2010

Fundamental Work Toward an Image Processing-Empowered Dental Intelligent Educational System

Grace Olsen
Virginia Commonwealth University

FUNDAMENTAL WORK TOWARD AN IMAGE PROCESSING-EMPOWERED DENTAL INTELLIGENT EDUCATIONAL SYSTEM

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Virginia Commonwealth University.

by

GRACE OLSEN

Director: SUSAN S. BRILLIANT
ASSOCIATE PROFESSOR, DEPARTMENT OF COMPUTER SCIENCE

Virginia Commonwealth University
Richmond, Virginia
April, 2010

Table of Contents

List of Tables

List of Figures

Abstract

FUNDAMENTAL WORK TOWARD AN IMAGE PROCESSING-EMPOWERED DENTAL INTELLIGENT EDUCATIONAL SYSTEM

By Grace Olsen, Ph.D. Student

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Virginia Commonwealth University.

Virginia Commonwealth University, 2010

Director: Susan S. Brilliant
Associate Professor, Department of Computer Science

Computer-aided education in dental schools is greatly needed in order to reduce the need for human instructors to provide guidance and feedback as students practice dental procedures. A portable computer-aided educational system with advanced digital image processing capabilities would be less expensive than current computer-aided dental educational systems and would also address some of their limitations. This dissertation outlines the development of novel components that would be part of such a system. This research includes the design of a novel image processing technique, the Directed Active Shape Model algorithm, which is used to locate the tooth and drilled preparation from a digital image, and also to measure the exact size, shape and location of the drilled preparation in relation to the expected preparation. The use of statistical measures taken from the digital images to provide feedback about the smoothness and depth of the dental preparation is also detailed. This research also includes the design and testing of a posture-monitoring component

for a portable educational system. Maintaining proper posture is critical for dental practitioners, because poor posture can affect not only the dental practitioner's health, but also the quality of the practitioner's work. The algorithms and techniques designed for use in the dental education support system could also be applied in the design of computer-aided educational systems for the development of procedural skills in many other fields, and in the design of systems to support practicing dentists.

Novelties and Contributions

A portable, computer-aided educational system with advanced digital image processing capabilities and an audio user interface would address the problems with current computer-aided dental educational systems. Clinical labs for the practice of dental procedures are dependent on the availability of instructors for feedback, and therefore students have only set lab hours to practice hands-on techniques. Computer-aided educational systems address this to some extent by allowing the students to see what they have drilled as a three-dimensional graphical model. Due to the limitations and expense of these systems, however, students at dental schools that use these systems still have only limited practice time. Another major drawback of existing dental intelligent tutoring systems (dental ITSs) is that they heavily rely on a graphical user interface. When students must look away from their work to consult a computer monitor to determine if they are doing a procedure correctly, they break their concentration. More importantly, none of the existing computer-aided educational systems provides students with real-time quantitative assessment of the work based on the processing of the images that can be captured during the process. An ideal dental ITS must be able to follow the progress of the dental student through the steps of a dental procedure, processing and evaluating in real time the quality of the student's work. This dissertation outlines a number of novel components that can be used in the design of such a system. This research has identified and addressed new and challenging problems in intelligent tutoring system design and image processing. Three major components of the system are described:

- A novel variation of the active shape modeling algorithm, called Directed Active Shape Modeling (DASM), is outlined in Chapter 3. This algorithm is used to identify within a color

digital image the drilled preparation on the surface of a tooth, and measure the change in size, shape and location of the preparation in relation to the tooth surface to provide quantitative feedback to the user. The DASM algorithm has advantages over the traditional ASM algorithm for this application, including:

  - Use of color intensity values instead of grayscale values when creating the gradient profiles for identifying best fit landmark locations within an image. This expedites the search for the preparation and the tooth surface within the image, and increases the chance of convergence to the actual shape of the preparation and the tooth surface.
  - Use of interactive machine learning, so that the algorithm interactively receives feedback from the user, improving the training speed and accuracy.

- The novel use of statistical and signal processing measurements of color digital images of the surface of the tooth and the drilled dental preparation to extract data concerning the smoothness and depth of the dental preparation are defined in Chapter 4. Machine learning techniques are used to classify the smoothness and depth of the preparation, using the data extracted from the images. The application of these techniques to identify caries on the surface of teeth increases accuracy and allows for the detection of caries that may go unnoticed with traditional detection techniques.

- The design and testing of a portable posture-monitoring system using on-body sensors and machine learning techniques is outlined in detail in Chapter 5. The use of on-body sensors to monitor posture in a setting for dental students and dental professionals is novel for this application, and may be used in other fields and for other applications.

The use of the proposed system in pre-patient-care dental training facilities will not only address the ongoing problem of decreasing numbers of faculty members in US dental schools, but also will give dental students more practical training with immediate feedback. The additional practical training can be expected to improve their technical skills when they enter a clinical setting and therefore improve the quality of their work, which will directly benefit their patients. The proposed system could also be used in a clinical setting, increasing the students' and patients' confidence level by providing a mechanism to prevent physical damage to biological structures. The use of the posture monitoring aspect of the system can reduce the risk of injury to the dental student.

The algorithms and techniques designed for this system could also be applied in the design of computer-aided educational systems for many medical fields in which the development of procedural skills is difficult, time-consuming and requires costly expert instructor time. They could also be applied in the design of systems to support dental practitioners in detecting and monitoring the progression of caries damage, and in maintaining healthful posture as they work.

CHAPTER 1 Introduction

1.1 Motivation

The purpose of this research is to address some of the major challenges that exist in the development of a dental intelligent tutoring system to assist dental students in developing the analytical and psychomotor skills needed to perform dental procedures. The need for computer-aided educational systems in dental education is a well documented problem [13], [55], [47], [54], [84].

Only a handful of dental schools currently make use of computer-aided systems. One of the most advanced systems currently in the market is DentSim. In school where DentSim is available, students begin their training using DentSim and then advance to a laboratory setting that provides no automated assistance. This state-of-the-art technology has been successful in increasing the amount of time students can practice and therefore in helping them to develop the basic technical skills needed for drilling teeth. However, surveys have shown that dental students have such a heavy classroom workload that they learn theoretical material through rote memorization, and state that they plan to learn the relevance of what they have memorized only when they get into clinical practice [70]. The educational system to which this research contributes would process digital images captured in real-time during a dental procedure and give auditory instructions, hints, and feedback to the dental students as they work. This would give dental students an earlier opportunity to integrate their classroom and theoretical knowledge with laboratory practice.

Many dental schools around the country and in Europe use automated virtual reality systems. One of the most advanced is DentSim, a system that allows dental students to observe a three-

dimensional image of their work as they practice cavity preparation. DentSim is a stationary system that has a patient mannequin with practice teeth, instrumentation for drilling preparations, and an infrared camera and diodes which measure the exact location in space of both the drill and the practice teeth and mouth. With this information, the system can calculate exactly where the student has drilled on the tooth. The system can show on its computer monitor a three-dimensional image of the tooth and drilled preparation, and give real-time graphical feedback to the students as they work. The system can compare the difference between the expected size, shape and location of the preparation the student is working on and the actual preparation the student is drilling, and show this difference to the student on the computer monitor with a graphical representation of both the actual and 'ideal' expected preparations. The work of the student can also be saved and graded, and this information can be used by the dental instructors to track the students' progress [13], [55], [47], [54], [84]. Although DentSim provides many advantages and benefits, there are some drawbacks to the system. For students to learn about the status of their preparation, they have to look away from their work to examine the graphical output on the computer monitor of DentSim. Although DentSim does give some auditory output when the student may drill outside the boundaries for a correct dental preparation, the auditory feedback to the students a digital beep, and does not descriptive information to the student. For the student to learn about the progress of their drilling, they have to look away from their work, which is disruptive and time-consuming.

The dental educational system proposed in this dissertation would not disrupt the work of the student as would a graphical user interface, because unlike the graphical user interface, it would not require the student to look away from the dental preparation. The proposed system would therefore act as a needed intermediary step between attention-disruptive instructional systems such as DentSim system and the clinical environment in which there will likely be no assistance at all.

The dental educational system outlined in this dissertation also contains a portable posture-monitoring system, which would address the need for additional posture monitoring and guidance in dental education. Chapter 2 outlines in detail the needs of dental schools for additional posture education and monitoring of the posture of dental students, as incorrect posture and positioning is a serious health issue for dental practitioners [97], [69].

1.2 Image Processing Enabled Dental Educational Systems

A practical design for a dental educational system would make use of a camera to record the student's progress and address one aspect of the image processing needed for such as system. To analyze digital images of the student's progress, a number of image processing techniques would be needed. To monitor the progress of dental students as they drill a preparation, the system would have to locate both the tooth and the preparation within the image, and quantitatively measure the size, shape and location of the preparation from frame to frame. To address this challenging problem, a modified version of an Active Shape Model (ASM) segmentation algorithm has been devised. This algorithm, referred to in this dissertation as a Directed Active Shape Model (DASM) algorithm, addresses some of the weaknesses in the ASM method with respect to the needs of this particular application.

The DASM is designed to be able to measure the difference between the dental preparation drilled by the student and the expected preparation, allowing the ability to give quantitative feedback to the student. The DASM also has a number of other advantages over the traditional ASM algorithm. First, it uses color pixel intensity values instead of grayscale intensity values to create a statistical model used to help locate the tooth and preparation in the color image. This additional information increases the accuracy of the DASM over the ASM.

The DASM also makes use of an interactive machine learning algorithm, used during the initial

creation of the training model for the algorithm. The ASM algorithm traditionally requires user input for the placement of every landmark used to help model the segmented shape, which can be tedious and time-consuming for the user and introduces the possibility of user error. The DASM requires less input from the user, partially automating the placement of the landmarks on training image to create the model. This decreases user time and increases accuracy through the reduction of user error. The outline of the DASM and the results of testing the new algorithm on multiple types of dental preparations is fully outlined in Chapter 3.

1.3 Image Processing of Color Images for Surface Texture in Teeth

Once the drilled preparation has been identified and the size and shape measured, another feature that can be extracted using image processing techniques is texture. The texture of the drilled preparation would be informative to dental students and instructors because the smoothness of the preparation is an important aspect of a correctly drilled preparation. Texture has also been used to help identify depth in single two-dimensional images, and this texture information could be used in this application to measure the depth of the preparation. To ascertain the ability of the novel image processing algorithm to measure texture on the surface of teeth from color images, it was tested by identifying carious regions within color images of the buccal surfaces of teeth. Research has shown that over 90% of all adults experience dental caries, and the early diagnosis of the carious lesion has become an important aspect of maintaining dental health. Advanced diagnostic and imaging devices can be used to identify tooth damage due to caries compensating for the low sensitivity (high false negative) rate of visual and visual-tactile inspection by dentists. However, existing systems have such a high false positive rate that dentists often do not rely on the results, instead relying on traditional visual or visual-tactile inspection. Of the existing computer-aided diagnostic systems, few use digital image analysis for detection and diagnosis. The feasibility of

using advanced image processing techniques and the use of multiple machine learning methods to identify caries accurately from digital images is demonstrated in Chapter 4.

1.4 Portable, Real-time Posture Monitoring Using Machine Learning Techniques

Along with work towards the development of image processing techniques for a computer-aided dental educational system, this paper outlines the development and initial testing of a novel system to measure and monitor posture in dental students, addressing another existing problem identified in dental education. Over 80% of dentists report having some type of back, neck or shoulder pain [97]. Research has identified significant costs linked to a very high rate of Work-Related Musculoskeletal Disorders (WMSDs) associated with poor ergonomic positioning in dentists. The annual costs of WMSDs across all occupations are estimated at between 13 and 54 billion dollars [30]. Little research has explored the design of portable, inexpensive, non-invasive and unobtrusive real-time systems to measure posture. The details of this system and the results of initial testing of the prototype are further explained in Chapter 5 of this dissertation.

CHAPTER 2 Previous Work

2.1 Computer-Aided Education in Dentistry

The most prominent system in computerized education is the Intelligent Tutoring System (ITS). An ITS can be considered an expert system that has a number of different models, each model representing either the knowledge base (or the "expert), the student, or the tutor [3]. The most basic ITSs simply compare the student model (based on the input from the student) to the expert model, and when the student deviates from the expert the system will provide information or hints based on what a real tutor would do, according to the tutor model. Recently the usefulness of the ITS has been called into question. In the 1980s and 90s most ITS systems focused on one model for teaching, that of an expert tutor correcting the work of a single student [3]. Others criticize ITSs for their expense and training time, both for instructors and students [48]. Other educational paradigms that include group work, students instructing each other, and student self-exploration and self-correction have not been widely explored [3]. Most traditional ITSs focus on correction, and only recently has work been done to explore aspects of education beyond simple correction of mistakes.

Another criticism of ITSs is that most do not address the important social aspects of tutoring; they do not deal with the social cues given by a real tutor nor do they pick up on the social cues given by the student. Research has shown that even people untrained in tutoring or teaching can be very effective tutors even if they don't follow any formal method. The ability to follow and use complex social cues would greatly increase the effectiveness of an ITS [23]. Some work has been done to address this problem; one interesting application is an ITS with auditory user interface

(both giving and receiving verbal input) that analyzes the speech of the student to measure for stress or frustration [31], [99]. Systems such as I-MINDS make use of multiple tutoring agents to facilitate group learning either within a classroom or for distance learning. The system uses both student and teacher agents to help the students form buddy groups and the student agent will even ask the teacher questions if the group is struggling [33].

Research has shown that ITSs are most useful for straightforward learning of procedural tasks (such as solving mathematical problems) [99]. Since the tasks the dental students would be performing are procedural in nature, an ITS would be a good fit. However, challenges arise in the fact that the tasks dental students learn in the laboratory are physical and three-dimensional, and do not involve a computer interface. Most of the existing research in ITSs has focused on systems with a relatively simple graphical user interface [31]. A computer-aided educational or tutoring system for dental education must be able to monitor the progress of a student through a procedure. Because there is no computer interface when students are practicing the skills of cavity preparation, there is no easy way to determine where the student is in the process of cavity preparation, nor is there a straightforward way to analyze the quality of the student's work. There are a number of existing systems that deal with learning physical tasks or learning in physical environments, such as DentSim for cavity preparation [55], "over the shoulder (OTS) instruction for anti-war craft simulation [33], and palpatory training with a virtual haptic back [43]. However, these systems have direct hardware connections from the student to the ITS. Most of these systems also involve complex and expensive hardware. Optoelectronic and electromagnetic systems have been used in dentistry (during dental surgery to monitor the positions of patients and drills) [13], [55], [47], [54], [84], and LED sensors and optical tracking have been effective in monitoring the location of the student's drill and the acrylic practice teeth in DentSim [55]. However the size and expense of these systems may be prohibitive for private practice or in smaller dental schools. The use of a digital camera

and image processing to monitor student progress is significantly less expensive and can be easily made portable.

2.2 Image Processing

Image processing plays an essential role in this research. To create the components of the computer-aided dental education system outlined in Chapter 1, there are a number of goals to be met with respect to the image processing of the digital images which serve as input to the system. These goals are:

- Locating and identifying the tooth surface on which the dental preparation is being drilled.
- Identifying and monitoring the location (in relation to the tooth surface), size and shape of the dental preparation being drilled, and classifying the preparation as being within or out of the bounds of an acceptable preparation in terms of the size, shape and location.
- Classifying the texture and depth of the dental preparation as being within or out of the bounds of an acceptable preparation in terms of the smoothness and depth.

To accomplish these image processing goals, a number of state-of-the-art image processing techniques are be used, and novel techniques to solve the problems outlined above are devised. The main image processing techniques used in accomplishing these goals are described below.

2.2.1 Registration

One of the main challenges of video processing is registration, which is the aligning of sequential images of the same scene. Image registration is used in a number of fields, including remote sensing (the gathering of information about an object through the use of a number of devices that are not in close proximity to the object), cartography, medical image analysis, computer vision

and video surveillance [35]. Some current techniques used in registration are cross-correlation, Fourier transforms, mutual information theory, clustering, Chamfer matching, invariant descriptors, relaxation techniques, and wavelet transforms [108], [64], [89].

2.2.2 Segmentation

An essential step in almost all image processing applications is the ability to find and isolate areas of interest within an image. Segmentation makes use of low level image processing techniques such as edge and line detection to determine boundaries between areas in an image. Segmentation allows for the identification and isolation of a feature within the image; it also may be used to determine the size and shape of a feature [35], [89], [72]. Color segmentation presents its own unique challenges. Although color images represent more information than gray-level, the algorithms for segmentation of color images are computationally much more complex than segmentation of gray-level images. For this reason, most research on color segmentation has been done within the past ten years as the computational ability of computers has increased [19]. Some of the major techniques used in color segmentation are histogram thresholding, splitting and merging techniques, region growing techniques [19], and supervised and unsupervised machine learning techniques such as clustering [61], neural networks [61], [22], and support vector machines [109]. There is no uniquely superior technique, as each application presents its own specific challenges. Segmentation of sequential images, as in video, is even more complex [53].

2.2.3 Active Contour Models (Snakes)

Active contour models (ACMs, or snakes) employ model-based methods that use a prior model to try to find the best match for the model within a test image [17]. In comparison to bottom-up image processing techniques, these use a top-down approach, which make use of the identification of local structures (edges, points and other low-level structures in the image) which are assembled into

groups to find objects [16]. The ACM algorithm creates, using training examples, a model of the shape that uses two opposing energy terms, an internal term which works towards smoothing the curve, and an external term which moves the curves towards image features, to locate the outline of an object. ACMs are good for amorphous objects like cells, but they tend not to perform well with objects that have a known shape. The weak constraints of the model tend to provide convergence towards local minima [16].

2.2.4 Active Shape Models

Traditionally, like snakes, active shape models (ASMs) are a top-down segmentation technique [16], used to locate a particular shape within an image. ASM has found popularity as a model-based method because it uses statistical methods to find the best match to the model, which has benefits over other types of model-based segmentation because it is flexible, and expert knowledge is held within the statistical model. Through the use of landmarks, ASM creates a model that is compact but still holds enough information about the shape so that it can be found in new images [16]. One of the main advantages of using ASM is that it is robust to noise. ASM works best with shapes where consistent and distinguishable landmarks can be defined; it does not work well with amorphous shapes. ASM is used often with complex image processing segmentation tasks (Cootes et al. uses images of faces and MRIs of the knee as examples to outline their original ASM model) because with these types of applications the shapes are particularly difficult to represent with primitives [16]. The exponential combination of the large number of primitives makes a bottom-up approach computationally challenging if not impossible [16].

The shape to be segmented from an image has a number of feature landmarks placed around the border of the shape. The landmarks are manually placed on the shape in a number of training images. The general model for the shape is created by training the system to learn the shape from

the training images based on statistical methods. With the training, the model learns the maximum deformation for each landmark that still fits within the model. This model can then be placed on a new image and deformed to try to fit to a shape in the new image. The growth or deformation of the model is repeated until convergence is reached [17], [16].

The ASM algorithm has a number of limitations. The landmarks that define the model must be identified manually in the training images, and the initial placement of the model within the test image is also done manually. The model itself is also limited by the number of training images. The training set must exhibit all of the variation that may be expected in test images, or else the model will be too narrow in scope and unable to correctly find shapes outside of the variation held within the statistical model [17], [16].

ASM differs from other types of model-based segmentation in a number of ways. A key difference is the definition of the fit function used to fit the model to the best fit shape within the test image [16]. Fitting the ASM model to the shape within the test image is based on the minimization of the cost function. For ASM, the cost function is defined as an error measure, which is the distance between each landmark and the best fit point along the strongest edge perpendicular to the landmark [16]. Other model-based methods, such as as Active Appearance Models, use different fit functions, looking at low level structures in the images instead of edges, or textures inside and/or outside of the shape [16]. There are some limitations to the use of this fitness function. Without prior knowledge of the general location of the target shape within the image, the landmarks identified by the ASM algorithm might be locally optimal landmarks for that area of the image, but not the globally optimal landmarks (where the shape is actually located in the image). With only locally optimal landmarks, the ASM would not find the correct location of the shape within the test image [16]. Local optimization can be used instead. There have also been extensions on the ASM algorithm to address issues such as warping of the image using thin plate splines [18].

Active shape modeling and its variations (including smart snakes, active contour modeling, and active appearance modeling) have been applied to medical image processing [39], [57], [103], [90], [12], [81], [6]. There has even been some research on the use of active shape modeling theory in dental applications [38], [100], [63], [2], [41], [95], [60]. This research has focused mainly on the development of dental prostheses [60] and forensic applications [42], but there has been some early research with x-ray images to segment caries lesions [39]. However, the focus has been on segmenting anatomical structures in medical images such as CTs, MRIs and x-rays, limiting image processing to gray-scale images [17], [16], [38], [100], [63], [2], [41], [95], [60], [42], [39]. Little research has been done on the use of active shape modeling to segment biological structures within color digital images [57].

2.2.5 Measuring Depth in 2D Images

One of the most challenging aspects of image processing is the analysis of three-dimensional aspects of a two-dimensional image. Three-dimensional characteristics of a dental preparation, such as depth and angle, can be estimated only when a two-dimensional image is properly analyzed. 3D modeling research has been done within such diverse fields as computer vision and robotics, remote sensing, virtual reality development, and archeology [90], [12], [81], [6], [85], [78]. Remote sensing research has created techniques to determine the angle and slope of digital elevation models [49], [101], [87]. What is more challenging, however, is the creation of 3D models from 2D images without any prior knowledge of the 3D geometry. There are two main approaches to 3D modeling from multiple 2D images. The first uses a series of calibrated images of the object to be modeled. This means that the location of the camera is known from one image to the next. The second approach uses uncalibrated images, but again requires multiple images taken from multiple angles of the object to be rendered [90], [87]. Recent work has been done to determine depth from multiple images (or video) without prior knowledge of camera location or images from multiple locations,

but it relies on identifying objects of known height within the images [87]. Another challenge is 3D modeling of rounded surfaces (such as teeth). However, work has been done on 3D modeling of smoother surfaces, such as skin and faces, with techniques such as non-rigid factorization [12]. Much research has focused on using specialized hardware to capture the 2D images for the extraction of 3D features; however using specialized equipment such as a plenoptic camera [68] is not practical for this application.

