \to k}^{[i]} = \frac{\varepsilon_k^{[i]} - \varepsilon_{k_0}^{[i]}}{\max_{<j>} (\varepsilon_j^{[i]}) - \varepsilon_{k_0}^{[i]}}$$

where $\{\varepsilon_k, k=1,..,4\}$ to denote the energy discrepancies (distance) between the image of a view-normalized tooth and the four classes. Accordingly we define the total cost $C_{S_F \to S_F^`}$ for changing the pattern $\mathbf{S_F}$ to $\mathbf{S_F^`}$ as $C_{S_F \to S_F^`} = \sum_{<i>} c_{k_o(i) \to k(i)}^{[i]}$. Of all the possible changes, if a change is required due to impossibility of matching $\mathbf{S_F}$ to $\mathbf{X}$ without errors, we seek $\mathbf{S_F^`}$ to minimize the cost (or *edit distance [72]*) $C_{S_F \to S_F^`}$. In case of having one don't care character 'd' at the i$^{th}$, we define the total cost $C_{S_F \to S_F^`} = \sum_{<i>} c_{k_o(i) \to k(i)}^{[i]}$ for $i \neq 1$; in other word there is no cost to change 'd' to any of the four classes.

There are instances where a match exists between $\mathbf{S_F^`}$ (or $\mathbf{S_F}$) and X at more than a unique shift, e.g. if $\mathbf{S_F^`}$ (or $\mathbf{S_F}$) $\in$ {'MM','MMM','*PP*'}. This situation specifically appears when $\mathbf{S_F^`}$ (or $\mathbf{S_F}$) exhibits no transitions between teeth classes. When we encounter such a situation with a periapical film, we declare a teeth-numbering tie. However, when we encounter such a situation with bitewing film, and if the sequence resulting from the opposite quadrant is unique, we attempt to determine a unique sequence. For example if we have one sequence as 'MMM' (can be 'L8-L7-L6' or 'R6-R7-R8'), and the opposite quadrant has a sequence like 'MMP' or 'PMM' this means that the tie is solved.

Moreover, with bitewing views we detect, and either correct or identify, instances where the resulting sequences of the upper and lower quadrants are inconsistent with one another, i.e. crisscrossed quadrants. For example having a case where is one sequence is 'PPM' corresponds to 'R4-R5-R6' and the other is 'MMM' corresponds to 'R6-R7-R8'. For these instances they can be corrected based on their cost functions, i.e. shifting one at a time and select the consistent sequence that has minimum overall cost. On the other hand films as the one shown in Figure 43 are declared as rejected bitewing film due to inverted upper and lower quadrants. The upper sequence is this case is labeled right ('R5-R6-R7'), while the lower

sequence is labeled left ('L7-L6-L5'). In other word one of them is labeled wrong that is why the system declared this as a bitewing tie.

3.3.7 Experimental Results

To setup the image subspaces for the four teeth classes, we prepared a data set of exemplars using the dental image database provided by the Missing and Unidentified Person Unit of Washington State Patrol [47]. We segmented and preprocessed 100 molars, 100 premolars, 30 canines, and 60 incisors from records. We also used the exemplars produced by horizontal flipping of these preprocessed teeth. We have a total of 580 teeth exemplars for training. In selecting these teeth exemplars, we tried our best to avoid biased data sets. We evenly used teeth from upper and lower jaws. We used teeth from the right and left sides of the mouth. We used teeth images of film having different intensity and contrast.

To select the optimal normalized teeth image size, we conducted a preliminary experiment based on the selected exemplars. We used 220 teeth for training and 130 teeth for testing. We found that moving from image size 16x16 to 32x32 enhances the classification performance. Moving to 64x64 didn't enhance the classification performance any more, while it takes more processing time. We returned this to the geometric properties of segmented teeth, as going to a higher dimension means that we will magnify most of the teeth images.

Based on the above mentioned setup, we conducted a second preliminary experiment to select K eigenvalues. To add up to 99% of the total eigenvalues, we have to use $K = 104$ eigenvalues. To add up to 95%, we have to use $K = 51$ eigenvalues. We tested the performance of these subspaces versus the reduced subspaces [66]. We varied the number of eigenvectors used as basis for the four subspaces and found that keeping the first 15 eigenvectors per subspace achieves the best teeth classification performance compared to keeping 5, 10, 20, 25, 51 or 104 eigenvectors. Ideally the performance should be better with the use of more eigenvectors, but this is not the case for the noisy teeth spaces. A second benefit of reducing the number eigenvectors is to reduce the classification time. The results reported herein are thus based on experiments using 15 eigenvectors per subspace.

Third: to decide whether we have a missed segment (tooth), we conduct some segments measurements on a hundred of sequences (each of them consists of three neighbor teeth). We found that 99% the distance between the first and the third (width of the missing tooth) was greater than 0.39 of the average width of the first and the third. We decide to have $c = 0.39$.

We started by a small experiment studying the initial classification performance for the Fisher teeth compared to Eigen teeth/reconstruction method. We partitioned the training samplers, where we used (18 Incisors, 32 Canines, 40 Premolars, and 40 Molars) for testing.

		Assigned Class			
		I	C	P	M
True Class	I	2	0	0	16
	C	0	28	0	4
	P	0	0	40	0
	M	0	0	0	40

(a) Eigen Teeth Classification

		Assigned Class			
		I	C	P	M
True Class	I	11	3	0	4
	C	0	32	0	0
	P	0	0	40	0
	M	1	0	0	39

(b) Fisher Teeth Classification

Table 5: Performance comparison

The experiment result (Table 5) shows that the error rate for the suggested method is 6.2% compared to 15.3% Eigen teeth/reconstruction method. Notice that the main advantage of the Fisher teeth method is in classifying Incisors which are rare in the available database.

To test the proposed approach for teeth classification and numbering, we prepared a disjoint test set with the data set of exemplars described earlier. 1675 teeth were correctly segmented and used for the teeth labeling experiment (814 molars, 806 premolars, 30 canines, and 25 incisors). In preparing the test set we used the CJIS ADIS project dental image database [48]. Table 6 lists the number of films and teeth used in our testing experiments categorized by film type.

	No. of films	No. of teeth
Bitewing films	269	1321
Periapical films	238	354
Total	507	1675

Table 6: Number of films and teeth used in testing

For the Initial classification, we also pursued Fisher Teeth method, but it achieves almost the same performance as Eigen teeth/reconstruction method; hence we find no merit showing results due to Fisher Teeth method. Using two classes (molar and premolar) only with bitewing films boost the initial teeth classification of bitewing films from 65% to 80%, and hence enhance the overall initial teeth classification accuracy by 10%.

The teeth classification confusion matrices for the three types films we tested are given in Table 7 (all films), Table 8 (bitewing films only), and Table 9 (periapical films only). We observe that overall, initial classification works with 75% accuracy, teeth class validation enhances the overall teeth classification accuracy to 87%. The accuracy of intra-quadrant numbering is 70%; however, 85% of assigned intra-quadrant numbers are within one position from the true value.

		Assigned Class			
		I	**C**	**P**	**M**
True Class	**I**	12	4	9	0
	C	4	2	21	3
	P	28	8	751	19
	M	5	2	311	496

Overall Performance 75.28%

(a) Eigen Teeth Classification

		Assigned Class			
		I	**C**	**P**	**M**
True Class	**I**	11	5	8	1
	C	3	10	16	1
	P	5	15	752	34
	M	4	3	128	679

Overall Performance 86.69%

(b) After Validation

Table 7: Confusion matrix for teeth classification in all films

		Assigned Class			
		I	**C**	**P**	**M**
True Class	**I**	0	0	0	0
	C	0	0	12	0
	P	0	1	648	14
	M	0	0	249	397

Overall Performance 79.11%

(a) Eigen Teeth Classification

		Assigned Class			
		I	**C**	**P**	**M**
True Class	**I**	0	0	0	0
	C	0	4	8	0
	P	0	3	638	22
	M	0	0	98	548

Overall Performance 90.08%

(b) After Validation

Table 8: Confusion matrix for teeth classification in bitewing films

		Assigned Class			
		I	**C**	**P**	**M**
True Class	**I**	12	4	9	0
	C	4	2	9	3
	P	28	7	103	5
	M	5	2	62	99

Overall Performance 63.84%

(a) Eigen Teeth Classification

		Assigned Class			
		I	**C**	**P**	**M**
True Class	**I**	11	5	8	1
	C	3	6	8	1
	P	5	12	114	12
	M	4	3	30	131

Overall Performance 74.01%

(b) After Validation

Table 9: Confusion matrix for teeth classification in periapical films

Figure 43 shows examples of successful teeth numbering of bitewing and periapical films respectively. Figure 43 –a) shows an example of incorrectly numbered in periapical films. Figure 43 –b) shows an example of a rejected bitewing film due to a quadrant inversion situation that could not be resolved by the system.
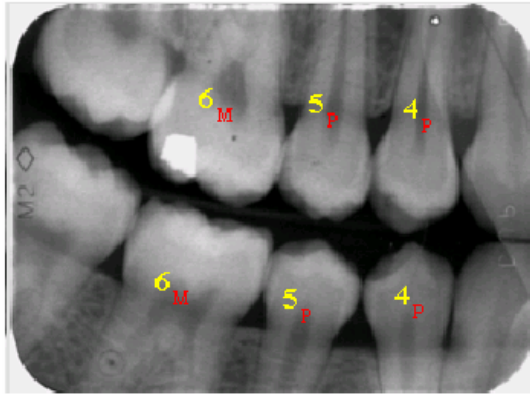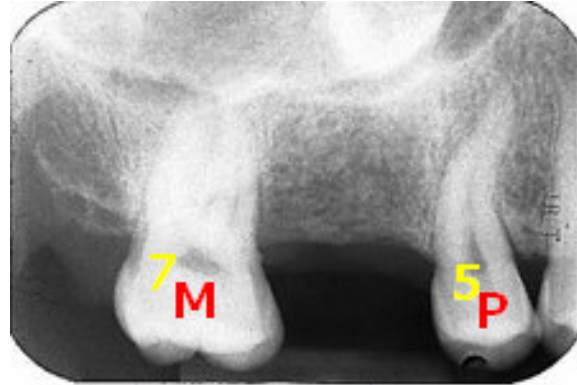
(a) Correct Bitewing Film   (b) Correct Periapical Film with missing tooth

Figure 42: Examples of films with correct teeth numbers
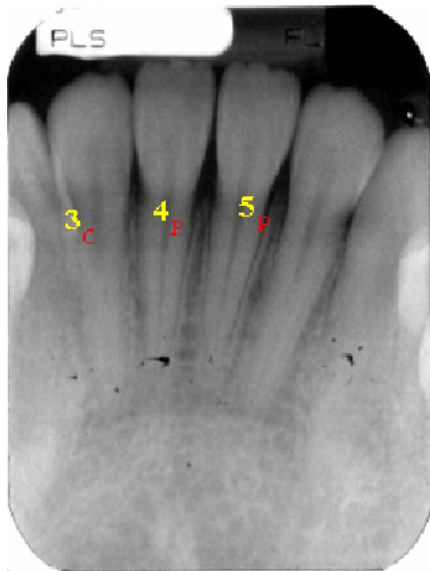
(a) Incorrect Periapical Film   (b) Tie Bitewing Film
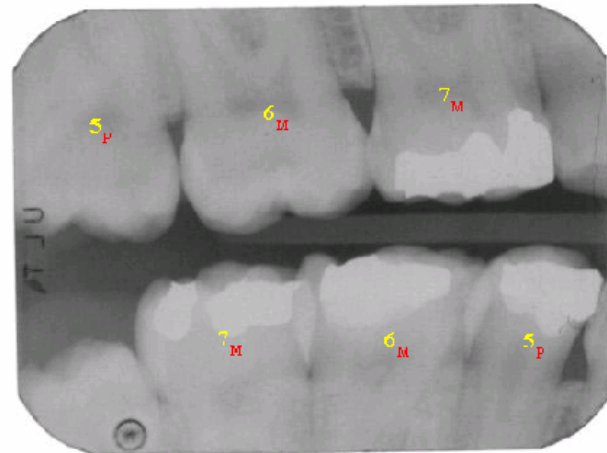
Figure 43: Examples of a) films with incorrect teeth numbers and b) unsolved tie

We also conducted a second experiment on the same dataset used by Mahoor and Abdel-Mottaleb [15], where the films are bitewing only (i.e. two class problem), and the background of segmented teeth are eliminated. The Pre-Classification performance of the proposed method is 91.67% for molars compared to the best performance 90.55% (average performance between mandible and maxilla for the posted results in [15]). While our method performance was 92.86% for premolars compared to the best performance 90.95% as shown in Table 10. The average tooth classification time for their method [15]is 0.0212 sec/tooth (without adding the processing time needed for contour extraction). While to classify a tooth using our proposed method, it takes 0.0115 sec/tooth. This means that our proposed method cuts the

tooth classification time by half approximately. On the other hand if we add 0.16 sec/tooth for the fastest tooth contour extraction [43], which means our proposed technique cuts the initial classification time to 7%.

Method	Molars Average	Premolars Average	Labeling Time
Complex Signature	89.6%	90.95%	0.0213 sec
Centroid Distance	90.55%	87.85%	0.0213 sec
Eigen Teeth	91.67%	92.86%	0.0115 sec

Table 10: Pre-classification performance comparison.

The approach that Mahoor and Abdel-Mottaleb used [15] for validation of sequences, didn't fully exploit neighborhood relations in bitewing films, particularly they didn't use inter-quadrant sequences for sequence correction and hence the numbers assigned to the different teeth on the upper/lower quadrants may include inconsistencies (e.g. quadrant inversion). Their approach isn't dealing with periapical films only, which represents 42% in the used CJIS ADIS project dental image database. They didn't deal with films that have missed teeth, 21% of the database.

The main advantage of the work presented by Jain and Chen [36] is that by using a Hidden Markov Model for registration of teeth to a human dental atlas, they have implicitly embedded the capability of learning the proper misclassification costs. They claimed 88% performance accuracy for their technique. However: 1) their technique classifies (canines and premolars) as (Bi)cuspid, in other word, they are doing a three class problem. Given that classifying to a smaller number of classes means better performance; 2) they added panoramic radio-graphed films to their test set, and they didn't specify the percentage of these films. Finally they didn't indicate that they used any inter-quadrant validation, which we think is important particularly in bitewing views.

We returned the lower performance for teeth labeling of periapical films to the following:

- For periapical radio-graphed films, it is initially a four class problem, compared to two in bitewing radiographed films,

- For bitewing films, we made use of the existence of upper and lower to detect and either correct or identify instances where the resulting sequences of the opposite quadrants are inconsistent with one another.

We didn't use panoramic films (According to Hajsaid et al [20], the only existed dental segmentation technique for the panoramic films are semi-automated). It appears as an advantage to include the panoramic radio-graphed films, but complete panoramic views are solved cases for the stage of teeth labeling. In other word as the complete panoramic views contain the 32 teeth

('MMMPPCIIIICPPMMM') in each jaw and no need for initial classification. Also the presence of panoramic views is rare (4%) in the used CJIS ADIS project dental image database [48].

We observe that premolars and molars are more often used for identification purposes. A careful look at the confusion matrix presented in Table III (a) reveals that, overall premolars are initially highly correctly classified 93%; we also observe that misclassified molars are predominantly seen as premolars. The performance of the class validation scheme, and hence the overall performance of teeth numbering and classification, depends on how well we assign the cost (or risk) of misclassification.

We notice that we can improve the overall teeth-numbering accuracy, by adjusting the misclassification risk, such that we take into consideration the observations we just made on the molars misclassification rate, and the accuracy of premolars classification. The adjustment, we made is such that it always costs less to reassign a premolar as a molar. Overall this led to approximately the same accuracy in the classification of premolars while enhance the accuracy in molars classification to 83%, overall this enhance the initial classification accuracy to 87%.

4   ADIS Potential Match Search

Recent disasters have brought the significance of automated dental identification systems (For example 9/11 attack, and Tsunami). For these mass disasters most of the victims were identified based on manual dental information [82]. As can be concluded that manual dental identification is very powerful postmortem identification but one of its main limitation is also the time factor. This is the main motivation for building the Automated Dental Identification System (ADIS), and for that system to have a very powerful searching engine (which we call it potential Match Search component).

4.1   Introduction

ADIS starts with dental record preprocessing component, followed by the Potential Matches Search component that manages archiving and retrieval of dental records from the database and produces a short candidate list. The final component of ADIS is Image Comparison which handles one to one comparison against the candidate list to produce a match list (as shown in Figure 1).

In the context ADIS, the potential Match Search component is a fast search engine that reduce the database to be searched by a more time expensive image comparison. Given that the record to record comparison encounters an average of 10 tooth to tooth comparisons (as shown in Figure 44). The image comparison component takes about 20 minutes to make record to record comparison. To search a database of 50 records directly using the image comparison, the system will run for 17 hours. We measured the processing time for running the image comparison component using MATLAB® implementation running on a 2.66 GHz 1.0 GB RAM Intel® Pentium IV PC platform. To directly address the scalability of ADIS, we need the potential Match Search component to achieve very high percentage of having the target reference in a relatively very short candidate list.

In my opinion, the problem of implementing a potential match search was not handled in a proper way in the context of ADIS [14][18][16][17][35][44]as all the present trial are based on contour alignment and matching to find the distance between two dental images. All these techniques made a tooth to tooth comparison; in other word they are implementing the image comparison stage, and not a fast potential match search.
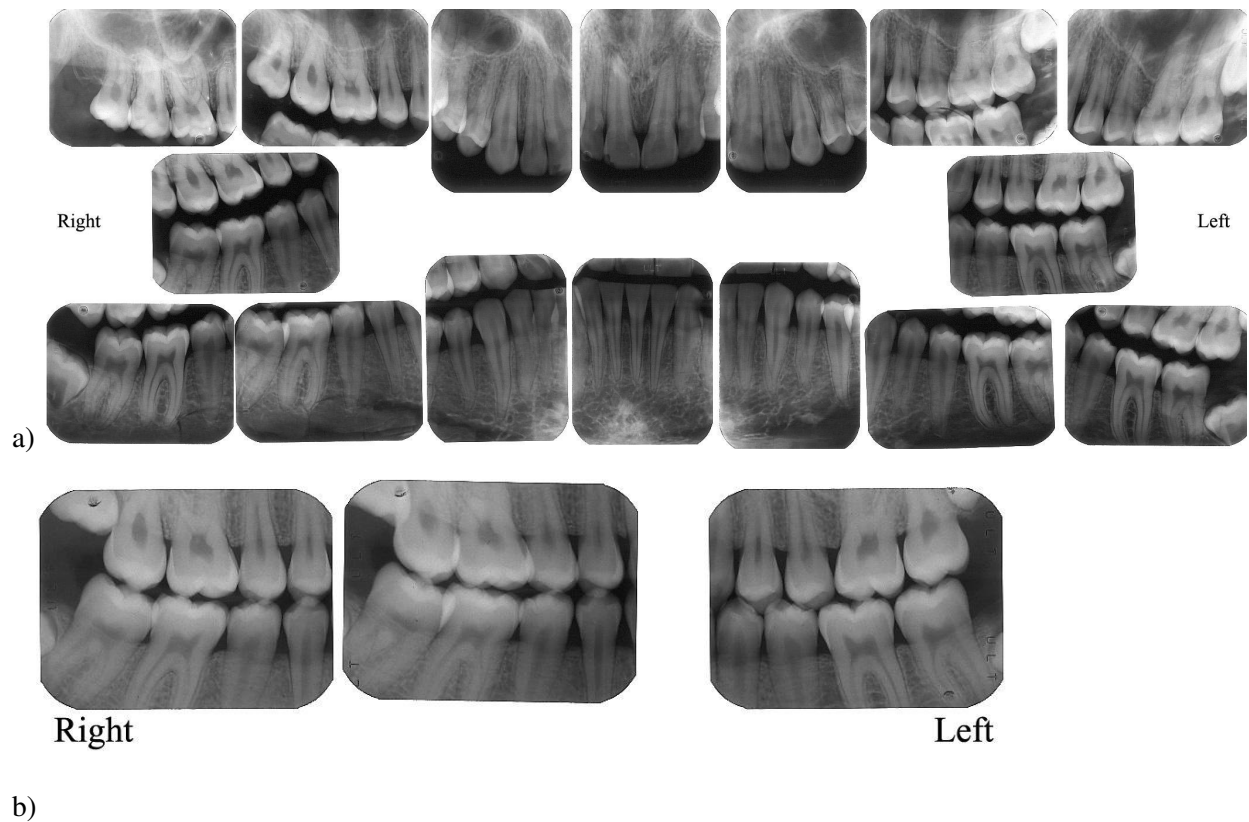
Figure 44: Dental records for the same person; a) AM record, b) PM record

Nomir and Abdelmottaleb [44] used a technique that is based on matching teeth contours using Hierarchical Chamfer distance. They proposed to use 3-4 distance to approximate the Euclidian distance with fast calculating method. They used a hierarchical edge matching algorithm at different resolution, in other world they found an approximate solution at lower resolution (lower calculation cost) that guides the solution at finer resolution. Nomir and Abdelmottaleb [44] was the first to consider the time factor, however their method would be more powerful, if they consider using the lower resolution matching for tiding their search domain while moving to finer resolution.

In this chapter, we introduce two approaches for fast retrieval from the dental database (as shown in Figure 45):

- The first approach relies on using color feature, namely Eigen images. These are low computational-cost, appearance-based features, as discussed in section 4.2. In section 4.2.1, we introduce the view normalization step; which we use to align all the segmented teeth vertically. In section 4.2.2, we introduce the theory behind the Principal Component Analysis (PCA), and then we elaborate on projecting each tooth in a lower dimensional space, where the tooth-to-tooth comparison will be with a very low computational cost. In section 4.2.3, we present the

experimental results on the performance of the proposed appearance-based method, followed by discussion on these results.

