

OBJECTIVE ASSESSMENT AND FEEDBACK GENERATION IN DENTAL  
SURGICAL SIMULATION: A FRAMEWORK BASED ON CORRELATING  
PROCEDURE AND OUTCOME

von Myat Su Yin

Dissertation  
zur Erlangung des Grades eines Doktors der  
Ingenieurwissenschaften  
- Dr.-Ing. -

Vorgelegt im Fachbereich 3 (Mathematik & Informatik)  
der Universität Bremen  
im Februar 2018

Datum des Promotionskolloquiums: 7. Februar 2018

Gutachter: Prof. Dr. Christian Freksa (Universität Bremen)  
Prof. Dr. Peter Haddawy (Mahidol University)

## ACKNOWLEDGEMENTS

I would like to express my deep gratitude to my advisor Prof. Dr. Peter Haddawy for his inspiring and invaluable guidance, infinite advice, motivation, and continuous support throughout my Ph.D. journey. I owe my gratitude and thanks to Prof. Dr. Christian Freksa for his advice and guidance in developing this thesis. I would like to express my appreciation to both advisors for providing me for the opportunity to join the dual degree program and for their time and endless efforts from my admission until graduation from the program. I would like to acknowledge with gratitude to Prof.Dr. Siriwan Suebnukarn for her expert's advice and extraordinary support in this thesis process.

I wish to express my appreciation to thesis committee members, Assoc.Prof Matt Dailey, Asst. Prof. Worapan Kusakunniran and Dr.Mores Prachyabrued for the valuable suggestions to improve my research work. I would like to thank Prof.Dr.Gabriel Zachmann, Prof.Dr.Rainer Malaka, Prof.Dr.Johannes Schoening for joining my defense presentation and for very constructive discussions. I would like to thank Dr.Pattanaphon Rhienmora for the wonderful collaboration.

This thesis would have been impossible without the scholarship and the financial aids from the Faculty of ICT. I would like to express my sincere gratitude to Prof.Dr. Supachai Tangwongsan, Assoc. Prof. Dr. Jarernsri L. Mitranont and Asst. Prof. Dr. Boonsit Yimwadsana for giving me with the opportunity to study in the program and their supports throughout my journey. My special thanks are extended to all the faculty members and the staffs from the Faculty of ICT. I would like to express my special thanks of gratitude to Dr.Songsri Tangsripairoj and Dr. Benjarath Pupacdi for their encouragement, support and advice on my early works. I also would like to thank the faculty members, the staffs and the students from the Faculty of Dentistry, Thammasat University for their assistance in experiments. In addition, I have to thank the undergraduate students group who were involved in the development of portion of feedback system in this thesis.

I would like to offer my special thanks to Holger Schultheis for his valuable advice and discussions for my research. I am extending my heartfelt thanks to all the members of Bremen Spatial Cognition Center and Cognitive Systems (CoSy) Lab for inspiring me to pursue a career in research. I am also grateful to Gracia Kanz, the staffs members from FB3 and International Office from the University of Bremen for their administrative supports. I would like to offer my special thanks to the Schroder and Fischer families for a home away from home while I was developing this thesis. Most especially, I must express my very profound gratitude to my parents, my sisters, and my brother for providing me with unfailing support, unconditional love and extraordinary care throughout my study. I would like to thank my husband for his continuous support and encouragement, and for standing by me. Also thank to my fantastic seven for brightening my days.

Myat Su Yin

OBJECTIVE ASSESSMENT AND FEEDBACK GENERATION IN DENTAL SURGICAL SIMULATION: A FRAMEWORK BASED ON CORRELATING PROCEDURE AND OUTCOME

MYAT SU YIN 5338194 ITCS/D

Ph.D. (COMPUTER SCIENCE)

THESIS ADVISORY COMMITTEE: PETER F. HADDAWY, Ph.D., CHRISTIAN FREKSA, Ph.D., SIRIWAN SUEBNUKARN, Ph.D., WORAPAN KUSAKUNNIRAN, Ph.D., MORES PRACHYABRUED, Ph.D.

ABSTRACT

Fine motor skill is indispensable for a dentist. As in many other medical fields of study, the traditional surgical master apprentice model is widely adopted in dental education. Recently, virtual reality (VR) simulators have been employed as supplementary components to the traditional skill-training curriculum, and numerous dental VR systems have been developed academically and commercially. However, the full promise of such systems has yet to be realized due to the lack of sufficient support for formative feedback. Without such a mechanism, evaluation still demands dedicated time of experts in scarce supply. With the aim to fill the gap of formative assessment using VR simulators in skill training in dentistry, this thesis presents a framework to objectively assess the surgical skill and generate formative feedback automatically. VR simulators enable collecting detailed data on relevant metrics throughout a procedure. Our approach to formative feedback is to correlate procedure metrics with the procedure outcome in order to identify the portions of a procedure that need to be improved. Prior to the correlation, the procedure outcome needs to be evaluated. The scoring algorithm designed in this thesis provides an overall score and identifies specific errors and their severity. Building upon this, we developed techniques to identify the portion of the procedure responsible for the errors. Specifically, for the errors in the outcome the responsible portions of the procedure are identified based on correlation of location of the error. For some types of feedback one mode may be more suitable than another. Tutoring formative feedback are provided using the video- and haptic- modalities. The effectiveness of the feedback systems have been evaluated with the dental students with randomized controlled trials and the findings show the feedback mechanisms to be effective and have potentials to use as valuable supplemental training resources.

KEY WORDS: SURGICAL SIMULATION / FORMATIVE FEEDBACK / AUTOMATED OUTCOME ASSESSMENT / FORCE FEEDBACK / VIDEO-BASED FEEDBACK

129 pages

# CONTENTS

	<b>Page</b>
<b>ACKNOWLEDGEMENTS</b> .....	<b>iii</b>
<b>ABSTRACT</b> .....	<b>iv</b>
<b>LIST OF TABLES</b> .....	<b>viii</b>
<b>LIST OF FIGURES</b> .....	<b>ix</b>
<b>CHAPTER I INTRODUCTION</b> .....	<b>1</b>
1.1 Scope of the Thesis . . . . .	3
1.2 Contributions . . . . .	4
<b>CHAPTER II BACKGROUND AND RELATED WORK</b> .....	<b>7</b>
2.1 Dental Surgical Skill . . . . .	7
2.2 Assessment and Feedback in Skill Learning Theories . . . . .	8
2.3 Dental Surgical Skill Training and Assessment . . . . .	10
2.4 Simulation based Skill Training and Assessment . . . . .	12
2.4.1 Assessment of Skill using Procedure Analysis . . . . .	15
2.4.2 Assessment of Skill using Outcome Analysis . . . . .	16
2.4.3 Assessment of Skill using Both Procedure and Outcome Analysis	18
2.5 Feedback . . . . .	19
2.5.1 Feedback in Motor Skill Training . . . . .	21
2.5.2 Structure and Timing of Feedback . . . . .	23
2.6 Root Canal Treatment . . . . .	26
<b>CHAPTER III OBJECTIVE ASSESSMENT AND FEEDBACK</b>	
<b>GENERATION FRAMEWORK</b> .....	<b>29</b>
3.1 Framework Overview . . . . .	29
3.1.1 Automated Outcome Scoring System . . . . .	31
3.1.2 Correlating Procedure and Outcome . . . . .	32
3.1.3 Feedback System . . . . .	34
3.2 Evaluation . . . . .	35