2.3 Medical and Dental Image Processing

2.3.1 Computer-Aided Caries Detection and Diagnosis

Ninety percent of all adults have carious lesions [105]. Research has shown that the rate of caries growth has changed due to advances in dentistry. Dentists are now able to diagnose and treat caries early, allowing the use of preventative, non-invasive measures [105], [44]. Research has also shown that visual inspection or visual-tactical inspection has very low sensitivity rate; that is, human inspection alone misses a high percentage of caries [105], [44], [46]. For these reasons a great deal of research has gone into the development of computer-aided caries detection and diagnosis systems. Most research has focused on developing systems that use advanced imaging tools, such as laser or light florescence [5], [4], [79], digital radiography [105], fiber-optics transillumination [105], and electrical resistance [51] to detect and measure tooth damage due to caries. Although these systems can visualize demineralization that cannot be seen visually, the low specificity rate (high rate of false positives) of all of these systems has led experts to agree that these tools cannot replace visual inspection for caries detection; instead they should be used to augment visual or visual-tactile inspection done by a dentist [105], [5], [4], [79], [51]. With florescence techniques, it is difficult to differentiate between natural demineralization and demineralization that is caused by a carious lesion. While dentists use these systems to augment their diagnosis, they are given the difficult

task of having to determine for themselves the accuracy of the detection and diagnosis feedback from these systems [105], [4], [82].

Most of the research done in computer-aided caries detection focuses on the development of advanced imaging techniques [105], [79]. Many of the systems provide only a different means of visualization, but some (such as DIAGNODent and DIFOTI) allow for data about the caries detection and tooth damage to be recorded digitally [79]. One reason DIAGNODent is popular in clinical settings is the ability of the system to rate the detected caries damage quantitatively. This allows the comparison of the numerical value of a diseased area on a tooth from one dental visit to the next. However, a numeric value representing the demineralization of the tooth area the device is placed on is the only feedback from the DIAGNODent system. Although it is popular, DIAGNODent's procedural use is complex and time-consuming [82]. DIFOTI, a system that makes use of fiber-optic transillumination, allows for the recording of digital images of the teeth being examined. As in many of the available techniques, DIFOTI allows for more in-depth visualization of teeth, but the dentists still have to interpret the images they see [79]. Currently, no computer-aided systems exist that can provide easily understandable quantified information about tooth damage due to caries. Existing systems mainly provide imaging information to dentists about carious lesions and suffer from lack of low specificity, and most have a high learning curve and complex procedures for their use, which limit their helpfulness [105], [5], [79], [82]. One of the goals of this research is to provide easily understood quantitative feedback about the presence and extent of the carious lesion, allowing dental professionals to be able to interpret and integrate data from the system in a quick and reliable manner.

2.3.2 General Image Processing in Dentistry

Image processing has been proven to be an effective tool in the field of dentistry. To date, the main applications of image processing in dentistry are either as navigational tools or for diagnostic purposes [26], [19], [7], [58], [102], [75], [11], [36], [59], [71], [24]. No research has been done on the use of image processing as input to an ITS or other computer-aided dental educational system. Image processing techniques such as edge detection [58], [35], modeling [58], [102], [11], [24], template matching [102], segmentation [58], and thresholding [24] have all proven to be effective and could also be used in a dental ITS to monitor student progress. One of the major goals for using image processing in a dental ITS is to determine what step the student is currently working on within the procedure. Determining whether the dental drill or other instruments are in the field (the mouth of the patient or the mouth of the practice head) is an important first step towards this goal. Edge detection has been used effectively in other dental applications of image processing [108], and can be used to determine the presence of dental instruments in an image of a dental procedure. The analysis of CT scans as a navigational tool for dental implantology (dental implant surgery) is a widely studied field, with many commercial software applications in existence [19], [7], [53], [17]. CT scans are used to create a 3D or 2D visualization for dental surgeons by using image processing, graphics and HCI (human-computer interface) techniques [108], [72], [59], [24]. Current research makes use of systems that are semi-automatic, that is, experts can manually extract the areas of interest for measurement [19]. All of this helps the expert figure out exactly where to place the implant or perform surgery on oral tissues and how to properly design the dental implant prosthesis. Recent work has shown the safety benefits of using computer-aided navigation. One teaching hospital analyzed the results of the use of such a system over nine years, and discovered that surgery was completed in less time and was safer, with no sleeping [submerged] implants [26]. There

are multiple commercial software applications for visualization for dental implants, including IGI, DentaScan, MedScannII, VISIT, VirtualImplant and SIM/Plant [26], [19], [7], [75], [36], [24], [14]. VISIT, for instance, uses CT scans and image analysis to help dental surgeons plan dental implant surgery, and can use post-op CT scans to determine success of placement of the implants [26].

The use of CT images to create a 3D model to guide dental implant surgery has been widespread in teaching hospitals within the past five to ten years, but this approach is not widely used in private practices [102]. Computer-aided navigation is an intensive process needing a navigation system, software and reference tools. Recently researchers have been trying to create systems to make image-guided implant surgery more widely accessible [19]. One recent paper has demonstrated the use of image processing to do registration of the mold of the mouth, with a reference frame to make a 3D model, so experts can plan the surgery. The system can also create templates with expert input. This research showed an increase in accuracy and a decrease in surgical preparation time [26], [102]. Systems such as CEREC [74] make use of infrared cameras to create a 3D model of a dental preparation, but require input from the dentist and have a high learning curve. Most of the current dental image processing systems make use of commercial software systems and require user input.

2.4 Interactive Machine Learning and User Interface Design

There are many challenges in image processing. Despite the large body of research in the application of image processing in medicine, there are few successful clinical applications in practice [56]. One of the reasons for the lack of successful medical image processing software systems is the challenge of creating accurate, repeatable, reliable, user-friendly, quick, generalized and robust systems. One of the main causes for the lack of generalization, speed and ease of use for image processing is the parameterization of image processing and machine learning techniques [56].

For image processing applications, machine learning is used to classify either low level aspects (such as pixel values) or high level aspects (such as objects or regions) of the image [35]. For example, an algorithm may be trying to classify pixels in a digital image as either inside of or outside of a region in the image. In classical machine algorithms, users would have to label training data (pixels in sample images, for example) that would be used to train the algorithm during a lengthy training session. The algorithm would then be tested against new testing data (again labeled by the user) to see if the algorithm could correctly identify which pixels were within the region in the image. If the algorithm was not accurate, the user would have to manually adjust learning parameters, and begin again. This example demonstrates why classical machine learning techniques are difficult to implement in clinical image processing applications. Interactive machine learning (IML) addresses these issues by changing the paradigm of classical machine learning [27]. Instead of having user input only in the beginning (labeling training data for the learning phase of the algorithm) and end (analysis of results, adjusting parameters to train the classifier again) of the machine learning process, interactive machine learning allows for iterative input from the user to train and adjust the algorithm quickly [27].

IML is a relatively new technique, and it has been used primarily in image processing applications. The "Crayons system, based off a decision tree classifier, allows users to create an image classification system interactively by "drawing on an image to help the system differentiate between objects or areas (classes) within the image [27]. There are no traditional training and testing phases to train the algorithm; if the system incorrectly classifies part of the image, the user can immediately draw more on the image to correct the system. Examples of IML include systems that have been used to identify roads in satellite images (using a version of support vector machines) [106], help non-programmers design computer vision programs (which has users test multiple machine learning algorithms via a quick iterative method to determine which is best for their particular

application) [67], design an email sorting classifier [91], design sensor applications [40], and create various systems for content-based image and video search engines [91], [29]. The idea of using a form of IML to help perform image processing in medical images goes back as far as 1999, but thus far evidence of the formal IML paradigm applied to the design of an image processing system for medical or dental clinical applications has not been found in the literature.

2.5 Posture Monitoring Using On-Body Sensors

Current research has shown that poor posture due to lack of training in correct ergonomic positions while working with patients is the greatest cause of work-related musculoskeletal disorders (WMSDs) and loss of productivity for dentists [94]. WMSDs include disk herniation, tension neck syndrome, muscle necrosis, and chronic back pain. Injuries can be costly, both in terms of medical expense and lost time, and in some cases WMSDs have been cited as the reason for early retirement or career change in the dental occupations [97]. Not only can WMSDs lead to pain and damage in the dentist, certain disorders such as carpal tunnel syndrome can lead to clumsiness, which could put patients at risk [86]. In general, posture has been shown to be associated with the quality of a dentist's work because improper posture can lead to stress, exhaustion, and clumsiness [86]. If the dentist's body is in an ergonomically correct position, the weight of the body falls naturally on the disks in the spine. However, if the body moves so that the natural curves of the spine are distorted, the weight of the body is moved to the muscles, tendons and tissue surrounding the spine. This is called a static posture, and staying in a prolonged static posture is a known cause of various WMSDs. One way to relieve the pressure on one muscle group is to change positions so that another muscle group can take over the task of supporting the weight of the body, giving the first group time to rest [98]. But studies have shown that the five main types of tasks performed by dentists (examination of the teeth, cleaning or polishing, scaling, drilling and filling) are very

similar in terms of movement of the body [28]. Because all of these tasks are performed in nearly identical body positions, dentists are unlikely to change position even when moving from one task to another and therefore are not resting these stressed muscle groups.

Over the past fifty years some changes have been made in the way dentists work. These changes were implemented in an attempt to help dentists, but some have in fact been harmful to their health. In the 1940s approximately 65% of dentists reported back pain. At that time most dentists worked standing up, and they did not work with assistants. Now almost all dentists work sitting down most of the time, and most practice what is called "four-handed dentistry where the dentist works with a dental assistant. These techniques were devised to help dentists to relieve back pain and to increase productivity. However, although four-handed dentistry does allow the dentist to be more productive, it has not been successful in relieving work-related pain, probably because it both increases working time and reduces the number of changes in the dentist's position [97]. Currently, over 80% of dentists report having some type of back, neck or shoulder pain [97]. The annual costs of WMSDs across all occupations have been estimated to be between $13 and $54 billion dollars [30]. The proposed system has the potential to be modified to help not only dental professionals, but practitioners of any occupation associated with a high risk for WMSDs.

Much research has been done in the field of dental ergonomics, but little new technology has emerged from this work. Three common approaches are used in an ergonomic assessment of safety issues: 1) subjective judgments such as those measured via surveys or questionnaires, 2) quantitative measures such as those resulting from use of an EMG (electromyography) to record muscle activity, and 3) subjective observation such as that of an ergonomic expert watching a recorded video of a dental clinician and noting changes in posture [28]. Most research into the ergonomics of dentistry has used subjective judgment and observation, because current methods for obtaining quantitative measurements are expensive and intrusive. The existing work involving qualitative

judgment has shown overwhelmingly that there is a connection between WMSDs and dental work. An early, yet compelling, study of dental hygienists in Ontario showed that 7% of hygienists reported via a questionnaire that they had been diagnosed with carpal tunnel syndrome (CTS), whereas the likelihood of contracting carpal tunnel syndrome in the general population is less than one percent [10]. Because the dental procedures performed by hygienists use the same posture as those done by dentists, this research has relevance to both dentists and dental hygienists.

One of the newest innovations in the study of the posture of dentists and dental hygienists is the Dental Operator Posture Assessment Instrument (PAI). The PAI is a scale that can be used by ergonomic experts or others to measure the posture of a dental professional over a set period of time (usually a five-minute time period) [10]. This study addressed an important question in ergonomic dental research. Although a proper position for dentists or hygienists is defined in textbooks and in dental school, the concept of a range of acceptable and unacceptable positions had not previously been investigated. Although research has been done for decades on the ergonomics of dentistry, most research has focused on identifying the causes of WMSDs. Research on correcting or preventing WMSDs has focused on the redesign of dental equipment and chairs [69], giving dentists instructions on how to stretch and strengthen their core muscles [98] and recommendations that the work of a dentist needs to be entirely restructured to include more breaks [97]. No work has focused on methods to correct the poor posture of dental professionals in real time as they work. Although correct and incorrect postures have been identified by researchers, dental students are not always trained effectively to recognize the difference. Furthermore, it can be difficult for a dentist to remember to stay in a neutral position during clinical practice even with proper training in recognizing appropriate positioning. In fact, research has shown that dentists are not likely even to take the time to adjust their own chairs or the height of the patients' chairs, two factors that are known to greatly affect the strain placed on the back [69].

There are two major approaches to studying the movement of the human body. One is the use of video recordings and motion capture. By identifying either visual markers placed on the body or the silhouette of a person, these systems analyze and classify different types of movement or placement of the body. The other main approach in human motion research is the use of on-body sensors. Since the goal of this research is to develop a portable and real-time system to measure posture, the focus is on the use of on-body sensors. These have the portability advantage over motion capture and video analysis because they do not require stationary systems with video cameras. They also require less computational time because the outline of the body does not have to be determined using computationally-intensive image processing techniques. However, these systems can become very complex, with increased computational time and invasive measuring systems; recent research into complex posture-measuring systems shows that attempts to increase performance speed had a significant impact on the level of error [76]. Most of the classification techniques in the current motion analysis research are complex, relying on a combination of machine learning theory, statistical analysis, and intensive preprocessing and feature extraction [104], [62], [76], [66], [107], [34], [65], [62], [45], [73], [25], [66]. Although some systems may have very high accuracy using these methods, in most cases they do not allow for real-time classification. Most research has focused on using complex filtering techniques, such as FFT (Fast Fourier Transformation) and DWT (Discrete Wavelet Transformation) to filter the data and to extract features [34], [62], [45], [25], [66]. Although these methods by themselves may not be computationally time consuming, a significant amount of data must be collected to apply them. Immediate classification of data cannot be performed on a single sensor reading if such techniques are used. Recent research has shown that new faster algorithms can be developed and are effective, but these algorithms have been shown to improve only the filtering of the data, and have not been shown to decrease classification time while maintaining accuracy [107], [34].

Current research has focused on studying movement for a number of applications within the medical field including gait analysis [66], [45], monitoring patients for tremors [25], seizure activity and falls, and studying ergonomics and posture [76], [104], [34]. Most of the systems are not designed to be used in real time, and they tend to be invasive or have interfaces that are designed for use only by either medical or ergonomic experts [76], [66], [107], [34], [65], [62], [45], [73]. Of the research that is being done for use by an end-user, very little has been done to design systems that are truly non-invasive and unobtrusive. Several systems are deemed to be non-invasive but they require the on-body sensors to be either taped to the skin of the user or attached with tight, inflexible straps [104], [62], [25]. At the time of this paper, only one study could be found with a truly unobtrusive system. In this study an accelerometer was attached to a cell phone that could be carried either in the breast pocket or side pocket of an individual [45]. The data collected by this sensor was sent via the cell phone to an off-site computer that would process the data and determine the gait of the user. Although this system has achieved the goal of creating a human motion system that is not obtrusive or invasive, it is not truly real-time, nor does it give immediate feedback to the user. This system also had a very different goal, that is, to distinguish different gaits, and had a wide range of success for different types of running and walking, with an identification accuracy rate ranging from 72% to 95% [45]. It has not been proven that such a non-invasive system can be successfully used to monitor posture.

There are a number of types of sensors that can be used to record the movement and position of the human body. The most commonly used sensors are accelerometers, gyroscopes, electro-goinometers, magnetometers, and smart sensing fabric [104]. Each of these has advantages and disadvantages. Accelerometers are the most commonly used, as they are relatively inexpensive and do not require a large power source. The accuracy of accelerometers depends on the quality of the sensor used, and although research has shown they can have a relatively high error rate, recent

advances in technology have resulted in the design of accelerometers that are highly sensitive to movement [104]. They can be used to measure acceleration and, if the acceleration is small in relation to gravity, they can also be used as inclinometers (used to measure the change in incline or angle) [104], [62]. Magnetometers can be used to measure the strength of the magnetic field in the vicinity of the instrument. Magnetometers can be disadvantageous as they may be affected by the Earth's magnetic field or metallic implants, and they may not be portable because some require an external receiver as well as an on-body sensor [104], [107], [34]. Gyroscopes are able to measure angular velocity due to their rotating reference frame. The main disadvantage with gyroscopes is that they have a tendency to drift and have to be re-calibrated [66]. Electro-goinometers, also called flexible angular sensors or strain gauges, have been used to determine the relative positions of sensors, but it has been shown that the movement of the user may compromise the sensor itself [104], [25]. Most current applications use a combination of sensors in a single system, so that the disadvantages of one type of sensor can be overcome by the use of another type of sensor.

Most multi-sensor systems combine accelerometers with either gyroscopes, magnetometers, or both [104], [76], [66], [107]. Electro-goinometers can be used to determine the position of different sensors in relation to each other [76]. However, these systems can become very complex, with increased computational time and obtrusive measuring systems. In a recent study, trunk posture was measured using a hybrid system with two sensor groups, each with accelerometers, magnetometers and gyroscopes [76]. One sensor grouping was placed on the center of the back of the research subject (over the T1 vertebra), and the other sensor grouping was attached to the lower back (over the sacrum); both were attached with tight straps. The two groupings were connected with a flexible rod potentiometer to track the relative positioning of the sensors. Although this system is highly accurate in correctly determining the position of the subject, the system is bulky, cannot be worn underneath clothing, and could not be used outside of a laboratory testing environment [76].

In summary, current research has shown that on-body sensors can be used to create highly accurate systems to monitor human movement or position. However, these systems are usually designed for use by medical or ergonomic experts and do not provide real-time results. Also, no unobtrusive and non-invasive system has been devised that can analyze posture accurately [104], [76], [25]. A commercial system based on this technology could be used in dental schools as an affordable means to teach and reinforce correct body positioning. It would also be an affordable way for practicing dentists and dental hygienists to monitor and correct their posture as they work with patients. The resulting improvement in body positioning would have many potential benefits in terms of student/practitioner health, productivity, reduction of error and risk to patients, and savings in the direct and indirect costs associated with WMSDs.

CHAPTER 3 Directed Active Shape Modeling

3.1 Introduction: The Original ASM Algorithm

As discussed in the previous chapter, the Active Shape Modeling (ASM) algorithm is an effective segmentation technique, and it has been used successfully in various medical applications [18], [17], [16]. Due in part to the benefits outlined in Chapter 2, the ASM algorithm is an appropriate method for the segmentation needs of the image-based dental educational system discussed in Chapter 1. However, there are significant limitations to the ASM algorithm. These limitations are addressed with the development of the Directed Active Shape Modeling (DASM) algorithm, a novel extension of the ASM created for the image-based dental educational system. To provide a foundation for explaining the differences between the DASM and the traditional ASM algorithm, the steps that comprise the ASM algorithm are outlined below.

The ASM algorithm consists of two main stages. First, in the training phase a statistical model representing the shape(s) to be segmented is created from a number of training images. The second stage, the testing phase, makes use of the statistical model to find that shape within a test image, using a cost minimization function to locate the best fit of the model in the new image, and Gaussian pyramid filtering to shorten the computational time of the algorithm.

3.1.1 Creating the Statistical Model

The first important aspect of creating a comprehensive statistical model of the shape is finding good landmarks to represent the shape. Good landmarks usually represent corners, edges, and junctions between lines that define the shape(s) to be modeled. However, research has shown that often these landmarks alone create a sparse model, and more than just these landmarks are needed to create

a model that can accurately find the shape in a new image [16], [18]. Cootes et al. [16] found that if landmarks are added along the edges between these defining points accuracy improves. The number of landmark points can vary and is user-defined depending on the application. The order of the landmarks that define the shapes is an important aspect of the model. One of the drawbacks to ASM is the fact that landmark identification in the training images is a manual task and can be time-consuming, and usually requires expert input. However, once these points are recorded the remainder of the ASM algorithm is automated (with the exception of the initial placement of the model within the test image) [16]. The outline for the training of the ASM algorithm is shown in Figure 3.1.1.1.

For the landmarks for each training shape, a vector is created (Equation 3.1.1.1) [16]:

$$X = (x_1, ..., x_n, y_1, ..., y_n)^T \tag{1}$$

where x_i and y_i represent the x and y coordinates for landmark i in the image, for a total of n landmark points within T images. The next important step after the landmarks have been defined is to align all of the training images into the same coordinate frame before creating the statistical model. This is done by rotating, scaling and translating each training shape so that the distance between each shape and the mean shape is minimized [16]. This is done through the following iterative approach:

1. Each training model is translated so that the center of gravity is at the origin.

2. Take a random training shape example and make it the initial estimate of the mean shape and scale, so that for that model's x and y values (stored in vector X of length i), as seen in Equation 3.1.1.2:

$$|\overline{x}| = \sqrt{\sum_i \overline{x}_i^2 + \overline{y}_i^2} = 1 \tag{2}$$

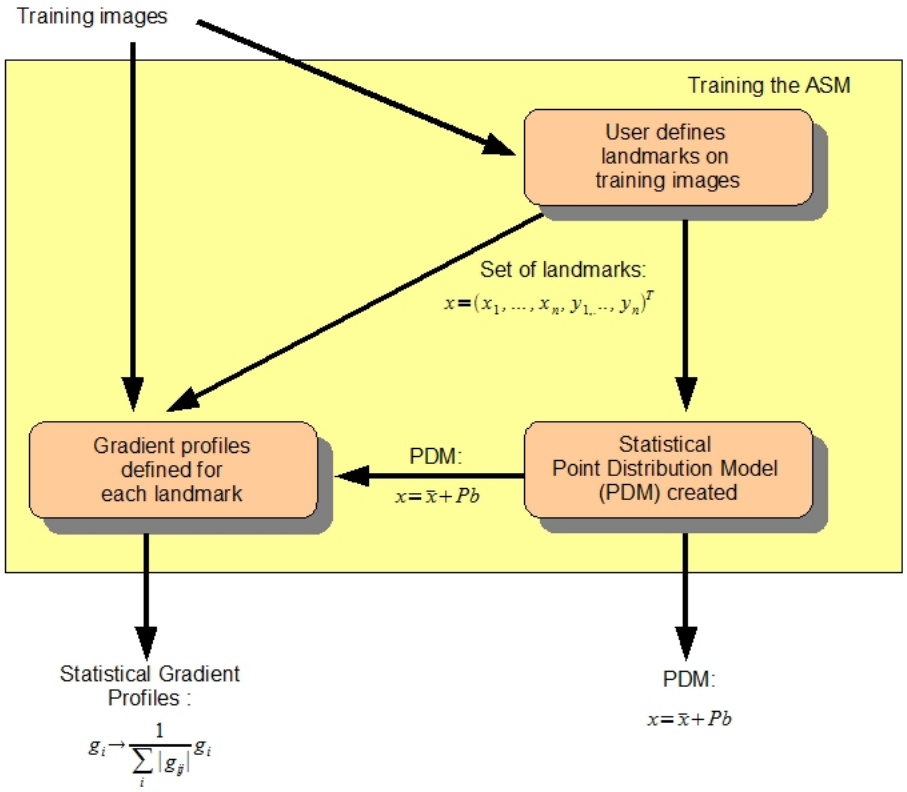


Figure 1: Outline of the algorithm for creating the statistical model.