- The second approach relies on teeth contours descriptors (shape-based approach), where we use the moment invariant and edge orientation histogram as features to describe the tooth contour, as discussed in sections 4.3.1 and 4.3.2 respectively. In section 4.3.3, we present the experimental results on the performance of the proposed methods, followed by discussion on these results.

- In section 4.4, we present a comparison experiment between the performance of the proposed appearance-based method and feature-based method, followed by discussion on these results.
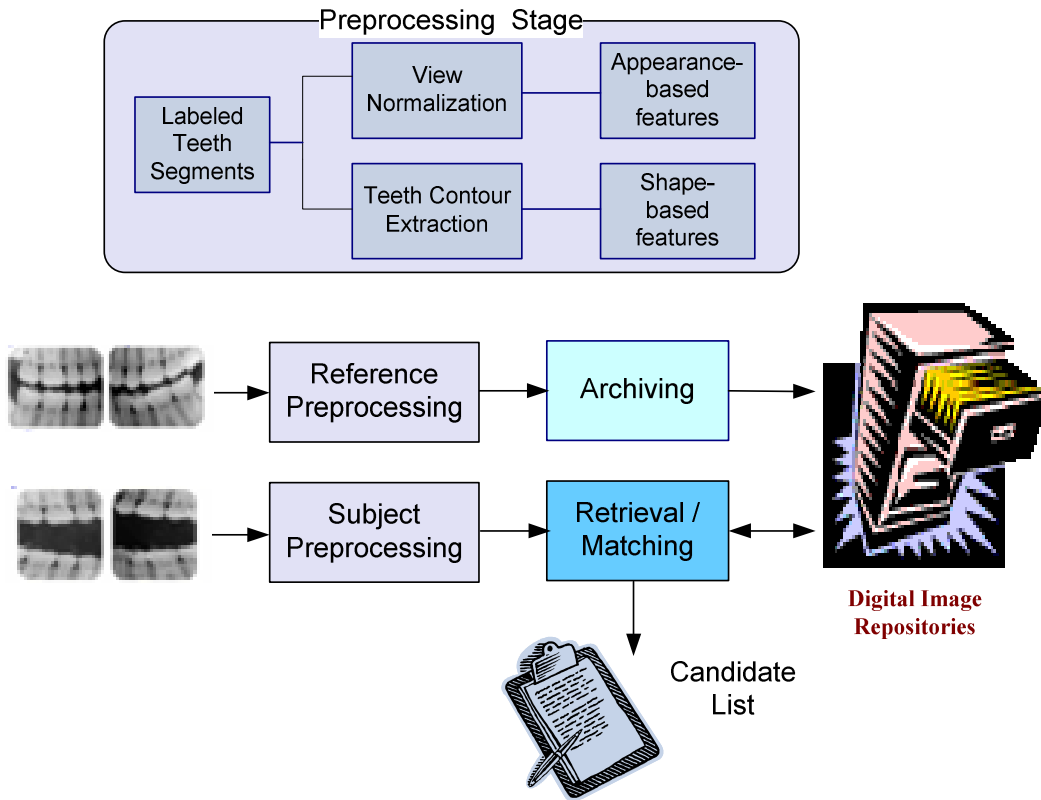
Figure 45: ADIS potential match search component block diagram

4.2   Eigen Image (Appearance-Based)

The proposed approach for potential match search using eigen images is illustrated in Figure 46, where the subject tooth and the reference tooth images are view-normalized to compensate for the geometric and intensity different. For the single class approach, all the teeth are project to a single class which is based on projection of all the four teeth training samples. In other word, it is a single teeth space, which we call

it EigenTeeth. While for the four classes approach, according to the teeth class, we project it to a lower dimension PCA sub-space (EigenMolar, EigenPremolar, EigenCanine and EigenIncisor). In other words, we have four teeth spaces. This Euclidian distance between the weights of the principal components represents the match score on a tooth level. Finally we integrate the different matching scores of subject and reference teeth to have one matching score between the subject and reference records, and hence retrieve the candidate records.
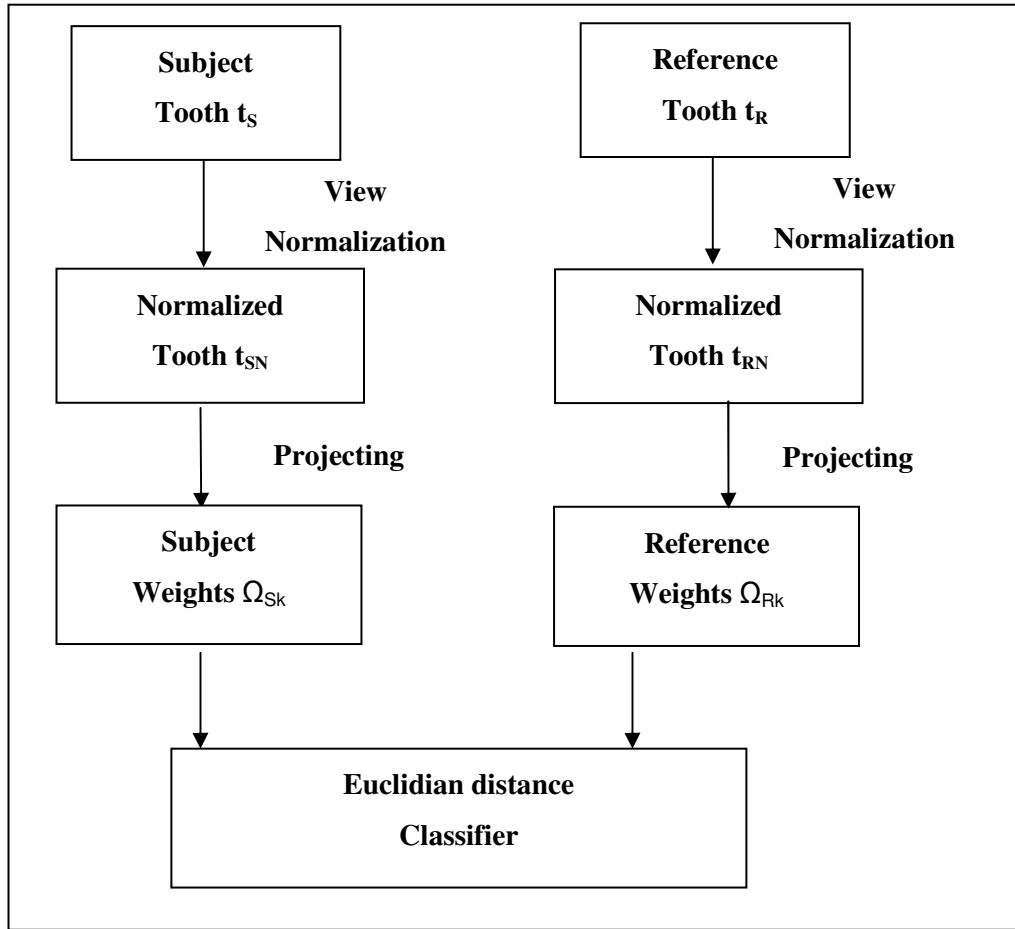
Figure 46: Block diagram for the Eigen image retrieval approach

4.2.1 View Normalization

In practice, dental radiographic films show teeth that need not comply with the standardizing assumptions on view that we made. On one hand, crowns need not appear on the topside of the segmented regions, as in the case with upright films showing teeth of the upper jaw (or flipped films showing teeth of the lower jaw). While some of the teeth segmentation algorithms annotate teeth as either upper or lower, and thus partly solve the first type of deviations from standard view [14][16][17], other teeth segmentation algorithms do not discriminate between upper and lower teeth [20]. To accommodate the different teeth

segmentation algorithms, a simple test is conducted for bitewing films (For Periapical films; the film type classification stage differentiates between upper and lower periapical films). This test advises based on whether the center of the segmented tooth is closer to the top or the bottom part of the bitewing film. On the other hand, the lateral surfaces of radiographed teeth do not necessarily appear as predominantly vertical lines. In view normalization, there are two types of transformations that we usually need to address. First is the photometric (or intensity) transformation.

The photometric transformation was addressed at earlier step of dental record preprocessing (film enhancement). Hajsaid et al. [20], proposed a grayscale contrast stretching transformation using a parametric sigmoid transform. Thus they mapped an input grayscale g to T_{Enh} (g, α, β, g_{th}) according to:

$$
T_{Enh}\left(g,\alpha,\beta,g_{th}\right)=
\begin{cases}
g_{TH}\cdot\dfrac{e^{\alpha\cdot\frac{(g-g_{min})}{(g_{th}-g_{min})}}-1}{e^{\alpha}-1} & g_{min}\leq g\leq g_{th} \\[4ex]
g_{TH}+\left(255-g_{TH}\right)\cdot\dfrac{e^{\beta\cdot\frac{(g_{th}-g)}{(g_{max}-g_{th})}}-1}{1-e^{\beta}} & g_{th}\leq g\leq g_{max}
\end{cases}
\tag{4.2.1}
$$

where α, and β designate compression and expansion factors, respectively, g_{th} is a threshold grayscale between the compression and the expansion regions of T_{Enh}, g_{TH} is the image of g_{th} under T_{Enh}, g_{min} and g_{max} are the minimum and maximum grayscales of the input image, respectively. We choose g_{th} such that its proportionality to the new grayscale range [0; 255] resembles that of g_{TH} to the grayscale range of the input image. Thus $g_{TH} = 255.(g_{th}-g_{min})/(g_{max}-g_{min})$.

The next step towards normalizing the view of a segmented tooth is to ensure that its lateral surfaces appear predominantly vertical. Starting with a tooth with an upward oriented crown as an input image (as shown in Figure 47-a), the view normalization task proceeds in the following steps:

- According the tooth class, the tooth is aligned to the corresponding models (as shown in Figure 47-b), using the circular projection method (as shown in Figure 47-c). More details on this method are given in section 4.2.1.1.

- For model alignment, as we didn't build one model to represent each teeth type (molar, premolar, canine, and incisors), we are registering against more than one vertical model for each teeth type. We are using the mutual information [105] as an indication to select the optimal alignment (as shown in Figure 47-b).

- The optimal vertical orientation is computed, and its negative is used as the rotation angle for the input image such that the lateral surfaces of the tooth appear predominantly vertical (as shown in Figure 47-d).

- An edge image is calculated from the input image. By analyzing the vertical and horizontal span of the tooth, the tooth ROI is confined to a square area, whose edge is just above the predominantly horizontal line segment belonging to the crown and such that the lateral sides of the tooth are also contained in the square area (as shown in Figure 47-e).

- The resulting square images are then resized to a fixed sized image of n.n pixels. Based on experimental study in section 3.3.7, we used n = 32.

(a) Tooth Image (b) Vertical Models (c) Angle Estimation

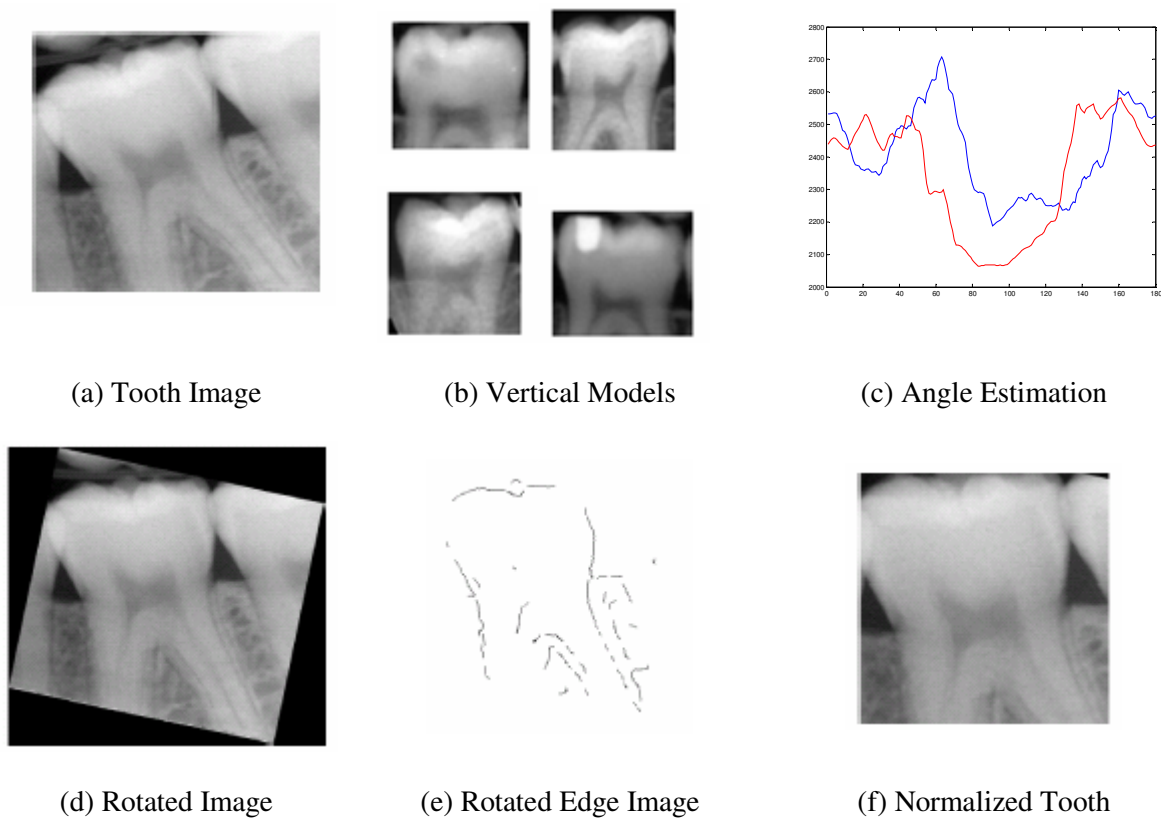(d) Rotated Image (e) Rotated Edge Image (f) Normalized Tooth

Figure 47: illustration of the orientation adjustment and modified view normalization steps of a segmented tooth.

4.2.1.1 Teeth Alignment

In this section a novel projection based approach is proposed to estimate the rotation angle between two images. We use this technique to align the teeth images to enhance the performance of view

normalization stage. The technique is based on circular projection followed by 1-D correlation to estimate the rotational angle between the two images.

Since information gained from two images (source and target acquired at different times, from different imaging modalities, or as consecutive volumetric slices) is usually of a complementary nature, proper integration of useful data obtained from these images is often desired. A first step in this integration process is registration [74].

Automatic image registration is a fundamental component in medical imaging systems, where a target image is to be aligned to a source image. According to Brown [75], registration in the context of medical imaging involves four key components:

(1) the feature space, where features (such as points, surfaces, or intensity values) to be matched are selected in the image pair;

(2) the search space, or range of transformations that best aligns the feature space;

(3) the search strategy for finding the transformation parameters within the search space; and

(4) the similarity metric to measure the optimality of a transformation for the selected features.

To carry out the image registration process automatically, several algorithms have been proposed. Reddy and Chatterji [79] grouped them into the following classes: (1) Algorithms that directly use image pixel values (2) Algorithms that operate in the frequency domain (3) Algorithms that use low-level features such as edges and corners and (4) Algorithms that use high-level features such as identified objects, or features.

Junck et al. [76] propose using correlation analysis to rotate functional brain images to a standard vertical orientation, and to align multiple brain images. Image rotation and centering are performed by determining the angle of rotation and centerline coordinate that results in maximal left-right correlation. Gonzalez, and Woods [49] defined correlation between two real images $I_1(x; y)$ and $I_2(x; y)$ as:

$$Corr(I_1, I_2) = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} I_1(x, y) I_2(x+m, y+n) \qquad (4.2.2)$$

Cain and Hayat [77] propose a projection based method which depends on the image pixel values for registration of image translation. In this method, the projections along the horizontal and vertical axes of the image are computed. Horizontal and vertical components of the translation are then computed using one-dimensional (1-D) cross correlation based estimator. They show that the method is robust in the

presence of temporal and spatial noise and computationally efficient compared to the two-dimensional (2-D) correlation based translation registration.

Khamene et al. [78] also proposed a projection method to reduce the 6 degrees of freedom (DOF) of the conventional volumetric rigid registration. They use successive registration of three (2-D) projection images to perform volumetric registration. In other words, they formulate the volumetric registration as iterative optimization with respect to parameters $(\theta,\ t_x$ and $t_y)$ which characterize the in-plane transformation matrix:

$$ T = \begin{pmatrix} \cos(\theta) & -\sin(\theta) & t_x \\ \sin(\theta) & \cos(\theta) & t_y \\ 0 & 0 & 1 \end{pmatrix} \tag{4.2.3} $$

Reddy and Chatterji [79] use a frequency domain approach. Moving to the frequency domain transfers the translation between two given images into a phase shift, which can be detected using phase correlation. Converting the magnitude of the Fourier transform of these images from rectangular coordinates to log-polar coordinates and by calculating the similarity ratio, they represent rotation and scaling errors also as shifts. These three parameters are used to establish the mathematical model, and then the target image is geometrically rectified with respect to the source image.

In the context of ADIS, Chen and Jain [18][25], Zhou and Abdel-Mottaleb [17] used teeth contour as a feature space. They assumed rigid body transformation and they used sequential quadratic programming as a search strategy. While Nomir and Abdel-Mottaleb [16] assumed quasi-Affine transformation and minimize the distance equation between teeth curves as a search strategy. The two techniques depend on teeth contour features (shape based). Automatic extraction of teeth contour is computationally intensive, and to the best of our knowledge, none of the techniques exploited in teeth contour extraction, so far, achieves this task reliably [43]. Nassar et al. [41] used extracted edges from the teeth regions, assumed Affine transformation and used genetic algorithm as a search strategy. Their process is very time consuming, average tooth-to-tooth alignment of 1 minute.
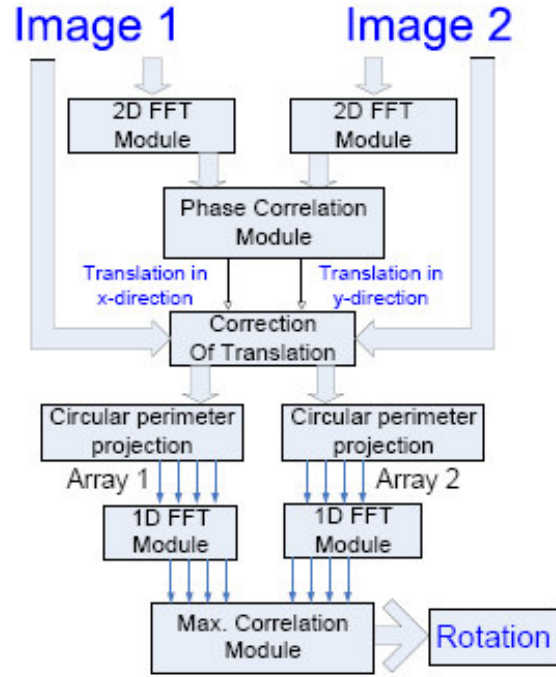
Figure 48: Block diagram of the registration technique

We use circular projection in order to recover the in-plane rotation, without having to do 2-D cross correlation. The first step in this algorithm is to align the two images for any in-plane translation t_x and t_y. A block diagram of the proposed algorithm is shown in Figure 48. We use a method that relies on the translation property of Fourier transform [79], which is referred to as the Fourier shift theorem. Let f_1; f_2 be the two images that differ only by a displacement $(t_x; t_y)$ i.e.

$$F_2(x, y) = F_1(x - t_x, y - t_y) \qquad (4.2.4)$$

Their corresponding Fourier transforms F_1 and F_2 are

$$F_2(\xi, \eta) = F_1(\xi, \eta) \exp(-j2\pi(\xi t_x + \eta t_y)) \qquad (4.2.5)$$

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

$$\frac{F_1(\xi, \eta) F_2'(\xi, \eta)}{\left| F_1(\xi, \eta) F_2'(\xi, \eta) \right|} = \exp(-j2\pi(\xi t_x + \eta t_y)) \qquad (4.2.6)$$

By calculating the inverse Fourier transform we will have an impulse that is approximately zero everywhere except at the displacement location (t_x, t_y). There will be a discernable peak (the maximum value) at this location.

4. ADIS Potential Match Search

For teeth model registration, we replace the translation correction module as follows: we start by zero padding of the two images to size (N.N), then we align the center of gravity of each tooth image to the center of the image (N/2 , N/2).
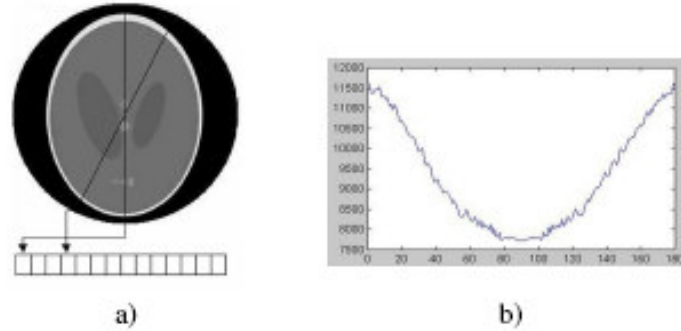
Figure 49: Registration using projection on circular perimeter:
(a) circular projections; (b) radial resampling.

The second step, after aligning any translation between the images, is to consider a hypothetical circular perimeter around the object. The diameter of this circle is equal to the maximum of the two image dimensions (in case of squared image (N . N), we use N). The projection of all diameters of this circle at uniform angular steps μ ranging from 0 to π are computed to make an array (as shown in Figure 49).

$$I(\theta) = \int_{-N/2}^{N/2} I(r,\theta)dr \qquad (4.2.7)$$

Each element in this 1-D array is an integral projection of the points on the corresponding diameter and represents the corresponding angle (what we call it radial re-sampling). Most of these points are not falling into the normal grid, so we calculate their corresponding value using bilinear interpolation. Our hypothesis is that the teeth regions are contained within the hypothetical circle and do not exhibit circular symmetry. The detection of the rotational angle between two images becomes a simple registration of two signals (1-D arrays). For these two resulting signals, the maximum correlation, or phase correlation, corresponds to the rotation angle between the two images.