## CONTENTS (cont.)

	Page
<b>CHAPTER IV AUTOMATED OUTCOME SCORING SYSTEM .....</b>	<b>39</b>
4.1 Outcome Scoring System . . . . .	40
4.2 Experimental Evaluation . . . . .	45
4.3 Results . . . . .	47
4.3.1 Bland-Altman plots . . . . .	47
4.3.2 Information-based disagreement . . . . .	49
4.3.3 Mean Absolute Errors . . . . .	51
4.3.4 Perceptual effects . . . . .	52
4.3.5 Discriminatory Power of the AOS Scores . . . . .	53
4.4 Discussion . . . . .	53
<b>CHAPTER V CORRELATOR: CORRELATING PROCEDURE AND</b>	
<b>OUTCOME .....</b>	<b>57</b>
5.1 Error Regions in the Outcome . . . . .	59
5.2 Types of Errors and Causes . . . . .	60
5.3 Procedure and Outcome Correlation . . . . .	63
<b>CHAPTER VI VIDEO-BASED FORMATIVE FEEDBACK SYSTEM .....</b>	<b>69</b>
6.1 Video-based Formative Feedback System . . . . .	70
6.1.1 User Interface Design . . . . .	70
6.1.2 Video control panel . . . . .	71
6.1.3 Simulation panel . . . . .	73
6.1.4 Mode control panel . . . . .	74
6.1.5 Viewing aspect control panel . . . . .	75
6.2 Experimental Evaluation . . . . .	77
6.3 Results . . . . .	82
6.4 Discussion . . . . .	86
<b>CHAPTER VII FORCE FEEDBACK .....</b>	<b>89</b>
7.1 Sensory Motor Integration . . . . .	93

## CONTENTS (cont.)

	<b>Page</b>
7.2 Haptic Formative Feedback . . . . .	95
7.3 Experimental Evaluation . . . . .	96
7.4 Results . . . . .	98
7.5 Discussion . . . . .	99
<b>CHAPTER VIII CONCLUSIONS AND FUTURE RESEARCH . . . . .</b>	<b>101</b>
8.1 Automated Outcome Scoring System . . . . .	101
8.2 Video-based Formative Feedback . . . . .	103
8.3 Haptic-based Feedback . . . . .	104
8.4 Correlator and Formative Feedback Framework . . . . .	105
8.5 Conclusion . . . . .	108
<b>REFERENCES . . . . .</b>	<b>110</b>

## LIST OF TABLES

<b>Table</b>	<b>Page</b>
2.1 Examples VR dental simulators . . . . .	12
2.1 Examples VR dental simulators (cont.) . . . . .	13
2.2 Metrics used for objective assessment of dental surgical skill . . . . .	15
4.1 Mean and standard deviation values for the scores of AOS and human raters in all experiments . . . . .	47
4.2 Information-based disagreement between AOS and human raters with confidence intervals . . . . .	49
4.3 Information-based disagreement between human raters R2 vs. R3, R4, R5 .	50
4.4 Information-based disagreement between human raters R3 vs. R4, R5, and R4 vs. R5 . . . . .	50
4.5 Mean Absolute Error of pairs of AOS and human rater scores . . . . .	51
4.6 Chi-square values from Kruskal-Wallis tests for R2, R3, R4, R5 (Significant results are indicated with the bold-faced fonts . . . . .	52
4.7 ROC AUC values for AOS scores over all grades in all experiments . . . .	54
6.1 Intra-class correlation coefficient (ICC) between two evaluators (95% confidence interval) . . . . .	82
6.2 Descriptive Statistics of Scores . . . . .	83
7.1 Pre-training and post-training outcome scores between control and experimental groups (mean $\pm$ SD) . . . . .	98



## LIST OF FIGURES

<b>Figure</b>	<b>Page</b>
2.1 A saggittal cross-section of a tooth with an infected pulp (Source:www.medindia.net) . . . . .	26
2.2 Tooth surfaces . . . . .	27
3.1 Objective assessment and formative feedback framework . . . . .	30
3.2 Framework components . . . . .	31
3.3 Automated outcome scoring . . . . .	32
3.4 Correlator overview . . . . .	33
3.5 Feedback system overview . . . . .	34
3.6 Simulator interface . . . . .	36
3.7 Evaluation plan . . . . .	37
4.1 Virtual optimal template creation(a) pulp chamber location (b) projection from pulp chamber (c) removal of tooth mass for projected area . . . . .	41
4.2 Portion of score cube between surface and Max template . . . . .	41
4.3 Score cube z-axis cross-sectional view . . . . .	42
4.4 Non-linear weight function . . . . .	43
4.5 Tooth walls . . . . .	44
4.6 Segmentation of tooth into four axial walls (a) Original tooth walls (b) Adjusted tooth walls . . . . .	44
4.7 Wall score computation . . . . .	45
4.8 Example views the experiments . . . . .	46
4.9 Min, Max and Optimal templates overlaid on student’s drilling area from the lingual wall and occlusal view . . . . .	46
4.10 Bland-Altman plots of scores between AOS and human raters in three experiments . . . . .	48
4.11 Mean Absolute Error between AOS and human raters . . . . .	51
5.1 The outcome scores and the error information for an outcome . . . . .	59

## LIST OF FIGURES (cont.)

<b>Figure</b>	<b>Page</b>
5.2 Operative errors . . . . .	60
5.3 Overcut example . . . . .	61
5.4 Undercut example . . . . .	62
5.5 Perforation example . . . . .	63
5.7 Outcome scores, errors and temporal information in procedure . . . . .	64
5.6 Procedure and outcome correlation flow . . . . .	65
5.8 Mapping between error information and collided voxels log . . . . .	66
5.9 Portion of the analysis of wall, error occurrence in time (Note: Occurrences of overcut errors are marked in red color.) . . . . .	67
5.10 Pseudo code of the correlator component . . . . .	68
6.1 Video playback user interface . . . . .	71
6.2 Video control panel . . . . .	71
6.3 Example view of video progress with different types of error . . . . .	72
6.4 Stage Borders on the video progress bar . . . . .	72
6.5 Example undercut error region on the tooth in the Simulation Panel . . . . .	73
6.6 Example kinematic graphs (top: Orientation, below: Force) . . . . .	74
6.7 Student and expert windows in the Replay mode . . . . .	75
6.8 Viewing aspects control panel . . . . .	75
6.9 The Buccal wall view . . . . .	76
6.10 Magnified view of the simulated tooth . . . . .	76
6.11 Flowchart of participants through trials . . . . .	80
6.12 Artificial teeth (a) Before preparation (b) Mounted on the artificial jaw (c) During preparation (d) After preparation . . . . .	81
6.13 Plot of the change [(post-training overall score) (pre-training overall score)] vs. pre-training overall score. . . . .	85
7.1 Opposite force training . . . . .	96
7.2 Comparison of force used in mesial wall between expert and novice in stage 3	97

## LIST OF FIGURES (cont.)

<b>Figure</b>	<b>Page</b>
8.1 Five aspects in fillet weld [208] . . . . .	103

## **CHAPTER I**

### **INTRODUCTION**

Precise psychomotor skills are essential for dentists in performing everyday tasks. Dental schools have allocated a significant portion of curriculum to skill training. Similar to the other disciplines of medicine, the mainstream approach to surgical skill training in dentistry is in the form of master-apprentice model along with practice and feedback. Trainees perform procedures under direct observation of an experienced surgeon who can give feedback and assess their psychomotor skills. The assessor usually identifies errors made by the student during the procedure and provides formative feedback on the outcome. A significant amount of time and resources are required in this classic approach to skill training. Usually, students use either the extracted teeth from patients or the plastic teeth in skill training laboratories. Due to anatomical variations in human teeth, the use of extracted teeth yields unstandardized training content which often leads to the different knowledge and skills acquired by the students. In contrast, artificial teeth facilitate the standardized training experiences, however, students have limited or no opportunities to train with rare cases as there is no pathological variation in plastic teeth. Regarding assessment, the assessor usually observes the students work provide the score and the verbal feedback on their preparation. With an increasing number of enrollment in dental schools and a large student to faculty ratio, students are not receiving the feedback as much supervised training as would be desirable.