Make this x_0 and define it as the orientation for the other examples to be oriented against.

3. Align all of the training shapes with this normalized and oriented current estimation of the mean shape. Aligning two shapes is done by choosing a scale (s) and rotation value (Θ) and then transforming the shape ($M_{s\Theta}(x_1)$) so that the sum of squared differences between the

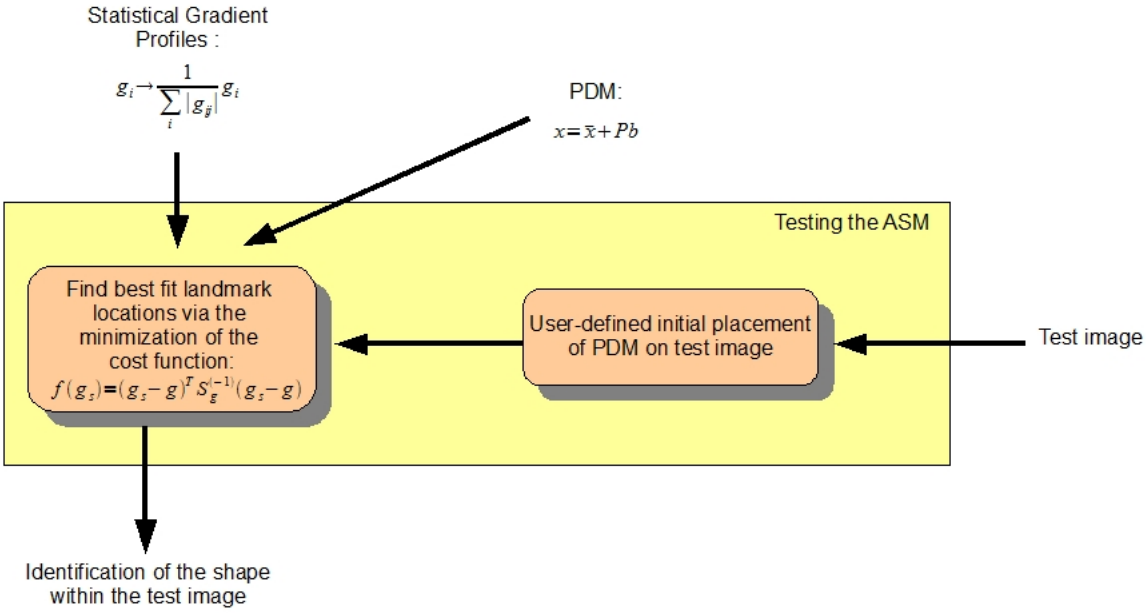


Figure 2: Outline of the algorithm for testing the statistical model.

two shape models is minimized, see Equation 3.1.1.3:

$$|M_{s\Theta}(x_1) - x_2|^2 \tag{3}$$

4. Re-estimate the mean from the aligned example models.

5. Constrain the scale and orientation to this current mean by aligning it with x_0 and $|\overline{x}| = 1$ (scale).

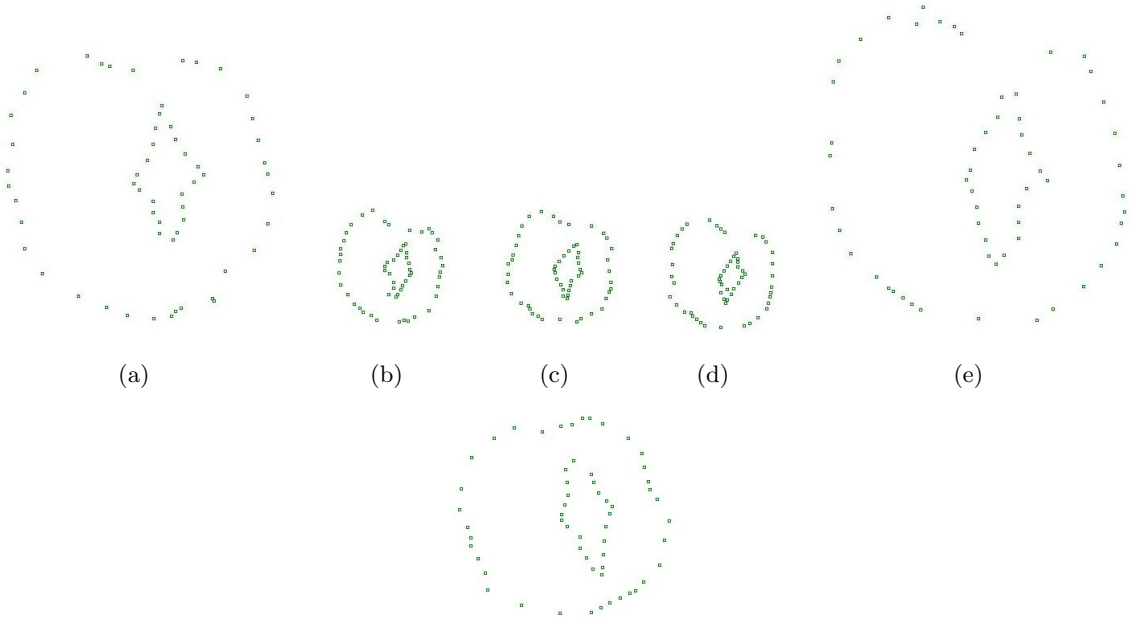


Figure 3: Example of six training landmark sets before alignment.

6. Repeat steps 4 and 5 until convergence. Convergence occurs when the mean does not change more than a set threshold from one iteration to the next.

After the training shape models (see Figure 3.1.1.3, note the differing scale, location, and rotation of the models) have been aligned, Principal Component Analysis is applied to the model. The purpose of applying PCA is to reduce data held in the model to create the most compact model that still holds enough information about the shape to accurately locate the shape in new images. PCA is a statistical technique that projects a highly dimensional set of data into a new, lower dimensional space [103]. PCA finds the eigenvectors and eigenvalues from the covariance of the data set; covariance is defined for the data set as seen in Equation 3.1.1.4.

$$S = \frac{1}{m-1}\sum_{m}^{i=1}(x_i - x)(x_i - x)^T \tag{4}$$

Each eigenvalue (λ_i) gives the variance of the data around the mean in the direction of each

eigenvector. The first t eigenvectors (p_i) are chosen to represent the model, with the percentage of variation of the model defined by the number of eigenvectors selected to represent the model [16]. The total variance for all the eigenvectors λ_i is computed from Equation 3.1.1.5:

$$V_T = \sum \lambda_i \tag{5}$$

where T is the number of training images [16].

With PCA applied to the data set, any model from the training set (x) can be approximated by Equation 3.1.1.6:

$$x \approx \overline{x} + Pb \tag{6}$$

where $\overline{x}$ is defined as the mean model of points in Equation 3.1.1.7

$$\overline{x} = \frac{1}{T}\sum_{i=1}^{T} x_i \tag{7}$$

P is the matrix of the first t eigenvectors ($P = (p_i, p_2, ..., p_t)$) , and b is a vector of parameters with T dimensions, see Equation 3.1.1.8:

$$b = P^T(x - \overline{x}) \tag{8}$$

P defines the coordinate frame in the original space, and the parameters, b, are the significant coordinates of the shapes within this frame. The model can be varied by changing the elements in b. The allowable variation of the parameters b is usually limited to the first three standard deviations from the mean, so that the shape maintains similarities to the training examples (see Figure 3.1.1.4 and Figure 3.1.1.5 [16] and note that in Figure 3.1.1.5 the mean shape model is shown in green, with one standard deviations of b shown in red, which is the allowable deformation of the model based off of the training set of images).

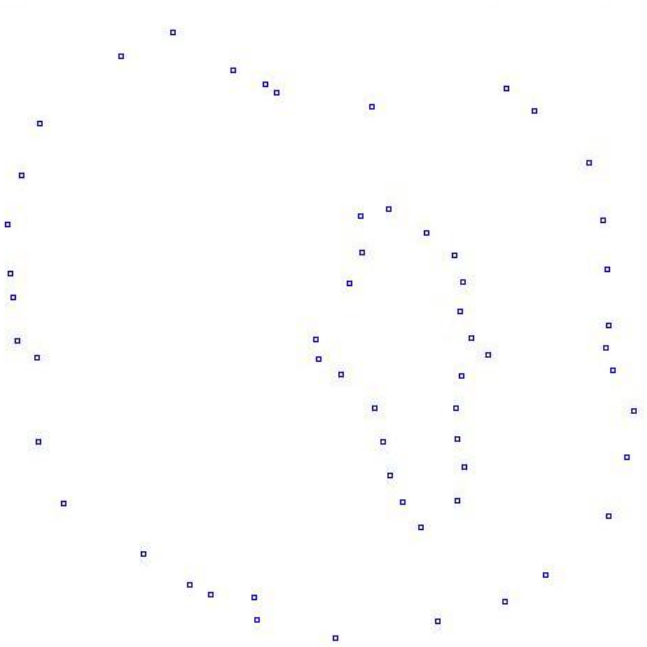


Figure 4: An example of a mean shape model.

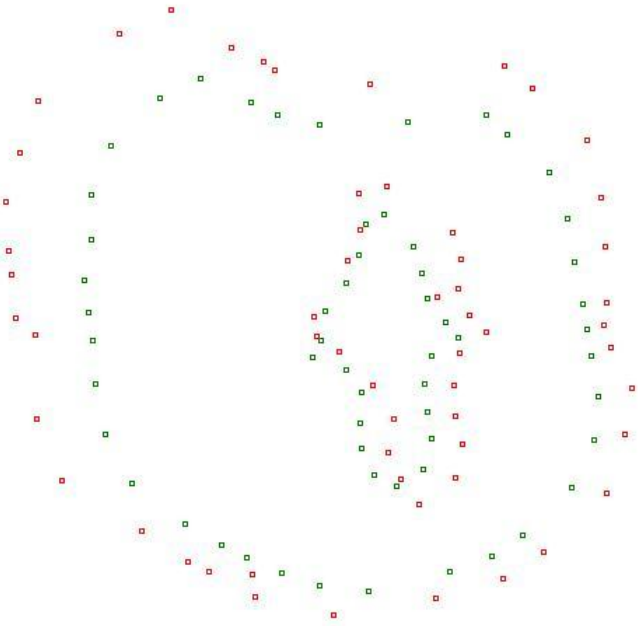


Figure 5: An example of a mean shape model with deviations from the mean.

3.1.2 Fitting the Model

After the model is created, it can be used to find the target shape in a new image in the testing phase of the ASM algorithm. The general outline for finding the shape is shown in Figure 3.1.2.2. Fitting the model is done through the optimization of the cost function [16]. The best fit model (which is found as described above) is compared to the potential shape in the image using a cost function, $F(c)$. $F(c)$ can be thought of as a measurement of error, so that if the model and shape are a perfect match $F(c)$ would be 0 [16]. The parameters of the function are selected so that $F(c)$ is as close to 0 as possible. *Cootes et al.* [16] define the optimal cost function as $P(c|I)$, where I represents the image, $P(c|I)$ is the probability that the model parameters describe the object within the image, and c is selected to maximize this probability. The parameters are the shape parameters (b) and the pose parameters (X_T, Y_T, s and Θ). The fit function can be defined in many ways. The fitness function for the ASM algorithm is defined as the distance between each landmark point and the closest strongly defined edge within the test image [16]. This distance can be defined as a measurement of error in Equation 3.1.2.9:

$$F(b, X_T, Y_T, s, \Theta) = |X' - X|^2 \tag{9}$$

To minimize the fitness function, the following iterative algorithm is used in Equation 3.1.2.3.1.2:

1. Search the neighborhood of each landmark point within X and find the best match for that point.

2. Update the pose parameters that best fit these new points identified for each landmark point, using the fitness function (error measurement function) outlined below.

3. Apply the constraints to each shape parameter, b_i, so that it is within three standard devia-

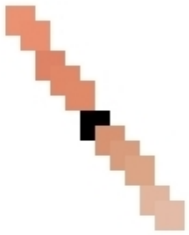


Figure 6: An example of the color pixels normal to the landmark (the landmark is in black).

tions of the mean (Equation 3.1.2.10):

$$|b_i| < 3\sqrt{\lambda_i} \tag{10}$$

4. Repeat until convergence, i.e. the change from one iteration to the next is below a predefined threshold.

To find the initial best match for each landmark, the pixels normal to each landmark are examined, and compared to statistical models for each landmark created from the training examples [16]. These statistical models are defined with Equation 3.1.2.11:

$$g_i \rightarrow \frac{1}{\sum_i |g_{ij}|} g_i \tag{11}$$

Each landmark has a profile, which is defined as vector g_i, that has the length $2K + 1$, as seen in Figure 3.1.2.6. The derivative is taken from each vector, and normalized by the division of the vector by the sum of the absolute values of the pixel intensity values. The mean and covariance for the set (across all training images) of profiles for each landmark is calculated, and this defines the statistical model for the landmark [16].

This fitness function is defined as the following Equation 3.1.2.12:

$$f(g_s) = (g_s - \overline{g})^T S_g^{-1} (g_s - \overline{g}) \tag{12}$$

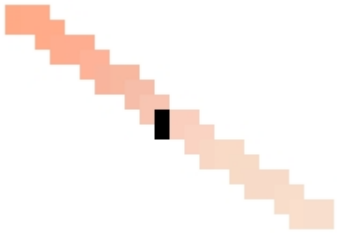


Figure 7: An example of the search vector of color pixels normal to the current landmark in the test image.

The fitness function uses mean ($\overline{g}$) and covariance (S_g) as defined above, and calculates the Mahalanobis distance between model landmark profiles and profiles of the best fit points for each landmark in the test image (see Figure 3.1.2.7). Again, to find the best fit shape for the model within the test image, this cost function $f(g_s)$ must be minimized. The training profile is compared along the profile for the test landmark, and the location with the lowest cost function is the best fit, see Figure 3.1.2.8 [16]. The vector of is searched to find the best fit for the statistical model for that landmark by minimizing the fitness function as the difference in the values in the statistical model are compared to the values of the potential matches in the search vector.

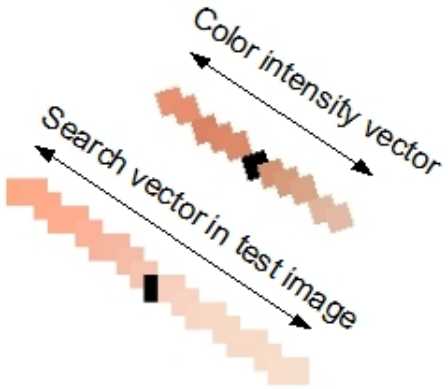


Figure 8: An example of vector of color pixels normal to the current landmark in the test image.

Once the cost function has been minimized, the statistical model created in the first part of the algorithm can be adjusted to fit to the new shape using Equation 3.1.2.13:

$$X = M_{X_c,Y_c,s,\Theta}(\overline{x} + Pb) \tag{13}$$

The model is adjusted by applying the transformations (M) to the model (similar to the alignment of the training shapes described earlier) by translating (X_c, Y_c), rotating (Θ) and scaling (s) each point in the model, transforming the size, rotation and location of the shape, seen in Equation 3.1.2.14 [16]:

$$M_{X_c,Y_c,s,\Theta}\begin{pmatrix} x \\ y \end{pmatrix} = (X_c Y_c) + \begin{pmatrix} s\cos\Theta & -s\sin\Theta \\ s\sin\Theta & s\cos\Theta \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} \tag{14}$$

When a new shape is presented to the model, the best variation of the model to fit to the new shape is determined through the minimization of Equation 3.1.2.15:

$$|Y - M_{X_c,Y_c,s,\Theta}(\overline{x} + Pb)|^2 \tag{15}$$

which represents the sum of the squared distance between the model (x) and new shape (Y) points. To minimize this equation, the following iterative algorithm is used:

1. Initialize the shape parameters to 0.

2. Create the model with Equation 3.1.2.16:

$$x = \overline{x} + Pb \tag{16}$$

3. Find the transformation (M) parameters that align the model (x) points to the new shape (Y) points.

4. Project the new shape's set of points (Y) into the model frame by inverting M, via Equation 3.1.2.17:

$$y = M^{-1}_{X_c,Y_c,s,\Theta}(Y) \tag{17}$$

5. Project y into the tangent plane of the mean model shape ($\overline{x}$) by scaling y in Equation 3.1.2.18:

$$y' = \frac{1}{y.\overline{x}} \tag{18}$$

6. Update the parameters to match to y' in Equation 3.1.2.19:

$$b = P^T(y' - \overline{x}) \tag{19}$$

7. Repeat until convergence, which occurs when there is no change in pose or parameters from one iteration to the next.

To decrease the computational time of the ASM algorithm, a multi-resolution version of the ASM algorithm can be used. This makes use of Gaussian image pyramids. The bottom level of the

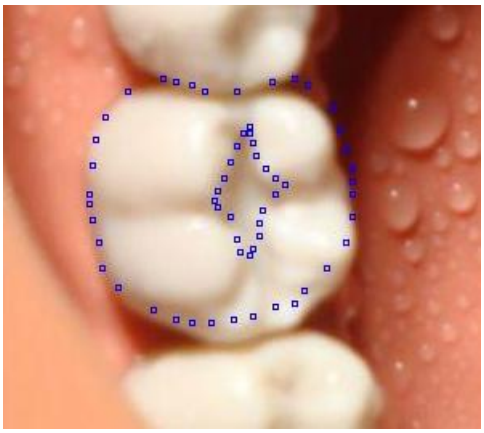


Figure 9: The initial placement of the model on the lowest pyramid level image.

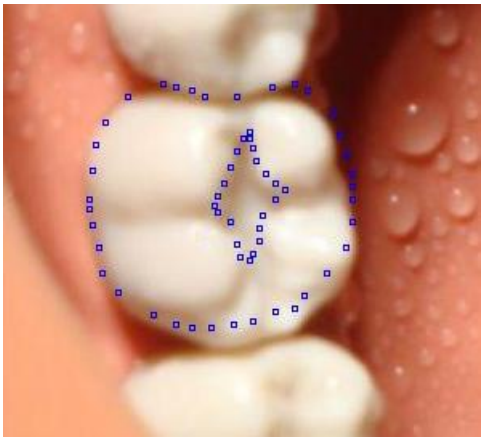


Figure 10: The model being fitted to the image.

pyramid is the original test image. Each subsequent (higher) level is a smoothed and subsampled version of the image, with the number of pixels halved from the previous level. The ASM algorithm is first carried out on the highest level of the pyramid (the most minimal version of the test image), which allows for a wide search area for the best fit for each landmark.

To determine when convergence for that level of the pyramid has been reached, the location of the best fit landmark along the search vector is examined. If the best fit landmark is close to the center of the vector (near the previous best fit landmark, which is located halfway along the search

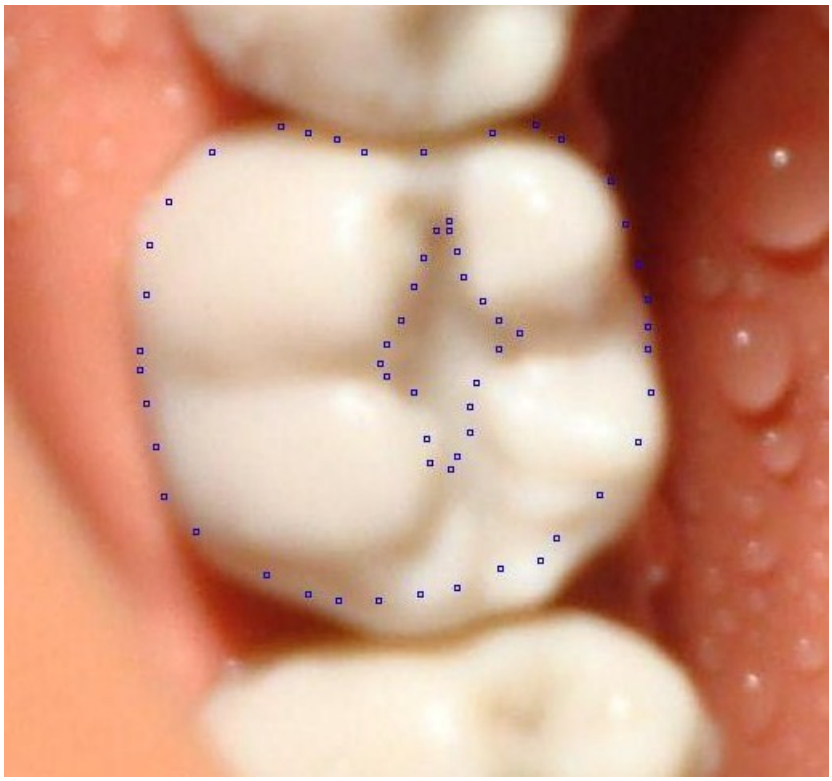


Figure 11: The next pyramid level, and further fitting of the model to the shape in the image.

vector), then it has not moved significantly from the previous location of the best fit landmark, and it can be assumed that the landmark has converged on the best fit. However, if the best fit landmark is near either extreme of the search vector (far away from the previous best fit landmark location) then the best location for that landmark has not been found, and convergence is not reached. The number of best fit landmarks that are close to the center of their search vectors is calculated. To define the length of the search vector that is considered close to the center, the vector is divided into four equal lengths. The two lengths that are in the center of the vector are considered the 'close' range, and if the new best fit landmark falls within that length, it is counted towards convergence, see Figure 3.1.2.13. If this number is above a user-defined threshold, convergence has been reached for that level, and the algorithm moves on to the next lower (finer) level of the pyramid [16].