The correlation between two signals (cross correlation) is a standard approach to feature detection. Gonzalez and Woods [49] describe the convolution theorem and the attendant possibility of efficiently computing correlation in the frequency domain (phase correlation) using the fast Fourier transform. However experimental we find the maximum correlation carried out in the spatial domain to have better performance compared to phase correlation especially with very small rotational angles.

4.2.2   The Eigen Image Methodology

One approach to deal with the problem of excessive dimensionality is to reduce dimensionality, by combining features, or by other methods like finding lower-dimension representative appearance-based features. According to Duda et al. [72], Linearly combining features are simple to compute, and can be tracked analytically. Principal component analysis (or Eigen Image) provides a roadmap for how to reduce a complex data set to a lower dimension to reveal the sometimes hidden, simplified structure that often underlie it. Principal component analysis (PCA), as a linear transformation method, seeks linear projection that best represent the data in a least-square error sense, as follows:

The average (Mean) of each of the four samples is defined as:

$$\underline{\mu}_k = \frac{1}{N_k} \sum_{<i>} \underline{x}_{ik}, k \in \{'I','C','P','M'\} \tag{4.2.8}$$

where N_k is the number of exemplars of class k, and $\underline{x}_{ik}$ is the vector representation of the ith k exemplar.

We define the squared-error criterion function $J_0(x_0)$ by

$$J_0(x_0) = \sum_{<i>} \left\| (x_0 - x_{ik}) \right\|^2, k \in \{'I','C','P','M'\} \tag{4.2.9}$$

and we seek the value of x_0 that minimize J_0. Duda et al. [73] showed that the solution to this problem is given by $x_0 = \mu_k$, but as the sample mean is a zero-dimensional representation to the data set, they obtained a more interesting, one- dimensional representation by projecting the data onto a line running through the sample mean. Through the coming we followed the derived equations by Duda et al. Let **e** be a unit vector in the direction of the line. Then the equation of the line can be written as:

$$x = \mu_k + ae_k, k \in \{'I','C','P','M'\} \tag{4.2.10}$$

where **a** represents the distance of any point x from the k class mean μ_k. We seek an optimal set of coefficients a_{ik} by minimizing the squared-error criterion function:

$$J_{1k}(a_1,...a_n,e_k) = \sum_{<i>} \left\| (\mu_k + a_{ik}e_k) - x_{ik} \right\|^2, k \in \{'I','C','P','M'\}$$

$$J_{1k}(a_1,...a_n,e_k) = \sum_{<i>} a_{ik}^2 \left\| e_k \right\|^2 - 2\sum_{<i>} a_{ik}e_k^t (x_{ik} - \mu_k) + \sum_{<i>} \left\| x_{ik} - \mu_k \right\|^2 \tag{4.2.11}$$

4. ADIS Potential Match Search

Recognizing that $\|e_k\| = 1$, partially differentiating with respect to a_{ik}, and setting the derivatives to zero, we obtain:

$$a_{ik} = e_k^{\ t}(x_{ik} - \mu_k), k \in \{'I','C','P','M'\} \tag{4.2.12}$$

This brings us to finding the best direction e_k for the line. Let the inter-class scatter be defined as

$$S_k = \sum_{<i>}(x_{ik} - \mu_k)(x_{ik} - \mu_k)^t, k \in \{'I','C','P','M'\} \tag{4.2.13}$$

Substituting (4.6) and (4.5) in (4.4), we obtain:

$$J_{1k}(e_k) = -e_k^t S_k e_k + \sum_{<i>}\|\mu_k - x_{ik}\|^2, k \in \{'I','C','P','M'\} \tag{4.2.14}$$

clearly the vector e_k that minimize J_1 also maximize $e_k^t S e_k$. We use the Lagrange multiplier. Letting λ be the undetermined multiplier, we differentiate

$$u_k = e_k^t S_k e_k + \lambda_k(e_k^{\ t} e_k - 1), k \in \{'I','C','P','M'\} \tag{4.2.15}$$

with respect to e_k to obtain

$$\frac{\partial u_k}{\partial e_k} = 2S_k - 2\lambda_k e_k, k \in \{'I','C','P','M'\} \tag{4.2.16}$$

Setting this gradient vector equal to zero, we see that e_k must be an eigenvector of the scatter matrix. In particular, it follows that to maximize $e_k^t S e_k$, we select the eigenvector corresponding to the highest eigenvalue of the scatter matrix. In other word to find the best one-dimensional projection of the data, we project the data onto a line through the sample mean and in the direction of the eigenvector corresponding to the highest eigenvalue of the scatter matrix. This result can be readily extended from one-dimensional projection to d dimensional projection, where we will have $e_{1k}...\ e_{dk}$ in the direction of the d eigenvector corresponding to the highest eigenvalue of the scatter matrix.

4.2.2.1 EigenTeeth Construction

In the following paragraphs, we will proceed with the construction of the four sub-spaces (as the single class can be considered as a special case). To construct the reduced subspaces, we proceed as follows:

- Compute the average of the four samples (equation 4.2.8)

- Let $A_k = [(x_{1k} - \mu_k).....(x_{Nk} - \mu_k)]$ denote the matrix of the average-adjusted sample of class k, where N_k is the number of exemplars of class k, and x_{ik} is the vector representation of the ith k exemplar. Then the eigenvectors u_{ik} of the scatter matrices of the four samples using the efficient technique proposed in [65], by first solving the eigenvalue problem:

$$A_k^T . A_k . v_{ik} = \lambda_{ik} . v_{ik} \tag{4.2.17}$$

where λ_{ik} is the ith eigenvalue, and v_{ik} is the ith eigenvector of matrix $(A^T_k : A_k)$. Finally, v_{ik} is linearly mapped to u_{ik}, using

$$u_{ik} = A_k . v_{ik} \tag{4.2.18}$$

- Select N eigenvectors for each set whose corresponding eigenvalues are the largest. The eigenvectors $u_k = [u_{1k}...u_{Nk}]$ are then arranged in a descending order according to their corresponding eigenvalues for each class. To maintain 99% of the image energy across the sample, we have to use the (N=104) most significant eigenvectors. With (N=51) we can maintain 95% of the image energy across the sample. We tested the performance of these subspaces versus the reduced subspaces [66]. A key benefit of reducing the number eigenvectors is to reduce the retrieval time.

4.2.2.2 Using EigenTeeth to classify tooth image

We use the EigenTeeth images calculated from the highest N eigenvectors to describe teeth images. We are investigating their usefulness as a tool for teeth identification and hence for dental image retrieval. Each tooth image (Γ) is transformed into its corresponding EigenTeeth components (projected into "teeth sub-space) by:

$$w_{jk} = u_{jk}^T . (\Gamma_k - \mu_k), j = 1...N \tag{4.3.11}$$

where μ_k is the mean, and u_{jk} are the principal Eigen vectors of each of the four classes (or the general class in case of using the single class method).

- The weights form a vector $\Omega_k = [w_{1k}, w_{2k},...w_{Nk}]$ that describes the contribution of each eigentooth in representing the normalized input tooth image, treating the EigenTeeth as a basis for teeth images. The resulted weights vector Ω_k is then used in a standard pattern algorithm to represent the similarity between the subject tooth and the reference tooth (as shown in Figure 46).

- We use the weights of the highest N principal component as representative for each tooth. We use the Euclidian distance between these two teeth after projecting to these lower dimensional spaces, as a matching score. The Euclidean distance between points P = (p_1,p_2,…p_n) and Q = (q_1,q_2,…q_n), in Euclidean n-space, is defined as:

$$E_d = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \ldots + (p_n - q_n)^2} = \sqrt{\sum_{i=1}^{n}(p_i - q_i)^2} \qquad (4.3.12)$$

  This calculated distance represents the match score on a tooth level. In other words, it represents the distance between reference tooth image and subject tooth image.

4.2.3 From Tooth to Record Matching

To move from the tooth-to-tooth matching level to record-to-record matching level, we have to address two problems:

- The same tooth may have multiple representations in either the reference record or the subject record or even both.

- Integrating the different matching scores of subject and reference teeth to have one matching score between the subject and reference records.

At the tooth-level fusion, we want to calculate the matching score between the different views of the same tooth t_{si} of a subject record (t_{S1}, .., t_{S32}) and the available views of reference tooth t_{Ri} of a reference record (t_{R1}, .., t_{R32}), such that we produce a single distance representing the matching score between the subject tooth t_{Si} and reference tooth t_{Ri}, namely $d_i(t_{Si}, t_{Ri})$. We use min fusion rule for multiple teeth representation fusion (refer to section 2.5.2).

With N teeth in common between the dental charts of a subject record and a reference record, we obtain N ($N \leq 32$) fused tooth-to-tooth distances. At the record-level fusion, we are looking for a scheme for combining these N {$d_i(t_{Si}, t_{Ri})$, $i = 1, .., N$} into a single record-to-record matching score $D(S, R)$, as shown in Figure 50. We study two alternative rules for fusing the match scores, namely the mean rule, and the min rule.
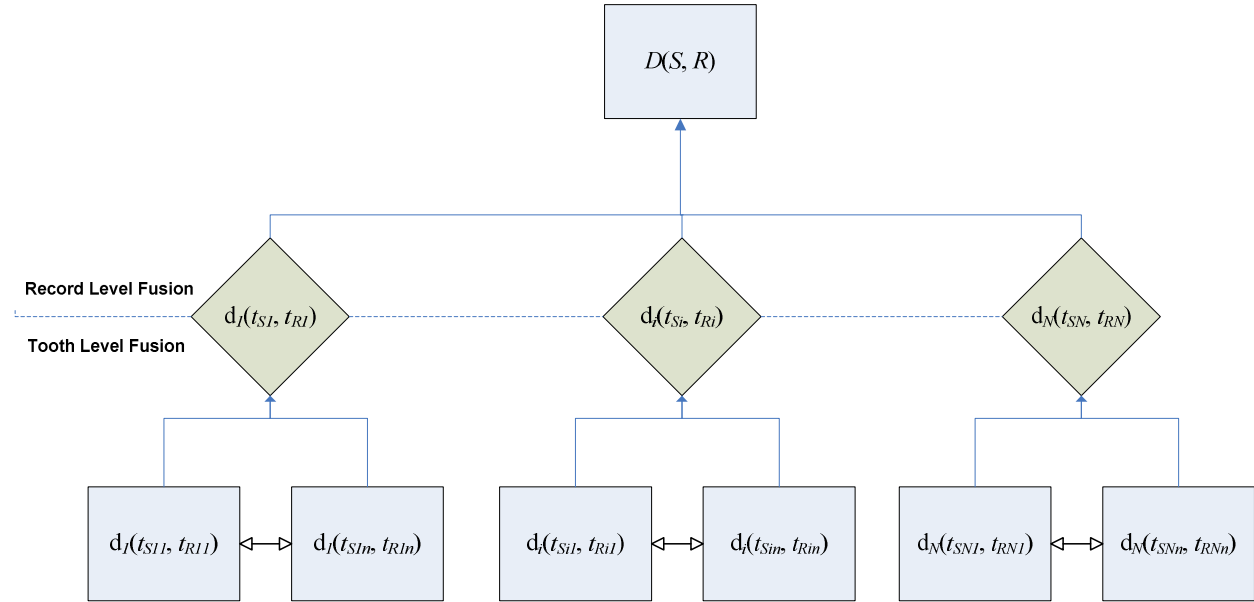
Figure 50: A hierarchical fusion scheme

4.2.4 Experimental Results

To test the different proposed approaches for Potential Match Search a test set was prepared using the CJIS ADIS project dental image database [48]. The test data-set involves 104 records containing about 500 bitewing and periapical films, involving more than 2000 teeth; 47 Antemortem (AM) records and 57 Postmortem (PM) records.

In a preliminary experiment, to test the proposed alignment technique, we used 80 segmented teeth images (representing molar, premolar, canines and incisors) where we introduced different angles between 1 and 40. We registered each segmented teeth image to the remaining aligned images of the same type. So for each molar image, we align it to the rest 19 molar models. We used the maximum mutual information as a performance evolution to select the best alignment angle. We found the error in the estimated angle to be as low as 2.82˚ on average.

To setup the image subspaces for the four teeth classes a data set of exemplars was prepared using the dental image database provided by the Missing and Unidentified Person Unit of Washington State Patrol [47] disjoint from the test set. 100 molars, 100 premolars, 30 canines, and 60 incisors from records were segmented and preprocessed. In selecting these teeth we tried the best we could avoiding biased data sets. So we evenly used teeth from upper and lower jaws, we used teeth from the right and left sides of the mouth, we used teeth images of different intensity contrast.

4. ADIS Potential Match Search

To setup the general the general teeth image space, we used the same mentioned data set of exemplars, but this time without labeling them. In other words, we are building a general (one) class, so all the teeth types are considers as exemplars for this one class.
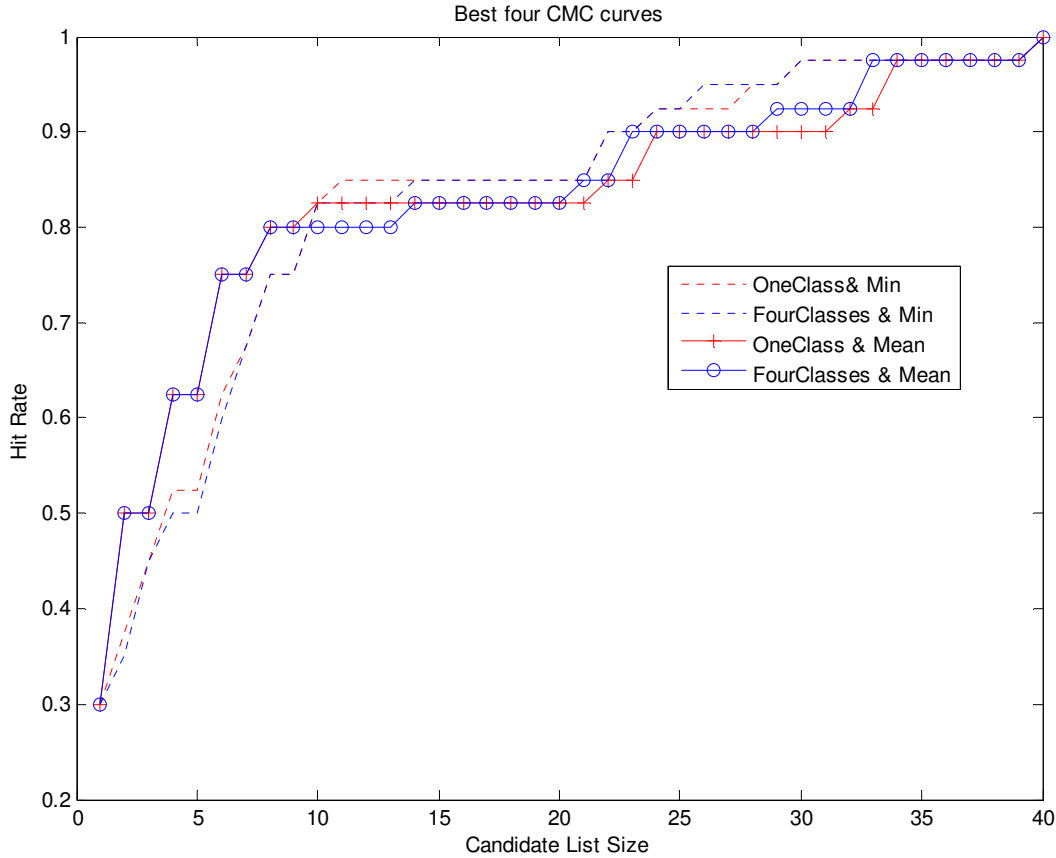
Figure 51: The best four CMC curves for Eigen image retrieval

We tested the performance using varied number of eigenvectors (15, 20, ... 55 or 100 eigenvectors) as basis for both method (the general class and the four classes), and using minimum and mean fusions criteria. We visually assessed the resulted 10 Cumulative Match Characteristic (CMC) curves for each method using one of the fusion criteria. Figure 51 shows the best 4 results as follows: (i) For the one-class method using minimum fusion (K=35), while for the four-classes (K=45). (ii) For the one class method using mean-fusion (K=50), while for the four classes (K=100). As can be induced from the curves in Figure 51:

- the mean fusion criteria behaves better than the minimum one, we verified this conclusion using paired t-test for rank 10 candidate list (p= 0.0012).

- the four-classes method needs to maintain more significant eigenvectors to achieve the same performance as the one-class. However we see that with the future of unbiased training and testing database the four-classes will perform better.

- 50% of the times the match record is within the top 3 ranks, and 75% is within top 6 ranks. On average, the record to record comparison takes less than 0.1 sec (using normalized teeth). To find a candidate list for a reference, it takes about 5.2 sec, which encountered about 52 record to record comparisons.

To verify our visual assessment, we perform a paired t-test, and calculate the p-values:

One-Class Min-Method	K=15	K=20	K=25	K=30	K=40	K=45	K=50	K=55	K=100
K=35	<0.01	<0.01	<0.01	<0.01	0.6232	0.5339	0.2099	0.1354	<0.01

Note: The statically result didn't yield significant results between K=35, K=40, and K=45. We picked to reserve the minimum number of principal components to achieve minimum retrieval time.

One-Class Mean-Method	K=15	K=20	K=25	K=30	K=35	K=40	K=45	K=55	K=100
K=50	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	Nan	0.5703

Note: The statically result didn't yield significant results K=50, K=55, and K=100. We picked to reserve the minimum number of principal components to achieve minimum retrieval time.

Four-Class Min-Method	K=15	K=20	K=25	K=30	K=35	K=40	K=50	K=55	K=100
K=45	<0.01	<0.01	<0.01	<0.01	<0.01	0.1684	<0.01	0.2099	<0.01

Note: The statically result verify the visual assessment that K=45 is the best choice.

Four-Class Mean-Method	K=15	K=20	K=25	K=30	K=35	K=40	K=45	K=50	K=100
K=100	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01

Note: The statically result verify the visual assessment that K=100 is significantly the best choice.

4.3 Contour Descriptors (Shape-Based Approach)

There are several methods to represent shapes. Most of these shape descriptors are typically translational and rotation invariant and some are also scale invariant. In this section we are presenting a brief overview of two shape descriptors, namely moment invariants and edge direction histogram, and how we use for dental image retrieval.

4.3.1   Method One: Moment Invariants

Invariant geometric moments have been widely used in shape analysis and pattern recognition [102]. Geometric moments of order $p + q$, defined as:

$$m_{p,q} = \int_y \int_x f(x, y) x^p y^q \, dxdy \tag{4.3.1}$$

The two-dimensional moment for a ($N . M$) discretized image, $I(x, y)$, is:

$$m_{p,q} = \sum_{y=0}^{M-1} \sum_{x=0}^{N-1} I(x, y) x^p y^q$$

Zeroth Order Moments: $m_{0,0}$, when computed for a binary segmented object, represents the total object area.

First Order Moments (Center of Mass): The two first order moments, $\{m_{0,1}, m_{1,0}\}$, are used to locate the *center of mass (COM) o*f the object. If an object is positioned such that its COM is coincident with the origin of the field of view, then the moments computed for that object are referred to as *central moments* and are designated by $\mu_{p,q}$. (Note that $\mu_{1,0} = \mu_{0,1} = 0$)

Second Order Moments: $\{m_{0,2}, m_{1,1}, m_{2,0}\}$, known as the *moments of inertia*, are used to determine several useful object features.

Prokop and Reeves [102] identified five principal moment-based invariant feature techniques. i) Moment invariants are based on non-linear combinations of low-order two-dimensional Cartesian moments that remain invariant under rotation. ii) Rotational moments, are based on polar image representations, are also proposed as a solution for their simple rotation properties. iii) Orthogonal Moments, is utilizing uncorrelated basis functions, were developed to reduce the information redundancy that existed with conventional moments. iv) Complex moments provide straightforward computation of invariant moments of an arbitrary order. v) Standard moments are unique in that they achieve invariance completely through image feature normalization in the moment domain rather than relying on algebraic invariants.

4. ADIS Potential Match Search

We selected to use moment invariants (2-D) as it was the first significant technique in this direction. If the object R is represented as a binary image, then the central moments of order $p+q$ for the shape of object R are defined as [103]:

$$\mu_{p,q} = \sum_{(x,y) \in R} (x - x_c)^p (y - y_c)^q \qquad (4.3.2)$$

Where (x_c, y_c) is the center of object. This central moment can be normalized to be scale invariant:

$$\eta_{p,q} = \frac{\mu_{p,q}}{\mu_{0,0}^\gamma}, \gamma = \frac{p + q + 2}{2}, p + q \geq 2$$

The first significant work considering moments for pattern recognition was performed by Hu [90]. Based on these moments, Hu derived a set of seven moment invariants to translation, rotation, and scale. These moments are combinations of second and third order normalized central moments. In term of the central moments, the seven moment invariants are given by:

$$\phi_1 = \mu_{2,0} + \mu_{0,2}$$

$$\phi_2 = (\mu_{2,0} - \mu_{0,2})^2 + 4\mu_{1,1}^2$$

$$\phi_3 = (\mu_{3,0} - 3\mu_{1,2})^2 + (\mu_{0,3} - 3\mu_{2,1})^2$$

$$\phi_4 = (\mu_{3,0} + \mu_{1,2})^2 + (\mu_{0,3} + \mu_{2,1})^2$$

$$\phi_5 = (\mu_{3,0} - 3\mu_{1,2})(\mu_{3,0} + \mu_{1,2})\left[(\mu_{3,0} + \mu_{1,2})^2 - 3(\mu_{0,3} + \mu_{2,1})^2\right]$$
$$+ (\mu_{0,3} - 3\mu_{2,1})(\mu_{0,3} + \mu_{2,1})\left[(\mu_{0,3} + \mu_{2,1})^2 - 3(\mu_{3,0} + \mu_{1,2})^2\right]$$

$$\phi_6 = (\mu_{2,0} - \mu_{0,2})[(\mu_{3,0} + \mu_{1,2})^2 - (\mu_{0,3} + \mu_{2,1})^2] + 4\mu_{1,1}(\mu_{3,0} + \mu_{1,2})(\mu_{0,3} + \mu_{2,1})$$

One skew invariant is defined to distinguish mirror images and is given by:

$$\phi_7 = (3\mu_{2,1} - \mu_{0,3})(\mu_{3,0} + \mu_{1,2})[(\mu_{3,0} + \mu_{1,2})^2 - 3(\mu_{0,3} + \mu_{2,1})^2]$$

This technique, and its subsequent variations, proved to be basis for the majority of the moment research to date [102]. Hu [90] used the first two moment invariants through a simple pattern recognition experiment. An unknown pattern could be classified by computing its first two moment values and finding the minimum Euclidean distance between the unknown and the set of well-known pattern

representations in feature space. If the minimum distance was not within a specified threshold, the unknown pattern was considered to be of a new class, given an identity, and added to the known patterns.