In response to these limitations, virtual reality (VR) simulators have been recently introduced into skill training and assessment areas of surgical curricula. As virtual procedures provide standardized training contents, VR simulators have been used in adjunct with extracted and artificial teeth benefiting to the students as they can be trained to learn from errors without consequences. VR simulators can collect all the kinematics variables while the users perform the procedure. A plethora of work has been established the objective assessment using metrics obtained from the simulators.

Existing studies mainly address the validity and reliability of simulators and metrics for assessment while paying less attention to the formative aspect of the assessment. With the summative assessment feedback currently available in most of the dental VR simulators, the skills training using VR simulators still demands the presence of instructors during the training and the trainees are still receiving the quality and quantity of feedback before the simulators are included in preclinical training laboratories.

The formative assessment identifies the gap between the target and student performance and provides feedback on where, when, and how elements of students' errors which can help the trainee to rectify or modify their performance on the next trial and increase the likelihood of achieving the desired outcome. Beyond the simple analysis of kinematics, the procedure analysis on VR simulators can be extended by incorporating the information from outcome analysis. The combined procedure and outcome analysis could be instrumental in generating the formative feedback for skill training. VR simulators are considered to be the perfect platform to investigate the formative assessment feedback as they facilitate the analysis of both procedure and outcome.

For a virtual procedure performed on a simulator, the resulting outcome is available at the end of a training task for further analysis. This outcome needs to be evaluated with an appropriate outcome assessment system to provide formative feedback. Ideal procedure outcome from experts using VR simulator is easily accessible, yet the objective assessment of outcome is not a trivial task for 3D voxel structure commonly used in VR dental simulators. The challenge lies in the assessment at a sufficient level of detail to allow for the identification of the performance gap in irregular 3D anatomy objects given the procedure specific evaluation requirements. Specifically, in the identification of the cause of the errors to determine which lead to the performance score. The content of the formative feedback could be generated from the combined analysis of procedure and outcome, other elements in the provision of feedback such as choosing the appropriate modality and the right timing to provide the feedback pose as challenges.

With the aim to fill the gap of formative assessment using VR simulators, we present a framework to objectively assess the surgical skill and generate feedback automatically using a VR simulator. To identify and localize error in a students perfor-

mance, we first obtain the flaws in the outcome by assessing outcome. In the subsequent step, we identify the portions of the procedure which are considered to be responsible for each fault and correlate them as the sources of errors. Finally, formative feedback is generated using the resulting correlation information on the error and the identified sources. The tutoring feedback generated enables students to learn to associate their actions with the resulting performance which is of importance in skill acquisition.

The framework addresses the problem of objective performance assessment and feedback generation in skill training using VR simulators. Relying solely on analyses performed on the data obtained from the simulator, the automated assessment framework achieves an objective assessment of surgical skill that is free from subjectivity. Objective formative assessment feedback is generated by establishing the correlation between the way the procedure is performed and the resulting outcome.

## **1.1 Scope of the Thesis**

This thesis focuses on the objective assessment of skill and automatic feedback generation for surgical skill training using a VR simulator. While we are interested in developing general techniques, a specific domain is required to formulate and validate the problem in a systematic way. We choose the area of surgical skill training using a dental VR skill simulator. The existing VR simulator is employed in this thesis to focus our attention mainly on the objective assessment and feedback and to a lesser extent on other technical and nontechnical aspects of simulations.

The dental simulator is selected because the evaluation scheme for the simulation outcome can be set up in a straightforward manner in accurately deformed hard-tissue objects in the VR simulator. The focus is further narrowed to the access opening procedure of the root canal treatment procedure in endodontics.

The root canal endodontic treatment is necessary when the dental pulp is infected. In the access opening stage, a small hole is drilled out from the tooth surface to get access to the pulp chamber, the root canals, and the roots. The access opening is considered as the most important stage in the root canal treatment as the successful completion of the treatment relies on the shape and the size of the drilled area in this

stage. Accurate shape and size of outcome demand the precise fine motor skills of the dentists. While the frameworks design is based on a dental surgical procedure, many of the techniques implemented should generalize well to other surgical areas.

Firstly, the components of the objective assessment and feedback framework are identified along with their input/output and define the structure of the framework (Chapter 3). Using a VR dental surgical skill training simulator, and the simulated procedure, an automated outcome scoring system is implemented. The outcome scores from the system are validated in comparison with the human experts (Chapter 4). The error information from outcome scoring system is combined with the procedural analysis information, to associate the errors in the outcome with the procedure portions that are responsible for them. The formative feedback output from the correlator component is provided to the student in the form of video-playback (Chapter 5). The effectiveness of video-based formative feedback is evaluated in comparison with the traditional training approach without using the simulator and training using the simulator without feedback. The extended use of correlation information in the formative feedback is demonstrated with the use of haptic feedback as a means to train the correct application of force in a dental procedure (Chapter 7)

## **1.2 Contributions**

The thesis contributes in a number of ways to the body of knowledge of the field. Effective formative feedback requires the ability to correlate procedure metrics with the procedure outcome in order to identify the portions of a procedure that need to be improved. The automated objective assessment and feedback generation framework in this thesis is based on correlating procedure and outcome. Indeed, determining the correlation between procedure and outcome has been identified as one of the important outstanding problems in surgical simulation [1,2]. We designed and developed a general purpose outcome scoring technique for dental surgery and implemented a prototype for scoring of outcomes in access opening for root canal procedure. This is the first scoring algorithm for endodontic surgery and can more precisely identify errors in the outcome than other commonly used techniques for outcome assessment in dental simulators. The

scoring algorithm provides an overall score and identifies specific errors and their severity. Building upon this, we identified the portions of the procedure responsible for the errors. Responsible portions of the procedure are identified based on correlation of location of the error and the procedure kinematics variables. This problem is an instance of the well-known credit assignment problem [3]. We have contributed a tractable solution to the credit-assignment problem provided that the immediate effect of the actions are available. Once the tooth is drilled, the effects of drilling actions are available immediately in the outcome. If the drilling actions lead to an error in the outcome, both spatial and temporal error information can be localized in a straightforward manner. Using the localized error information, we correlate the portion of procedure and error in the outcome, and the integrated information is translated into the precise and objective language using dental terminology and notation with which the trainee dentists are familiar as formative feedback.

The video-based formative feedback system presented in this study visualizes the errors in the outcome and provides the trainee with an opportunity to review their performance to comprehend when and what aspects of the errors in the outcome. To the best of our knowledge, this is the first use of replay with visualization techniques to provide feedback in dental surgery. The video playback provides a large amount of information to the user in an easily intelligible form. We show the effectiveness of formative feedback using video-modality in motor skill training.