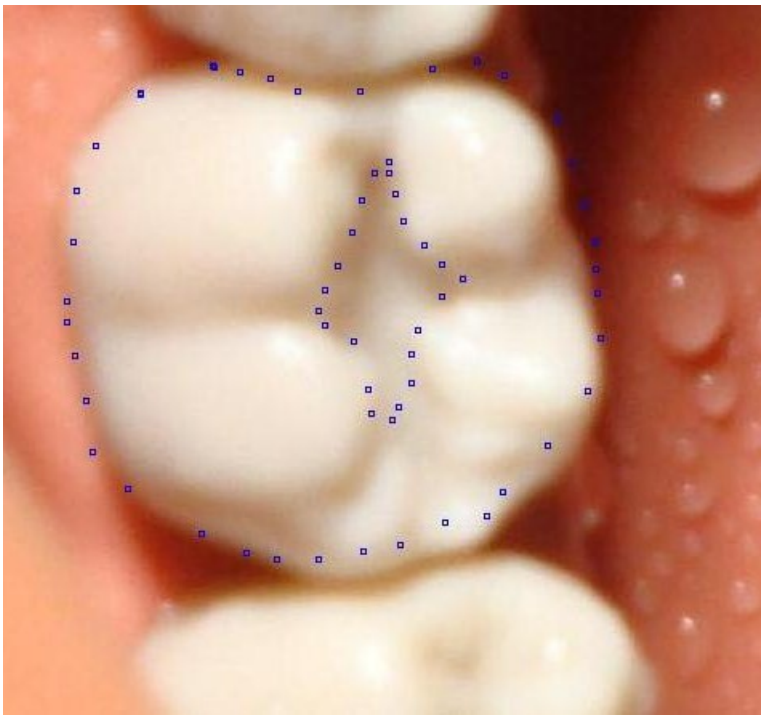


Figure 12: Further fitting of the model to the image.

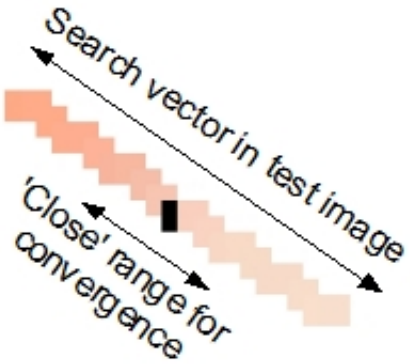


Figure 13: Comparison of the statistical model to the search vector.

The ASM algorithm has a total of seven parameters that can be adjusted. For model creation, the number of landmarks, the variation of the statistical model (the number of modes), and the number of pixels used to create the statistical profiles for each landmark are all determined by the creator of the particular model. For the fitting of the model to the target shape, the parameters include the number of pixels searched for the best fit point for each landmark, and for the use of the Gaussian image pyramid, the number of levels in the pyramid, the threshold for converging at each level and a maximum number of iterations at each level. It is important to note that the location of landmarks for each training model is also user-defined [16].

3.2 Methods: Directed Active Shape Model

To effectively apply the ASM segmentation algorithm to locate and monitor drilled dental preparations as outlined in the proposed image-based dental educational system, a number of significant modifications need to be made to the original ASM algorithm. The resulting algorithm is the novel Directed Active Shape Modeling (DASM) algorithm. The three significant modifications are listed below:

1. An interactive machine learning algorithm is used (along with the use of edge detection image processing techniques) to assist the user in more easily identifying the landmarks in the training image, reducing time and increasing the accuracy and consistency of the design of the statistical model.

2. Color intensity information is used instead of gray scale intensity values for the normalized gradient profiles for the landmarks in the model. This modification triples the amount of information held within the statistical model because values for all three colors (RGB) are used for the files in place of a single grayscale intensity value.

3. The testing portion of the algorithm is modified to allow for full deformation of the drilled preparation shape model, even if it does not conform to the shapes seen in the training images. This modified testing algorithm is run concurrently with the traditional version, allowing for the expected drilled preparation shape to be placed on the tooth surface as well. This modification will allow for full distortion of the model (outside of the allowed variation in the model) and therefore allow for quantitative measurement of the difference between the expected and actual preparation shape.

These novel modifications are also shown in the flowcharts Figures 3.2.14 and 3.2.15; the changes from the original ASM algorithm are highlighted in red.

The DASM algorithm is used within the image-based dental educational system to located to related areas in the image. First, the DASM algorithm will need to be able to correctly identify the exact location of the tooth the student is working on. After locating the tooth within the image, the DASM algorithm will then locate and measure the size and shape of the dental preparation the student is drilling. Two shapes (the tooth surface and the drilled preparation) are segmented the from digital images of the dental field. Once both shapes are located, the quantitative information about the segmented shapes is then used to measure the progress of the shape of the preparation and monitor the appropriateness of the location of the preparation on the tooth. It should also be noted that for the intended application of the DASM, the testing phase (locating the shapes within in the test image) does not need to be completely automated. The educational system will be designed to be interactive with the student or practitioner. The start of the testing phase of the DASM algorithm requires input from the user as to the initial placement of the model. The practitioner or student can point with the drill or another instrument to the appropriate tooth, and the system can use this visual input to determine the starting location for segmentation.

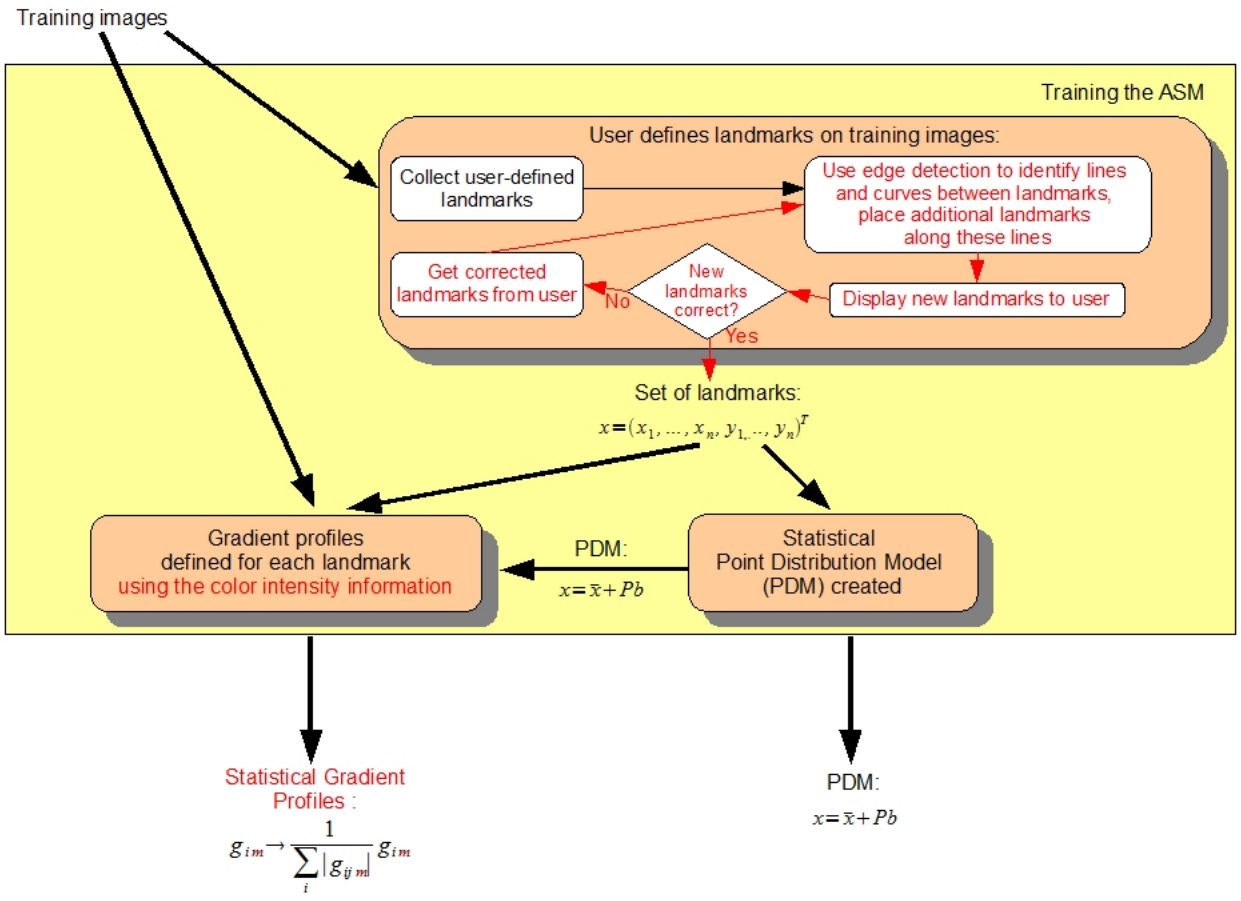


Figure 14: Outline of the algorithm for fitting the statistical model in a new image for the DASM algorithm.

3.2.1 Use of IML for training the DASM

The first major modification of the ASM algorithm is the addition of Interactive Machine Learning (IML) to the first step in creating the statistical shape model. In the traditional ASM algorithm, landmarks must be manually placed on each training image to properly identify the shape(s) that will define the ASM model. Traditionally landmarks are placed in clearly identifiable locations along

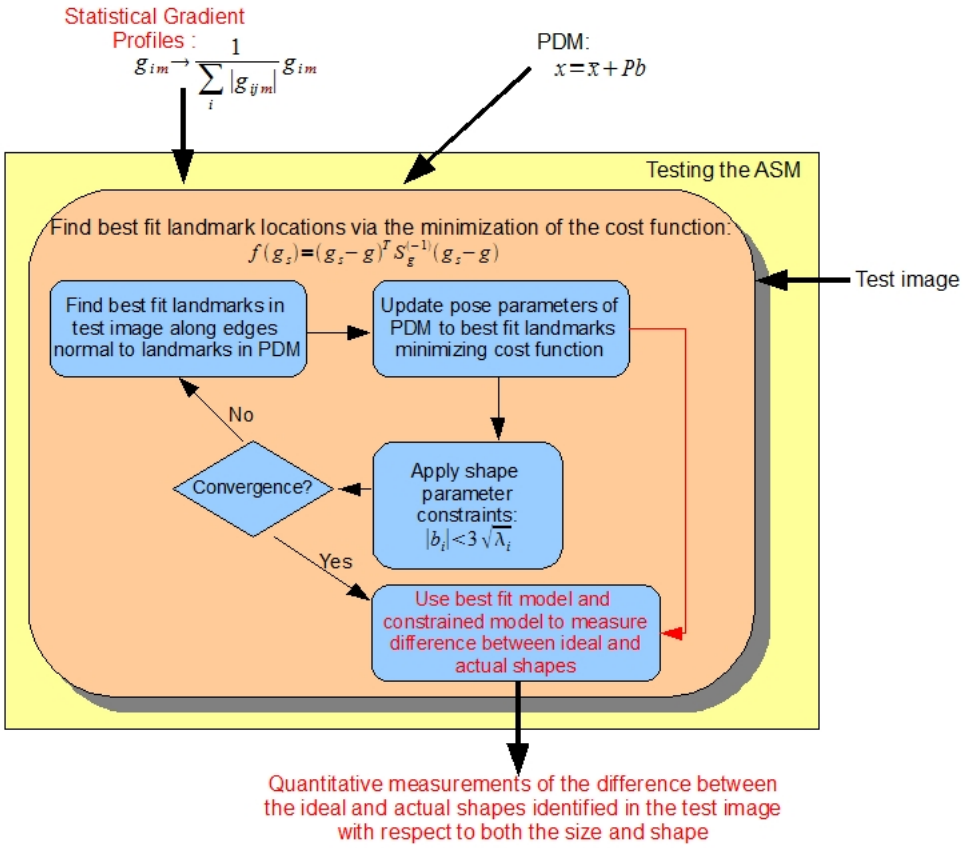


Figure 15: Outline of the algorithm for fitting the statistical model in a new image for the DASM algorithm.

the outline of the shape, such as corners and intersecting lines. However, Cootes et al. [16], [18] found that to improve the accuracy of the statistical model, additional landmarks should also be placed along the edges of the shape in between the easily defined landmarks. The placement of a large quantity of landmarks is time-consuming and introduces the possibility of user errors, especially in the placement of additional landmarks between the easily identified landmarks. It is particularly difficult to place landmarks for organic, amorphous shapes (such as the teeth and

Figure 16: Initial four user-placed landmarks on a training image.

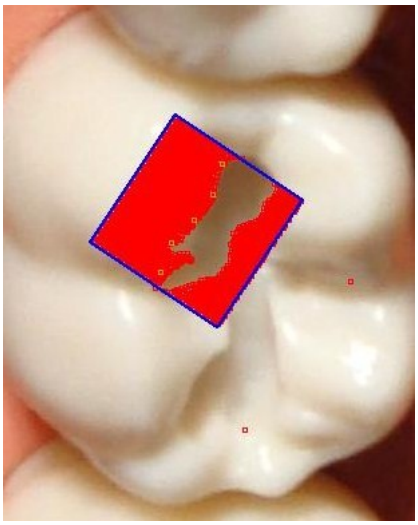


Figure 17: Edge detection during the IML placement of landmarks.

preparation areas for this application).

To address these problems with the ASM algorithm, an Interactive Machine Learning (IML) approach is used. As explained in detail in Chapter 2, IML represents a paradigm shift from classical machine learning by allowing for quick iterative input from the user to train and adjust a

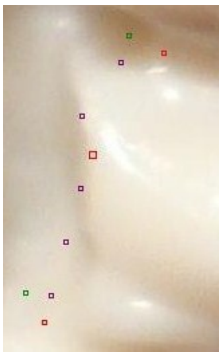


Figure 18: Refinement of landmarks during IML landmark placement.

machine learning algorithm quickly [27]. In the first step of the DASM algorithm, instead of having the user manually place all of the landmarks, the user initially places only a handful of landmarks in pertinent locations. In this application, the user initially places size landmarks for the tooth shape, marking the four locations where the edge of tooth surface meets the two adjacent teeth, as well as two additional landmarks marking the center of the two exposed sides of the tooth surface, as seen in Figure 3.2.16 and Figure 3.2.17. In this example, just the drilled preparation is being segmented. Four initial user-placed landmarks were placed at the extreme top, left, bottom and right of the tooth surface (seen here in red). A search area is defined between each pair of initial landmarks (seen outlined in blue in which to run the edge detection algorithm. The edges can be seen highlighted by the red within the blue box. If more than one edge was detected (as seen in this example), then the one closest to the middle of the search area is used, as it is closer to the initial points placed by the user. Then, following the IML paradigm, a simple edge detection algorithm (a Homogeneity edge detection algorithm [35]) is applied to the regions of the image between each of the initial user-placed landmarks to quickly identify the rest of the tooth edge.

If the edge detection algorithm has misidentified the edge of the tooth or the drilled preparation, the user can simply click on the image along the correct edge of the shape. The edge detection

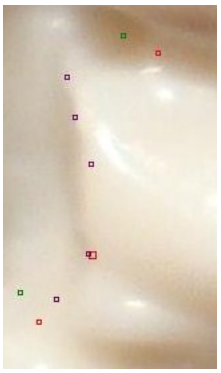


Figure 19: Further refinement of landmarks during IML landmark placement.

algorthim is modified by narrowing the search area within the image, increasing the likelihood that the detection algorithm will correctly identify the edge of the shape. This process can be repeated as many times as necessary until all of the edges of the shape(s) are correctly identified. This is demonstrated in Figures 3.2.18 and Figure 3.2.18. The system has taken in the placement of the additional user-placed landmark seen in Figure 3.2.18 and has refined the landmarks for that edge (seen in purple). The small red landmarks are the initial user-placed landmarks. The green landmarks are the initial DASM-placed additional landmarks. The purple landmarks are those that are being corrected by the user. The large red landmark was placed by the user to inform the system of the correct edge. One additional user-placed landmark on that edge has resulted in the correct placement of the landmarks along the edge of the preparation.

Once the outline of the shape(s) is correctly identified by the edge detection algorithm, a large number of additional landmarks can be placed automatically by the DASM algorithm between the initial user-placed landmarks. This process has two major benefits. First, it greatly decreases the input needed from the user by requiring only the initial landmarks and then a relatively small number of additional points to be identified by the user. The second benefit is the even spacing of the landmarks along the edge of the shape in each training example. This cannot be guaranteed

with user-placement of additional landmarks, when there is nothing along the edge of the shape in the image that can assist the user to make sure he or she is maintaining equal spacing between the additional landmarks.

3.2.2 Use of Color Intensity Profile Models in DASM

The second major addition to the ASM algorithm is the use of color pixel information in the creation of the statistical intensity profiles for the landmarks. For the ASM algorithm to be able to identify the best locations for each landmark of the shape in the new image, a statistical model is created for each landmark, based on the intensity values of the pixels surrounding the landmarks in each training image. To increase the accuracy of the DASM algorthim, instead of using only the grayscale intensity values with the traditional ASM algorithm, the color intensity values are used (see Equation 3.2.20).

$$g_{im} \rightarrow \frac{1}{\sum_i |g_{ijm}|} g_{im} \tag{20}$$

Each landmark's intensity profile is defined by the vector g_{im}, that has the length $2K+1$. The addition of m, which takes a value from $1-3$, represents the intensity values of R,G and B for the color pixels that make up the vector in the intensity profile. The Mahalanobis distance that is used in the cost function (as described in the ASM testing above) between model landmark profiles and profiles of the best fit points for each landmark in the test image in this case is the averaged distance for each color aspect of the pixels in the search vectors.

3.2.3 Quantitatively Comparing the Expected and Actual Drilled Preparation using DASM

The addition of the ability to monitor the differences between the actual and expected drill preparation is the most important modification of the original ASM algorithm. One of the defining characteristics of the ASM algorithm is the restrictions in the variation of the point distribution

Figure 20: Example of the color pixels making up the search vectors normal to the landmarks for the color intensity profile for the landmarks in the DASM.

model. For DASM, the ASM algorithm is run as usual. However, the algorithm is run a second time with the modified version of the iterative function used to find the best fit for the target shape of the drilled preparation. Step three in this function in the ASM algorithm,

- Apply the constraints to the shape parameters, b, so that it is within three standard deviations (see Equation 3.2.21):

$$|b_i| < 3\sqrt{\lambda_i} \tag{21}$$

is left out of the function in the DASM version of the algorithm. The best fit landmark points identified by the modified algorithm and the original algorithm may differ. This difference is used as a quantitative measure to compare the expected shape and the actual shape, as well as to compare the expected and actual location of the preparation on the tooth.

Although in theory the fitting of the landmarks for the drilled preparation shape should be completely unrestrained, a heuristic was added to keep all of the best fit locations for the shape preparation landmarks within the boundary of the tooth edge (as determined by the concurrently running ASM algorithm segmenting the tooth shape and the expected shape preparation). This allows for the unrestrained segmentation of the preparation within the boundary of the tooth, but

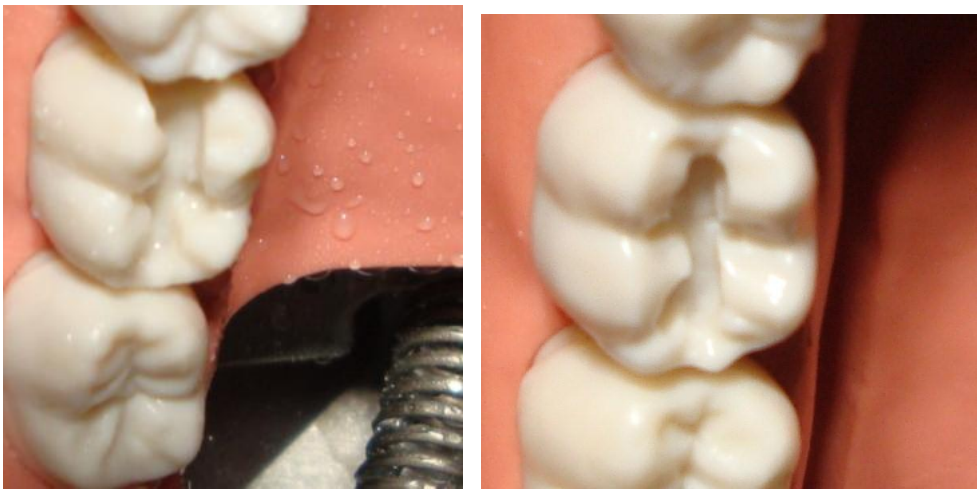
ensures that no landmarks identifying the preparation could be located outside of the tooth surface area.

3.3 Testing of the DASM Algorithm

The effectiveness of the DASM model was tested using images of dental preparations drilled in practice teeth at the Virginia Commonwealth University School of Dentistry. The preparations were drilled by a dental technician. Preparapations were drilled on two different teeth, 19 and 18, and a total of three preparations were drilled, a Class 1 preparation on tooth 19, a training Rectangle preparation on tooth 19, and Class 2 preparation on tooth 18 (see Figure 3.3.21). There was little to no variation in color or texture of the plastic teeth used in these preparations, as they are manufactured and there is no visible disparity of the tooth surfaces.