We use Hu [90] seven moment invariants through our dental recognition experiments. The contour of the subject tooth could be matched to that of the reference tooth by computing their first seven moment values and finding the Euclidean distance between the subject and the subject in that feature space. To move from the tooth-to-tooth matching level to record-to-record matching level, we use min and mean fusion rules as described in section 4.2.3.

4.3.2   Method Two: Edge direction histogram:

Edge histogram captures the general shape information in the image. Because the feature is local, it is robust to partial occlusion and local disturbance in the image. A histogram of the edge directions is used to represent the shape attribute. The edge information contained in the database images is generated in the preprocessing stage using edge operator (as Canny [64] edge, Sobel edge operators). A histogram intersection technique is used for dental shape-based retrieval. The histogram of the edge directions is invariant to translation in the image. Thus, the positions of the objects in the image have no effect on the edge directions.

The use of edge directions is inherently not scale invariant. Two images identical in every respect except their size will yield different number of edge points and hence different histograms. In order to have invariance to scale, we normalize the histograms with respect to the number of edge points in the image. A drawback of normalized histograms is its inability to match parts of images. Let *H(i)* be a histogram of an image, where the index i represents a histogram bin. Then, the normalized histogram I is defined as follows:

$$I(i) = \frac{H(i)}{\sum_i H(i)} \tag{4.3.3}$$

The similarity and distance between the subject image edge histogram S and a stored reference image edge histogram I in the database, $S_e^{HI}(R, S)$, and $D_e^{HI}(R, S)$ are given by the following equations [104]:

$$S_e^{HI}(R, S) = \frac{\sum_e \min(R_e, S_e)}{\min(|R|, |S_e|)} \tag{4.3.4}$$

$$D_e^{HI}(R, S) = \sqrt{\frac{\sum_i (R(i) - S(i))^2}{2}} \tag{4.3.5}$$

A histogram of the edge directions is also not invariant to rotation. A shift of the histogram bins during matching partially takes into account rotation of the images. A rotation of the image shifts each of the edge directions by the amount of rotation. However, due to the effect of quantization of the edge directions into bins, the effect of rotation is more than a simple shift in the bins. Rotation also affects the membership in the bins. To reduce this effect of rotation, Jain and Vailaya [104] smooth the histograms. Any histogram can be treated as a 1-D discrete signal. Smoothing can then be defined as:

$$I_s[i] = \frac{\sum_{j=i-k}^{i+k} I[j]}{2k+1}$$

(4.3.6)

where I_s is the smoothed histogram, I is the original histogram and the parameter k determines the degree of smoothing.

The matching results depend on the bin size. By choosing a larger bin size, the matching speed is increased. However, this also reduces the accuracy in case of an arbitrary rotation of the image. A rotation not only causes a shift in the histogram bins but also affects the bin membership. Use of a very small bin size reduces the matching speed (since the number of bins increases). Use of a small bin size also requires that the edge directions be found to a very high degree of accuracy [104].

We use histogram of the tooth contour edges directions through our dental recognition experiments. The contour of the subject tooth could be matched to that of the reference tooth by computing their histogram of edges directions and finding the Euclidean distance between the subject and the subject in that feature space. To move from the tooth-to-tooth matching level to record-to-record matching level, we use min and mean fusion rules as described in section 4.2.3.

4.3.3 Experimental Results

To test the different proposed approaches for Potential Match Search a test set was prepared using the CJIS ADIS project dental image database [48]. The test data-set involves 104 records containing about 500 bitewing and periapical films, involving more than 2000 teeth; 47 Antemortem (AM) records and 57 Postmortem (PM) records.

To test the performance of the two proposed method for shape features image retrieval, we used the teeth contours [43] to represent these features. Figure 52 shows the performance curve for the moment invariant features; while Figure 53 shows the performance curve for the edge direction histogram features (Norm and smoothed respectively). As can be induced from the curves:

- the min fusion criteria behaves better than the mean one for all the presented shape features. To verify this conclusion, we perform paired t-test, which yields (p<0.01, i.e. significant test) for the three shape features (moment invariant, edge direction histogram, and smooth edge direction histogram).

- the performance of the moment invariant features is lower than the edge direction histogram. However to find a candidate list for a reference, it takes about 18 sec for the moment invariant compared to 6.5 sec for the edge direction histogram, which encountered about 52 record to record comparisons.
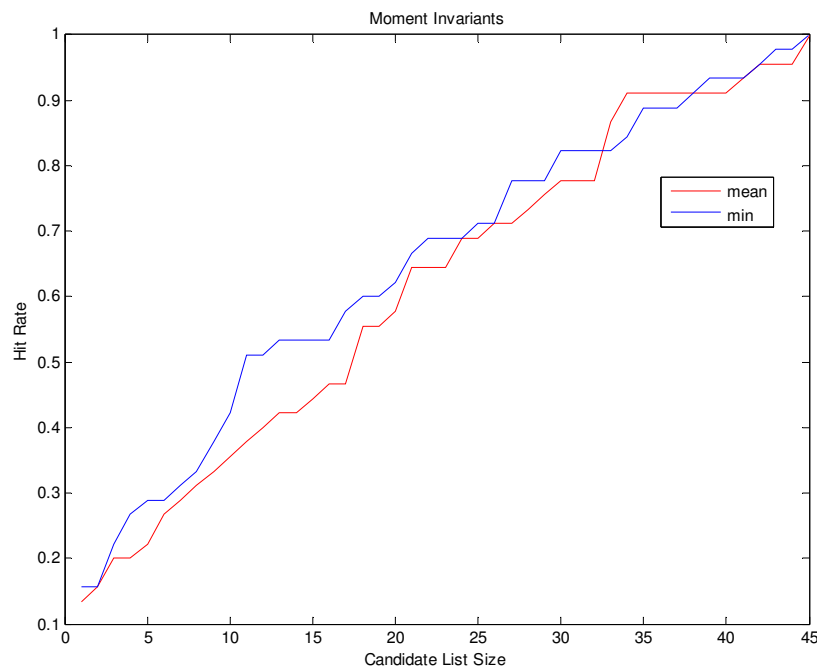
Figure 52 CMC curves for moment invariants method
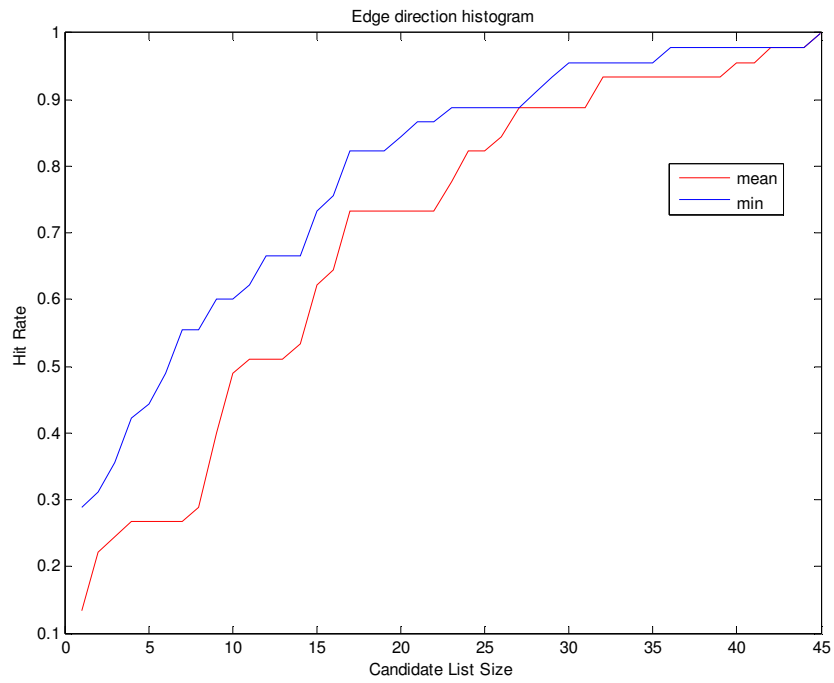
a)
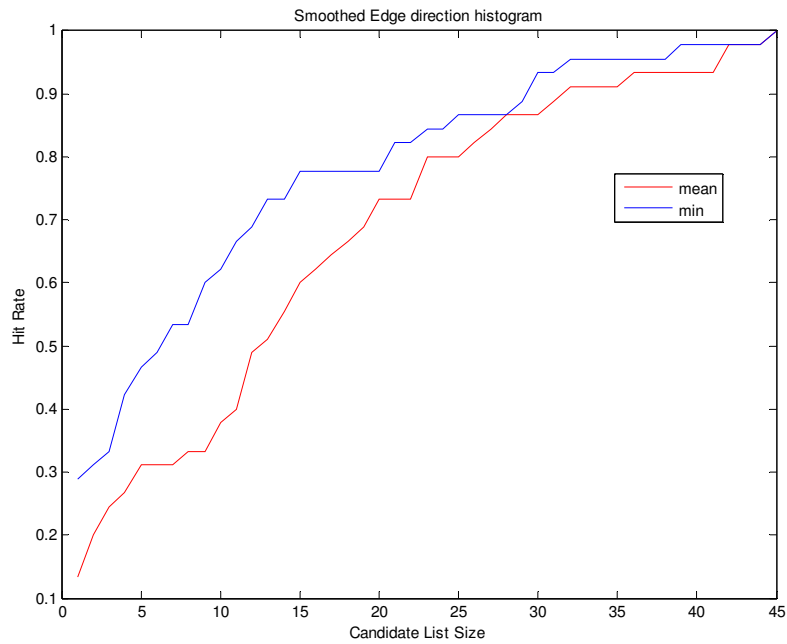
b)

Figure 53: CMC curves for edge direction histogram method
a) normal histogram and b) smoothed histogram

Figure 54 shows the performance curve for the edge direction histogram features using min fusion method (Norm and smoothed respectively). As can be induced from the curves that the performances are almost the same (p=0.0511 for the paired t-test, i.e. not significant test), we returned this to the large number of bins that we are using.

Most systems use a hybrid combination of features. Apparently the shape features used for matching are not the most effective ones, however we see dental image retrieval system that can utilize both color and shape feature will achieve better performance.
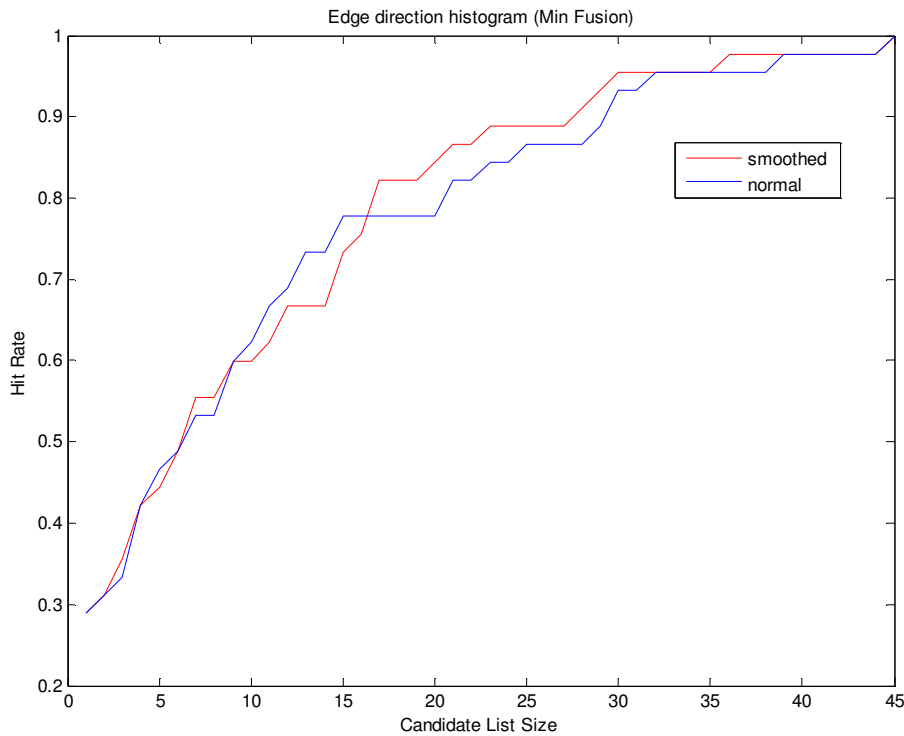
Figure 54: Best performance CMC curves for edge direction histogram image retrieval using min fusion

4.4 Experimental Results

To test the different proposed approaches for Potential Match Search a test set was prepared using the CJIS ADIS project dental image database [48]. The test data-set involves 104 records containing about 500 bitewing and periapical films, involving more than 2000 teeth; 47 Antemortem (AM) records and 57 Postmortem (PM) records.

4. ADIS Potential Match Search

In this experiment, we are comparing the performance of appearance-based (one class and four classes) versus shape-based (edge direction histogram). We are using minimum fusion, as it proved to work better for shape-based.

shows the performance curve for the edge direction histogram features (representing shape-based) and eigen images, one teeth class and four teeth classes respectively (representing appearance-based). As can be induced from the curves that the appearance-based performs better for short candidate list, however the edge direction histogram achieves the same performance for slightly longer candidate list. To verify our visual assessment, we perform paired t-test, which yield (p<0.01, i.e. significant test) for both methods of appearance-based versus the smooth edge direction histogram (shape based).

Figure 55: CMC curves for appearance-based versus shape-based image retrieval using min fusion

5   Automated Ear Identification System (AEIS)

Ear biometrics refers to the automatic measurement of distinctive ear features with a view to identify or confirm the identity of the owner. It has received scant attention compared with the more popular face, iris, or fingerprint recognition techniques. However, the potential of the human ear for personal identification was recognized from 1890 by the French criminologist Alphonse Bertillon [106].

5.1 Introduction

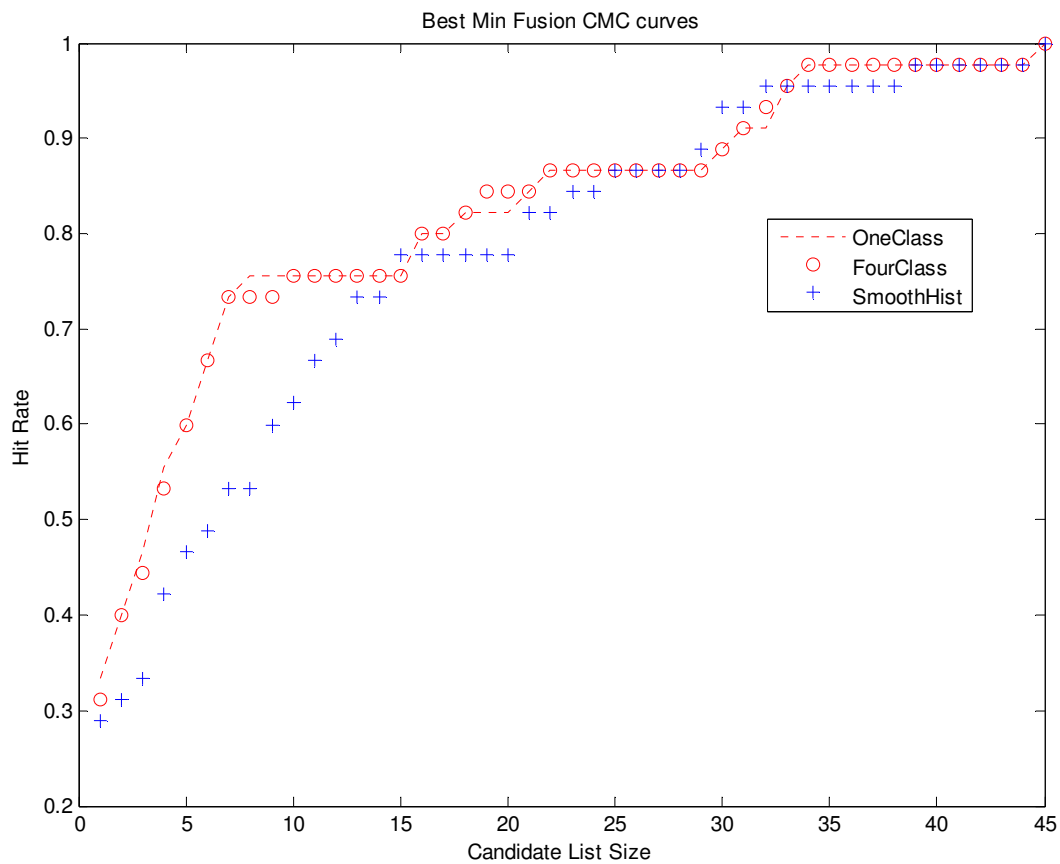The ear does not have a completely random structure; it is made up of standard features just like the face. As shown in Figure 56, the shape of the ear tends to be dominated by the helix rim, and the lobule shape. The anti-helix roughly runs parallel to the outer helix. Note also the crus of helix where the helix intersects with the anti-helix and the tarus.

An individual's ear structure is a unique characteristic that remains fairly stable over time. As with the fingerprint, it is impossible to prove that an individuals' ear structure is unique, however studies have supported this hypothesis. One such study done in California compared thousands of ears [116], in the preface to his book Iannarelli states:

"Through 38 years of research and application in earology, the author has found that in literally thousands of ears that were examined by visual means, photographs, ear prints, and latent ear print impressions, no two ears were found to be identical – not even the ears of any one individual. This uniqueness held true in cases of identical and fraternal twins, triplets, and quadruplets"

In response to the US appeals court ruling, a large-scale study involving 10,000 subjects has been proposed by Prof. Andre Moenssens to determine the variability of the ear across the population [119]. This study proved that in all cases there were no two ears which were indistinguishable. Another study examined the difference in ear structures between fraternal and identical twins, who share similarities in many physiological features [118]. It was proven that even though their ear structures were similar, they were still clearly distinguishable. In 1906, Imhofer studied a set of 500 ears and noted that he could clearly distinguish between each ear based on only 4 features [117]. The results of these studies support the hypothesis that the ear has a unique physiological structure.

There is also evidence available to show that ear structure is both stable and permanent. The medical literature reports that ear growth after the first four months of birth is highly linear. It turns out that even though ear growth is proportional, gravity can cause the ear to undergo stretching in the vertical direction. The effect of this stretching is most pronounced in the lobe of the ear and measurements show that the

change is nonlinear only in that part. The rate of stretching is approximately five times greater than normal during the period from four months to the age of eight after which it is constant until around the age of seventy when it again increases [118].

- Helix Rim: is the outer frame of the auricle, it is a rolled up edge.

- Lobule: fleshy lower portion of the ear.

- Anti-helix: folded "Y" shaped part of the ear. Elevated ridge of cartilage between the concha and the scapha.

- Concha: the hollow bowl like portion of the outer ear next to the canal. An enlarged concha forces the outer ear away from the scalp.

- Tragus: small projection just in front of the ear canal.

Figure 56: **External anatomy of the ear [116]**

- Anti-tragus: lower cartilaginous edge of the conchal bowl just above the fleshy lobule of the ear.

- Crus of Helix

- Triangular Fossa

- Incisure Intertragica

Our objective is to develop components of a viable automated method for Ear-Recognition and to evaluate their performance based on statistical analysis. The proposed system is an automation of the Iannarelli System, which has been manually used for more than 40 years. Our selection to Iannarelli System specifically, because:

1- This system has been used for years, where it works with an excellent accuracy.

5. Automated Ear Identification System (AEIS)

2- To the best of our knowledge, this system is the only one that the human can evaluate each and every step of it. For the automated Iannarelli, the experts can easily verify the extracted distances (features) as evidence in the court.

3- We found it to be a challenging problem, as Burge and Burger [108] mentioned that:

"This method is additionally not suited for machine vision because of the difficultly of localizing the anatomical point which serves as the origin of the measurement system"

We developed the automated ear identification system as a series of modular components. This allowed us to evaluate the performance of each module independently and then improve the poorly performing modules, in order to improve the overall performance of the system. Each module was evaluated using various possible implementations to determine which one performed the best. Also, for future research one or more modules could be used in different applications.