Additionally, we have demonstrated the extended use of formative feedback to train correct application of force in endodontic surgery with haptic feedback modality. The force-feedback represents the first time that haptics have been used to teach correct application of force in dental surgery. The haptic training is one of the current research issues in human computer interaction studies involving computer-based simulators and haptic devices [4–8]. A commonly used approach to haptic training consists of two phases. In the first phase, the expert's movement is recorded in terms of position, velocities, and force patterns. In the subsequent phase, the recorded movements are haptically and visually displayed to learners during the playback mode training. Playback can be either passive or active mode. In the passive playback mode, the trainees have to grasp the end-effector of the haptic device and are physically guided through the ideal motion



through a desired trajectory to acquire the kinesthetic understanding of what is required. In contrast, the trainee moves the end-effector through a desired trajectory at his/her speed in the active playback mode. Our approach deviates from the traditional haptic training approaches in utilizing the correlated information (formative feedback) from the correlator component. For the procedure stage containing portions of the procedure which are labeled as the portion responsible for a certain error region in the outcome, trainee has to undergo the haptic training. Using the expert's force in the identified procedure stage, the expert's force is rendered to the student via a haptic device and the student has to cancel it with the opposite force. Training on the proper use of force is more meaningful with the sense of touch (haptic cues) involved in the countering force. The initial evaluation with dental students revealed significant changes in the applied force in post training performance.

Evaluation of both feedback components show that the methods significantly improve learning outcomes. While our work is implemented using a VR dental surgical simulator, many of the components generalize beyond the domain of dental surgery, making the work of importance to many problems that involve training of complex psychomotor skills.

## **CHAPTER II**

### **BACKGROUND AND RELATED WORK**

This chapter includes background on relevant concepts from dental surgery, the existing approaches to skill training and assessment in medical and dental education with or without using computer-based simulators.

#### **2.1 Dental Surgical Skill**

Surgery is the branch of medicine that is concerned with diagnosis and treatment of injuries or diseases using operative procedures. Surgical competency consists of a combination of cognitive capacity, technical skill and non-technical skills such as communication, decision-making and leadership skills. The term, technical skill refers to the manual/physical skill that requires the individuals capacities in visual and haptic perception, fine-grained physical and temporal movement, and manipulation of instruments, while actively monitoring the overall task at hand [9]. Psychomotor skill is of paramount importance to the trainees as it serves as the basis for the successful completion of adequately planned surgery. Spencer et al. [10] reported that 25% of a skillfully performed operation is thought to be attributed to manual dexterity. This 25% cannot be learned or improved by studying textbooks or visiting lectures [11] and has to be acquired by practicing the related movements numerous times.

Not limited to surgery, but almost all activities performed by dentists involve the use of instruments, like high-speed handpieces that can cut through and potentially harm any tissue in contact with the tool. Usually, these instruments are used in intraoral environments that have limited access challenges, offer less than optimal light and are often obscured by blood and saliva. Lastly, most procedures performed on teeth are irreversible, and harm to the patient can occur if the procedures are performed incorrectly. Given these circumstances, a significant proportion of dental education is dedicated to

training psychomotor skills.

## **2.2 Assessment and Feedback in Skill Learning Theories**

To provide training feedback, it is essential to understand how trainees learn to acquire skills and progress from novice to expert. Various learning theories exist, aimed at explaining and understanding how people acquire knowledge and skills such as those of Dreyfus and Dreyfus [12], Fitts and Posner [13], and Bloom [14]. The five-stage skills acquisition model by Dreyfus [12] has been applied to many disciplines in medical education [15–18]. The five stages, which individuals progress through in their acquisition of skills, are novice, advanced beginner, competent user, proficient user, and an expert. As a novice, an individual follows the given rules or plans in approaching the task mechanistically and needs supervision to complete them. To move to advanced beginner, individuals must have practice by applying the facts and rules to real situations along with proper feedback on the performance. After having considerable experience, competence develops in advanced beginners, and they learn to organize principles to access the particular rules that are relevant to the specific given task. Proficient individuals use intuition in decision-making and develop their own rules to formulate plans. At the expert stage, individuals act intuitively and produce a fluid performance that happens unconsciously, automatically, and no longer depends on explicit knowledge.

Fitts and Posner [13] described three stages that individuals undertake when learning a new skill as cognitive, associative and an autonomous. The learner gains a better understanding of the skill, and build the ability to execute the skill as he progresses through the phases. The cognitive phase involves the identification and development of the components and mechanics of the task. During this stage, the novice gathers and brings together reasoning abilities and experiences, which appear to relate to the performance of the task. In particular, the learner places great emphasis on the required responses and their ordering. In the subsequent associative phase, the prior cognitive activities begin to fade out. The learner links the parts into a smooth action via skill practices with feedback to perfect the skill. Major errors are significantly reduced as the learner refines responses. The learner focuses on better coordination and integration

by identifying redundant or inefficient responses. During this phase, feedback plays a significant role. In the automated stage, the essential elements of the skills have become so highly integrated that they are retained as an intact unitary skill. The acquired skill is deeply ingrained in the learner, and it could appear automatic or second nature to him. At this stage, errors have been greatly reduced, and less attention is required for the learner to perform the task. Regarding assessment, Millers pyramid conceptual model [19] provides a framework for assessing clinical competence in medical education. At the lowest level of the pyramid is cognitive levels of knowledge (knows), followed by application of knowledge/competence (knows how), demonstrated performance (shows how), and how a doctor (or dentist) performs in practice with patients/action (does). In the context of dentistry, consider the following as an example of the student passing through the stages of Millers pyramid. A dental student may first learn tooth anatomy and morphology, then learn how to recreate this knowledge using a material such as composite cast models. Afterward, they demonstrate that they can perform this skill in a simulated setting such as a plastic tooth or extracted patient tooth followed by being able to complete the same task for a patient.

Miller's pyramid model has been used to match assessment methods to the competency being tested. It helps in formulating the objectives for a particular training session with consideration on the achievement target. As an example, consider a training session on application of local anesthesia stage in a root canal treatment. The following are the examples of objectives and the level of Millers pyramid being represented: (a) Understanding what is meant by cardiac risk from anesthesia and why its crucial (knows) (b) Knowing what to do if the risk is too high (knows hows), and (c) Being able to demonstrate the application of local anesthesia on a patient (shows). Millers pyramid model for medical competence overlaps Blooms taxonomy of educational objectives [14] which was revised by Anderson et al. [20]. Traditionally assessment methods have been developed to test learning objectives within the cognitive domain and less commonly the psychomotor domain [21]. Bloom and colleagues [14] classified learning objectives into one of three domains cognitive, affective and psychomotor. The stages in Blooms concept of the Psychomotor domain which were further developed by Dave [22] include imitation, manipulation, precision, articulation, and naturalization. In

the context of dentistry, William et al. [9] view that it is possible to reach precision by training using phantom heads while the fourth articulation level would refer to training in a clinical setting. The naturalisation stage relates to the skills of the competent practitioner which could be equivalent to the shows level of Millers model, while it would seem to be more likely to be observed as part of does or performance [21]. Theories of acquisition of motor skills differ, however, one unifying aspect is the importance of feedback in skill mastery [23].

### **2.3 Dental Surgical Skill Training and Assessment**

In traditional curricula, dental students acquire the required psychomotor skills in the early stage by consciously following detailed textual and verbal instructions, observing from demonstrations and simulated clinical activities showing how to carry out motor skill tasks. Then the students go through preclinical activities by assisting mentors in providing care to patients in the clinic. As the trainee gains experience, training involves performing irreversible operative procedures on patients under the supervision of experienced clinicians [24]. At this point, the training process typically centers on the provision of patient care; it still includes several practice sessions with supervisors in both laboratory and clinical sessions until they can perform independently. Repetition of clinical procedures to achieve clinical competence is widely accepted and adopted in dental education [25].