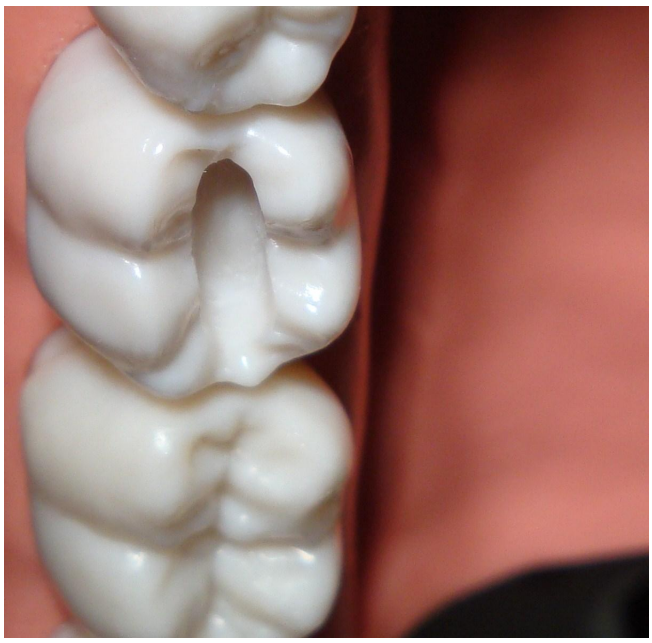
For all of the dental preparations images were recorded using a hand-held commercial off-the-shelf digital camera (a Sony Cybershot DSC-W170/R 10.1 MP Digital Camera) at the highest image resolution (10.1 megapixels) during the drilling process. For each preparation, the drilling process was divided into three sessions. The digital photographs taken during the first session recorded images of the tooth at various points from preparation through the points at which the preparation was just smaller than an acceptable preparation (ie. the size and the shape of the preparation were not large enough to be graded as a correct preparation). Images recorded during the second phase show the tooth with a preparation within an acceptable range. And finally, digital images from the third session show the tooth with a drilled preparation beyond the acceptable range (ie. the drilled preparation is is one or more dimensions larger than acceptable). While the images were being recorded, the lighting and angle and relative positioning of the camera were changed manually throughout the drilling sessions.

In addition to these three phases for each preparation, the Class 1 preparation on tooth 19 was



(a)

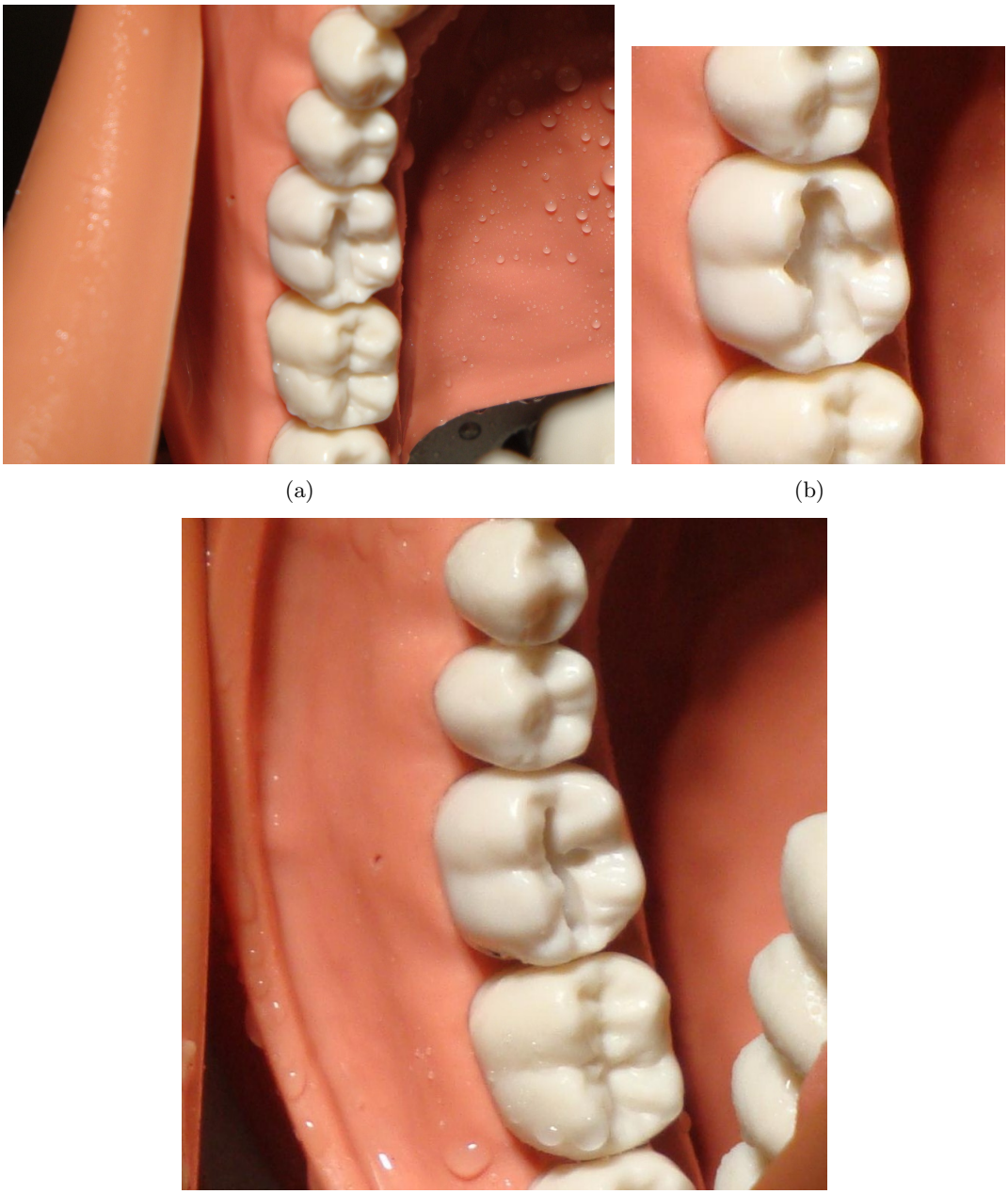
(b)



(c)

Figure 21: Examples of the three different types of tooth preparations used for testing the DASM. a) A Class 2 preparation on tooth 18, b) A Class 1 preparation on tooth 19, and c) A Rectangle preparation also on tooth 19.

repeated so that the preparation was in the wrong position on the tooth surface. All of the images recorded during the drilling of this preparation are incorrect; even if the shape of the preparation is acceptable, the fact that it is not placed in the correct location on the tooth surface renders it



(a) (b)

(c)

Figure 22: Examples of the three incorrect preparations on tooth 19 a) A Class 2 in an incorrect location on the tooth surface, b) A Class 2 preparation that is too large, and c) A Class 2 preparation that is too small.

unacceptable; see Figure 3.3.22.

All of the images used for testing are considered "pre-labeled" because the dental technician

clearly defined each of the three different phases of each drilling of the preparations.

To demonstrate the robustness of the DASM algorithm, the test images were recorded from a variety of angles and with varied lighting, simulating the types of images that would be recorded by an automated image-based dental educational system (for example, if a surgical camera attached to the loupes of the dental student recorded the student's field of view). A small percentage of the testing images contained teeth that were partially obscured by the edge of the mouth (although the drilled preparation was not obscured). Approximately 10%, of the edge of the tooth shape was obscured in these particular test images. Research has shown that ASM algorithm can be robust to partially obstructed images [16]. For this application and set of test images, the DASM algorithm shares the robustness of the ASM algorithm when the shape to be segmented is partially obscured (see Section 3.4).

Two types of testing for the DASM algorithm were completed using the images described above. First, testing was done to see if the addition of the IML to the landmark placement and the use of color intensity values in the statistical modeling for the landmarks improved the accuracy of the DASM over the ASM algorthim. This was done by using the DASM algorithm to locate the correct tooth and preparation on a subset of the images described above. This testing made use of just the images that contained preparations that fell within the "acceptable" range, as the third part of the DASM algorithm was not being tested yet.

For this subset of images, the DASM algorithm was run first with just the modified semi-automatic landmark placement, then with just the color intensity profiles and then with both modifications. Table 3.4.3.4 outlines the results of this testing. The testing was done using 10-fold cross validation, meaning that the training set was divided into 10 'folds', then trained 10 times, each time leaving a different 'fold' out for testing purposes. To confirm accuracy, the placement of each landmark by the DASM algorithm was compared to the correct location expected for

the landmarks in each testing image (as placed by the user). The difference was measured using Euclidean distance; then averages of the error distances were taken.

The purpose of the second testing trial was to confirm the ability of the DASM algorithm to properly locate the drilled preparation in the image (even if it was not in the expected location on the surface of the tooth or if it did not have the expected size and/or shape). This was done again by comparing the location of the landmarks placed by the DASM algorithm with correct location of the landmarks (as placed by the user), with images that contained "incorrect" preparations as described above.

3.3.1 Setting the Parameters for Testing and Training

As stated earlier, there are a number of parameters that can be adjusted for the training and testing of the traditional ASM algorithm, and the same parameters can be adjusted for the DASM algorithm. The number of landmarks used to segment the tooth shape and the drilled preparation totaled 60, 36 for the tooth outline and 24 for the preparation. As mentioned in the IML landmark placement subsection of this chapter, the user places only the initial relevant landmarks. For the tooth outline that was a total of six landmarks; the user identified the point at which the tooth edge touched the edge of an adjoining tooth (four landmarks) and two additional landmarks identified the widest point of each exposed side of the tooth. Four initial landmarks were placed by the user on the drilled preparation; one each at the highest, lowest, and widest points of the preparation. The IML algorithm then automatically placed an additional five landmarks between adjacent user-placed landmarks, resulting in a total of sixty landmarks for the entire model. Limited testing was done with both greater and fewer landmarks. A greater number did not greatly increase accuracy (and did greatly increase training and testing time), and fewer landmarks decreased the accuracy of the model.

Another parameter for the DASM algorithm is the size of the statistical intensity profiles for the landmarks. A vector of ten pixels was selected, because this size was able to capture the most pertinent information about the location of each landmark in the training images (mainly that each landmark was located on an edge). The length of the search vectors to locate the best fit for the intensity profiles within the test image was set at 30 pixels. A smaller search vector resulted in a model that was not always able to locate the best fit landmark based off of the initial placement of the model in the test image. A larger search vector decreased accuracy, as the model searched far outside the parameter of the tooth shape in the image, distorting the results. The highest pyramid levels available (based on the size of the training images) were used.

For testing the DASM, the percentage of convergence for finding the best fit landmarks (the percentage of best fit landmarks that are within the closest fifty percent of the search vector to the center) was 80%. Through visual inspection of the search vectors and the landmark placement during testing, this percentage of convergence resulted in almost every landmark correctly locating either the edge of the tooth or the edge of the preparation for this application. A higher percentage did not result in any increased accuracy of the model.

One standard deviation was allowed for b in the trained model. This resulted in a highly constrained model during testing to insure that the proper shape and relative positioning of the model was retained during the search process. This was especially important in searching for the drilled preparation in an image that contained an incorrect preparation, as the goal was to maintain a correct preparation model to overlay on the test image.

3.4 Results and Discussion

As seen in Table 3.4.1, the addition of the the DASM modifications to the ASM algorithm increases the accuracy of the segmentation of the tooth and drilled preparation within test images. In

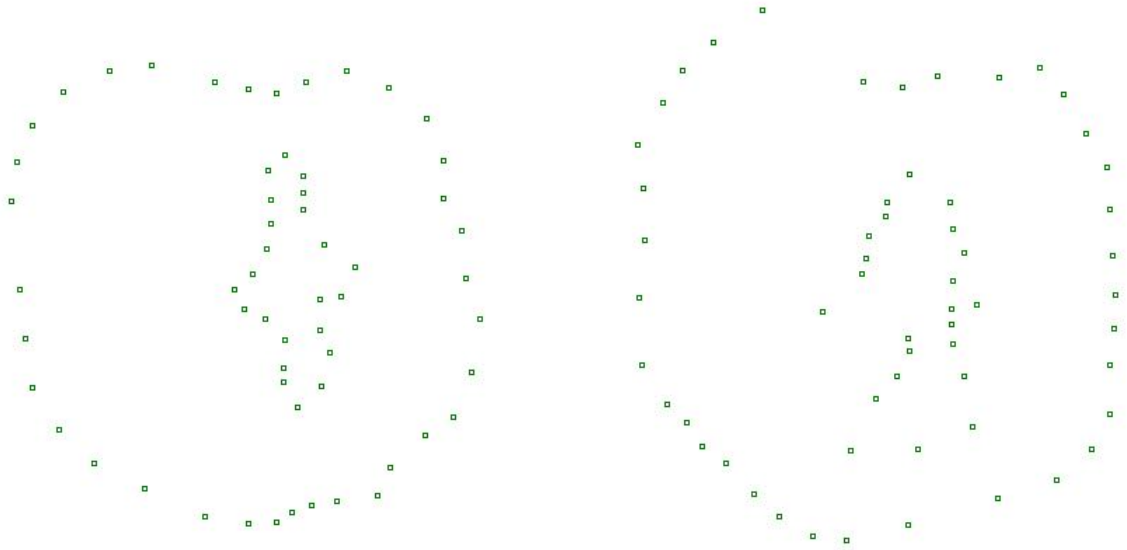
Table 3.4.1 the mean distance of error (number of pixels) is the averaged Euclidean distance between each user-placed landmark (the correct location of the shape) and its corresponding system-placed landmark. Significantly, the use of color intensity pixel values greatly increases the level of accuracy. The combined effect of both additions to the traditional ASM algorithm shows decreased mean distance error of over 80% (see Table 3.4.1). It should be noted that the approximate width and height of the tooth in the digital images was over 700 pixels. Therefore, for instance, the error of the placement of the landmarks when both the IML and color profiles were used was less than 5% of the width or height of the tooth.

Table 1: Accuracy of IML and Color Intensity Profiles on Training Images(Using 10-Fold Cross Validation)

	Traditional ASM	IML	Color Profiles	IML and Color Profiles
Mean Distance of Error	202.56	210.51	56.05	39.7

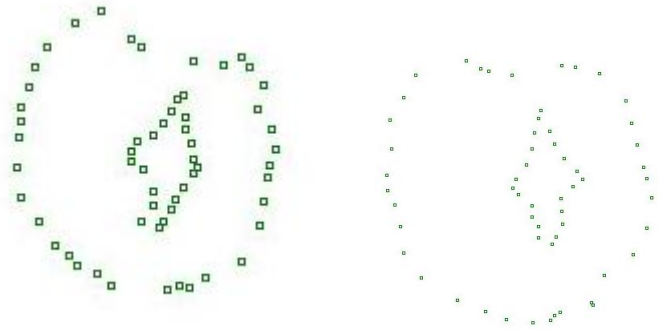
As hypothesized above, the semi-automated placement of landmarks by the IML algorithm, when combined with the use of color pixel intensity profiles, creates a more accurate and consistent training set of landmarks. From image to image, the relative position of the landmarks stays the same, which is not the case in the ASM-created training set, where user error results in inconsistent spacing, as seen in the Figures 3.4.23 and 3.4.24. As a result, IML combined with color pixel intensity profiles produces a more robust model and increased accuracy in the placement of the model in the test image.

The effect of using color pixel values versus grayscale pixel values when creating the statistical intensity models for the landmarks results in an even more dramatic increase in accuracy of the DASM algorithm. The use of the ASM algorithm in this application presents a very difficult segmentation problem. The variation in contrast and color between the teeth in the image and the



(a) (b)

Figure 23: Two examples of the landmarks placed manually with the ASM algorithm. a) Well placed landmarks b) Poorly placed Landmarks



(a) (b)

Figure 24: Two examples of the landmarks placed semi automatically with the DASM algorithm. a) Well placed landmarks b) Poorly placed landmarks

surrounding gum and tissue, is very subtle. When just using a grayscale intensity model for the landmarks, the ASM algorithm has difficultly in locating the correct placement for the landmarks when minimizing the cost function during testing. With the addition of the color intensity values,

the amount of information about the neighborhood surrounding each landmark triples, and the error decreases by 72% percent.

The addition of these modifications to the ASM algorithm is not detrimental in terms of the amount of time or effort for the user, or for the amount of time for the model to train or test. The average amount of time for a user to manually place landmarks was 27 seconds, in comparison to the average time of 26 seconds with the IML algorithm. Although the IML takes approximately the same about of time for the user, the number of clicks needed from the user was decreased by over a third (from the needed 60 clicks to place all of the landmarks, to an average of 38 clicks), which is a significant improvement. The addition of the color intensity values had minimal impact on the time for training or testing for the DASM algorithm in comparison to the ASM algorithm. The average time for training of the DASM with color intensity values was 68 seconds, and the average training time for the ASM algorithm with grayscale intensity values was 39 seconds. The average testing time with color intensity values was 72 seconds, and the average testing time with grayscale intensity values was 47 seconds. The accuracy of the DASM in finding the actual

Table 2: Accuracy of DASM in Locating Incorrectly Drilled Preparations in Averaged Euclidean Distance

Averaged Error Distance	Class 1 Tooth 18	Class 2 Tooth 19	Rectangle Tooth 19
Start of Preparation	64.82	38.91	48.89
Past Correct Preparation	65.3	38.48	36.05
Offset Preparation	–	80.19	–

location of drilled preparation is shown in Table 3.4.2. The error distance in the table is again the averaged Euclidean distance between each landmark placed by the DASM algorithm and its matching landmark hand-placed by the user on the actual preparation. The results demonstrate the ability of the DASM algorithm to accurately locate the drilled preparation. An example of the DASM results are seen in Figure 3.4.25. The yellow landmarks show the location and shape of

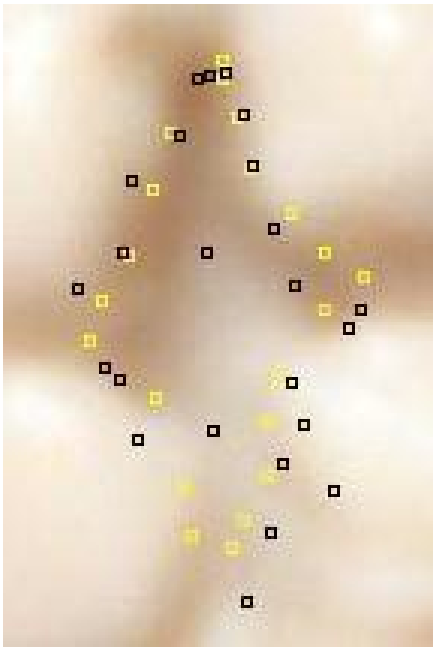


Figure 25: Example of the DASM locating the actual and expected drilled preparation.

the expected drilled preparation. The black landmarks are the actual preparation that the DASM algorithm identified. Via visual inspection, it can be seen that some landmarks are correctly placed and some are not. The level of accuracy is reflected in the Tables 3.4.1 and 3.4.2, by the average error distance. One black pixel in the middle is a landmark placed incorrectly by the DASM system, reflected in the averaged error distance in Table 3.4.2.

CHAPTER 4 Image Processing of Color Images for Texture Detection and Depth Classification in Teeth

4.1 Introduction

Currently, no computer-aided systems exist that can provide user-friendly and easily understandable quantified information about tooth damage due to caries. Existing systems, including DIAGNODent and DIFOTI, also suffer from low specificity, a high learning curve, and complex operational procedures, all of which limit their applicability [105], [5], [79], [82]. One of the goals of this research is to provide an easy way to understand quantitative feedback about the presence and extent of the carious lesions, allowing dental professionals to interpret and integrate data quickly and reliably. This chapter outlines the use of advanced image processing techniques to identify carious lesions on the surfaces of teeth.

Extensive research has also been done in the use of image processing techniques (specifically the use of information about color and texture in images) to determine 3-D depth in 2-D digital images [1], [20], [21], [96]. The same image processing techniques used to identify carious regions in images of the surface of teeth (statistical measurements of texture and color) were also used to attempt the classification of the depth of drilled dental preprations from digital images of plastic teeth (the same types of preparations and teeth as seen in Chapter 3).

For both of these applications, a wide variety machine learning classification techniques were used to classify the digital images containing images of teeth with carious lesions, and digital images of drilled dental preparations (for depth measurement). The same statistical measures were used as feature vectors for both classification applications.

This chapter is organized as follows. First, the machine learning classification techniques used in both applications (and used in the posture-monitoring system outlined in Chapter 5. The methods and results for the classification of carious regions are outlined next, and the chapter is concluded with the methods and results for the depth classification in digital images of drilled dental preparations.

4.2 Classification: Machine Learning Techniques

Before outlining the results, the classification methods used for both of these applications (as well as the posture-monitoring application in Chapter 5 are described. A number of different machine learning techniques were used for all of the classification applications and compared. The techniques are described below.

The decision tree C4.5 was selected as it has been proven to work well within the framework of interactive machine learning [80]. This decision tree method was used in the classification of images with carious and non-carious regions. As an extension of the ID3 decision tree algorithm [80], C4.5 creates a decision tree that splits the data into classes using the attributes that have the highest normalized information gain (the normalized difference in entropy). At each level of the tree, a node is created by calculating what the normalized information gain would be if the node was split using each feature. The feature that has the highest information gain is used to create that node. This is repeated until a predetermined threshold is reached (a maximum depth of the tree has been reached, or all of the nodes contain only samples from a single class, for instance) [80]. C4.5 has been shown to work well with image processing techniques [35].

Another classification technique used for all three classification applications outlined in this paper is the artificial neural network (ANN). ANNs, based on the biological concept of learning through the connection of neurons, are one of the most well known classification methods [9]. Many

variations of ANNs exist; artificial neural networks are very robust for noise, and work best when classifying data from a complex system. In general, an ANN takes a relatively long time to train, but a very short time to test or run [83]. One of the main drawbacks of classification with an ANN is their tendency towards over fitting. For that reason, more noise may actually help with the classification of ANNs as the algorithm will create broader classes that may encompass more data.

Adaptive boosting (AdaBoost) is an adaptive classification technique that is used in conjunction with other algorithms. AdaBoost strengthens a weak learning algorithm by calling the weak learning algorithm multiple times; the classification of the weak learning algorithm for the training data is given a distribution of weights. In comparison with ANNs, AdaBoost is more sensitive to noise, but not as prone to over fitting. Like ANNs, AdaBoost can be time-consuming to train, depending on the number of iterations with the weak learning algorithm (T) [32].

Support vector machines (SVMs) are supervised linear classifiers that map input data into classes through the creation of hyperplanes. The SVM model makes use of training data, which is divided by hyperplanes that are separated by the maximum distance possible such that no points are between them. The samples that fall along the hyperplanes are called support vectors [77], [50]. Although the SVM algorithm is robust, like many machine learning techniques it suffers from high sensitivity to parameter tuning, which makes accuracy computationally expensive.