The proposed system has an input data, as a video image sequence of the subject. The first step is to pre-process this input data and produce segmented ear image with enhancement, followed by feature extraction and then according to the mode (enrollment or matching) the last step will be either achieving/retrieving images to/from the database (as shown in Figure 57). The first step starts by face and ear localization and tracking the different input frames, followed by an enhancement step of super resolution, which is used for the time being to help the forensic expert to visualize a magnified image; this enhancement step also includes contrast and brightness adjustment, size (or dimension) normalization, and background elimination. The third and main step is feature extraction which starts by the

Figure 57: Process breakdown

origin point detection followed by different feature point's measurements. According to the operational mode either these measurements are stored as a feature representation of the case; or the distance between the subject features and each reference case is used to retrieve the top matched cases from the database.
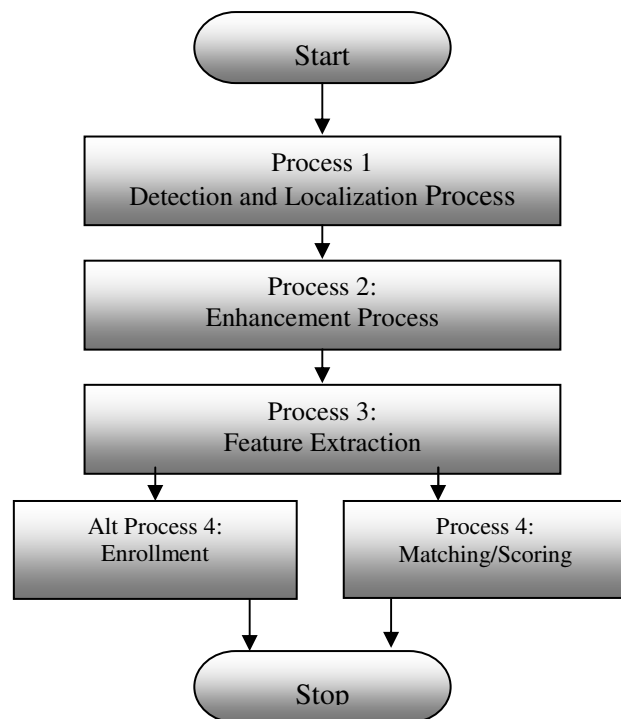
5. Automated Ear Identification System (AEIS)

This chapter is organized as follows; section 5.3 introduces literature review about different techniques used for face and ear detection. In section 5.4, we introduce different ear image enhancement steps. In section, we presented automated Iannarelli feature extraction process. In section 4.4, we present the experimental results on the performance of the proposed system and discussion on these results.

5.2 Face and Ear Detection

Videos are actually sequences of images, each of which called a frame, displayed in fast enough frequency so that human eyes can percept the continuity of its content. It is obvious that all image processing techniques can be applied to individual frames. Besides, the contents of two consecutive frames are usually closely related.

Visual content can be modeled as a hierarchy of abstractions. At the first level are the raw pixels with color or brightness information. Further processing yields features such as edges, corners, lines, curves, and color regions. A higher abstraction layer may combine and interpret these features as objects and their attributes. At the highest level are the human level concepts involving one or more objects and relationships among them.

Object detection in videos involves verifying the presence of an object in image sequences and possibly locating it precisely for recognition. In our case, we know that the face and ear are there in the frame. Even, we know that there is only one face image. Object tracking is to monitor an object's spatial and temporal changes during a video sequence, including its presence, position, size, shape, etc. This is done by solving the temporal correspondence problem, the problem of matching the target region in successive frames of a sequence of images taken at closely-spaced time intervals. These two processes are closely related because tracking usually starts with detecting objects, while detecting an object repeatedly in subsequent image sequence is often necessary to help and verify tracking. In this work, we assume only face and ear detection across frames.

Images containing faces are essential to intelligent vision-based human computer interaction and research efforts in face processing include face detection, face localization, face recognition, and face tracking. However, many reported methods assume that the faces in an image (or an image sequence) have been identified and localized. To build fully automated ear identification system that analyze the information contained in the ear segment of the face images, robust, efficient, and fast face detection algorithms are required [121]. In order to localize human ear from side face range images, we need first to extract the face from the background. In our experimental setup (section), we have a controlled environment, where

we can define two main classes of objects in the input images; face that map to areas with variant grayscale, and background that maps area with smooth grayscale.

Given a single image, the goal of face detection is to identify the image region which contains the face regardless of its three-dimensional position, orientation, and lighting conditions. Such a problem is challenging because faces are non-rigid and have a high degree of variability in size, shape, color, and texture. Numerous techniques have been developed to detect faces in a single image.

The challenges associated with face detection can be attributed to the following [120]:

- Pose. The images of a face vary due to the relative camera-face pose (e.g. frontal, 45 degree, profile), and some facial features such as an eye or the nose may become partially or wholly occluded.

- Presence or absence of structural components. Facial features such as beards, mustaches, and glasses may or may not be present and there is a great deal of variability among these components including shape, color, and size.

- Facial expression. The appearance of faces is directly affected by a person's facial expression.

- Image orientation. Face images directly vary for different rotations about the camera's optical axis.

- Imaging conditions. When the image is formed, factors such as lighting (spectra, source distribution and intensity) and camera characteristics (sensor response, lenses) affect the appearance of a face.

HajSaid [122] tackled the problem of face and ear segmentation using a Mathematical Morphology (MM) approach. First, the technique starts by face detection, where two main classes of objects in the facial images are defined; face that map to areas with variant grayscale, and background that maps area with smooth grayscale. Figure 58 shows the main steps of face extraction from the background. HajSaid [122] proposed to subtract the subject image from its erosion, and then to threshold the resulted image using basic thresholding. At the last stage, binary dilation and holes is applied filling to ensure that the whole face is comprised.

For Ear detection, HajSaid [122] started by applying opening top hat transformation, which is defined by subtracting morphological opening of the image from itself Figure 59-b, and then he applied basic thresholding on the transformed image Figure 59-c. The last stage is a refinement stage to analyze the connected components (Figure 59 –c), whether it corresponds to an ear or any other part of the face, based

on their geometric properties including (area, position and dimension), and then to eliminate the unqualified objects (Figure 59 –d).

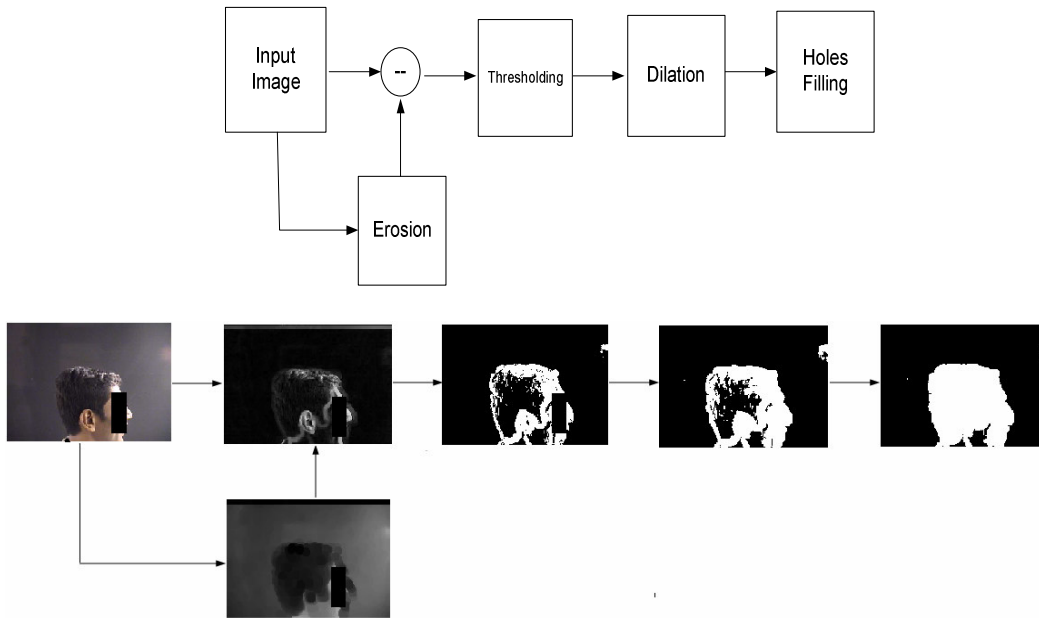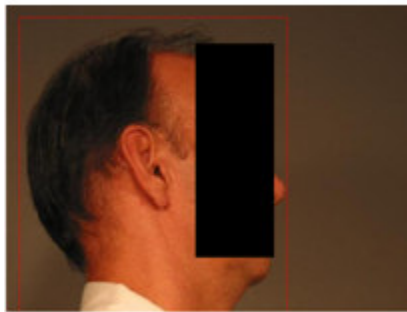**Figure 58: Main stages of face detection**

(a) Detected Face

(b) Transformed image

(c) Thresholded Images

(d) Detected ear

Figure 59: Stages of ear segmentation for different facial image [121]

5.3 Ear Image Enhancement

The second step of ear identification system is the ear image enhancement, which starts with a step to enhance the video image resolution, followed by normalization module which normalizes the segmented ear image with respect to intensity, size, contrast, and various other criteria. Further processing would locate and remove the hair from the image.

5.3.1 Super Resolution

For enhancing the visual assessment of the proposed ear identification system, we need ear-images with high resolution (HR) or super resolution (SR); while the video sequence that we are acquiring is having a limited resolution (LR) of 480 * 640 pixels per frame. SR means that pixel density within an image is high, and therefore a SR image can offer more details that may be vital for the accuracy of feature extraction stage, Building 3-D ear stage, and the identification stage. We are working toward inexpensive SR digital images close to that of an analog image that almost has no visible artifacts when magnified. One promising approach is to use image processing techniques to obtain a SR image from observed multiple (LR) images, as shown in Figure 60. Recently, such a resolution enhancement approach has been one of the most active research areas, and it is called super resolution (SR) or (HR) image reconstruction or simply resolution enhancement [123].
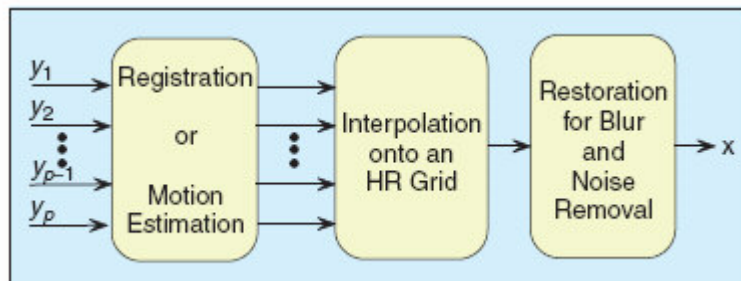
Figure 60 Block diagram for the super resolution process

In SR, typically, the LR images represent different "looks" at the same scene. That is, LR images are sub-sampled as well as shifted with sub-pixel precision, which is the case of successive frames in the video sequence that we are acquiring.

The first step toward a super resolution image is to make a precise Image Registration for the LR images (frames), as shown in Figure 61. This is a well-established area in image processing applications. Vandewalle et al. [124] used a frequency domain technique to precisely register a set of aliased images, based on their low-frequency, aliasing-free part; while Lucchese et al. [125] presented another method for estimating planar roto-translations that operates also in the frequency domain and, as such, is not based on

features. Since the proposed technique uses all the image information, it is very robust against noise, and it can be very accurate; estimation errors on the rotational angle range from a few hundredths to a few tenths of a degree, depending on the noise level.
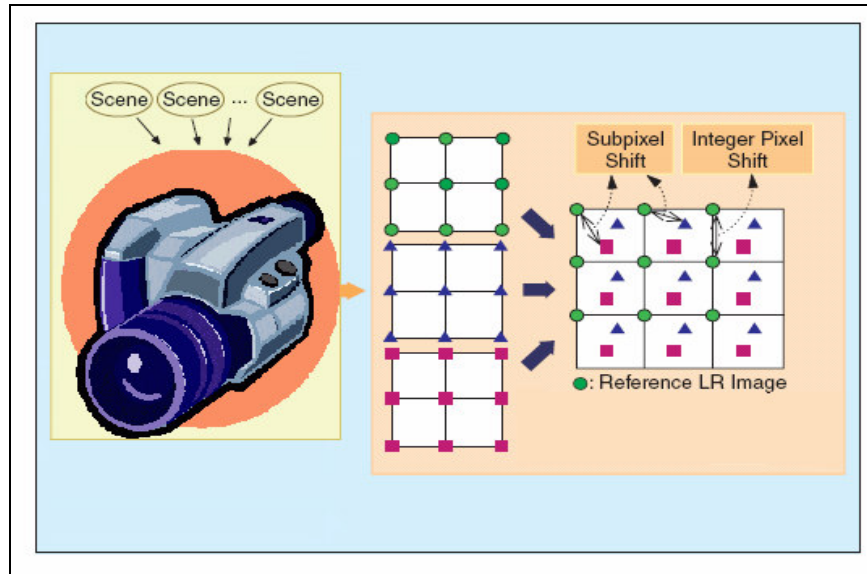
Figure 61 Sub-pixel sift between successive frames

The second step is interpolation into the SR grid, to help executing this step; we set the ear segments across the frames to specific size (n.n). In other words, in the ear detection step we segment a window of size (n.n) around the ear.
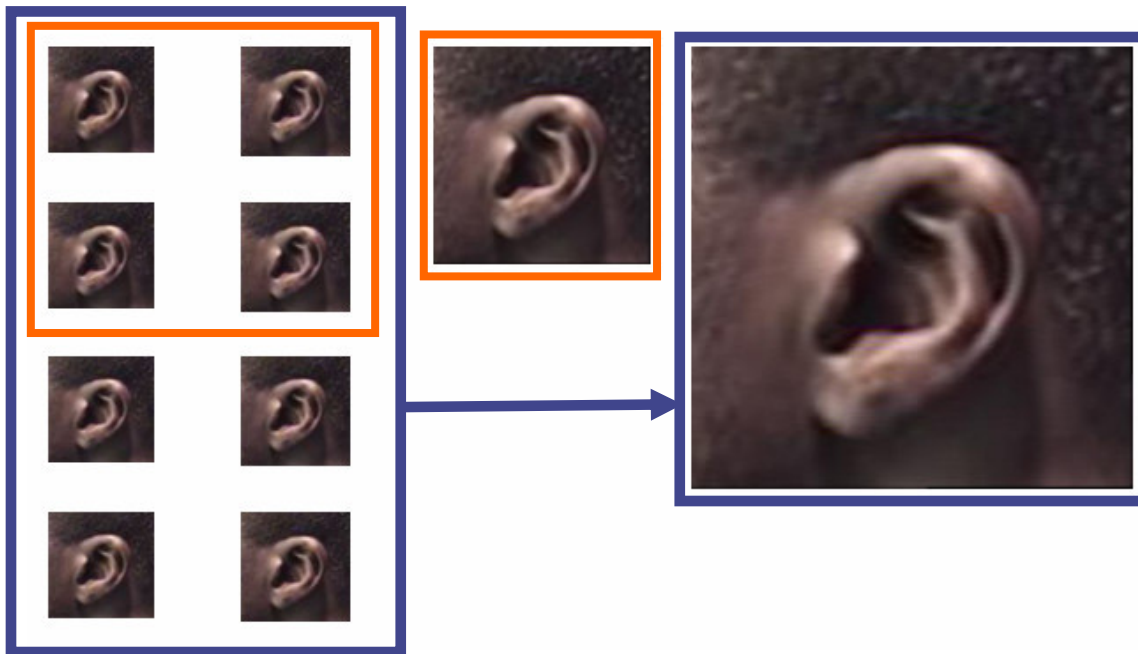
Figure 62: Ears segments from four (eight) successive frames, and their corresponding super-resolution image.

The third (and last) step for super resolution is image restoration, where the main objective is blur and noise removal. This is the main concern, which makes us to use this step as an optional step that can help visual (or manual) assessment. Figure 62 shows an example of super resolution ear segment, using 4 and 8 successive ear segments. Notice that the blurring problem starts to appear with the second one. This is the main reason that this stage is not useful for the proposed automated ear identification. However, we see that further improvement on the suggested technique may help in the future.

5.3.2 Ear Image Normalization and Background Elimination

The edging process operates on the brightness and contrast levels within the image. For this reason we implemented a function to normalize the brightness and contrast in an attempt to ensure consistent performance of the system. For manual comparison, Iannarelli used to magnify the ear image to fixed dimensions. In the automated Iannarelli system, we can achieve this by either normalizing the segmented ear image or normalizing all the extracted distances. For the implementation of super resolution, we adjust the segmented ear image during the localization process (as shown in Figure 62).

Initially we used to subtract the hair from the image. As the hair usually has uniform color in the image, it manifests itself in the histogram. Our hair removal function adjusted the intensity of pixels in the vicinity of this detected peak to a level closer to that of the skin intensity. The preliminary results show that implementing the module in this way decreases the performance of the edging module. This is because the locally adjustment of the grayscale again means, we are diminishing the performance of edge detection, which is a mandatory step for our feature extraction stage. Later, we moved to another method that depends on closing the external curve of the ear (as shown in Figure 63). We developed a better method to separate the ear from the background using a mask. This mask applies a convex hull around the detected ear edges. The convex hull of a set of points is the smallest convex set that contains the points [126], as shown in Figure 63. We use this mask to eliminate the background.

Figure 63: a) ear extracted curves; b) convex hull separating the segmented ear from background

5.4   Feature Extraction Process

The third process in our proposed system (Process 3) is the Feature Extraction process. We are developing a modified automation of the Iannarelli System. In general, this method requires that the system detect specific edges in the ear image (at horizontal, vertical and diagonal lines) and calculate their distance to a specific reference point (Crus of Helix). So we have two main goals in this stage, the first is to determine a reference point. The second is to extract the internal and external structure of the ear.

5.4.1   Calculating the Reference Point

Iannarelli calculates 12 measurements with respect to the tip of the Crus of Helix. This reference point suits his manual method as it is an anatomical feature of the ear. This insured consistence and repeatability in his system. For our automated system we have investigated other criteria for specifying that reference point.

Our first implementation relies on calculating the center of gravity (centroide) of the masked ear image. The center of gravity (centroide) is defined as:

$$C_x = \frac{\sum_i m_i \ x_i}{\sum_i m_i} \qquad C_y = \frac{\sum_i m_i \ y_i}{\sum_i m_i}$$

As a refinement of finding this reference point, we come to define this reference point as the centroide of the external part of the ear. We define the external part of the ear as the difference between the segmented ear curves, and the convex hull surrounding the ear (as shown in Figure 64).
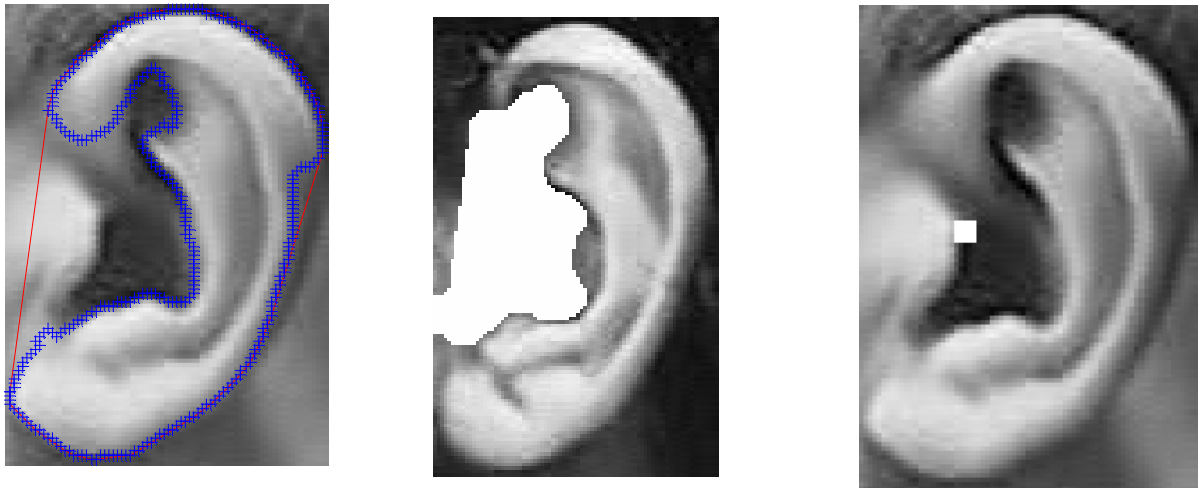
Figure 64: a) segmented ear image; b) mask of the external ear part; c) reference point.

5.4.2 Finding Ear Curves

In the literature, there exist many edges detection methods. We evaluated the performance of five of the most commonly used methods. By visually assessment of the results of this comparison study, we selected Canny edge detection method [64]. Figure 65 shows an example of how canny edge method performs compared to the other methods.
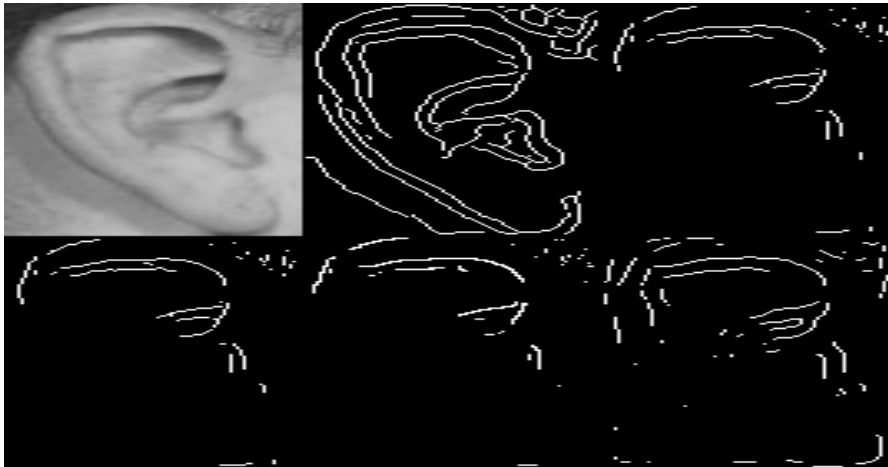
Figure 65: edge methods performance comparison;
a) original image, b) Canny method, c) Sobel method,
d) Prewitt method, e) Roberts method, f) Laplacian of Gaussian

The performance of the edging operator is dependent on selecting the optimal edging threshold. The effect of using the canny edging operator at 2 different (non optimal) edging thresholds on a grayscale image is shown in Figure 66.
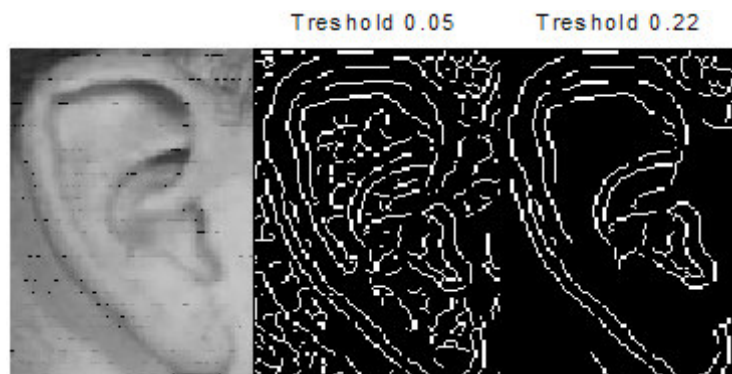
Figure 66: Original and edged images

The optimal edging threshold is different for each gray-scale ear image, and therefore must be recalculated for each image processed. In order to automate this process, we developed a function to determine the optimal threshold automatically. The function operates by implementing the edging operator at a very low threshold and then analyzing the resulting edged image.