Laboratory training includes a wide range of exercises that require the integration of theoretical knowledge to practice. Procedures are practiced on benchtop models such as typodont plastic teeth, extracted teeth from patients on manikin heads, and animal jaws (porcine or ovine) which closely resemble the dentistry of the human jaw. As there are no two identical teeth, variable external and internal anatomy when using the extracted teeth is an obstacle to deliver the standardized means of training and assessing surgical performance. Little or no mechanism to control anatomic variation or the presence of specific pathology means training is often done in a non-uniform fashion. Inconsistent learning experiences [26] and inconsistent feedback from different tutors [27] are the known issues associated with the use extracted teeth in dental skill

training.

Another drawback of plastic teeth and extracted teeth is that the practice cannot be repeated. Additionally, the look and the feel of artificial teeth are different from natural extracted teeth and usually have neither simulated anatomy nor pathology; those that do exhibit either of them are expensive [28]. With limited internal anatomical and pathological variations available in plastic teeth, the trainees opportunity for exposure to a range of examples with anatomical variation during training, an essential aspect of the skill training [29, 30], is restricted. In addition, the simulated activities in the laboratory are associated with high costs and are time-consuming to set up. Students interact with the supervising dentist on a one-to-one basis at the chairside while providing patient care; the presence of the patient in clinical learning environment adds extra stress and pressure to the students [31].

Assessment using typodont teeth is an almost universal competency assessment method of dental students before allowing them to perform procedures on a patient. Dental educators still rely upon the assessment of directly observed clinical procedures carried out under supervision (glance and mark) in assessing skill [21, 32, 33]. Inconsistencies can occur between and within examiners which is a well-recognized problem in dental education [34]. The findings from studies by Satterthwaite and Grey [35] and Goepferd and Kerber [36] highlight the presence of inter-rater variations in dental pre-clinical laboratories. Satterthwaite and Grey [35] found the intra-examiner agreement of two experienced assessors to be 0.53 when assessing typodont preparations. Other studies relating to clinical and laboratory assessments in dentistry, inter-examiner agreement scores are found to be ranged between 0.012-0.94 [35–44]. Jenkins et al. [32] found that assessment scores of Class II cavities vary up to seven marks out of a thirteen point grading scheme. They also found out when the assessors are reluctant to give high grades if they think the preparation is the students work [34].

Another observation based assessment method that has been widely used in clinical skill evaluation in general medicine, albeit with lesser frequency in dentistry, is assessment via video recordings. In the video-based assessment, a student's performance is video recorded as he performs the tasks and the examiners evaluate the performance later from the videos. Hassanpour et al. [45] compared video observation

of procedural skills (VOPS) to direct observation of procedural skills (DOPS) using a 10-point Likert scale. VOPS and DOPS scores are highly correlated and high intra-observer reliability is observed in their study. As the assessment can be done at a later time, multiple assessors can be included in VOPS conveniently.

## 2.4 Simulation based Skill Training and Assessment

In dentistry, a handful of VR simulators are dedicated to teaching psychomotor skills. Table 2.1 lists a few examples of the pedagogical use of virtual reality simulations (both commercial products and academic prototypes) in various branches of dentistry. One advantage of the simulator is the facility to point out to trainees their errors after every procedure or allow them to assess their procedural mistakes.

Table 2.1: Examples VR dental simulators

No.	Studies	Branch	Procedures	Evaluation
1	VRDTS [46]	Endodontics	Drilling, caries removal and cavity preparation. Filling cavities	
2	IDSS [47]	Endodontics	Detection of carious lesions	Y
3	VDP [48]	Periodontics	Pocket probing, calculus probing, and removal	
4	Voxel-Man Dental [49]	Endodontics	Cavity preparation, Carious lesion removal	
5	Forsslund [50]	Endodontics	Drilling, Wisdom teeth extraction	Y
6	Simodont [51]	Endodontics,	Drilling, decay removal, cavities filling	Y
7	Periosim [52]	Prosthodontics Periodontics	Crown and bridge removal Pocket probing, calculus detection and removal, scaling and root planing	Y
8	HapTel [53]	Endodontics	Drilling, caries removal and cavity preparation	Y
9	VirDenT system	Prosthodontics	Tooth preparation (crown and bridges), teeth grinding	Y
10	iDental [54]	Endodontics, Peri- odontics	Drilling, decay removal  Pocket probing, calculus detection, and removal	

*Continue on the next page*

Table 2.1: Examples VR dental simulators (cont.)

	<b>Studies</b>	<b>Branch</b>	<b>Procedures</b>	<b>Evaluation</b>
11	VirTeasy [54]	Odontology, En- dodon- tics, Prosthodon- tics, Implan- tology	Drilling, caries removal, Im- plant	Y
12	Simulator used in this study [55]	Endodontics	Drilling, access opening, crown preparation	

The Virtual Reality Dental Training System (VRDTS) prototype [46] allows the trainee to practice cavity preparation and virtual restoration of teeth using the haptic devices. The system is at an early stage of development, and the detailed technical information is not available [56]. The Iowa Dental Surgical Simulator (IDSS) [47] supports the detection of carious lesions (cavities) on the surface of teeth using a haptic probing system. Originally, it was designed to support the virtual simulation of the clinical evaluation of tooth decays using a probe, and then the focus changed on prosthodontic restorations (crowns). The virtual Dental Patient (VDP) by Marras et al. [48] aims to assist trainees in familiarizing themselves with tooth anatomy and handling of instruments used for drilling as well as challenges associated with the drilling procedure [57]. The trainee can perform virtual tooth drilling within the oral cavity which is constructed using anatomical data [57]. The Voxel-man dental simulator [49] supports training manual dexterity and problem-solving skills in using virtual teeth models with carious lesions. The Forsslund Dental system [50] supports wisdom teeth extraction training using the haptic devices. The MOOG Simodont Dental Trainer [51] is an immersive virtual reality unit that allows dental students to be trained in operative dental procedures with haptic, visual, and audio sensory feedback. Teeth are simulated with pathological dental conditions and it includes simulation of drilling, removal of tooth decay, restoring cavity preparations, crown and bridge preparations as well as a mirror reflection. In addition to practicing manual dexterity skills, the system allows users to select virtual



patient profiles, supports to perform diagnosis, treatment planning and provides automatic user evaluation [58]. The PerioSim [52] simulator is the first VR simulator that includes training and assessment of performance in periodontics procedure. Students can perform periodontal pocket probing, calculus detection, and calculus removal using the haptic device. The HapTel [53] consists of a haptic unit adapted from a computer gaming device, includes two screens that enable the user to look down onto a simulated jaw as if they were treating a real patient. The simulator supports caries removal cases with different difficulty levels. The Virident [59] is designed for training users on restorative dental procedures, particularly in how to prepare teeth (crowns and bridges) in fixed prosthodontics [59–61]. No additional technical information is available regarding feedback and evaluation for the Virident simulator. The iDental simulator uses the voxel-based modeling based on the linear list for drilling simulation [54], and performance metrics are defined for the quantitative assessment of three periodontics procedures. VirTeasy simulator [54] is designed for teaching implantology. Trainees can obtain an objective evaluation of an action and review them at any time while performing the procedure and is also possible to save the information for the subsequent review and assessment. Automatic skills assessment is said to include process and result of a student preparation, objective evaluation is based on predefined standards such as a reference preparation of a cavity or a crown.

Recently, robotic patients (DENTAROID [62]) have been introduced into dental education. Automatic dialogue features enable the robotic patient to communicate with the trainee just like with an actual patient [63]. Medical emergencies can be simulated with robotic patients [64] permitting the trainees to gain clinically realistic experiences [62].