Learning Vector Quantization (LVQ) can be considered to be a special case of an ANN. LVQs are most closely related to self-organizing maps (SOMs) (they are considered a precursor of SOMs), which is a type of unsupervised neural network. Like neural networks, LVQs consist of layers of neurons that are connected via weights. Through the use of reference, or codebook, vectors the neuron that most closely matches the expected output is strengthened via updating of its weights. Like SOMs, the LVQ algorithm makes use of competitive learning. The LVQ technique also has similarities to k-Nearest Neighbor, because the winning neuron is determined by calculating the

distance between a codebook vector and the input vector [88], [93]. LVQ can be a time-consuming method of classification, as the network needs to be trained in a method similar to an ANN.

k-Nearest Neighbor can be considered as the most basic machine learning technique. As a sample is given to the algorithm, it is compared to all of the training data that is pre-labeled into their respective classes. The distance between all of the sample data and each new training sample in computed. The training sample that is "closest to the testing sample is found, and the testing sample is then labeled with that training sample's label. Usually Euclidean distance is used to compute the distance between testing and training samples, but other distances can be used. In comparison to many other machine learning techniques, k-NN has no required "training period, but testing or running the algorithm can be computationally expensive. As the training data set increases in size, the running time can increase greatly for k-NN, although optimization techniques do exist [8].

Another method used for both image processing applications is the radial basis function network. A radial basis function network is a type of artificial neural network that uses radial basis functions as activation functions. In general, artificial neural networks are well suited for problems like image classification because they can deal with complex problems. Radial basis function networks are especially well suited as they have a faster training time than other artificial neural networks, and, unlike some artificial neural networks, they can find a global solution to a classification problem. A radial basis function network maps the data into a higher dimensional space so that the problem is easier to solve [103], [15].

All of these machine learning techniques are applied to all of the classification problems in this chapter and Chapter 5.

4.3 Methods: Classification of Carious Regions

4.3.1 Feature Extraction Using Statistical Measurements

To measure the presence and extent of caries, the algorithm takes as input digital photographs of the tooth surface to be diagnosed. These images are processed by first extracting features to identify carious regions on the surface of the tooth, and then applying classification techniques to identify which pixels in the digital image are part of the carious region(s).

Extraction techniques are used to identify and evaluate quantitatively caries on the surface of the tooth. The input to the system is the image of the surface of the tooth. The main visual cues used by dentists to identify the demineralization characterizing early caries include cavitation, texture and roughness, discoloration and opacification of the surface of the tooth [95]. These visual cues are quantitatively expressed through the feature extraction process. A feature vector is created for each pixel and its neighboring pixel within the segmented tooth region of the image. This feature vector describes the visual cues listed above.

The feature vector contains 7 features extracted from the pixel and a 7x7 pixel sub-image comprised of the neighboring pixels. The 7 features are:

- Gradient: The magnitude of the gradient (the rate of change) of color of the image [35] and
- Statistical and signal processing measures: six texture measures are extracted from the image for the red color channel [35]:
    1. Average intensity level
    2. Average contrast
    3. Measure of smoothness
    4. Third moment (the skewness of the intensity histogram)

5. Measure of uniformity

6. Entropy

The statistical and signal processing measures are calculated from the histogram of each pixel and its 7x7 pixel neighborhood. The first two measures, the average intensity level and the average contrast, are, respectively, the mean and the standard deviation of the intensity values of the histogram. The third measurement is a measure of smoothness, which is calculated in Equation 4.3.1.22:

$$R = 1 - \frac{1}{1 + \sigma^2} \tag{22}$$

where σ is the standard deviation. An R value approaching 0 would be a smooth area of the image, and a value approaching 1 would have a high value of variety in intensity of pixel values, and therefore would appear very textured. The fourth measurement of texture is the third moment of the histogram,which measures the "skewness of the intensity histogram. It is defined in Equation 4.3.1.23

$$\mu_3 = \sum_{i=0}^{L-1} (z_i - m)^3 p(z_i) \tag{23}$$

where L is the number of possible intensity values (256 for the images used), $p(z_i)$ is the probability of z_i (an intensity value, between 0 and 256) occurring in the histogram, and m is the mean for the histogram. This third moment describes statistically the symmetry of the histogram, and the resulting value is 0 if the histogram is symmetrical, a positive number if the histogram is skewed to the right, and a negative number if the histogram is skewed to the left. This gives information about the overall trend of intensity values in the histogram [35]. The fifth texture measure is a measure of uniformity. It is defined in Equation 4.3.1.24:

$$U = \sum_{i=0}^{L-1} p^2(z_i) \tag{24}$$

The sixth measure of texture is the entropy of the histogram, measuring the randomness of the intensity values [35], and defined in Equation 4.3.1.25:

$$\varepsilon = \sum_{i=0}^{L-1} p(z_i) log_2 p(z_i) \tag{25}$$

where $p(z_i)$ is the probability of a specific intensity value occurring in the histogram.

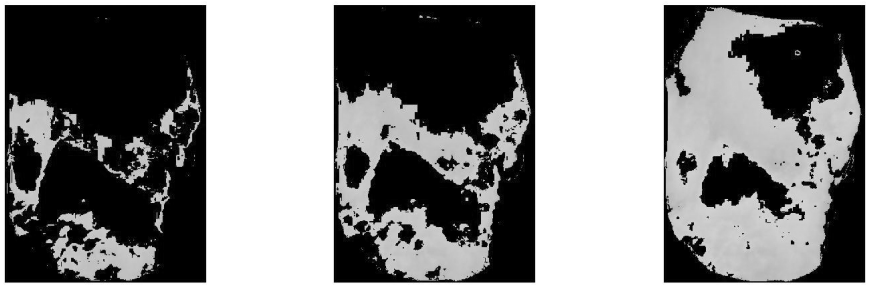
4.3.2 Segmentation of the Tooth to Identify Carious Regions for Classification Training

To test the ability of the system to correctly identify carious regions in the images of the teeth, a color-based segmentation method is used to segment training images of the teeth into carious and non-carious regions.

As shown in Figure 4.3.2.26, segmented images of teeth were used for this step. First, an area of the tooth surface that contains caries is manually highlighted. Since all of the carious regions on the surface of a tooth may not be continuous, color segmentation was applied to identify other carious areas. Using the input from the user that identifies the carious region, the average color of the pixels in that region is calculated. Each pixel in the image of the tooth is compared to this average color, measuring the similarity between the colors using Euclidean distance. Each pixel in the image is then labeled as either "carious or "not carious based on the similarities between the color values of the pixels. The cut-off for classifying the pixels is based on a threshold value, T. For each image containing caries, an optimal value of the threshold T is determined using a set of training examples which are visually inspected and segmented. The value of T ranged between 10 and 30 depending on the image (see Figure 4.3.2.27).

The segmented images are then used to create a training set of pixels to test the classification algorithm.

Figure 26: Example image of segmented tooth containing caries.



(a) The segmented region using T = 10.

(b) The segmented region using T = 15.

(c) The segmented region using T=30.

Figure 27: Examples of color segmentation of carious regions of the tooth from Figure 4.3.2.26 with different values for threshold T.

4.3.3 Classification of Caries Using the Pixel Feature Vectors

To identify pixels that represent areas of the tooth surface that are damaged by caries, a supervised method was applied, using images with known and labeled areas of caries are used to train the

system prior to the use for caries identification. Principal Component Analysis (PCA) was tested as a filtering technique to reduce the redundancy among attributes by comparing classification results with and without using PCA. Recall that PCA is a statistical technique used to reduce the number of features describing an image by removing any features that contain little to no novel information about the image. PCA projects the data into a lower dimensional space, reducing the dimensionality of the data through a linear combination of the features [103].

4.3.4 Results and Discussion

The classification techniques were tested using images of six teeth. These images were segmented from six different images of carious and non-carious teeth in stock educational images from the Virginia Commonwealth University School of Dentistry. Thirty-three images of teeth showing no carious regions and nineteen images of teeth containing carious regions were used for testing. Feature vectors were created for all pixels in each image, and approximately 40,000 pixels representing a random sampling from all of the segmented teeth images were used for classification. Due to the relatively small number of carious regions in the images, there were 38,505 samples of non-carious pixels and 1,532 samples of carious pixels.

Table 3: Results of Classification of Pixels as Carious or Non-Carious

Classifier		Accuracy	Precision	Recall
RBF	With PCA	96.62%	0.97	0.11
	Without PCA	96.86%	0.97	0.23
SVM	With PCA	96.17%	0.96	0.98
	Without PCA	96.62%	0.97	0.98
k-NN	With PCA	94.63%	0.97	0.97
	Without PCA	96.34%	0.98	0.99
LVQ	With PCA	96.20%	0.96	0.99
	Without PCA	96.31%	0.96	0.04
ANN	With PCA	96.17%	0.96	0.07
	Without PCA	96.17%	0.96	0.07

The feature vectors were used for classification using 10-fold cross validation, as shown in Table 4.3.4.3. For example, with the RBF network classification, without PCA, 96.86% of pixels were correctly classified and 3.14% of pixels were incorrectly classified. With PCA applied, and the first 5 transformed features used, 96.62% of pixels were correctly and 3.38% of pixels were incorrectly classified. In addition to accuracy, the precision and recall for the different machine learning techniques were calculated. Where accuracy is simply the precentage of true results, precision is the number of true positives divided by the number of true positives plus the number of false positives. That is to say, precision is the number of samples correctly labeled into one class divided by all samples that are labeled as belonging to that particular class (correctly or incorrectly). Recall is the number of true positives divided by the number of true positives plus the number of false negatives; it can also be defined as the number of samples correctly classified into one class divided by all of the samples that should be classified as that class. If precision or recall has a value of 1.0, then all of the samples were correctly labeled by class.

Looking at Table 4.3.4.3, it is interesting to note that although all of the machine learning techniques showed a high level of accuracy, ANN, LVQ without PCA, and the RBF network had a very low recall rate. As stated earlier, there were 38,505 samples of non-carious pixels and 1,532 samples of carious pixels. For these three machine learning techniques, a large number of samples that were carious were erroneously labeled as non-carious. Due to the large disparity in the sample sizes for the two classes, accuracy was still shown as over 90% although the majority of the carious pixels were misclassified. This is an indication of overfitting for these techniques. It is also interesting to note that LVQ with PCA had a high recall rate, and without PCA LVQ had a low recall rate. This may be due to the effect of PCA lowering the number of features used for classification, reducing the problem of overfitting of the classifier to the training data.

The high level of accuracy by all of the machine learning techniques testing and the high precision

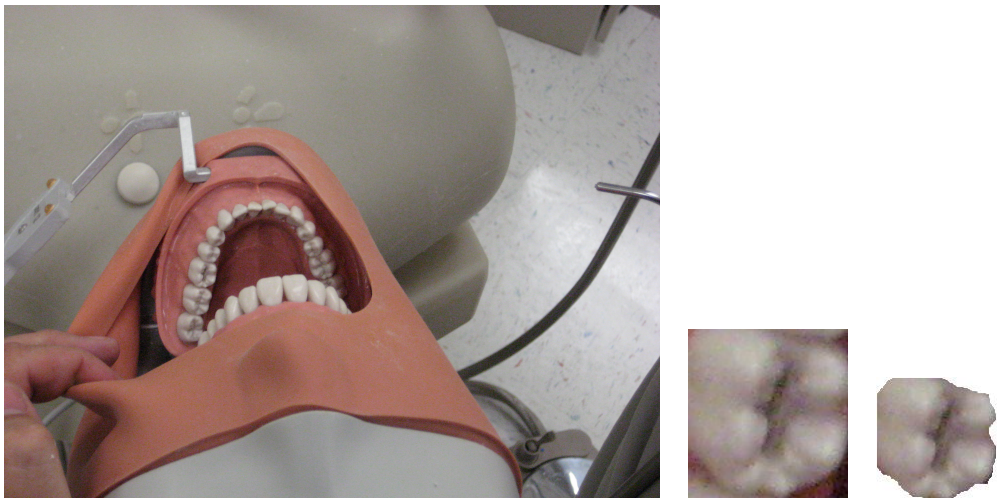
and recall values for the majority of the techniques used demonstrates the effectiveness of using statistical measurements to differentiate between pixels representing carious and non-carious regions on the tooth surface. This demonstrates the promising performance of the system in accurately identifying damaged areas on tooth surfaces using digital color images.

4.4 Methods: Classification of Depth in Dental Preparations

4.4.1 Classification of Depth of Drilled Dental Preparations in Color Digital Images

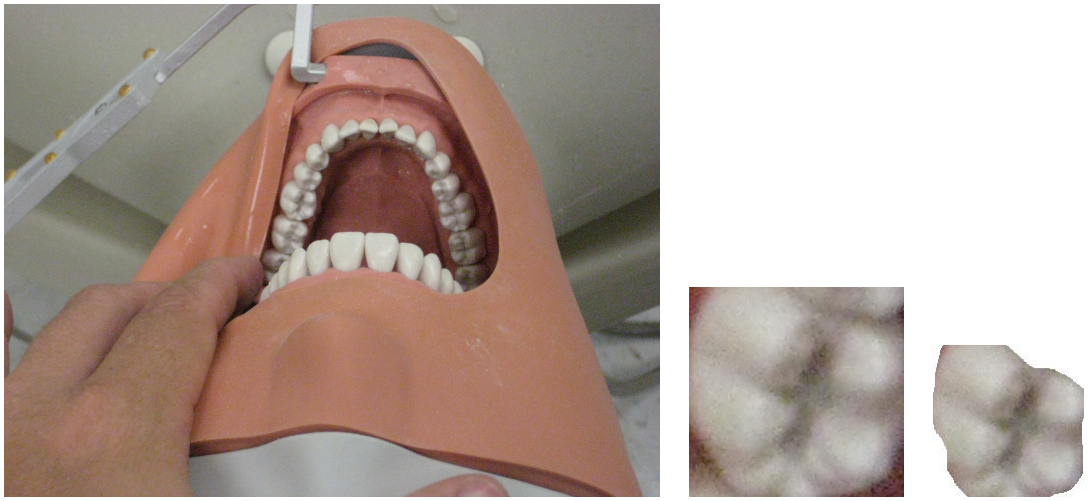
The same statistical measures outlined in 4.2.1 were used to test the accuracy of depth classification in images of drilled dental preparations. Research has shown that monocular cues such as variations in texture and color can be used to successfully determine depth from 2D images [1], [20], [21], [96]. As in Section 4.2.1, 7 features were calculated from the pre-segmented images of the drilled teeth. However, all three color channels were used for the statistical and signal processing measurements, giving a total of 19 features:

- Gradient: The magnitude of the gradient (the rate of change) of color of the image [35], and
- Six statistical and signal processing measures of texture were extracted from the image for each color channel, for a total of 18 features (6 for each of three color channels) [35]:
    1. Average intensity level
    2. Average contrast
    3. Measure of smoothness
    4. Third moment (the skewness of the intensity histogram)
    5. Measure of uniformity
    6. Entropy



(a) (b) (c)

Figure 28: Digital images of too deep preparation. (a) Original image of the preparation. (b) and (c) are images of tooth 19, identified and isolated.



(a) (b) (c)

Figure 29: Digital images of a preparation with acceptable depth. (a) Original image of the preparation. (b) and (c) are images of tooth 19, identified and isolated.

The feature vectors were created using three steps, the third being an optional optimization step:

1. First, the image was divided into 15x15 pixel patches.

2. Second, the texture and color features were extracted for each patch of the image. This helped

to maintain information about location of changes in depth within the image.

3. Third, PCA was applied to the feature vectors. For this application, PCA was used to reduce the number of features. For machine learning techniques, the number of samples used for training needs to be larger than the number of relevant features, and the values of the features in the training set must be representative of all the data to be classified in order to avoid overfitting [103]. Due to the small sample size (150 images) available for training and testing, the number of features was reduced to try to increase the accuracy and generalizablity of the classification model. By using PCA, the number of features could be reduced without losing much relevant information [103], [15].

4.4.2 Results and Discussion

To test the ability of these statistical and signal processing measurements to classify depth in color digital images of dental preparations, the feature vectors outlined in Section 4.4.1 were applied to a total of 150 images of dental preparations done at the DentSim laboratory with an artificial practice head and tooth at the Virginia Commonwealth University School of Dentistry. The images were recorded using a digital camera while an expert cut a Class I Amalgam preparation on the occlusal surface of tooth 19 (universal numbering system) [92]. This preparation was done twice with two different practice teeth. Each tooth was first prepared to an acceptable depth, and then each tooth was drilled past an acceptable depth as determined by the dental expert. Of the 150 images, 78 images were taken when the preparation was at an acceptable depth, and 72 images were taken past an acceptable depth. A feature vector was extracted from each image using MATLAB (2007a, The MathWorks Inc., Natick, MA) and these vectors were then classified using WEKA's implementation of a radial basis function (RBF) network and the other machine learning techniques outlined in Section 4.2 [103].

Of all of the feature vectors created from the images, 75% were used for training and 25% were used for testing. The results of the classification are seen in Table 4.4.2.4.

Table 4: Results of Classification of Depth

Classifier		Accuracy	Precision	Recall
RBF	With PCA	84.21%	0.78	0.88
	Without PCA	73.68%	0.61	0.69
k-NN	With PCA	30.67%	0.34	0.38
	Without PCA	36.00%	0.35	0.40
ANN	With PCA	58.0%	0.56	0.61
	Without PCA	46.67%	0.56	0.61
LVQ	With PCA	44.0%	0.42	0.42
	Without PCA	46.0%	0.43	0.38
SVM	With PCA	55.67%	0.57	0.22
	Without PCA	56.0%	0.55	0.44

These results demonstrate the difficulty of classifying depth in a 2-dimensional image. However, the level of accuracy reached using the RBF network classifier shows that depth can be ascertained from 2-dimensional images. Since the data was not easily classified with the other machine learning techniques, it can be concluded that the data is complex. k-NN had the worst results which implies that the data is most likely not linearly separable, as research shows k-NN does not do well with non-linearly separable data [103]. The other machine learning classifiers (SVM, ANN and LVQ) that were tested can classify non-linearly separable data, but RBF had the highest level of accuracy. By mapping the data into a higher dimensional space (which is not done with the other machine learning classification techniques tested) the RBF was able to achieve a much higher accuracy, demonstrating it to be a good fit for this application. Another advantage of the RBF was the training time. On average, it took the SVM classifier 98 seconds to train. It took the k-NN 0.05 seconds, the LVQ classifier 11 seconds, and the ANN 40 seconds. The RBF network took only 0.1 seconds, faster than all other classifiers except the k-NN (which had the worst performance).

Taking into consideration of the use of a commercial off-the-shelf digital camera instead of an intraoral or surgical camera (which would provide much higher magnification and detail) and a limited number of sample images, these preliminary results demonstrate the feasibility of the use of image processing for the classification of depth in a 2D image. With more extensive sampling, the use of higher resolution cameras and possibly exploration into other image processing techniques, it is likely that this accuracy can be increased substantially.

CHAPTER 5 Posture Monitoring

5.1 Introduction

Another need in dental education is computer-aided posture monitoring, as research has shown that there is limited instruction in correct positioning during dental education, and incorrect posture is a substantial problem in dental professions [86]. Few if any Schools of Dentistry have technological resources to assist in teaching ergonomically correct posture to dental students. Currently ergonomic issues are addressed in a one-hour lecture before students begin working on mannequins, followed up by instruction and feedback during practice time. Despite this background and the use of magnification devices (loupes) that allow dental clinicians to observe the mouth of the patient in detail without having to bend over, the students are nonetheless observed bending over to view the mouth at a closer range, working in positions that are physically harmful to themselves. The school has inadequate human instructional resources to provide sufficient one-on-one real time feedback to correct these problems.

Current research, as described below, corroborates the need for appropriate training in body positioning. Yet, according to a recent study, 60% of dentists say they did not receive sufficient training on the ergonomic aspects of dentistry [94]. It would be cost-prohibitive to employ sufficient staff trained in the ergonomics of dentistry to assess and correct students' postures in real time as they practice cavity preparation. A relatively inexpensive automated system, developed in consultation with experts in ergonomics as well as dental education, would respond to the need to monitor students' posture during training, and would also be useful to assist practicing professionals in maintaining proper posture.

5.2 Methods

The main goals in the design of this ergonomic positioning prototype are to insure that the system is portable and unobtrusive. To meet these goals the system proposed is based around the laboratory coat, an article of clothing that is worn by most dentists. A number of small, lightweight inclinometers are attached to the laboratory coat so that the dental practitioner can wear the prototype without having sensors strapped or taped to his or her body. The inclinometer sensors are SCA121T 2-Axis inclinometers manufactured by VTI Technologies. They are attached via unobtrusive cables to a small analog-to-digital converter, which in turn is attached to an off-the-shelf Pocket PC. This system is relatively inexpensive as compared to other technologies used for similar applications. It is important to note that the inclinometers were affixed solely to the laboratory coat, and therefore were not stationary on the body of the user. The coat had the ability to shift and move during the testing of the system, and the relative location of the sensors varied from user to user as only a single extra-large laboratory coat was used for all users, regardless of their size or body type.

There are currently three inclinometers used in the prototype, one placed over each shoulder blade, and one placed in the middle of the lower back (approximately over the L1 vertebra). The placement of the inclinometers were determined as optimal through initial testing. The system was first tested with only one inclinometer affixed on the middle of the back (approximately over the T10 vertebra) with poor results, and through initial trial and error the placement of the three inclinometers for the system was determined. Further testing of this system may reveal more specified locations for higher accuracy, or the need for an additional inclinometer sensor affixed to the head of the dental practitioner in order to measure additional aspects of the ergonomic position. This may be accomplished by affixing an inclinometer to the required loupes (magnification glasses)

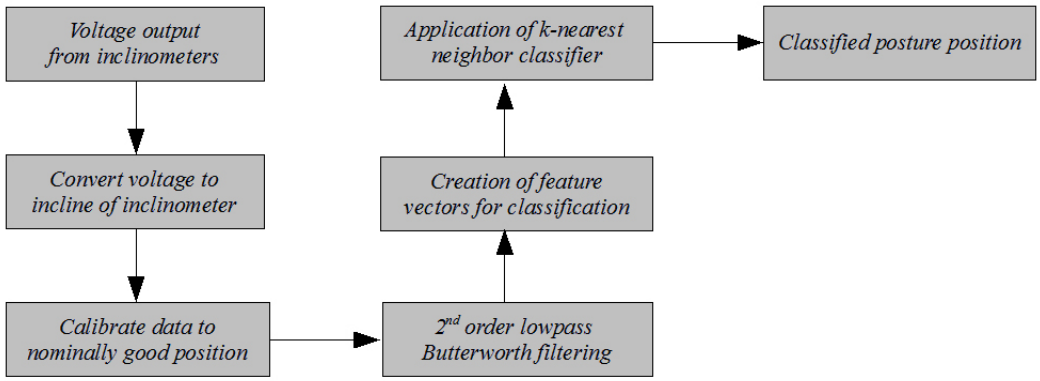


Figure 30: Signal processing outline for posture classification.

of the dental practitioner. The design includes the planned use of an audio user interface, which are anticipated to be unobtrusive to the dental practitioner. The output from the inclinometers goes through a series of signal processing techniques to classify the data into one of a number of predetermined posture positions.