Connected pixels (edges) are identified as long or short components and counted. This edged image resulting from edging with a low threshold will contain the maximum number of long components. The edging threshold is increased and the edging operator is reapplied to the gray-scale image so that the components can be recounted. The process is repeated so as to maximize the count for long and minimize the count for short segments (edges). Finally, the function returns the optimal threshold.

5.4.3 Measuring Feature Points

The masked image is used to calculate the distance from the centroide to the outermost curve of the ear. The distances between the centroide and the first black pixel of the mask image are calculated for every multiple of a specified angle from the horizontal. Our implementation uses 22.5° (shown in Figure 67-A) increments to gain more information as compared to Iannerelli's method which makes use of 45°.

For simplicity in explaining the directions we will refer to the angle directions using a geographical description with north being the top of the image. The distances that we calculated from these measurements were based on the external edge of the ear. After superimposing these measurements onto the edged image, the distances between the centroide and the internal edge components were calculated as shown in Figure 67-b).

**Figure 67: a) ear mask showing 22.5° angles centered at reference point;
b) edged image with measurements superimposed**

5.5 Experimental Results

The training data set contains images of 110 different cases. It is part of the ear data set collected at West Virginia University, where the users' right, frontal and left profiles were captured by a video camera. The raw data is a video sequence 180˚ for the case face in about 100 secs. We converted the video sequence to image sequence of rate 3 frames per seconds. According to our assessment to this input data, this law frame rate helps eliminate any motion artifacts. The resulted image sequence has the case represented by about 300 frames; each of these frames is about 0.6˚ apart. The size of each frame is 640*480, 24-bit RGB. In this experimental setup, we used 10 images 3˚ apart from each others to represent each case. These ten images cover the left profile of each case. Figure 68 shows two images from a case, the starting frame and the ending frame 27˚ apart.
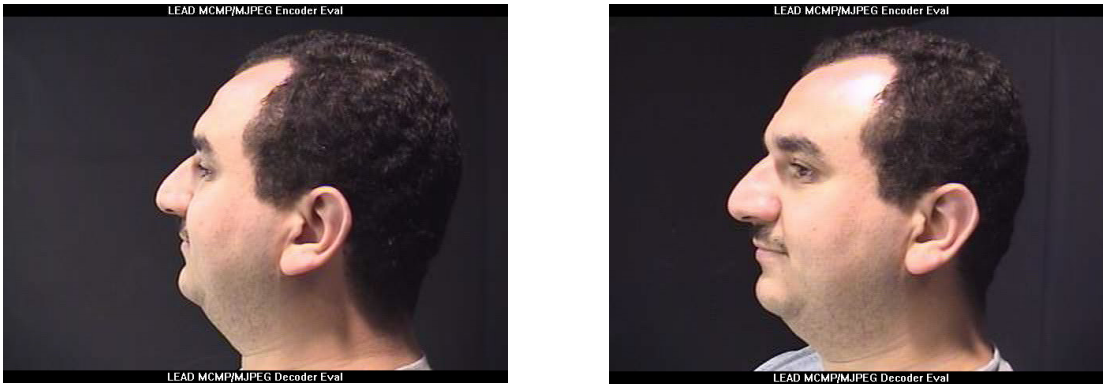
Figure 68: Sample image of a case from the database

We use the Euclidian distance between the extracted features as a matching score. The Euclidean distance between points P = $(p_1, p_2, \ldots p_n)$ and Q = $(q_1, q_2, \ldots q_n)$, in Euclidean *n*-space, is defined as:

$$E_d = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \ldots + (p_n - q_n)^2} = \sqrt{\sum_{i=1}^{n}(p_i - q_i)^2}$$

5.5.1 Verification Experiment

In this experiment, we used all the training set for both training and testing. In testing we depends on eliminate one technique; so to test an image x of case y, we eliminate this image from the database. In other words we do not count the distance between the image and itself.

Figure 69 shows the ROC curve between the false accept rate (FAR) and false reject rate (FRR) based on the above experimental setup. Where the equal error rate (ERR=22.9239, when FAR = FRR).
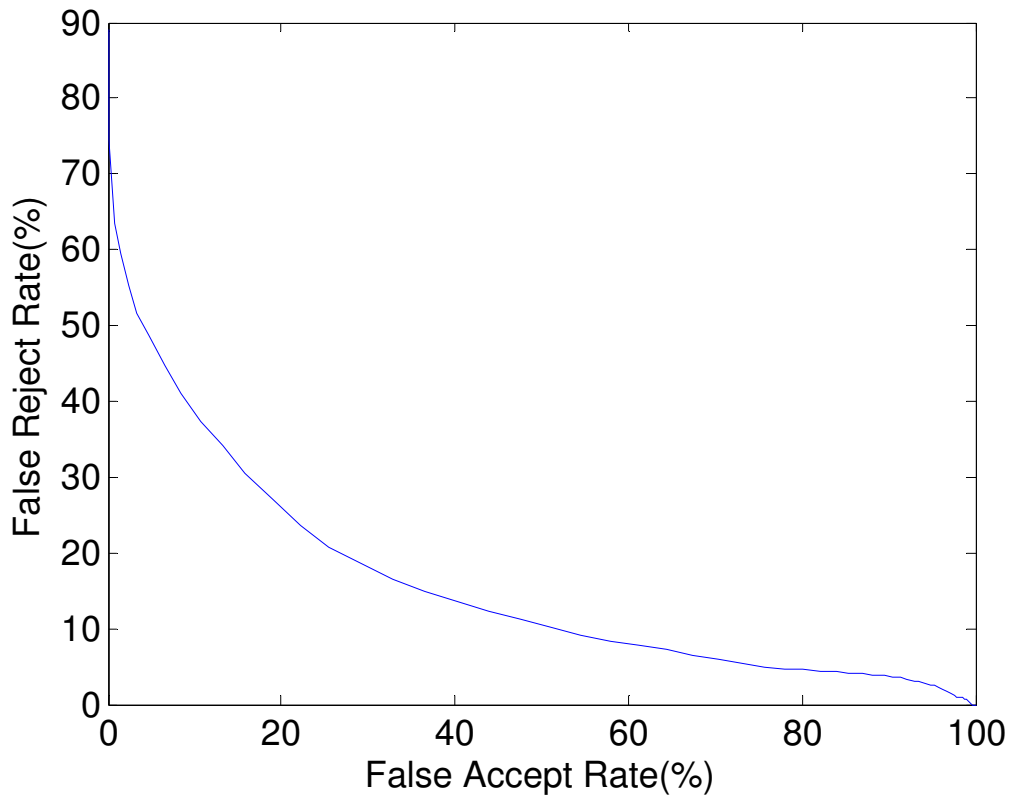
Figure 69: ROC curve

5.5.2 Identification Experiments

For the first identification experiment, we used the same training samples, and 23 different videos representing 23 subjects from the training dataset. This is a repeated (second) video sequence; we used to have for some subjects in the original dataset. Each case is represented by 10 images as previously described.

Figure 70 shows the results for the above mentioned identification experiment where we have 90% rank 1, in other words, 90% of the times the correct identity is represented by the first retrieved image. Further analysis for the curves implies that by retrieving the first ten images, we can achieve 94% retrieval accuracy.
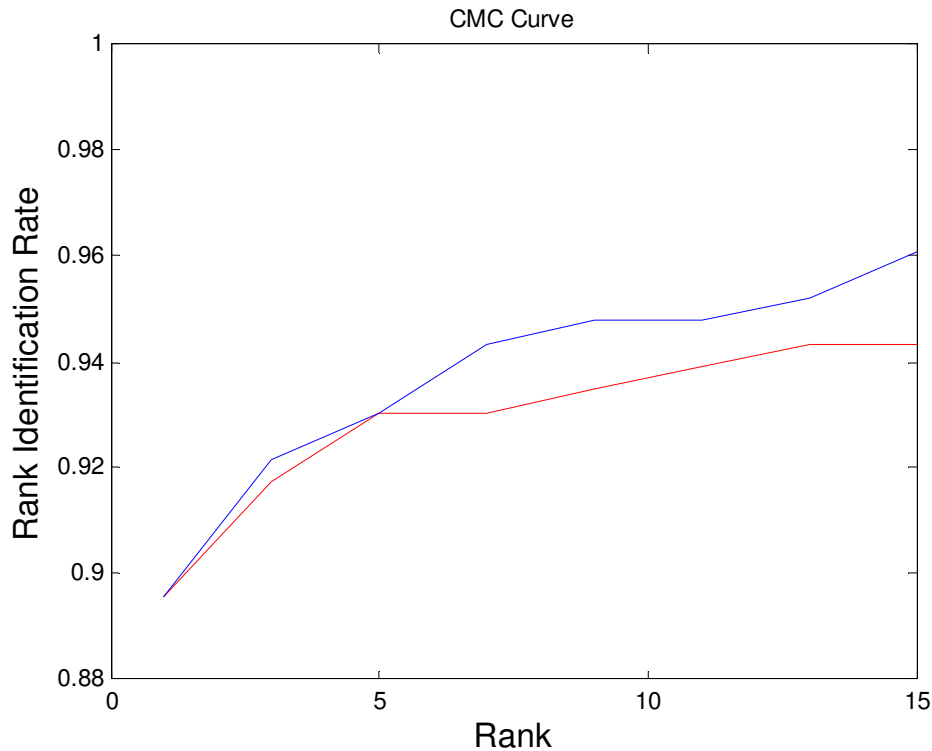
Figure 70: The blue curve for the first retrieved k cases; and the red for k images

For the first identification experiment, we used the same training samples, and 23 different videos representing 23 subjects from the training dataset (representing true matched cases). We also used a third set of 243 subjects different from the training dataset (representing unmatched cases). To plot True Hit Rate versus False Hit Rate:

- Let the training dataset gallery consist of images $G = \{G_1, G_2, ..., G_m\}$, m=1100 for this experiment.

- Let the probe consist of two sets of images:

$P = \{P_1, P_2, ...,P_n\}$: these are images which do not have any matching entry in G,

$Q = \{Q_1, Q_2, ...,Q_l\}$: these are images which have at least one matching entry in G,

for this experiment, n=2430, l=230.

- Let D (A, B) represent the Euclidian distance (matching score) generated when comparing A against B, and let t be an arbitrary matching threshold.

- For each P_i in P, i=1,2,...n, we compute s= min_j $\{S(P_i, G_j)\}$, j=1,2,...m. If d< t then this is a false hit. Thus, the false hit rate F(t) can be estimated as the number of false hits divided by n.

---

5. Automated Ear Identification System (AEIS)

- For each Q_i in Q, i=1,2,...m, we compute s = min_j{$S(Q_i, G_j)$}, and r = arg min_j{$S(Q_i, G_j)$}, j=1,2,...m (rank one images). If (d < t && identity of G_r is the same as identity of Q_i), then this is a true hit. Thus, the true hit rate H(t) can be estimated as the number of true hits divided by m.

- We plot H(t) against F(t) for varying values of t, as shown in Figure 71.

From Figure 71, as can be induced from the curves, by fixing the False Hit Rate to 5%, and use the corresponding threshold, we find the True Hit Rate to be 87.5%.
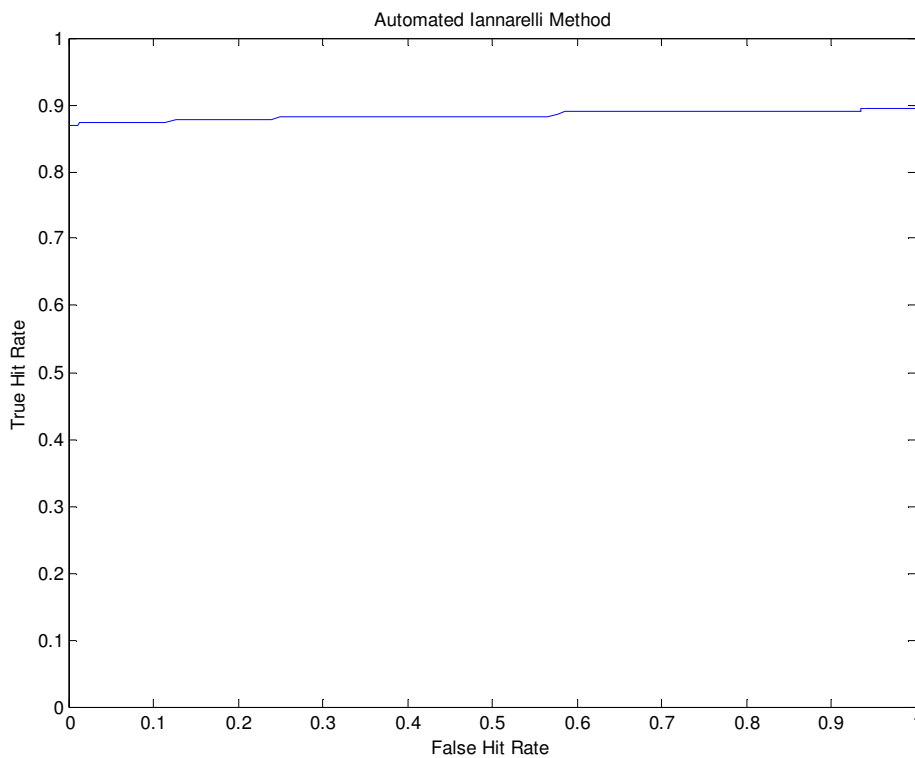
Figure 71: True hit rate versus false hit rate

5.5.3 Distance Measurements Accuracy

One of the limitations of our edging module is that while it can capture the external edges of the ear reliably (detection rate ranging from 97% to 99%), it is less consistent in capturing the edges within the internal structure of the ear as shown in Figure 72.
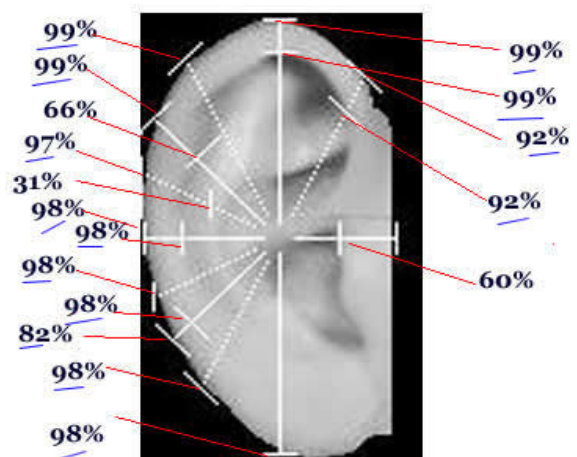
Figure 72: Detection rate

6 Conclusion and Future Work

We presented different problems regarding building a prototype for an Automated Dental Identification System (ADIS), as well as an Automated Ear Identification System (AEIS).

Automated Dental Identification System (ADIS)

ADIS can be used by law enforcement agencies for resolving cases of Missing and Unidentified Person (MUP). The main contribution is the proposed ADIS architecture that can handle a large database of images. Our proposed architecture is split in two main stages, where the first stage search the database for the candidate images, and the other stage reduces the number of candidate images to a match list of a few cases to be given to the forensic expert for a final decision. Samples of the performance results of this architecture were presented along with the description of the methodologies proposed. More extensive results of each component can be found in the list of ADIS publications. This is the first attempt to present a comprehensive ADIS in the computer vision/ image processing and pattern recognition areas. ADIS system is important to law enforcement and digital government officers managing mass disasters (e.g. Tsunami). To directly address the scalability of ADIS, we need the potential Match Search component (a very fast searching engine) to achieve very high percentage of having the target reference in a relatively very short candidate list.

- For the record cropping problem, we presented a global segmentation technique of cropping the dental films from the dental X-ray records. We started by using the rectangular film property and separating the background that have various colors and textures. Then we classify the connected components according to the corners being right or round. Cutting the round corner components depending on whether they are 90° or 180° V-corners; and cutting the right corner ones by viewing the boundary films under a graph theoretic framework, where it is always possible to reduce the graph by successively removing unit vertexes without affecting its connectivity. This technique performance can be summarized as follows: 73.7% of the films were perfectly cropped, 23.8% were under segmented, and only 2.5% films developed into erroneous segments. We measured the record cropping time of our algorithm using an uncompiled MATLAB$^{®}$ implementation running on a 2.0 GHz 512 MB RAM Intel$^{®}$ Pentium IV PC platform. The average cropping time is 30 kpix/sec and it varies depending on the number of films in the record and the amount of separation between films.

6. Conclusion and Future Work

- For the tooth contour extraction problem, an active contour without edges algorithm is used to extract the contour of teeth from dental images. This is an important component of ADIS. As digital X-ray images of the teeth usually do not exhibit sharp edges, the proposed algorithm is used to extract the tooth contour, after segmenting the tooth region, using a technique that depends on the intensity of the whole tooth region and not only on their edges. Experimental results suggest the efficacy of the contour extraction using the proposed technique over others described in the literature. The proposed contour extraction technique extracted a smooth and accurate teeth crown for 87% of the examined 340 teeth. We measured the tooth contour extraction time using our algorithm using an uncompiled MATLAB$^®$ implementation running on a 2.0 GHz 512 MB RAM Intel$^®$ Pentium IV PC platform. The average tooth contour extraction is 0.16 seconds.

- For the teeth labeling problem, we presented a dual-stage approach for automatic construction of dental charts based on appearance-based features and string matching with don't care possibility. Initially, each segmented tooth (of a bitewing/periapical film) is independently classified based on teeth reconstruction (and Fisher teeth) in two/four image subspaces. Including this previous knowledge about film type class enhances the overall initial classification by 10%. Using teeth-neighborhood rules the initial teeth class assignments is validated and possibly corrected even in the presence of missed tooth. Finally, if the resulting teeth class sequence is unique, a number for each tooth is assigned corresponding to its position in the dental quadrant it belongs to. Otherwise, if the resulting teeth class sequence is non-unique we call a reject option. Experimental results based on a sizable dataset suggest that appearance-based features achieve fast classification with an average film classification time of 0.4 seconds with competitive accuracy compared to methods proposed in the literature. Overall around 75% of the teeth are initially assigned to their proper classes, 86% of the teeth are assigned proper class labels after validation, and around 70% of the teeth assigned proper intra-quadrant numbers.

- For the potential match problem, we presented fast searching engine techniques based on color features (Eigen Images) as well as shape features (moment invariant and edges orientation histogram). By fusing multiple teeth matching scores, we calculate the Euclidean distance between the subject record and all the reference records of the database. Then we retrieve records based on minimum distance. Testing results based on a sizable data-set suggest that appearance-based features achieve fast (average record retrieving time of 5 seconds) and accuracy of retrieving 75% for rank 6.

6. Conclusion and Future Work

Challenges with building ADIS:

The main challenge to the current ADIS system is the poor quality of several dental images. This poor image quality can be defined as law resolution and law gray scale contrast of many dental radiographic films. Other challenges occur when the subjects have dental work, fillings, etc, which limits the effectiveness of the different proposed preprocessing methodologies.

The research on achieving higher levels of automation in the dental identification process, beyond manual feature extraction and coding, is fairly recent and there are many interesting problems that have to be addressed in order to achieve higher level of automation, and to enhance the overall performance regarding both accuracy and time. In the following we list some of these problems:

- Automatic construction of dental charts: for some cases, films are organized in a manner that does not immediately reflect the location teeth in the dental human atlas, for example, films may be flipped horizontally. In practice, dentists rely on a physical impulse on films to determine the front face of the film and accordingly determine whether teeth are right side of left side teeth. Detection of this impulse may be a feasible problem from an image processing point of view. The remaining problem would be to determine whether the impulse surface touching the scanner was convex or concave.

- Fillings and partial teeth: the relatively high likelihood of physical transformation is a characteristic of dental features. Restorative fillings and partial fracture of teeth, present two possibilities of physical transformations that manifest as change in the shape and content of teeth projections. Image painting is a candidate approach to work with teeth that were transformed by either one of the mentioned transformations.

- High-level features for efficient potential match search: so far all the studied features, in the context of the ADIS research prototype, do not possess invariance properties that that are needed for efficient retrieval. Also detection of fillings is essential for ruling out potential mismatches due to impossibility of evolution from an AM to a PM situation (e.g. a filled AM molar may not evolve to a virgin PM molar). Moreover, detection of fillings and description of the restored surfaces, and hence automatic annotation of teeth with dental codes, present a significant improvement opportunity that uplifts the automation level in existing systems like CAPMI and WinID[®].

- Matching strategies: as we mentioned above, films might be flipped horizontally, and unless these transformations are detected we cannot be very certain about the proper orientation of the dental

chart. One way around surviving with this situation is form four hypothesis regarding the proper orientation of a subject dental chart, perform matching (not necessarily exhaustive matching) under each of these hypotheses. Then select the strongest hypothesis for a more comprehensive case-to-case matching. We also recall that in some cases the number of tooth-pair views that we compared was unreasonable high, and in order not compromise on timeliness we may limit the number of views to be compared per tooth-pair.