Using simulators, dental students technical skill can be evaluated from the way the procedure is executed (procedure analysis) and from the quality of procedure outcome. In this regard, we categorized existing objective assessment using VR simulators into three groups as (a) assessment of skill using procedure analysis (b) assessment of skill using outcome analysis (c) assessment of skill using both procedure and outcome analysis.

### 2.4.1 Assessment of Skill using Procedure Analysis

The assessment of skill by procedure analysis using a VR simulator usually begins with the identification of interesting metrics, which are likely to distinguish expert behavior from novice behavior and can be acquired from the simulators while operators perform surgical tasks. Statistical models are then created using the acquired metrics to characterize the factors that constitute the difference between expert, competent, and novice. The assessment of student performance is carried out by computing the statistical similarity between the student's model and the prior ones. While the primary function of metrics is to provide the novice with objective feedback on performance, they also allow the trainer to provide formative feedback to aid the trainee in acquiring the skill [65]. A number of metrics useful for objective assessment can be extracted from simulators. Metrics commonly used in the objective assessment of dental surgical skill are summarized in Table 2.2.

Table 2.2: Metrics used for objective assessment of dental surgical skill

<b>Metrics</b>	<b>Description</b>
Traveling distance	Total traveling distance with respect to handpiece movement
Force	Force utilization in the x, y, and z-axes (x: mesio-distal direction; y: facio-lingual direction; and z: long axis of the tooth)
Task completion time	The total time taken to complete the task
Outcome score	Based on errors found in the incisal, labial-incisal, labial-lingival, and marginal regions of the tooth.
Tooth mass loss	The difference between the amount of tooth mass measured in grams before and after procedure
Mirror position	A measure to quantify the extent of bimanual dexterity through evaluation of the direct relationship between the mirror position and the handpiece

Objective feedback in skill assessment in dentistry using VR simulator was first introduced by Buchanan et al. [66] using the DentSim simulator [67]. The DentSim simulator consists of a phantom head with plastic teeth, a set of dental instruments, infrared sensors and an overhead infrared camera with a monitor and two computers [57]. As the student performs a procedure, the location of the instrument and the resulting outcome are displayed on the monitor while real-time evaluation is done in comparison

with the optimal performance kept in the repository. With limited anatomical and pathological variations available in plastic teeth, trainees are exposed to the general cases only. As the sensors are used, the tracking can be interfered from the way the handpieces are handled. In the Voxel-man dental simulator [49], the percentage of caries removed, the portion of healthy tissue removed are taken into consideration in computing the scores. To evaluate the wisdom teeth extraction the Forsslund Dental system [50] measures the removed amount of bone, enamel, dentin, and pulp on the virtual tooth. Likewise, Haptel simulator [53] measures the percentage of caries removed and the percentage of hard tissue removed and provide as a performance score [57].

The haptic dental VR simulator developed by Rhienmora et al. [55] is used in this study. The simulator supports the simulation of the dental procedures which involves the drilling acts. The simulator collects various kinematics variables while the user performs the procedure. The collected variables have been used to assess the quality of the performed procedure using hidden Markov models [55]. The details of the simulator is provided in Chapter 3. Rhienmora et al. [55] also presented a mechanism for objective feedback in skill assessment using an intelligent dental skill-training simulator. In their approach, once the student finished performing the procedure, the performance is classified as either novice or expert. User performance is modeled using the hidden Markov model (HMM) built from the kinematic variables such as the force applied, tool position, and tool orientation. The findings indicate their approach achieved high accuracy in classifying users into novice and expert and the generated feedback also received a high acceptance from experts.

#### **2.4.2 Assessment of Skill using Outcome Analysis**

In skill laboratories, students observe demonstrations of a task and perform supervised practice with mentors. To assess the performance, the mentor observes the students actions throughout a procedure. Assessment of the procedure outcome is usually carried out at the end of training session, and the mentor provides feedback regarding errors present in the result and suggestions to improve the procedure.

The idea of using the operative product as a measure of technical competence using bench model representation was first presented by Szalay et al. [68]. Their

study evaluated the quality of final product across six benchtop stations. The authors conclude that the method possessed construct validity. Additionally, they found a correlation between OSATS (Objective Structured Assessment of Technical Skills) and product analysis based assessment. Their finding is later reinforced by the study in Datta et al. [69]. Datta and colleagues [69] assessed the outcome of a vascular anastomosis on a bench model by measuring the leakage across the anastomosis and the cross-sectional area at the narrowest point of anastomosis. They found a significant correlation between these outcome measures and surgical dexterity.

Wierinck et al. [70] published the first study involving the evaluation of preparation using the dental VR simulator. In their research, Class II amalgam preparation on the lower left second premolar was performed, evaluated and graded on DentSim simulator [67]. An initial overall preparation score of 100 was assigned to each tooth preparation. The preparation was assessed by three parameters, namely, outline shape, cavity floor, and cavity walls. The error score of these parameters was later deducted from the initial maximum score. The passing grade for this task was defined at 60 by the grading system.

The E4D Compare software in its Beta version presented by Renne et al. [71, 72] evaluated students all-ceramic preparations compared to three calibrated clinicians. The gold standard preparation was agreed upon by the calibrated faculty members involved in prosthodontics course. The ideal and the student's preparations were scanned into the program with a laser scanner resulting in high-quality 3-D models. The two digital models were aligned based on common anatomical features of the adjacent teeth. Using the aligned teeth, the faculty member marked the finished line of the student preparation and the ideal standard, utilizing intuitive automatic margin finding tools and further refinement with manual tools. The areas of discrepancy in reduction (over-reduction or under-reduction) between two outputs were presented as color-coded regions, and the percent surface area of each color was calculated and responded back as the numerical value. The distance threshold between the ideal and the student's preparation was defined as 300 , the student's work varied from the ideal preparation within this range was considered as acceptable. The findings indicated that E4D software is significantly more precise method than the hand-grading method.

Zou and colleagues [73] presented a laser-scanning Cavity Preparation Skill Evaluation System (CPSES) for cavity preparation evaluations. The system scans evaluated and provided the scores on the outline form and depth of cavity preparations against a theoretical ideal. Firstly, the ideal cavity preparation was scan into the program, which was subsequently used to compare with the work prepared by the student. The CPSES measured over drilling or under-drilling (insufficient drilling) and displayed the discrepancies using different colors. The outline form score calculation is mainly based on the measurements associating with the width and depth of the prepared cavity and desired cavity. The student outline form score was computed from the outline form over drilling, and outline form drilling insufficiency score. And the depth score was calculated from over drilling and drilling insufficiency of the preparation depth data. The total score consisted of the score combined with form and depth scores. The authors noted the need for training before usage and the cost as drawbacks.

Currently, several computer-based evaluation systems are used in preclinical training, including the NISSIN Fair Grader 100 (Japan), KaVo PrepAssistant (Germany) [74], [75]. These systems can rapidly scan students cavity preparation samples, accurately compare them with standardized cavity preparation models, and then generate objective 2D or 3D feedback data on screen for the student.

### **2.4.3 Assessment of Skill using Both Procedure and Outcome Analysis**

The assessment of clinical skill performance of experts and novices using both process and outcome is found in a study Suenkarn et al. [76]. In their study, the assessment of performance process was based on the couples of kinematic measures gathered from the simulator while the user performs the procedure. The experiment was set up for the tooth preparation of a metal-ceramic crown on the upper left central incisor. Suebnukarn et al. [76] performed a comprehensive evaluation of all process measures including angulations of the instrument; force used, and task completion time. The results were saved and were graded by an experienced expert using a haptic VR simulator as the measurement can be made using the virtual bur and adjacent tooth. All tooth surfaces (incisal, facial, mesial, distal and lingual) were evaluated and graded using three evaluation parameters: depth, inclination, smoothness. Three-point scales (with 2 on

the scale defined as "reference", 1 as "acceptable", and 0 as "unacceptable") were employed. Thus, the maximum score for each surface was 6, and the total maximum score was 30. The findings indicated that task completion time and differences in preparation outcome successfully distinguished between novice and expert performance.