The steps are outlined in Figure 5.2.30. First, the raw data (sampled at 20Hz) is converted from voltage to the angles of the inclinometers. The sampling rate was selected based on initial testing, which showed that 20 Hz captured enough information to accurately classify posture positions. An increased sampling rate did not increase accuracy, and would potentially slow down the system. The SCA 121T Series inclinometers record any change in incline relative to gravity by showing a change in output voltage. This can be measured directly using a voltmeter and applying the following formula to determine the angle, A, in Equation 5.2.26:

$$A = \frac{\sin^{-1}(V - O)}{S} \tag{26}$$

where V is the voltage output from the inclinometer; O is the voltage output of the inclinometer at 0 degrees and S is the sensitivity of the device (given by the manufacturer of the SCA121T Series inclinometers). At the next stage, calibration, (as seen in the outline of the algorithm in

Figure 5.2.30), after the angle is determined for the x and y axis for each inclinometer, the data is calibrated, seen in Equation 5.2.27:

$$S'_{ij} = N_{ij} - S_{ij} \tag{27}$$

where S_{ij} is the original sample for axis i and inclinometer j, N_{ij} is the angle of the first sample recorded in the nominally good position for axis i and inclinometer j, and S'_{ij} is the resulting calibrated position for that axis and inclinometer. Because the relative position of the coat and user may change with each use, calibration is done so the movement is relative for each user from session to session.

Next, a filter designed in the Fourier domain is applied in the form of a second order Butterworth FFT filter with a cutoff frequency of 5Hz. A feature vector is created consisting of the filtered data from the x and y axes from all three inclinometers.

After filtering, the sample data is classified using a k-NN algorithm. k-NN is a supervised classifier, meaning it uses a set of labeled data representing the different classes (in this case, the different posture positions) to classify new, unknown data. The k-NN algorithm will classify a new data sample xi that has feature set where $(x_{i1}, x_{i2}, ..., x_{i6})$ represents the values for x and y axis for each of the 3 inclinometers (for a total of 6 features) for sample i. The classifier measures the Euclidean distance between the new data sample and the labeled data, seen in Equation 5.2.28:

$$d(x_i, x_1) = \sqrt{(x_{i1} - x_{l1})^2 + (x_{i2} - x_{l2})^2 + ... + (x_{i6} - x_{l6})^2} \tag{28}$$

for all l labeled samples in the feature space. The k-NN algorithm then takes the k samples that are closest, that is, those that have the smallest distance to the new unknown data sample. The class to which a majority of those closest samples belong is then assigned to the new data sample. To reduce ties, the k that is selected is usually an odd number. For this application, a number of different k values were tested.

To demonstrate the advantages of using a k-NN algorithm, the classifier was tested against a number of other machine learning techniques. The techniques that have been tested are a feed-forward back-propagation artificial neural network (ANN), AdaBoost with C4.5, a support vector machine (SVM), and learning vector quantization (LVQ). These classification methods were chosen because they are at the forefront of machine learning research and cover a broad range of diverse theories and techniques. Another reason these particular techniques were chosen is because they have a wide range of advantages and disadvantages; some have been proven to work well with the classification of position and movement via on-body sensors, and some have not previously been tested for this application. Testing shows that k-NN compares favorably with the range of machine learning algorithms.

5.3 Results and Discussion

The system outlined above was tested on data generated by the posture measuring system during posture-recording sessions with eleven different subjects (6 male subjects and 5 female subjects). One of the goals of this research is to design a system that is customizable for each individual user. For each subject, the laboratory coat was calibrated by having the subject sit in an ergonomically correct position and recording one set of data from the inclinometers. All further data from the system is recorded as the difference between the current posture position and the first calibration data. The range of posture of the user is then measured by recording data while the users sit in an ergonomically correct position and nominally harmful positions: leaning left, leaning right, leaning forwards and backwards, and slouching. Given the initial calibration and identification of a normal range of posture for the user, the processing techniques were used to infer the subject's nominally normal range of motion.

Figure 5.3.31 shows the data recorded from one user during a single session. There are three

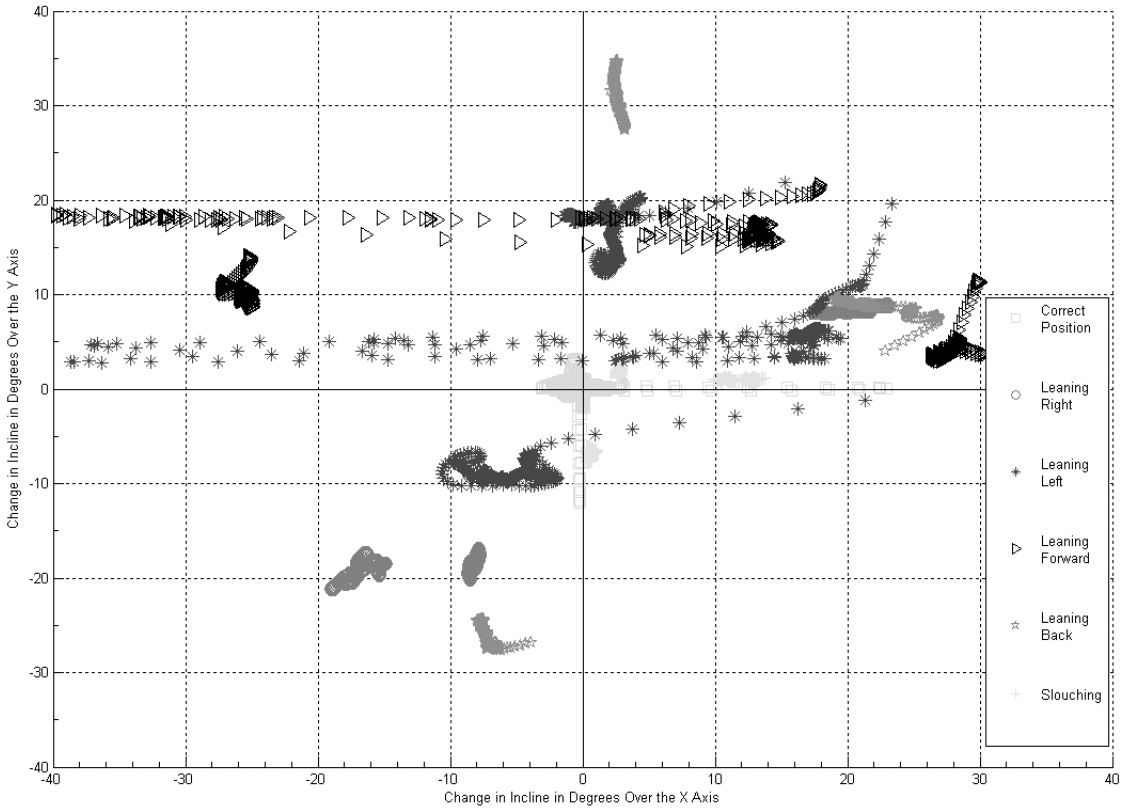


Figure 31: Calibrated and filtered output from three inclinometers showing change in x and y axis of incline.

groupings of data from each inclinometer for each position held. The data is calibrated in relation to the first recorded data sample from all three inclinometers while the user held the nominally good position. This allows for the sensors to move from session to session without a negative impact on classification accuracy. Figure 5.3.32 also shows the variation in data in each posture position. The three lines show the change in incline from the correct posture position for the three inclinometers. The variation in the data shows the movement and variation even while the subject is in one posture position. The difference between posture positions varied greatly from one subject to another, and the data recording the difference between the posture positions was not trivial. Because the system will be used by dentists while they perform dental procedures, the goal is to create a system that is robust to noise. The subjects who tested the system were instructed not to remain perfectly

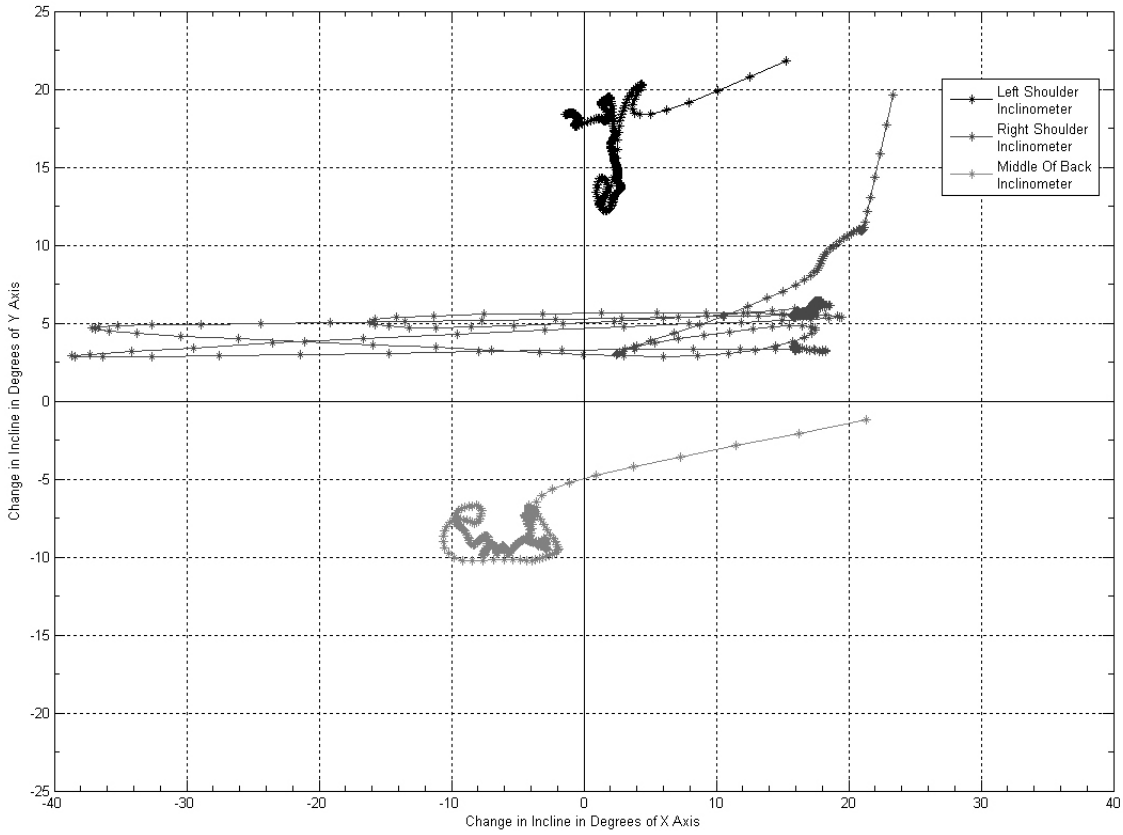


Figure 32: Calibrated and filtered output from three inclinometers during the time period when the subject leaning to the left.

still while each posture position was recorded; instead they were encouraged to talk, move their arms, etc. while holding each posture position. This movement within each position can be noted by the variation of data points in Figure 5.3.32. Each user was instructed to move to the different posture positions, and the system recorded their position for approximately one minute. Each user determined by how to interpret each position, for example, when leaning "left" each person chose how far to lean to the left and what angle their body would take (in addition to any movement within that position over the minute-long recording session), so there was a great deal of variation from user to user of the different nominal positions.

k-NN and the four other machine learning classification techniques were trained and tested with

the data sets from all of the eleven users' posture-measuring sessions, using 10-fold cross validation. The results are presented in Table 5.3.5. The k-NN algorithm was tested with neighborhood sizes of 1, 2 and 3. The other machine learning techniques were tested as follows. A standard feed-forward back- propagation neural network was used, with one hidden layer consisting of a number of hidden nodes (the number of hidden nodes was varied in testing), and a bipolar sigmoid activation function for both the hidden and output layers. The LVQ algorithm used was LVQ2.1, which uses two sets of codebook vectors, with one associated with the "correct class for the current sample that is being classified, and one associated with the incorrect class [52]. The AdaBoost technique was used with a decision tree, C4.5, as the weak learning algorithm. For the SVM algorithm, the training data was normalized, and a polynomial kernel was used, which allows for classifications that are not linear.

As seen in Table 5.3.5, all of the classifiers that were tested performed relatively well, with the LVQ algorithm having the lowest average accuracy at 92.31%. As mentioned earlier, all of the methods were tested using 10-fold cross validation. This tests for the overfitting of classifiers. In Table 5.3.5, the precision, recall, false positive rate and accuracy of the results are averaged over all 11 users. Table 5.3.5 also demonstrates the effect of changing the parameters of the ANN algorithm. A wide range of hidden nodes, from 2 to 8, as well as 50 and 75, were used to test the accuracy of the ANN classifier and the subsequent results show that for this application the neural network is saturated. Changing either the number of hidden nodes or the learning rate has a dramatic effect on the accuracy of the classifier. This demonstrates one of the main issues with parametric classifiers; by choosing the incorrect parameters the accuracy of the system is compromised, but finely tuning multiple parameters to create the most robust classifier is time intensive and may result in overfitting. These results demonstrate that with the correctly chosen parameters, all of the machine learning methods are relatively accurate in classifying different posture positions.

The k-Nearest Neighbor algorithm, although more accurate than the majority of other machine learning classifiers that were tested, was not statistically different than any other classifier. k-NN does, however, have distinct advantages over the other techniques tested. k-NN has a very low computational time, requires little to no training time, is simple and straightforward, is robust and generalizable, and as shown here gives the same or better accuracy than a number of other machine learning techniques.

Table 5: Averaged Results Over All 11 Users

Machine Learning Technique	Parameters		Precision	Recall	False Positive Rate	Accuracy
k-NN	k:					
	1		0.99	0.99	0.01	99.94%
	2		0.99	0.99	0.01	99.92%
	3		0.99	0.99	0.01	99.90%
ANN	Number of Hidden Nodes:	Learning Rate:				
	2	0.10	0.63	0.62	0.10	61.69%
	3	0.01	0.68	0.66	0.07	65.93%
	3	0.10	0.77	0.75	0.06	74.81%
	3	0.50	0.69	0.65	0.07	64.58%
	4	0.10	0.87	0.76	0.03	85.11%
	5	0.01	0.95	0.93	0.02	92.64%
	5	0.10	0.88	0.87	0.03	86.84%
	5	0.50	0.79	0.77	0.04	76.08%
	6	0.10	0.89	0.88	0.02	88.08%
	7	0.10	0.90	0.85	0.02	89.81%
	8	0.01	0.99	0.99	0.01	99.84%
	8	0.10	0.92	0.93	0.02	90.88%
	8	0.50	0.82	0.80	0.04	80.47%
	50	0.01	0.91	0.92	0.02	92.31%
	50	0.10	0.95	0.95	0.01	95.33%
	50	0.50	0.85	0.85	0.03	84.88%
	75	0.10	0.96	0.96	0.01	95.96%
SVM	–	–	0.96	0.96	0.01	99.16%
LVQ	–	–	0.92	0.92	0.02	92.31%
AdaBoost with C4.5	–	–	0.99	0.99	0.01	99.73%

CHAPTER 6 Conclusions and Proposed Future Work

6.1 Conclusions

Through the design and testing of the novel image processing techniques outlined in this dissertation, the feasibility of a portable, automated computer-aided dental educational system in demonstrated. One of the most challenging aspects of such as a system is the need for a complex image processing system to interpret the images of the student's work on the dental preparation. Identifying and monitoring the progression of the drilled dental preparation from 2D color digital images is not trivial. The primary problems to be solved in identifying and monitoring the preparation are determining the size, shape and location of the preparation, and assessing the depth and smoothness of the preparation. To address these two issues, two novel image processing algorithms have been designed and tested:

- To identify the location of the preparation in relation to the surface of the tooth, and to monitor the progression of the size and shape of the preparation as it is being drilled, a novel version of the Active Shape Model technique has been created. This algorithm, the Directed Active Shape Modeling algorithm (DASM), can give quantitative feedback to the user about the size, shape and relation to the surface of the tooth of the dental preparation.

- To monitor the depth and smoothness of the preparation, statistical measurements of the pixels in the image are used as features for classification of the preparation as either being within or outside of the range of acceptability for smoothness and depth. The results outlined in the above chapters demonstrate the ability of the image processing and machine learning

techniques to access accurately correctness of a dental preparation as it is being drilled.

Through the use of a computer-aided education system with image processing algorithms, dental students could receive immediate auditory feedback while working on drilled preparations of artificial teeth. This system would provide multiple benefits. It would help make more efficient use of faculty time, allowing the instructor to focus on weaker students that need additional instruction and allowing the more skilled students to advance on their own pace. Typically the instructor inspects the students' work after the procedure is completed and points out any deficiencies. Provision of immediate feedback would help the students to recognize the point within a procedure at which a problem arises, and allow them to focus on enhancing their skills in that specific area. This will improve the use of class time, require less faculty feedback, and enhance the self evaluation skills of the students, impacting clinical treatment of patients. The system could allow the students to continue individual training with immediate and consistent feedback at times when the instructors are not available (such as evenings and weekends), allowing them to develop their psychomotor skills more quickly. The system would also help dental students develop the analytical skills needed for assessing their own work, which is the foundation of correct clinical decision making when working with live patients. The system could also be modified to be used in a clinical setting, increasing the students' and patients' confidence level by providing a mechanism to prevent physical damage to biological structures.

This paper also outlines the design and proof of concept of a portable posture-monitoring system for dental students. By using machine learning techniques, change in body position can be accurately classified using unobtrusive on-body sensors. By monitoring posture and providing feedback in real-time, work-related musculoskeletal disorders may be prevented, lowering the risk of injury for dental students and dental practitioners.

6.2 Proposed Future Work

This dissertation provides the groundwork for the development of a comprehensive portable dental educational system. Novel image processing algorithms are outlined, and the ability of these techniques to accurately classify the size, shape, location and depth of drilled dental preparations is demonstrated. The development of a portable posture monitoring system is also described, and its accuracy is demonstrated as well. This work shows that the fundamental underlying techniques and algorithms needed to create a comprehensive dental educational system are feasible. Additional work needs to be done to complete the design and implementation of a fully functional dental educational system prototype. The future work is listed as follows:

- The three machine learning techniques that were developed in this research (the DASM algorithm for measuring the drilling of dental preparations, the use of statistical measurements for the classification of depth and texture on tooth surfaces, and the development of the posture monitoring system) were tested and shown to be accurate for the specific applications outlined in this dissertation. Further testing needs to be done for all three methods. Other testing techniques, such as measuring the error in area instead of Euclidean distance with the DASM testing, for example, could be used to further demonstrate the robustness of these techniques. Additional testing for the DASM could be done using images of human teeth instead of the plastic training teeth used in dental schools. In the testing done for this research, there was no variation of color or texture for the tooth surfaces, or the synthetic gum and mouth areas surrounding the teeth. More extensive testing could be done with images of human teeth to test the robustness of the algorithm when variations in color and texture were introduced. Further testing would also be needed for the image processing techniques used to classify depth in color digital images, as outlined in Chapter 4.4.1. In the testing used

for this dissertation, 75% of the images were used for training and 25% were used for testing. Another testing metholodgy, such as 10-fold cross validation (used in this dissertation for the testing of the DASM and the classification of carious regions on tooth surfaces), would be more robust and may give a more comprehensive estimation of the accuracy of the image procesing technique used to classify depth.

- The image processing techniques developed in this research could also be tested with images in other medical applications, to demonstrate the usability of the methods in other applications and further test their robustness.

- The majority of the future work for this project would entail the development of a comprehensive working prototype of the dental educational system. Hardware for a portable system would be required, including a surgical camera that could be mounted on the loupes of a dental student or practitioner, and give the system the "over-the-shoulder" view of the student's work.

- In addition to the integration of hardware, the different machine learning and image processing techniques have to be integrated, and some need expansion. For example, the needs of a real-time system would require that the DASM algorithm be expanded to included registration of images (shown in previous work as an application of ASM algorithms [18]) to follow the location of the tooth from image to image as the camera records the student's work.

- After the different machine learning and image processing techniques were integrated so that they could measure the change in size, shape, location, texture and depth of drilled dental preparations from sequential digital images, a user interface component needs to be designed to provide real-time informative audio feedback to the student. The input from the posture-

monitoring aspect of the system would also be integrated in the user interface component, to give real-time feedback on incorrect body positioning. With the audio-only interface of the portable image-based educational system outlined in this research, the students using the system would be able to work uninterrupted, not having to look away from their work as they received informative audio feedback from the system. This is in comparison to the DentSim system, where the student must look away from their work to get detailed information about the progress of their work. Unlike DentSim, this system would allow students to immediately associate feedback from the system with their work, instead of students relying on feedback from a graphical representation of the tooth on a computer monitor. With this advantage, the students would not develop a reliance on graphical feedback, and potentially develop their self-assessment skills more quickly than students relying on the graphical feedback of existing educational systems, such as DentSim.

- This user interface component would be the final major aspect of the system to be developed. With the audio-only interface of the portable image-based educational system outlined in this research, the students using the system would be able to work uninterrupted, not having to look away from their work as they received informative audio feedback from the system. This is in comparison to the DentSim system, where the student must look away from their work to get detailed information about the progress of their work. Unlike DentSim, this system would allow students to immediately associate feedback from the system with their work, instead of students relying on feedback from a graphical representation of the tooth on a computer monitor. With this advantage, the students would not develop a reliance on graphical feedback, and potentially develop their self-assessment skills more quickly than students relying on the graphical feedback of DentSim.After the prototype is fully developed and integrated,

comprehensive testing could begin, first with dental instructors to provide feedback on the design of the user interface component and the accuracy of the image processing components.