- Error detection capabilities: each of the steps involved in processing of dental records is prone to errors that are likely to affect the overall performance of the system. The aspect ratio and relative areas were used to exclude cropped regions that have apparent violations to expected values of these properties. We think that many segmentation errors may as well be detected by identifying a set of geometric and textural features possessed by true teeth. Also in alignment, we may use the resulting overlap area as basis to accept or reject registration.

- Performance limits based on matching multiple teeth: the intra-subject variability observed and encountered in teeth appearance features reflects typical characteristics of a biometric. However, teeth multiplicity is also a distinctive characteristic of dental records, which we hypothesize, can raise the upper limits on performance of dental identification based on appearance features. It would be interesting to conduct an empirical study to validate our hypothesis. Such a study requires a large database of genuine matched teeth.

- In addition to the listed problems, we think that advocating imaging and digitizing standards in dental imaging will certainly aid automation the dental identification process.

Automated Ear Identification System (AEIS)

We developed an automated ear identification system, based on automating the manual Iannerelli's System. We presented different challenging problems starting with face and ear segmentation, ear image enhancement and ending with different point detection, and matching. Testing the performance of the proposed automated Iannarelli system shows good identification rate 90% for rank one image and works extremely fast. We returned the reason that the automated system showed lower performance compared to the manual Iannerelli's System to the difficulty of extracting internal features.

Our plan for future work for ADIS includes:

- For record cropping: exploiting more geometric properties of films to develop additional post-processing rules. This will identify segments that require further processing by a complementary, more computationally expensive cropping.

- For tooth contour extraction: exploiting more teeth boundaries properties based on the human experts domain of knowledge, so as to be able to correct many of the encountered neighboring errors.

- For teeth labeling: (i) Investigating whether filling detection and substitution using image-in-painting techniques will help; (ii) Taking into consideration the relative pulp size, which is smaller in premolars as compared to molars, to improve the classification accuracy of premolars and molars; (iii) Learning the misclassification risks, to improve teeth-numbering accuracy; (iv) Fusion between the appearance based and the feature based techniques for better labeling.

- For potential match problem: (i) Adding weights to teeth matching score fusion, for example based on teeth type like more or less weights to the molars; (ii) Fusion between shape features based approach and color features based approach.

Our plan for future work for AEIS includes:

(i) Develop better edge detection method for internal edges to improve internal feature extraction; (ii) Automatically align images before calculating distances; (iii) Automatically detect and repair broken edges, to decrease Failure Edge Acquiring Rates; (iv) Adjust weights to different features, according to variation (distinguishability) and detection rate; (v) Fusion between this 2-D technique and other 3-D techniques.

I.   Dental Glossary

Bitewing: A kind of dental x-ray which is taken with the teeth bite together. The main function of this kind of x-ray is to detect cavity in between teeth and height of bone support.

Canine: The third tooth from the middle of the jaw. There are totally 4 of them. They are the longest teeth in human.

Crown: A crown is almost like a "cap" on a tooth. It covers the tooth partially or totally above the gum to restore its function and outlook.

Dentistry: A branch of medicine that involves diagnosis, prevention, and treatment of any disease concern about teeth, oral cavity, and associated structures.

Filling: A restoration places on a tooth to restore its function and appearance.

Implant: A device (usually "screw-like") put in the jaw bone to support a false tooth, a denture or a bridge.

Incisor: The four upper and lower front teeth.

Molar: The last 3 upper and lower teeth on both side of the mouth.

Occlusal: The biting surface of the back teeth.

Panoramic Radiograph: An x-ray film to obtain the wide view of upper and lower jaw and their associated structures.

Periapical: The surrounding of the bottom of the root of a tooth.

Permanent teeth: Adult's teeth. The first permanent tooth usually comes in around 6 years old.

Premolar: The two teeth located in front of the molar.

Primary teeth: Baby teeth.

Radiograph: An x-ray picture.

Restoration: An item a dentist uses to restore the normal function of a tooth or an area in the mouth. It can be a filling, an artificial crown, etc.

Root: The bottom part of tooth. It anchors the tooth to its supporting units.

Wisdom tooth: The eighth (also the last tooth) tooth from the middle of the jaw.

II.   Ear Database Collection

The ear data set is collected at West Virginia University. A video sequence is captured for each case right, frontal and left face profiles. The raw data is a video sequence 180° for the case face and ear taking about 100 seconds. To collect this data, we are following Institutional Review Boards (IRB) protocol. This appendix is organized as follows: II-A IRB protocol and instructions, II-B IRB Consent Form, II-C system description, and II-D Database distribution.

6.1   A. IRB protocol and instructions

The principles of protection of human participants in research were established in the Belmont Report in 1979. The Belmont Report was prepared by the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research at the request of the Secretary of the Department of Health and Human Services (DHHS). The report identified three principles essential to the ethical conduct of research with humans: respect for persons, beneficence, and justice. These three principles formed the foundation for the conduct of research, including guidelines for obtaining informed consent, respect for privacy and confidentiality, and risk/benefit assessment.

Research participants are essential to the conduct of research, enabling researchers to make progress and discoveries in the fields of medicine and health. As such, the relationship between researchers and participants is critical and should be based on accurate information, trust, and respect.

History carries many examples of awful research involving unwilling human participants; the experiments conducted by Nazi physicians during World War II were unprecedented in their scope and the degree of harm and suffering to which human beings were subjected. The most notorious example in the United States of prolonged and knowing violations of the rights of a vulnerable group of research participants was the long-term study of black males conducted at Tuskegee by the United States Public Health Service. This study was initiated in the 1930s as an examination of the natural history of untreated syphilis; it continued until 1972. These examples as well as others were the motivation behind the development of codes of research ethics. Followed by the Declaration of Helsinki, CIOMS Guidelines, and then Belmont Report (1979). The Belmont Report can be found at the following site:

http://ohsr.od.nih.gov/guidelines/belmont.html.

The three fundamental ethical principles that guide the ethical conduct of research involving human participants are:

II. Ear Database Collection

1- Respect for Persons (autonomy)

2- Beneficence

3- Justice

In addition to ethical considerations, the Federal Government mandates certain legal standards for protection of humans in research. These standards are set forth in the *Code of Federal Regulations*, Title 45 CFR Part 46 (see http://www.hhs.gov/ohrp/humansubjects/guidance/45cfr46.htm). Subpart A, the basic policy for protection of human research subjects, is referred to as the "Common Rule." The regulations were enacted in 1991 and apply to all federally funded human research. Once a research activity is deemed human subject research, the Department of Health and Human Services requires review by Institutional Review Boards (IRB) and imposition of certain standards for informed consent.

Research is a systematic investigation (including development, testing, and evaluation) designed to discover or contribute to a body of generalizable knowledge. Not all research involves human participants, but when they are involved, researchers and their teams are legally and ethically obligated to protect them.

As data collection involves the participation of human subject, we need to be familiar with the regulations governing the rights and safety of research participants. In other words, we have to ensure that:

- The study is properly designed, scientifically sound, and yields valid results.

- Participants meet selection and eligibility requirements.

- The study is approved by the IRB and conducted according to the protocol.

- Informed consent is appropriately obtained.

- Protocol changes and adverse events are reported to the appropriate boards and authorities.

- The rights and welfare of participants are monitored throughout the trial.

- All members of the research team are qualified and trained in research methods and human participant protections.

Research team members (key personnel) whose responsibilities are delegated by the researcher also have a role in ensuring integrity of the study by consistently applying procedures and ensuring that the rights of participants are safeguarded.

Research team members are typically responsible for:

II. Ear Database Collection

- Day-to-day protocol decision making related to study conduct.

- Participant recruitment, selection, and eligibility.

- Clarification of the complexities of the protocol to the participant and others.

- Collection of participant information and entry of data using procedures to maintain privacy and confidentiality.

- Ensuring that the rights and welfare of participants are monitored throughout the study.

To address the above requirement, and to protect the collected data, we had the following plan:

1- Physical access to the lab (G75D) of the Engineering Science building is controlled through hand geometry.

2- We lock the door after entering, and unauthorized persons are not allowed to enter with you without giving his hand geometry.

3- Collected data is to be protected by appropriate firewall and password security limited access to the machine located in (G75D) of the Engineering Science building.

4- We never distribute this data, or copy it to any other machine.

5- Control of access to the collected data and results that come out of this project, as well as descriptions of the developed software that are discussed at meetings held at WVU will be achieved through limiting those in attendance to the pre-approved members.

6- Data and results stored on PC hard drives are to be deleted from the storage media when they are no longer needed. Hard drives that might be replaced will first have all data and results files deleted, and then the hard drive will be reformatted, and finally, the physical hard drive will be destroyed.

6.2   B. IRB Consent Form

CONTACT PERSONS

In the event you experience any side effects or injury related to this research, you should contact Dr. Arun Ross at 304/293-0405 ext 2556. For more information about this research and about research-related risks or injury, you can contact Dr. Hany Ammar at 304/293-0405 ext 2514.

INTRODUCTION

You,_____,   have been asked to participate in this research study, which has been explained to you by Dr. Arun Ross, Ph.D. This study is being conducted by [Dr. Hany Ammar, Ph.D., and Dr. Arun Ross, Ph.D.] in the Lane Department of computer Science and Electrical Engineering at West Virginia University with funding provided by the Technical Support Working Group.

PURPOSE OF THE STUDY

The purpose of this study is to improve biometric research techniques for ear and face recognition. WVU expects to enroll approximately 500 subjects in this study.

DESCRIPTION OF PROCEDURES

This study involves capturing your ear and facial images, using commercial digital video cameras and will take approximately five – ten minutes to complete. You have a period of thirty days during which you can withdraw your data from the study with no consequences.

RISKS AND DISCOMFORTS

You should not experience any inconvenience or discomfort beyond that typically associated with your face and ear pictures being taken. The only potential risk from participating in this study is the distribution of your images to other agencies and other researchers, due to its inclusion in the dataset. However, no identifying information (such as name or social security number) will be associated with your data image. The different biometric data will be saved on secure personal computers in the following location: Engineering Science Building, room G75D, Evansdale Campus, and is coded with specific number so that there will be no means of associating name with the corresponding biometric data.

ALTERNATIVES

BENEFITS

You may not receive any direct benefit from this study. The knowledge gained from this study may eventually benefit others.

II. Ear Database Collection

FINANCIAL CONSIDERATIONS

You will incur no costs for participating in this study. If you participate in this study, you will be given $5 gift card for the session. If you withdraw from the study before the session is completed, there will be no payment for this session.

CONFIDENTIALITY

Any information about you that is obtained as a result of your participation in this research will be kept as confidential as legally possible. Your research records and test results, just like hospital records, may be subpoenaed by court order or may be inspected by federal regulatory authorities without your additional consent. In addition, there are certain instances where the researcher is legally required to give information to the appropriate authorities. These would include mandatory reporting of infectious diseases, mandatory reporting of information about behavior that is imminently dangerous to you or to others, such as suicide, child abuse, etc. Data collected from participants is stored with an identification number generated for the purpose of this data collection. No identifying information such as name or social security number will be stored with the data in the dataset. In any publications that result from this research, neither your name nor any information from which you might be identified will be published without your consent.

VOLUNTARY PARTICIPATION

Participation in this study is voluntary. You are free to withdraw your consent to participate in this study at any time. Refusal to participate or withdrawal will involve no penalty or loss of benefits and will not affect your grades or class standing. You have been given the opportunity to ask questions about the research, and you have received answers concerning areas you did not understand. In the event new information becomes available that may affect your willingness to participate in this study, this information will be given to you so that you can make an informed decision whether or not to continue your participation. Upon signing this form, you will receive a copy.

6.3 C. System Description
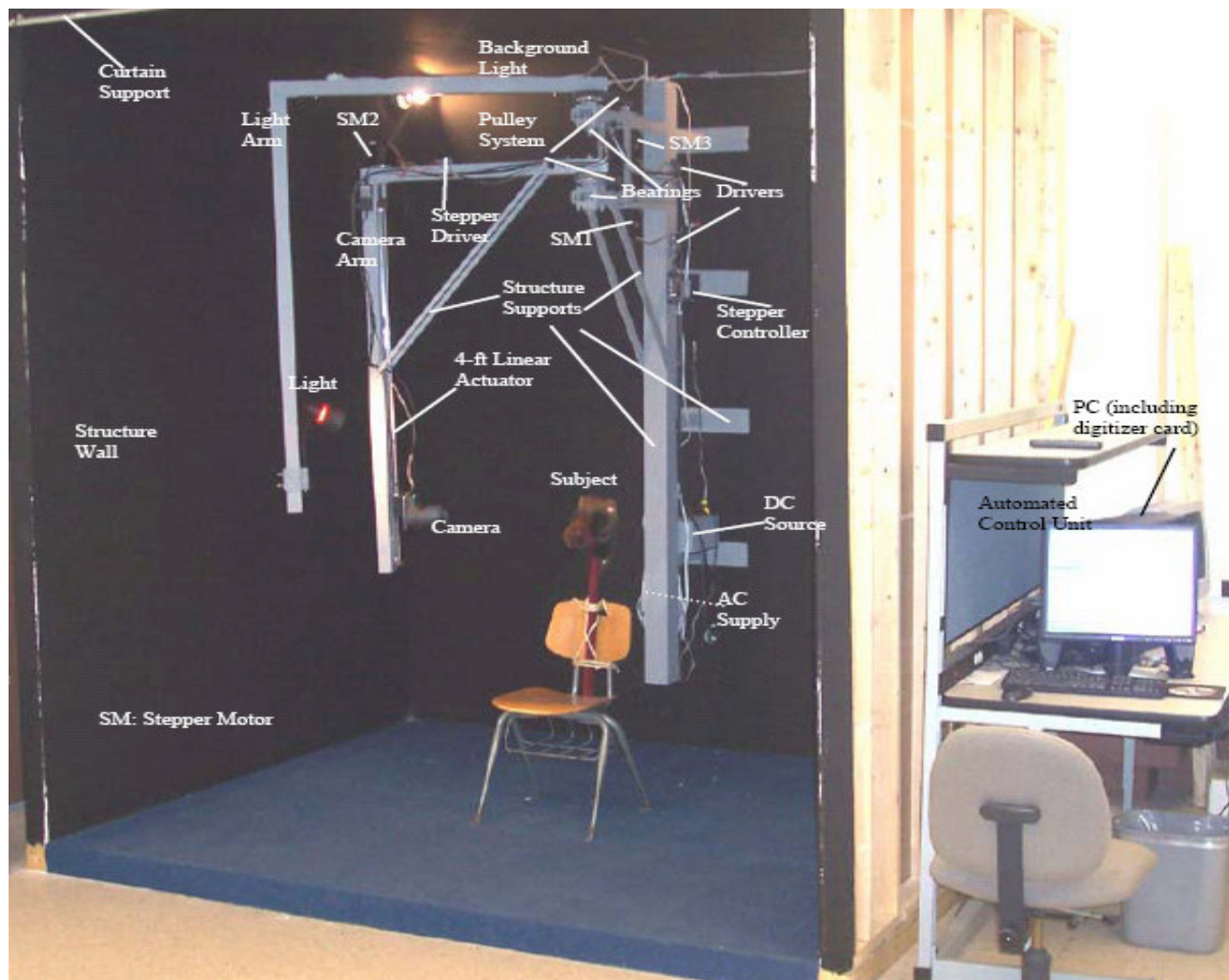
Figure 73: Structure and mechanical components for the lab

Figure 73 describes the system used in ear data set collection. This system was developed by Dr. Gamal Fahmy and Dr. Ahmed El-Sheerbiny from West Virginia University. This system had undergone various design, assembly, and implementation. The main hard ware components for this system include:

- PC: serves to provide complete control of the motion and image capturing.

- Camera: allowing capturing of video, it is attached to the camera arm, and the latter is controlled by a stepper motor.

- Linear Actuator: This is a unique custom made device (by Netmotion, Inc.) has 4-ft span and allows smooth, vertical (up or downward) translation. We use this to adjust the height of the camera according to the subject height.

II. Ear Database Collection

- Light: For our experiment, we fixed the light to same arm of the camera. In other words, we didn't use the light arm.

- Structural Framework: consisting of tinted wall, rotating arms, and other structural supports. We added uniform black board behind the chair, to serve as a uniform background during video capturing.

There are various software packages that we use:

- Posteus IPE: stepper system control. We used it to adjust the camera height, and then to move the camera around the object.

- EVI Series Demonstration Software: Camera control, to adjust zoom, tilt, and focus of the used camera.

- IC Capture: Video capturing, to record the subject different poses during the camera rotation.

6.4 D. Database Distribution

We collected video sequences for 402 different cases over a period of 2 years. We collected multi-sequence for 54 cases out of those 402. The database has the ethnicity and gender diversities shown in Figure 74.
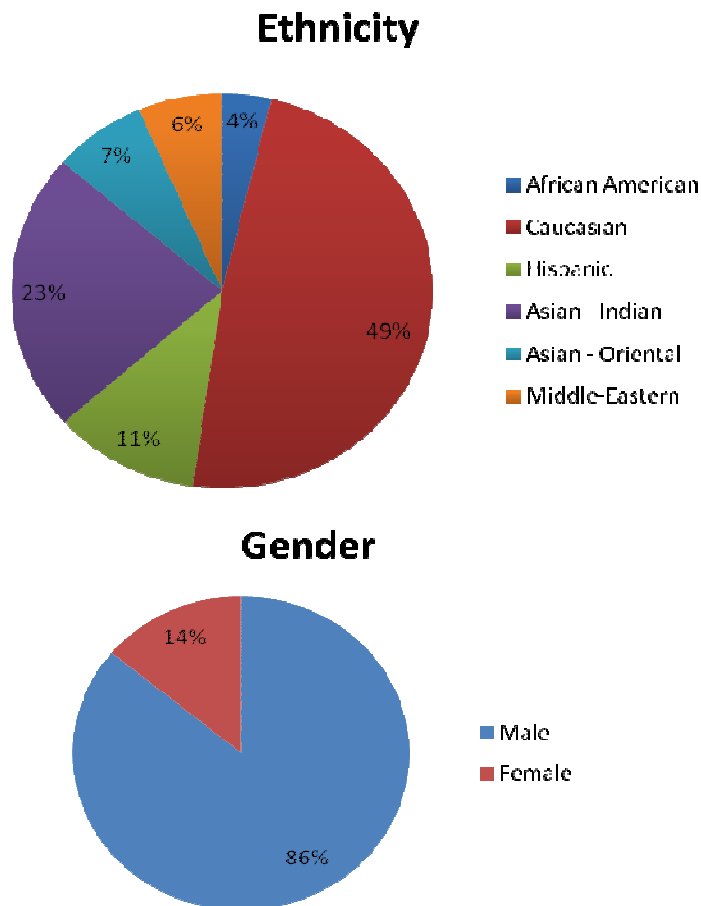

Figure 74: Ethnicity and Gender distribution

Additional Statistics include:

- 55 subjects with eyeglasses

- 42 subjects with earrings

- 38 subjects with partially occluded ears

- 2 fully occluded ears, in other words, not useful for the current study.

III. Bibliography

[1] Wikipedia, the free encyclopedia, http://en.wikipedia.org/wiki/Main_Page.

[2] About Inventors, http://inventors.about.com/od/fstartinventions/a/forensic.htm

[3] P. Stimson, and C. Mertz, "Forensic Dentistry", CRC Press 1997.

[4] . M. Bowers, "Forensic Dental Evidence: An Investigator's Handbook", Academic Press 2004

[5] I. Pretty, and D. Sweet, "A look at forensic dentistry - Part 1: The role of teeth in the determination of human identity", British dental journal, vol. 190, no. 7, pp. 359-66, 2001.

[6] E. Ndiokwelu,J. Miquel; N. Coudert, "Identification of victims of catastrophes: introduction to the role of forensic Odontology", Odonto-stomatologie tropicale (Tropical dental journal), vol. 26, no. 104, pp. 33-36, 2003.

[7] S. Keiser-Nielsen, "A six-digit code for computer-aided dental identification**,** Journal of forensic sciences (JFSCA), vol. 21, no. 1, pp. 85-89, 1983.

[8] G. Jones, "A simple overlay system for data comparison in dental identification", Journal of forensic sciences (JFSCA), vol. 33, no. 1, pp. 254-259, 1988.

[9] L. Lorton, M. Rethman, and R. Friedman, "The Computer-Assisted Postmortem Identification (CAPMI) System: A Computer-Based Identification Program," Journal of Forensic Sciences (JFSCA), vol. 33, no. 4, pp. 977-984, July 1988.

[10] James McGivney, WinID3, http://www.winid.com/

[11] C. Lewis, "WinID2 versus CAPMI4: Two Computer-Assisted Dental Identification Systems" Journal of forensic sciences (JFSCA), vol. 47, no. 3, pp. 536–538, 2002.

[12] National Dental Image Repository website, http://foia.fbi.gov/ndirpia.htm

[13] Integrated Automated Finger Identification System, http://www.fbi.gov/hq/cjisd/iafis.htm

[14] A. Jain, and H. Chen, "Matching of Dental X-ray images for Human identification", Pattern Recognition, vol. 37, pp. 1519-1532, July 2004.

[15] M. Mahoor, and M. Abdel-Mottaleb, "Classification and Numbering of Teeth in Dental Bitewing Images", Pattern Recognition, vol. 38, pp.577-586, April 2005.

[16]   O. Nomir and M. Abdel-Mottaleb, "A system for Human Identification from X-Ray Dental Radiographs", Pattern Recognition, vol. 38, pp. 1295-1305, August 2005.

[17]   J. Zhou, and M. Abdel-Mottaleb, "A Content-based System for Human Identification based on Bitewing Dental X-Ray Images", Pattern Recognition, vol. 38, pp. 2132-2142, Nov 2005.

[18]   H. Chen and A. Jain, "Dental Biometrics: Alignment and Matching of Dental Radiographs", IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), vol. 27, no. 8, pp. 1319-1326, August 2005.

[19]   G. Fahmy, D. Nassar, E. Haj Said, H. Ammar, M. Abdel-Mottaleb, H. Chen, A. Jain, "Toward an Automated Dental Identification System", Journal of Electronic Imaging, vol. 14, no. 4, 2005.