Their work was later expanded to assess outcome variables with the inclusion of bimanual coordination and force utilization in [77]. The study task was an access opening procedure on the upper right molar with a haptic VR system [76]. An expert graded the outcome of the preparation. The four-point scale used in their study is with four on the scale defined as "minimally extended cavity affording unimpeded access to/and visibility of the orifices of all canals presents"; three as "a coronal cavity permitting effective debridement of the canal system without prejudice to subsequent restoration; two as "incomplete removing of pulp chamber roof and/or inadequate retention form for the maintenance of an effective dressing"; and one as "unidentified canals and/or perforation". The referenced templates in the outcome scoring system presented in this thesis are designed based on their grading scheme which we will discuss in detail in Chapter 4.

## **2.5 Feedback**

Feedback is the information a trainee receives about his performance of a motor skill during or after the performance. Van de Ridder and colleagues [78] defined clinical feedback as specific information about the comparison between a trainees observed performance and a standard, given with the intent to improve the trainees performance. Evidence in motor learning indicates that feedback is essential and some form of feedback is necessary for learning to take place [79–85], and it has been suggested that its absence can prevent progress [86]. Among several possible channels used in the provision of feedback, the verbal feedback given during the surgical task is of prime importance in skill training. Based on the findings of Kannappan et al. [87], both positive and negative verbal feedback could be the potent stimulants for improved performance and motivation. Porte et al. [88] argue that the verbal input from an expert instructor could lead to lasting improvements in technical skill performance. Despite the known

benefits of feedback, instructors often hesitate in giving feedback, specifically in the verbal format of the sensitivity of verbal feedback aspects such as the tone, the timing, and the composition which could alter the positive feedback into negative one instantly [89].

Feedback, whether summative or formative, is an essential component of skill acquisition [90]. Summative feedback describes how well a student achieves a result in the form of scores or grades. With its nature as a measure for an assessment of learning, the advice the expert would make to the trainee is not adequately captured with the summative feedback. Formative feedback is considered as an assessment for learning, and it is the information communicated to the learner in response to some action on the learners part. Learners can use the formative feedback to modify their thinking or behavior to improve their skills. The various information can be conveyed as formative feedback (e.g., verification the action, explanation of the correct answer, hints) and can be administered at various times during the learning process (e.g., right after an answer, at the particular interval). Formative assessment in the form of paper feedback after the procedure is currently the gold standard in providing feedback to surgical trainees [89]. Although this approach is cheap, fast, and easily reproducible, due to its retrospective post-procedural timing, the precision and accuracy of the content heavily rely on the information being retrieved from experts memory.

In dental education, practice teeth prepared by students in clinical laboratory sessions are usually evaluated by experienced dentists primarily using visual inspection. During clinical laboratory sessions, the trainees require feedback on their work to move onto the next procedure. Consistent and accurate feedback from the faculty is vital for the students to achieve a higher level of performance before advancing to the clinics. However, the degree of consistency and accuracy of feedback could vary considerably, with many sources contributing to disagreement about student work, including the raters understanding on the grading scale, rater calibration, training for assessment, and personal influences [91].

### **2.5.1 Feedback in Motor Skill Training**

The acquisition of motor skill is a specialized process involving many receptors. Information received through sensory receptors is referred to as somatic sensation

which includes both exteroception and proprioception [92]. Thus, two types of feedback for motor skill acquisition are task-intrinsic and task-extrinsic feedback. Intrinsic feedback is sensory information inherently produced as a natural consequence of a movement. This feedback could originate from sources outside of the body (exteroception), or from within the body (proprioception). Extrinsic feedback is provided to the learner by sources external to the learner but from the environment. It is usually the information about the outcome of the movement that performers cannot obtain on their own. Extrinsic feedback is divided into the knowledge of their results (KR) or knowledge of performance (KP). KR provides a performer with information about the success in reaching the desired outcome. Without KR, it is difficult for students to know if their performance is close or far from the desired outcome. The error information is obtained from a comparison of the desired and actual outcomes and then may be provided by a trainer as KR (e.g., the pulpal depth was 1mm too deep) [76]. Feedback provided by an instructor when the student has completed all or part of the dental task, such as cavity preparation [93, 94] as an example of KR. The availability of KR feedback during simulated practice has been identified as one of the most critical factors that lead to effective motor learning [93, 95–98]. KP provides the performer with information about the quality and quantity of their actions. It represents information about the kinematics (pattern or speed). This information is gathered from the comparison between the desired parameters of the movement with the actual parameters (e.g., was the hand-piece held perpendicular to the long axis of the tooth when it should have been) [76]. Information about the quality of performance and movement characteristics known as knowledge of performance (KP) [93, 95–98]. With KP and KR, the performer can make appropriate adjustments to the movement on the next stage and increase the likelihood to achieve the desired outcome. On this basis, we can consider proprioceptors and exteroceptors as sources of data that the performer can access to regulate and modify their performance during the learning process. In general,

- *proprioceptors* provide sensory feedback that accompanies the movement of one's own body, such as signals of body and limb movement, and position.
- *exteroceptors* provide sensory feedback about the movement of the ob-



jects on the environment, mainly from the organs such as the eyes and ears.

Each time a motor response is made, the relationships between its contained sensory components (proprioceptive and exteroceptive) consequences and the initial conditions (e.g., tooth anatomy, the location of the tooth in the mouth, the posture of the operator, etc.) are established. This relationship is registered along with the performance outcome in the operator. Through practice, the operator comes to associate the movement with particular sensory consequences and the output product. During training, the availability of (KR) and (KP) enable connecting motor response to the output and accelerate the formation of an optimal association between them.

Feedback is the most crucial instructional component for simulation-based medical education, which centered at procedural skills training at its core. The simulation setting is considered as an ideal environment for providing feedback as trainees can practice the essential physical movements without risking the patient safety. Without the feedback, the trainee trained with the simulator might be unaware of committing an error and persists in this error, which might be, repeated or deliberately practice leading to the reinforcement of undesirable behavior rather than corrected [99]. With VR simulators, the delivery of feedback no longer relies on the availability of the expert during or after the skill training. It can be more explicit and more objective than human expert feedback. Several studies have investigated the impact of feedback on skill training using simulators; and the results are reportedly mixed between studies.

The role of feedback in dental skill training has been investigated using conventional phantom heads [94], computer-assisted simulators [100], and VR simulators [101, 102]. In operative dentistry training, using the DentSim simulator [67], Buchanan [66] shown the effectiveness of simulators in providing objective formative evaluation and in enhancing the rate of acquisition. Using the plastic teeth and dental handpiece, the students perform procedures on the Dentsim. Augmented visual feedback is generated based on the information from the motion tracking sensors and the students work compared to the standard preparation. In a study by Wierinck and colleagues [100], using DentSim simulator [67], they reported that when only the visual feedback from the simulator is provided to novices, short-term enhancement in performance is observed which is not found in retention. In the subsequent study by the same

authors [103] found that standardized feedback from an expert during the training session found to be more effective for retention and transfer of skill in relative to the VR feedback alone. In endodontic education, Suebnukarn et al. [101] demonstrated that augmented kinematic feedback from a simulator on movement patterns during the access opening for root canal treatment resulted in an enhanced performance at the early stages of training. Amid having the mixed results, feedback from the VR simulators is found to be useful as a means of improving performance and the combined feedback from experts and the simulator during the training are found to be useful in the early stages of skill learning and assisted in retention to a certain level. The findings from existing study suggest that a combination of instructor and visual display (VR)-driven feedback method could result in faster skill acquisition relative to VR alone [93], which leads to the studies to investigate the use VR dental simulators as adjuncts to traditional training approaches [44, 104–106].