- Testing would then continue with students in the VCU School of Dentistry, and comparison of the effectiveness of the system to the DentSim system and the traditional training of students on drilling dental preparations could be done. The quality of the students' preparations and the rate at which the quality of their preparations improved over the course of a semester could be used as metrics for comparing the educational systems.

A fully realized prototype of the dental educational system described in this research would address issues with existing computer-aided dental educational systems. The machine learning and image processing methods developed and tested in this dissertation could also be used in other medical applications as well, or other image applications where images need to be tracked over time.

Bibliography

Bibliography

[1] S.H. Chung A. Saxena and A.Y. Ng. 3-d depth reconstruction from a single still image. *International Journal of Computer Vision*, 76(1):53–69, 2008.

[2] P. Abolmaesumi and M.R. Sirouspour. An interacting multiple model probabilistic data association filter for cavity boundary extraction from ultrasound images. *IEEE Transactions on Medical Imaging*, 23(6):772–784, 2004.

[3] J. Andriessen and J. Sandberg. Where is education heading and how about ai? *International Journal of Artificial Intelligence in Education*, 10:130–150, 1999.

[4] D.C. Attrill and P.F. Ashley. Occusal caries detection in primary teeth: A comparison of diagnodent with conventional methods. *British Dental Journal*, 190(8):440–443, 2001.

[5] J.D. Bader and D.A. Shugars. A systematic review of the performance of a laser florescence device for detecting caries. *The Journal of the American Dental Association*, 135:1413–1426, 2004.

[6] A. Bartoli, E. von Tunzelmann, and A. Zisserman. Augmenting images of non-rigid scenes using point and curve correspondences. *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR04)*, 2004.

[7] P.E. Benson, N. Pender, and S.M. Higham. Enamel demineralisation assessed by computerised image analysis of clinical photographs. *Journal of Dentistry*, 28:319–326, 2000.

[8] K. Beyer, J. Goldstein, R. Ramakrishnan, and U. Shaft. When is nearest neighbor meaningful. *Proceeding of the 7th International Conference on Database Theory ICDT*, page 217235, 1999.

[9] C.M. Bishop. *Neural Networks for Pattern Recognition*. Oxford University Press, 1995.

[10] B.G. Branson, K.B. Williams, K.K. Bray, S.L. McIlnay, and D. Dickey. Validity and reliability of a dental operator posture assessment instrument (pai). *Journal of Dental Hygiene*, 76(4):255–261, 2002.

[11] E.-M. Brick, H. Rudolph, J. Arnold, and R.G. Luthardt. Analysis of three-dimensional sinter shrinkage of copings made from alumina in an innovative direct shaping process. *Computerized Medical Imaging and Graphics*, 28:159–165, 2004.

[12] M.Z. Brown, D. Burschka, and G.D. Hager. Advances in computational stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(8):993–1008, 2003.

[13] J. Buchanan. Experience with virtual reality-based technology in teaching restorative dental procedures. *Journal of Dental Education*, 86(12):1258–1265, 2004.

[14] N. Casap, A. Wexler, and E. Tarazi. Application of a surgical navigation system for implant surgery in a deficient alveolar ridge postexcision of an odontogenic myxoma. *Journal of Oral and Maxillofacial Surgey*, 63(7):982–988, 2005.

[15] K.J. Cios, W. Pedrycz, R.W. Swiniarski, and L.A. Kurgan. *Data Mining: A Knowledge Discovery Approach*. Springer, 2007.

[16] T. Cootes. *An Introduction to Active Shape Models*, pages 223–248. Oxford University Press, 2000.

[17] T.F. Cootes, C.J. Taylor, D.H. Cooper, and J. Graham. Active shape models their training and application. *Computer Vision and Image Understanding*, 61(1):38–59, 1995.

[18] T.F. Cootes, C.J. Taylor, and Manchester M Pt. Statistical models of appearance for computer vision, 2000.

[19] R. Cucchiara, E. Lamma, and T. Sansoni. An image analysis approach for automatically re-orienteering ct images for dental implants. *Computerized Medical Imaging and Graphics*, 24:185–201, 2004.

[20] A.J. Davison. Real-time simultaneous localization and mapping with a single camera. In *Proceedings of the Ninth IEEE International Conference on Computer Vision*, volume 2, page 1403 1410, 2003.

[21] E. Delage, H. Lee, and A. Ng. Automatic single-image 3d reconstruction of indoor manhattan world scenes. *12th International Symposium of Robotics Research*, 2005.

[22] G. Dong and M. Xie. Color clustering and learning for image segmentation based on neural networks. *IEEE Transactions on Neural Networks*, 16(4):925–936, 2005.

[23] B. du Boulay and R. Luckin. Modeling human teaching tactics and strategies for tutoring systems. *International Journal of Artificial Intelligence in Education*, 12:235–256, 2002.

[24] N. Efeoglua, D. Woodb, and C. Efeoglu. Microcomputerised tomography evaluation of 10% carbamide peroxide applied to enamel. *Journal of Dentistry*, 33:561–567, 2005.

[25] M. Engin, S. Demirag, E. Engin, G. Celbi, F. Ersan, E. Asena, and Z. Colakoglu. The classification of human tremor signals using artificial neural network. *Expert Systems with Applications*, 33:754–761, 2007.

[26] R. Ewers, K. Schicho, M. Truppe, R. Seemann, A. Reichwein, M. Figl, and A. Wagner. Computer-aided navigation in dental implantology: 7 years of clinical experience. *Journal of Oral Maxillofacial Surgery*, 62:329–334, 2004.

[27] J.A. Fails and D. Olsen. Interactive machine learning. In *Proceedings of the 8th International Conference on Intelligent User Interfaces*, pages 39–45, 2003.

[28] L. Finsen, H. Christensen, and M. Bakke. Musculoskeletal disorders among dentists and variation in dental work. *Applied Ergonomics*, 29(2):119–125, 1998.

[29] J. Fogarty, D. Tan, A. Kapoor, and S. Winder. Cueflik: Interactive concept learning in image search. In *Proceeding of the Twenty-Sixth Annual SIGCHI Conference on Human Factors in Computing Systems*, pages 29–38, 2008.

[30] National Institute for Occupational Safety and Health. National occupational research agenda for musculoskeletal disorders, 2007.

[31] K. Forbes-Riley and D.J. Litman. Modeling user satisfaction and student learning in a spoken dialogue tutoring system with generic, tutoring, and user affect parameters. In *Proceedings of the Main Conference on Human Language Technology Conference of the North American Chapter of the Association of Computational Linguistics*, pages 264–271, 2006.

[32] Y. Freund and R.E. Schapire. Experiments with a new boosting algorithm. *Procedures International Conference on Machine Learning*, pages 148–156, 1996.

[33] W. Fu, D. Bothell, S. Douglass, C. Haimson, M. Sohn, and J. Anderson. Toward a real-time model-based training system. *Interactions in Computing*, 18(6):1215–1241, 2006.

[34] A. Gallagher, Y. Matsuoka, and W. Ang. An efficient real-time human posture tracking algorithm using low-cost inertial and magnetic sensors. In *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems*, volume 3, page 2967 2972, 2004.

[35] R.C. Gonzalez, R.E. Woods, and S.L. Eddins. *Digital Image Processing Using MATLAB*. Gatesmark Publishing, 2004.

[36] R. Haak and M.J. Wicht. Gray-scale reversed radiographic display in the detection of approximal caries. *Journal of Dentistry*, 33:65–71, 2005.

[37] N.K. Haden, P.L. Beemsterboer, R.G. Weaver, and R.W. Valachovic. Dental school faculty shortages increase: An update on future dental school faculty. *Journal of Dental Education*, 64(9):657–673, 2000.

[38] G. Hamarneh and T. Gustavsson. Combining snakes and active shape models for segmenting the human left ventricle in echocardiographic images. *Computers in Cardiology*, 27:115–118, 2000.

[39] P. Hammond, T. Hutton, S. Maheswaran, and S. Modgil. Computational models of oral and craniofacial development, growth and repair. In *Dental Informatics and Dental Research: Making the Connection*, 2003.

[40] B. Hartmann, L. Abdulla, M. Mittal, and S. R. Klemmer. Authoring sensor-based interactions by demonstration with direct manipulation and pattern recognition. In *CHI Proceedings*, pages 145–154, 2007.

[41] K.B. Hilger, R. Larsen, S. Kreiborg, S. Krarup, T.A. Darvann, and J.L. Marsh. Active shape analysis of mandibular growth. In *Medical Image Computing and Computer-Assisted Intervention MICCAI*, pages 902–909, 2003.

[42] T.J. Hutton, P. Hammond, and J.C. Davenport. Active shape models for customised prosthesis design. *AIMDM*, pages 448–452, 1999.

[43] R.L. Williams II, J.N. Nowell, and D.C. Eland. The virtual haptic back for palpatory training. In *Proceedings of the 6th International Conference on Multimodal Interfaces*, pages 191–197, 2004.

[44] A.I. Ismail. Visual and visuo-tactile detection of dental caries. *Journal of Dental Restoration*, 83:C56–C66, 2004.

[45] T. Iso and K. Yamazaki. Gait analyzer based on a cell phone with a single three-axis accelerometer. In *Proceedings of the 8th conference on Human-computer interaction with mobile devices and services*, volume 159, pages 141–144, 2006.

[46] A. Jablonski-Momeni, V. Stachniss, D.N. Ricketts, M. Heinzel-Gutenbrunner, and K. Pieper. Reproducibility and accuracy of the icdas-ii for detection of occlusal caries in vitro. *Caries Research*, 42:79–87, 2008.

[47] T.R. Jasinevicius, M. Landers, S. Nelson, and A. Urbankova. An evaluation of two dental simulation systems: Virtual reality versus comtemporary-non-computer assisted. *Journal of Dental Education*, 68(11):1151–1162, 2004.

[48] L. Jerinic and J. Devedzic. The friendly intelligent tutoring environment. *ACM SIGCHI Bulletin*, 32(1):83–94, 2000.

[49] K.H. Jones. A comparison of algorithms used to compute hill slope as a property of dem. *Computers and Geoscience*, 24(4):513–323, 1998.

[50] S.S. Keerthi, S.K. Shevade, C. Bhattacharyya, and K.R.K. Murthy. Improvements to platt's smo algorithm for svm classifier design. Technical report, 1998.

[51] J. Khnisch, R. Heinrich-Weltzien, M. Tabatabaie, L. Stsser, and M.C.D.N.J.M. Huysmans. An in vitro comparison between two methods of electrical resistance measurement for occlusal caries detection. *Caries Research*, 40:104–111, 2006.

[52] T. Kohonen, J. Hynninen, J. Kangas, J. Laaksonen, and K. Torkkola. Lvq pak: the learning vector quantisation program package – version 3.1, 1995.

[53] I. Koprinska and S. Carrato. Temporal video segmentation: A survey. *Signal Processing: Image Communication*, 16(5):477–500, 2001.

[54] M.A. Lackey. One year's experience with virtual reality preclinical laboratory simulation at the university of tennessee. *International Journal of Computerized Dentistry*, 7:131–141, 2004.

[55] V. Le Blanc, A. Urbankova, F. Hadavi, and R. Lichtenthal. A preliminary study in using virtual reality to train dental student. *Journal of Dental Education*, 68(3):378–383, 2003.

[56] T.M. Lehmann and J. Brendo. Strategies to configure image analysis algorithms for clinical usage. *Journal of American Medical Informatics Association*, 12(5):497–504, 2005.

[57] H. Li and O. Chutatape. Automated feature extraction in color retinal images by a model based approach. *IEEE Transactions on Biomedical Engineering*, 51(2):246–254, 2004.

[58] S. Li, T. Fevens, A. Krzyzak, and S. Li. An automatic variational level set segmentation framework for computer aided dental x-rays analysis in clinical environments. *Computerized Medical Imaging and Graphics*, 30:65–74, 2006.

[59] S.-H. Liao, R.-F. Tong, and J.-X. Dong. Influence of anisotropy on peri-implant stress and strain in complete mandible model from ct. *Computerized Medical Imaging and Graphics*, 32:5360, 2008.

[60] C.P. Loizou, C.S. Pattichis, M. Pantziaris, T. Tyllis, and A. Nicolaides. Snakes based segmentation of the common carotid artery intima media. *Medical and Biological Engineering and Computing*, 45(1):35–49, 2007.

[61] L. Lucchese and S.K. Mitra. Color image segmentation: A state-of-the-art survey. *Proceedings of the Indian National Science Academy (INSA-A)*, 67-A:207221, 2001.

[62] H.J. Luinge and P.H. Veltink. Inclination measurement of human movement using a 3-d accelerometer with autocalibration. *IEEE Transactions on Biomedical Engineering*, 53:1385–1393, 2006.

[63] A. Madabhushi and D.N. Metaxas. Combining low-, high-level and empirical domain knowledge for automated segmentation of ultrasonic breast lesions. *IEEE Transactions on Medical Imaging*, 22(2):155–169, 2003.

[64] J.B.A. Maintz and M.A. Viergever. A survey of medical image registration. *Medical Image Analysis*, 12(1):1–36, 1998.

[65] M.J. Mathie, J. Basilakis, and B.G. Celler. A system for monitoring posture and physical activity using accelerometers. In *Proceedings of the 23rd Annual EMBS International Conference*, pages 3654–3657, 2001.

[66] R.E. Mayagoitia, A.V. Nene, and P.E. Veltink. Accelerometer and rate gyroscope measurement of kinematics: an inexpensive alternative to optical motion analysis systems. *Journal of Biomechanics*, 35:537–542, 2002.

[67] D. Maynes-Aminzade and T. Igarashi. Eyepatch: Prototyping camera-based interaction through examples. In *Proceedings of the 20th annual ACM Symposium on User Interface Software and Technology*, pages 33–42, 2007.

[68] L. McMillian and G. Bishop. Plenoptic modeling: An image-based rendering system. In *Proceedings of the 22nd Annual Conference on Computer Graphics and Interactive Techniques*, page 39 46, 1995.

[69] M.Dougherty. Ergonomic principles in the dental setting: Part 1. Technical report, 2001.

[70] R.J. Murphy, S.A. Gray, S.R. Staja, and M.C. Bogert. Student learning preferences and teaching implications. *Journal of Dental Education*, 86(8):859–866, 2004.

[71] T.C. Murphy, D.R. Willmot, and H.D. Rodd. Management of postorthodontic demineralized white lesions with microabrasion: A quantitative assessment. *American Journal of Orthodontics and Dentofacial Orthopedics*, 131:27–33, 2007.

[72] K. Najarian and R. Splinter. *Biomedical Signal and Image Processing.* CRC Press, 2008.

[73] R. Nevins, N. Durdle, and V. Raso. A posture monitoring system using accelerometers. In *Proceedings of the 2002 IEEE Canadian Conference on Electrical and Computer Engineering*, pages 1087–1092, 2002.

[74] T. Otto and S. De Nisco. Computer-aided direct ceramic restorations: A 10-year prospective clinical study of cerec cad/cam inlays and onlays. *International Journal of Prosthodontics*, 15(2):122–128, 2002.

[75] J.E.A. Palamara, P.R. Wilson, C.D.L. Thomas, and H.H. Messer. A new imaging technique for measuring the surface strains applied to dentine. *Journal of Dentistry*, 28:141–146, 2000.

[76] A. Plamondon, A. Delise, C. Larue, D. Brouillette, D. McFadden, P. Desjardins, and C. Lariviere. Evaluation of a hybrid system for three-dimensional measurement of trunk posture in motion. *Applied Ergonomics*, 38:697–712, 2007.

[77] J. Platt. *Fast Training of Support Vector Machines using Sequential Minimal Optimization. Advances in Kernel Methods: Support Vector Learning.* MIT Press, 1998.

[78] M. Pollefeys, R. Koch, M. Vergauwen, and L. Van Gool. Automated reconstruction of 3d scenes from sequences of images. *ISPRS Journal of Photogrammerty and Remote Sensing*, 55(4):251–267, 2000.

[79] I.A. Pretty and G. Maupome. A closer look at diagnosis in clinical dental practice: Part 5. emerging technologies for caries detection and diagnosis. *Journal of the Canadian Dental Association*, 70(8):540a–540i, 2004.

[80] J. R. Quinlan. Improved use of continuous attributes in c4.5. *Journal of Artificial Intelligence Research*, 4:77–90, 1996.

[81] M. Pollefeys R. Koch and L.V. Gool. Realistic surface reconstruction of 3d scenes from uncalibrated image sequences. *The Journal of Visualization and Computer Animation*, 11:115–127, 2000.

[82] D. Ricketts. The eyes have it: How good is diagnodent at at detecting caries. *Evidence Based Dentistry*, 6(5):64–65, 2005.

[83] B.D. Ripley. *Pattern Recognition and Neural Networks*. Cambridge University Press, 1996.

[84] H. Rosenberg, H. Grad, and D. Matear. The effectiveness of computer-aided, self-instructional programs in dental education: A systematic review of the literature. *Journal of Dental Education*, 67(5):524–532, 2003.

[85] V. Sequeira, K. Ng, E. Wolfart, J.G.M. Gonalves, and D. Hogg. Automated reconstruction of 3d models from real environments. *ISPRS Journal of Photogrammerty and Remote Sensing*, 54(1):1–22, 1999.

[86] M. Simmer-Beck and B. Branson. Posture perfect. *Dimensions of Dental Hygiene*, 3(14), 2005.

[87] A.K. Skidmore. Evolution of methods for estimating slope gradient and aspect from digital elevation models. In *Classics from IJGIS: Twenty Years of the International Journal of Geographical Information Science and Systems*, pages 111–116. CRC Press, 2006.

[88] P. Somervuo and T. Kohonen. Self-organizing maps and learning vector quantization for feature sequences. *Neural Processes*, 10:151–159, 1999.

[89] M. Sonka, V. Hlavac, and R. Boyle. *Image Processing, Analysis and Machine Vision*. Thompson Learning, 2008.

[90] I. Stamos and P.K. Allen. Geometry and texture recovery of scenes of large scale. *Computer Vision and Image Understanding*, 88:94–118, 2002.

[91] S. Stumpf, E. Sullivan, E. Fitzhenry, I. Oberst, W.-K. Wong, and M. Burnett. Integrating rich user feedback into intelligent user interfaces. In *Proceedings of the 13th international Conference on Intelligent User Interfaces*, pages 50–59, 2008.

[92] C.M. Sturdevant, T.M. Roberson, H.O. Heymann, and J.R. Sturdevant. *The Art and Science of Operative Dentistry*. Mosby-Year Book Inc., third edition, 1995.

[93] Y. Sun, M.S. Kamal, A.K.C. Wong, and Y. Wang. Cost-sensitive boosting for classification of imbalanced data. *Pattern Recognition*, 40:3358–3378, 2007.

[94] L.J. Thornton, C. Stuart-Buttle, T.C. Wyszynski, and E.R. Wilson. Physical and psychosocial stress exposures in us dental schools: the need for expanded ergonomics training. *Applied Ergonomics*, 35(2):153–157, 2003.

[95] C. Tobon-Gomez, C. Butakoff, S. Ordas, S. Aguade, and A.F. Frangi. Comparative study of diverse model building strategies for 3d-asm segmentation of dynamic gated spect data. In *Proc. SPIE Medical Imaging 2007: Physiology, function, and structure from medical images*, 2007.

[96] A. Torralba and A. Oliva. Depth estimation from image structure. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(9):1–13, 2002.

[97] B. Valachi and K. Valachi. Mechanisms leading to musculoskeletal disorders in dentistry. *Journal of the American Dental Association*, 135(10):1344–1350, 2003.

[98] B. Valachi and K. Valachi. Preventing musculoskeletal disorders in clinical dentistry. *Journal of the American Dental Association*, 134:1604–1612, 2003.

[99] E. van den Broek. Empathic agent technology. In *Agent-based Systems for Human Learning*. AAMAS Workshop, 2005.

[100] B. van Ginneken, A.F. Frangi, J.J. Staal, B.M. ter Haar Romeny, and M.A. Viergever. Active shape model segmentation with optimal features. *IEEE Transactions on Medical Imaging*, 21(8):924–933, 2002.

[101] S.D. Warren, M.G. Hohmann, K. Auerswald, and H. Mitasova. An evaluation of methods to determine slope using digital elevation data. *CATENA*, 58(3):215–233, 2004.

[102] G. Widmann, R. Widmann, E. Widmann, W. Jaschke, and R.J. Bale. A new concept of image guided template production reducing costs and efforts. *International Congress Series*, 1268:1209–1214, 2004.

[103] I.H. Witten and E. Frank. *Data Mining: Practical Machine Learning Tools and Techniques*. Elsevier, 2005.

[104] W.Y.Wong, M.S. Wong, and K.H. Lo. Clinical applications of sensors for human posture and movement analysis. *Prosthetics and Orthotics International*, 31:62–75, 2007.

[105] A.F. Zandona and T. Zero. Diagnostic tools for early caries detection. *The Journal of the American Dental Association*, 137:1675–1684, 2006.

[106] J. Zhou, L. Cheng, and W.F. Bischof. Online learning with novelty detection in human-guided road tracking. *IEEE Transactions on Geoscience and Remote Sensing*, 45(12):3967–3977, 2007.

[107] R. Zhu and Zhou Z. A real-time articulated human motion tracking using tri-axis inertial/magnetic sensors package. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 12:295–302, 2004.

[108] B. Zitova and J. Flusser. Image registration methods: A survey. *Image and Vision Computing*, 21:977–1000, 2003.

[109] A. Zou, Z.G. Hou, and M. Tan. Support vector machines (svm) for color image segmentation with applications to mobile robot localization problems. *Advances in Intelligent Computing*, 3645:442–452, 2005.

VITA

Grace Ferguson Olsen was born on January 10, 1980, in Richmond, Virginia, and is an American citizen. She received her Bachelor of Science in Biochemistry from Mary Baldwin College, Staunton, Virginia in 1998. She received a Master of Science in Computer Science from Virginia Commonwealth University in 2003.