[20]   E. Haj Said, D. Nassar, G. Fahmy, and H. Ammar, "Teeth Segmentation in Digitized Dental X-Ray Films Using Mathematical Morphology" IEEE Transactions on Information Forensics and Security, vol. 1, no. 2, June 2006.

[21]   D. Nassar, and Hany Ammar, "A Neural Network System for Matching Dental Radiographs", Pattern Recognition, vol. 40, pp. 65-79, Jan 2007.

[22]   D. Nassar, and H. Ammar, "A Prototype Automatic Dental Identification System", Proc. of the 4th National Conference on Digital Government Research (dg.o.), Boston – MA, 2003.

[23]   D. Nassar, E. Haj Said, G. Fahmy, R. Howel, and H. Ammar, "Challenges of developing an Automated Dental Identification Systems", Proc. of the Biometric Consortium Conference (BCC), Crystal City, Arlington – VA, 2003.

[24]   A. Jain and H. Chen, "Dental Biometrics: Matching Dental X-rays for Human Identification", Proc. of the Biometric Consortium Conference (BCC), Crystal City, Arlington – VA, 2003.

[25]   A. Jain, H. Chen and S. Minut, "Dental Biometrics: Human Identification Using Dental Radiographs", Proc. of the 4th International Conference on Audio- and Video-Based Biometric Person Authentication (AVBPA), pp. 429-437, Guildford - UK, Jun 2003.

[26]   M. Abdel-Mottaleb, O. Nomir, D. Nassar, G. Fahmy, and H. Ammar, "Challenges of Developing an Automated Dental Identification System", Proc. of the IEEE mid-west symposium for circuits and systems, Cairo - Egypt, Dec 2003.

[27]   J. Zhou, and M. Abdel-Mottaleb, "Automatic Human Identification Based on Dental X-ray Images", Proc. of SPIE Technologies for Homeland Security and Law Enforcement conference, Orlando - FL, Apr 2004.

[28]   E. Haj Said, D. Nassar, G. Fahmy, and H. Ammar, "Dental X-ray Image Segmentation", Proc. of SPIE Technologies for Homeland Security and Law Enforcement conference, Orlando - FL, Apr 2004.

[29]   G. Fahmy, D. Nassar, E. Haj Said, H. Chen, O. Nomir, J. Zhou, R. Howell, H. Ammar, M. Abdel-Mottaleb and A. Jain, "Automated Dental Identification System (ADIS)", Proc. of the 5th National Conference on Digital Government Research (dg.o.), May 2004.

[30]   H. Chen, A. Jain, "Tooth contour extraction for matching dental radiographs", Proc. of the International Conference on Pattern Recognition (ICPR), vol. III, pp. 522-525, Cambridge - U.K., Aug 2004.

[31]   G. Fahmy, D. Nassar, E. Haj Said, H. Chen, O. Nomir, J. Zhou, R. Howell, H. Ammar, M. Abdel-Mottaleb and A. Jain, "Towards an Automated Dental Identification System (ADIS)", Proc. of the 1st international Conference on Biometric Authentication (ICBA), Hong Kong, Aug 2004.

[32]   H. Chen and A. Jain, "Dental Biometrics: Matching of Dental Radiographs for Human Identification", Proc. of the Biometric Consortium Conference (BCC), Crystal City, Arlington – VA, 2004.

[33]   M. Mahoor, M. Abdel-Mottaleb, "Automatic Classification of Teeth in Bitewing Dental Images", Proc. of the international Conference on Image Processing (ICIP), Singapore, Oct 2004.

[34]   D. Nassar, F. Chaudhry, and H. Ammar, "On Performance Evaluation of Image Segmentation Algorithms", Proc. of 1st International Computer Engineering Conference, Cairo - Egypt, Dec 2004.

[35]   H. Chen and A. Jain, "Dental Biometrics: Alignment and Matching of Dental Radiographs", Proc. of Workshop on the applications of Computer Vision (WACV), pp. 316-321, Breckenridge, Colorado, Jan 2005.

[36]   A. Jain and H. Chen, "Registration of dental atlas to radiographs for human identification", Proc. of SPIE Conference on Biometric Technology for Human Identification II, vol. 5779, pp. 292-298, Orlando, Florida, Mar 2005.

[37]   D. Nassar, E. Haj Said, A. Abaza, S. Chekuri, J. Zhou, M. Mahoor, O. Nomir, H. Chen, G. Fahmy, H. Ammar, M. Abdel-Mottaled and A. Jain, "Automated Dental Identification System (ADIS)" Proc. of the 6th National Conference on Digital Government Research (dg.o.), Atlanta - GA, May 2005.

[38]   S. Chekuri, D. Nassar, A. Abaza, A. Bahu, H. Ammar, and G. Fahmy, " A web-based Automated Dental Identification System (webADIS)", Proc. of the 5th IBIMA International Conference on Internet & Information Technology in Modern Organizations, Cairo - Egypt, Dec 2005.

[39]   X. Li, A. Abaza, D. Nassar, and H. Ammar, "Fast and Accurate Segmentation of Dental X-ray Records", Proc. of the 2nd international Conference on Biometric Authentication (ICBA), Hong Kong, Jan 2006.

[40]   E. HajSaid, D. Nassar, and H. Ammar, "Image segmentation for automated dental identification", Proc. of IS&T/SPIE Electronic Imaging, San Jose - CA, Jan 2006.

[41]   D. Nassar, M. Ogirala, D. Adjeroh, and H. Ammar, "An Efficient Multi-Resolution GA approach to Dental Image Alignment", Proc. of IS&T/SPIE Electronic Imaging, San Jose - CA, Jan 2006.

[42]   S. Chekuri, D. Nassar, A. Abaza, E. HajSaid, A. Bahu, U. Qurashi, G. Fahmy, and H. Ammar, "webADIS: A Flexible web-based Enviroment for the Automated Dental Identification System" Proc. of the 7th Annual International Conference on Digital Government Research (dg.o.), San Deigo – CA, May 2006.

[43]   S. Shah, A. Abaza, A. Ross and H. Ammar, "Automatic Teeth Segmentation Using Active Contour Without Edges" Proc. of the Biometric Consortium Conference (BCC), Baltimore - MD, Sep 2006.

[44]   O. Nomir and M. Abdel-Mottaleb, "Hierarchical Dental X-Ray Radiographs Matching" Proc. of the International Conference on Image Processing (ICIP), Atlanta - GA, Oct 2006.

[45]   S. Chekuri, D. Nassar, A. Abaza, A. Bahu, U. Qureshi, and H. Ammar, "webADIS: A Flexible web-based Environment for the Automated Dental Identification System", Proc. of the 2nd International Computer Engineering Conference (ICENCO), Cairo - Egypt, Dec 2006.

[46]   Quereshi, U., "Classification of dental x-ray films", Masters Thesis, Department of Electrical and Computer Engineering - West Virginia University, August 2006.

[47]   CJIS Division – Washington State Patrol Missing and Unidentified Person Unit, Digitized Dental Images (Database), May 2000.

[48]   CJIS Division – ADIS project, Digitized Radiographic Images (Database), August 2002.

[49]   R. Gonzalez, and R. Woods, "Digital Image Processing, 2nd edition", Prentice Hall, 2002.

[50]   P. Maragos and R. W. Schafer, "Morphological Systems for Multidimensional Signal Processing", *Proc. IEEE,* vol. 78, pp. 690-710, April 1990.

[51]   S. White, and M. Pharoah, *Oral Radiology Principles and Interpretation*. Mosby, Inc. Fourth Edition 2000.

[52]   M. Kass, A. Witkin, and D.Terzopoulos, "Snakes: Active contour models", International Journal of Computer Vision, vol. 1, no. 4, pp. 321-331, 1987.

[53]   R. Malladi, J. Sethian, and B. Vemuri, "Shape Modeling with Front Propagation: A Level Set Approach", IEEE Trans. on Pattern Analysis and Machine Intelligence (PAMI), vol. 17, no. 2, Feb. 1995.

[54]   J. Sethian, "A review of recent numerical algorithms for hypersurfaces moving with curvature dependent speed", Journal of Differential Geometry, vol. 31, pp. 131–161, 1989.

[55]   S. Osher and J. Sethian, "Fronts propagatin with curvature-dependent speed: algorithms based on hamilton-jacobi formulations", Journal of Computational Physics, vol. 79, no. 1, pp. 12–49, 1988.

[56]   G. Chen, Y. and S. Goto, "Uniqueness and existence of viscosity solutions of generalized mean curvature flow equations", Journal of Differential Geometry, vol. 33, pp. 749–786, 1991.

[57]   V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic active contours", International Journal of Computer Vision, vol. 22, no. 1, pp. 61–79, 1997.

[58]   T. Chan and L. Vese, "Active contours without edges", IEEE Trans. on Image Processing, vol. 10, no. 2, pp. 266–277, 2001.

[59]   D. Mumford and J. Shah, "Optimal approximation by piecewise smooth functions and associated variational problems," in Communications on Pure and Applied Mathematics, vol. 42, pp. 577–685, 1989.

[60]   D. Nassar, "Automated Dental Identification: A Micro-Macro Decision-Making Approach", Ph.D. Dissertation, Lane Department of Computer Science and Electrical Engineering, West Virginia University, Dec. 2005

[61]   American Medical Association, http://www:medem:com

[62]   Dental Resource Net, http://www.dentalcare.com/soap/ce_preclin/ce105/pg07.htm

[63]   S. Bunn, http://www:drbunn:com/number.htm

[64]   J. Canny, "A Computational Approach to Edge Detection", IEEE Trans. on Pattern Analysis and Machine Intelligence (PAMI), vol. 8, no. 6, pp. 679-698, 1986.

[65]   M. Turk and A. Pentland, Face Recognition Using Eigenfaces, Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Maui, Hawaii, June 1991.

[66]   P. Belhumeur, J. Hespanha, and D. Kriegman, Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection, IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), vol. 19, no. 7, pp. 711-720, July 1997.

[67]   R. Boyer and J. Moore, "A Fast String Searching Algorithm", Communications of the Association for Computing Machinery (ACM), vol. 20, no. 10, pp. 762-772, 1977.

[68]   R. Wagner and M. Fischer, "The String-to-String Correction Problem", Journal of the Association for Computing Machinery (JACM), vol. 21, no. 1, p.168-173, 1974.

[69]   T. Cormen, C. Leiserson, and R. Rivest, Introduction to Algorithms, New York: McGraw-Hill, 1990.

[70]   R. Baeza-Yates, Algorithms for string searching: A survey, ACM Special Interest Group on Information Retrieval (SIGIR) Forum, vol. 23, no. 3-4, pp. 34-58, 1989.

[71]   S. Needleman, and C. Wunsch, A general method applicable to the search for similarities in the amino acid sequence of two proteins, Journal of Molecular Biology, vol. 48, no.3, pp. 443-453, 1970.

[72]   A. Jain, R. Duin, and J. Mao, "Statistical Pattern Recognition: A Review", IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), vol. 22, no. 1, pp. 4-37, 2000.

[73]   R. Duda, P. Hart, and D. Stork, "Pattern Classification, 2nd edition", Wiley-Interscience, 2000.

[74]   J. B. Maintz and M. A. Viergever, "A survey of medical image registration", IEEE Transactions on Image Processing, vol. 2, no. 3, pp. 1–36, 1998.

[75]   L. G. Brown, "A survey of image registration techniques", ACM Computing Surveys (CSUR), vol. 24, no. 4, pp.325–376, 1992.

[76]   L. Junck, J. G. Moen, G. D. Hutchins, M. B. Brown, and D. E. Kuhl, "Correlation methods for the centering, rotation, and alignment of functional brain images", Journal of nuclear medicine, vol. 31, no. 7, pp. 1220-1276, 1990.

[77]   S. Cain and M. Hayat, "Projection-based image registration in the presence of fixed-pattern noise", IEEE Transactions on Image Processing, vol. 10, no. 12, pp. 1860–1872, 2001.

[78]   A. Khamene, R. Chisu, W. Wein, N. Navab, and F. Sauer, "A novel projection based approach for medical image registration.," in Workshop on Biomedical Image Registration (WBIR), pp. 247–256, 2006.

[79]   B. S. Reddy and B. N. Chatterji, "An FFT-based technique for translation, rotation, and scale-invariant image registration", IEEE Transactions on Image Processing, vol. 5, no. 8, pp. 1266–1271, 1996.

[80]   Anil K. Jain, Patrick Flynn and Arun A. Ross, "Handbook of Biometrics: Automatic Forensic Dental Identification" Springer US, 2007.

[81]   Forensic Identification of 9/11 Victims Ends, NEW YORK, Feb. 23, 2005

http://abcnews.go.com/WNT/story?id=525937&page=1

[82]   Dental records beat DNA in tsunami IDs, NewScientist.com news service, 08 September 2005, http://environment.newscientist.com/channel/earth/tsunami/mg18725163.900-dental-records-beat-dna-in-tsunami-ids.html

[83]   R. Veltkamp, M. Tanase. A Survey of Content-Based Image Retrieval Systems. In: O. Marques, B. Furht, Content-Based Image and Video Retrieval, Kluwer Academic Pub., pp. 47-101, 2002.

[84]   Plato. Meno. Perseus Encyclopedia, Tuft University, 380 B.C

http://www.perseus.tufts.edu/Texts/chunk_TOC.grk.html

[85]   S. Brand, J. Laaksonen, E. Oja, "Statistical Shape Features in Content-Based Image Retrieval," Proc. the 15th International Conference on Pattern Recognition (ICPR), vol. 2, 2000.

[86]   F. Long, H. Zhang, D. Feng, "Fundamentals of Content-based Image retrieval", in "Multimedia Information Retrieval and Management - Technological Fundamentals and Applications", by D. Feng, W. Siu, and H. Zhang., Springer, 2002.

[87]   D. Zhang and G. Lu, "Review of Shape Representation and Description", Pattern Recognition, vol. 37, no. 1, pp. 1–19, Jan 2004.

[88]   D. Zhang, and G. Lu, "Content-based shape retrieval using different shape descriptors: a comparative study" International Conference on Multimedia and Expo, pp. 1139-1142, Aug. 2001.

[89]   Y. Rui, A. She, T. Huang, "A modified Fourier descriptor for shape matching in MARS", Proc. of Image Databases and Multimedia Search, pp. 165–177 Singapore, 1997.

[90]   M. Hu, "Visual pattern recognition by moment invariants", Proc. of the Computer Methods in Image Analysis, Los Angeles, CA, 1977.

[91]   P. Miranda, R. Torres, and A. Falcão, "TSD: A Shape Descriptor Based on a Distribution of Tensor Scale Local Orientation", Proc. the In XVIII Brazilian Symposium on ComputerGraphics and Image Processing, pages 139–146, Natal, RN, Brazil, October 2005.

[92]   Y. Li and L. G. Shapiro, "Consistent Line Clusters for Building Recognition in CBIR," Proc. of the International Conference on Pattern Recognition, August 2002.

[93]   Ross and A. Jain, "Multimodal Biometrics: An Overview", Proc. of the 12th European Signal Processing Conference (EUSIPCO), pp. 1221-1224, Sept. 2004.

[94]   A. Ross and A. Jain, "Information fusion in biometrics," Pattern Recognition Letters, vol. 24, pp. 2115–2125, Sep 2003.

[95]   L. Kuncheva, C. Whitaker, C. Shipp, and R. Duin, "Is independence good for combining classifiers?," Proc. of the International Conference on Pattern Recognition (ICPR), vol. 2, pp. 168–171, 2000.

[96]   Y. Zuev and S. Ivanon, "The voting as a way to increase the decision reliability," in Foundations of Information/Decision Fusion with Applications to Engineering Problems, pp. 206–210, Aug 1996.

[97]   A. Ross and R. Govindarajan, "Feature Level Fusion Using Hand and Face Biometrics", Proc. of SPIE Conference on Biometric Technology for Human Identification II, (Orlando, USA), pp. 196-204, Mar 2005.

[98]   A. Jain, K. Nandakumar and A. Ross, "Score Normalization in Multimodal Biometric Systems", Pattern Recognition, vol. 38, no. 12, pp. 2270-2285, Dec 2005.

[99]   A. Jain, A. Ross, and S. Prabhakar, "An Introduction to Biometric Recognition", IEEE Trans. on Circuits and System for Video Technology, vol. 14, no. 1, Jan 2004.

[100]   D. Maltoni, D. Maio, A. Jain, S. Prabhakar, "Handbook of Fingerprint Recognition: Multimodal Biometric System" Springer US, 2003.

[101]   J. Kittler, M. Hatef, R. Duin, and J. Matas, "On Combining Classifiers", IEEE Transaction on Pattern Recognition and Machine Intelligence (PAMI), vol. 20, no. 3, pp. 226-239, Mar 1998.

[102]   R. Prokop, and A. Reeves, "A survey of moment-based techniques for unoccluded object representation and recognition", Graphical Models and Image Processing (CVGIP), vol. 54, no. 5, 1992.

[103]   L. Yang, and F. Algregtsen, "Fast computation of invariant geometric moments: A new method giving correct results", Proc. of the International Conference on Image Processing, 1994.

[104]   A. Jain and A. Vailaya, "Image retrieval using color and shape", Pattern Recognition, vol. 29, no.8, pp. 1233-1244, 1996.

[105]   P. Viola, and .Wells III, "Alignment by Maximization of Mutual Information", International Journal of Computer Vision, vol. 24, no. 2, pp. 137–154, 1997.

[106]   D. Hurley, B. Arbab-Zavar, and M. Nixon, "The Ear as a biometric" in A. Jain, P. Flynn and A. Ross, "Handbook of Biometrics", pp. 131-150, Springer US, 2007.

[107]   K. Pun, Y. Moon, "Recent advances in ear biometrics", Proc. of 6th IEEE International Conference on Automatic Face and Gesture Recognition, pp. 164- 169, 2004.

[108]   M. Burge, and W. Burger, "Ear Biometrics in Computer Vision", the 15th International Conference of Pattern Recognition (ICPR), pp. 826-830, 2000.

[109]   D. Bamber, "Prisoners to appeal as unique 'earprint' evidence is discredited", Telegraph Newspaper (UK): http://portal.telegraph.co.uk/news/main.jhtml?xml=/news/2001/12/02/nearp02.xml

[110]   K. Chang, K. Bowyer, S. Sarkar, and B. Victor, "Comparison and Combination of Ear and Face Images in Appearance-Based Biometrics", IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), vol. 25, no. 9, pp. 1160-1165, Sept. 2003.

[111]   D. Hurley, M. Nixon, and J. Carter, "Automatic Ear Recognition by Force Field Transformations", Proc. of the 5th International Conference on Audio- and Video-Based Biometric Person Authentication (AVBPA), pp. 386-394, 2005.

[112]   H. Zhang; Z. Mu; W. Qu; L. Liu; and C. Zhang; "A novel approach for ear recognition based on ICA and RBF network", Proc. of the 4th International Conference on Machine Learning and Cybernetics, vol. 7, pp.4511–4515, Aug. 2005.

[113]   M. Abdel- Motallleb, J. Zhou, " Human Ear Recognition from Face Profile Images", Proc. of the 2nd International Conference on Biometrics, Hong Kong, pp. 786-792, Jan. 2006.

[114]   H. Chen and B. Bhanu, "Human Ear Detection from Side Face Range Images", Proc. of the 17th International Conference on Pattern Recognition (ICPR), vol. 3, pp. 574 – 577, Aug. 2004.

[115]   H. Chen, B. Bhanu, "Shape Model-Based 3D Ear Detection from Side Face Range Images", Proc. of the Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), vol. 3, pp.122 – 127, June 2005.

[116]   A. Iannarelli, "Ear Identification, Forensic Identification Series", Paramont Publishing Company, Fremont, California, 1989.

[117]   A. Hoogstrate, H. Van den Heuvel, E. Huyben, "Ear Identification Based on Surveillance Camera's Images", Forensic-Evidence.com: http://www.forensicevidence.com/site/ID/IDearCamera.html

[118]   M. Burge, and W. Burger, "Ear Biometrics" in A. Jain, R. Bolle, and Sharath Pankanti, "Biometrics: Personal Identification in Networked Society", pp. 273-285, Kluwer Academic, 1998.

[119]   A. Moenssens, "Ear Identification Research Proposal", Forensic-Evidence.com: http://forensic-evidence.com/site/ID/ID00004_4.html

[120]   P. Viol, and M. Jones, "Rapid object detection using a boosted cascade of simple features", IEEE Trans. Computer Vision and Pattern Recognition, vol.1, pp. I-511-I-518, Dec. 2001.

[121]   M. Yang, D. Kriegman, and N. Ahuja, "Detecting faces in images: A survey" IEEE Trans. on Pattern Analysis & Machine Intelligence (PAMI), vol. 24, no. 1, pp. 34-58, Jan. 2002.

[122]   E. Haj Said, "Image Segmentation for Biometric Identification Systems", PhD Dissertation in WVU, 2007.

[123]   S. Park; M. Park; and M. Kang; "Super-resolution image reconstruction: a technical overview", IEEE Trans. on Signal Processing Magazine, vol. 20, no. 3, pp. 21–36, May 2003.

[124]   P. Vandewalle, S. Susstrunk and M. Vetterli, "A Frequency Domain Approach to Registration of Aliased Images with Application to Super-Resolution", EURASIP Journal on Applied Signal Processing, vol. 2006, article ID 71459 14 pages, 2006.

[125]   L. Lucchese and G. M. Cortelazzo, "A noise-robust frequency domain technique for estimating planar roto-translations", IEEE Trans. on Signal Processing, vol. 48, no. 6, pp. 1769-1786, June 2000.

[126]   C. Barber, D. Dobkin, and H. Huhdanpaa, "The Quickhull Algorithm for Convex Hulls," ACM Trans. on Mathematical Software, vol. 22, no. 4, p. 469-483, Dec. 1996.