### **2.5.2 Structure and Timing of Feedback**

Feedback can be in the form of textual, verbal or tactile based on the context in which it is generated. Content presented in textual format could be the simplest form of feedback. However, feedback consisting of a significant amount of textual information could distract the student. Also, it may not be adequate for motor skills, as learners do not receive visual cues on task-relevant movement. During the training process, while the primary display is rendering the simulated environment (such as operating site in surgery), the appearance of a large number of text messages on screen during training could also distract the operator. To enhance the student's direct perception of errors and the desired outcome, and to retain the knowledge longer in memory, graphics information can be incorporated in the feedback. In addition to the textual and graphical feedback, auditory and haptic feedback are available in computer-based simulations. It is particularly suitable for VR dental simulators which are equipped with haptic devices as the haptic feedback can be provided efficiently without distracting the trainee's attention from the working area. In DentSim simulator [67], instant feedback messages are delivered for critical, non-correctable errors throughout the procedure. In addition, the user can request feedback at any point of execution. At the end of the procedure, an

extensive evaluative feedback report is presented in graphical and textual forms, including a numerical grade on the performance outcome. Rhienmora et al. [55] describe a mechanism for providing objective skill assessment and tutoring feedback during crown preparation procedure. Unique characteristics of force, position, the orientation of instrument observed by the simulator during student's preparation are compared to gold standards to generate useful feedback. The quality of the produced tutoring feedback is comparable to the feedback provided by human tutors.

Another essential factor to determine is the timing of the feedback. According to Shute [107], the feedback timing may independently influence its effectiveness. Typically, it is feedback given less than 100% of the time is considered optimal while there is no such the ideal combination of when and how feedback should be delivered [108]. Existing evidence suggests that efficacy and timing should be determined by the training objectives and the difficulty/complexity level of the task [109] and the competency level of the trainees themselves. Recently, the feedback related studies are moving towards the study of the optimal frequency and type of feedback [110].

In general, feedback can be immediate, delayed, and on-demand [111]. Instant feedback is given at the end each step, delayed feedback is provided after the task is completed, and on demand, feedback is presented in response to a request by the student. In motor skill training, Catania [112] argues that for optimal learning and practice, the operator needs feedback on procedure performance proximate to task execution, particularly for metrics errors [112]. Gallagher et al. [113] also considered that proximate feedback is particularly helpful in motor skill training because when it is provided closer to the event, it has more clarity and relevance for the recipients over the delayed feedback.

Concurrent/Immediate/Proximate feedback is given while the task is being carried out may be advantageous for a novice learner. Based on the analysis of learning curve, novice learners can gain advantages from the feedback provided while the task is being carried out. As the trainee can immediately correct the mistakes or misconceptions [114], the learners can directly improve their performance [115]. Immediate feedback is considered useful in short-term and for supporting the development of procedural skills [109]. Novices are found to be benefited from the feedback given during

the execution of the procedure at the early stages of the training [43]. McGahie et al. [133] argue that direct improvement in clinical performance from immediate feedback might result from its nature of the instant correction of misperceptions [43]. Otherwise, immediate/proximate feedback can interrupt trainees from the task at hand [43], and it could also lead another issue that the trainees become dependent on the feedback during procedure execution. Boyle et al. [116] assessed the effect of proximate or immediate feedback laparoscopic colectomy procedures. They concluded that the provision of standardized feedback during training was associated with significantly fewer errors and an improved learning curve.

On the other hand, immediate delivery of feedback as soon as a slight deviation or error is detected can interrupt trainees from the task at hand [108], and can lead to the reliance on feedback [117] which may detrimentally affect task resilience as well [95, 118]. Hence, immediate feedback should gradually be fade out or reduced in favor of delayed feedback given at the specific point of the procedure (such as the end of a stage/task) and terminal feedback presented at the end of the performance [108, 119].

Being postponed until the end of the procedure, the learners are supported with fewer interruptions during the task with delayed feedback. Trainees who already mastered the necessary skills may benefit from delayed feedback. Whether this is proximate or terminal, it is important to note that both the way the feedback is delivered and the way the reception of feedback determines its effectiveness. The feedback that is given in a standardized and structured manner could lead to in an improvement in trainee performance [116]. On the other hand, the feedback can not necessarily promote learning if the recipients do not receive it mindfully [120]. Archer [121] stated the circumstances that could lead to the mindless feedback including when the feedback (answers) is provided before the recipient has had time to think, when the challenge is too easy or too complicated, or when the process is random or inconsistent.

## **2.6 Root Canal Treatment**

As described in the scope of the thesis in Chapter 1, this thesis focuses on the area of motor skill training in endodontic surgery in dentistry. We will briefly describe

the background on the endodontic root canal treatment in this section.

Root canal treatment is an endodontic procedure used to treat infection at the root canal system inside the tooth due to the disease and injuries related to dental pulp. The treatment requires attention and precision while working in a confined space such as the root canal system of the tooth [122]. Learning a fine motor skill in endodontics requires the trainee to establish control of instruments as well as the integration of posture, motion, and muscle stimulation that, in turn, allows them to execute a variety of motor behaviors that are controlled by a range of task requirements [123] [124]. In the field of dentistry, endodontics is regarded as a stressful and challenging discipline by dental students. The complex anatomical diversity of teeth and the reliance on indirect vision to visualize the root canal, the dependence on tactile feel, and the array of instruments and materials tends to make some students feel inadequately prepared to deal with endodontic treatment procedures [125].

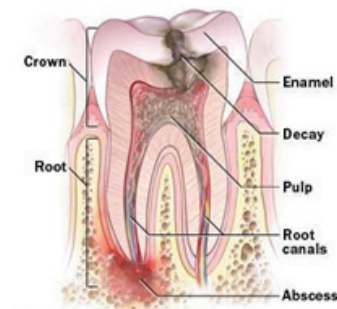


Figure 2.1: A sagittal cross-section of a tooth with an infected pulp

(Source:www.medindia.net)

The root canal system primarily consists of the pulp chamber and the root canals. The pulp chamber comprises blood vessels, nerves, and connective tissues. As shown in Figure 2.1 when the pulp is injured or infected, the pulp tissues die. If the necrotic pulp is not removed, the tissues around the root of the tooth can become infected. During the root canal treatment, the pulp tissues are removed, the pulp chamber is cleaned and sealed using a root filling. As the roots are located at a lower portion of the tooth concealed by the crown, to perform root canal treatment, the route to access the roots needs to be prepared first. This preparation step is commonly known as the access opening phase. In this stage, the endodontist drills a small access hole through

the surface of a tooth crown to gain access to the pulp chamber and root canals for treatment. The expected result of access preparation is to create an unobstructed passageway to the pupal space and the apical portion of the root canal.

Inadequately prepared access could create access related mishaps including perforation and ledge formation in the subsequent stages. A well-designed and thoroughly executed access preparation is necessary for the successful root canal treatment. Typically, errors may arise in later stages of the treatment when access opening is prepared in either under-extended or over-extended manner [126]:

- *Under-extension* occurs when the access cavity is not opened up across the width of the root sufficiently, and as a result, some canals may be left unidentified. If the cavity is not adequately extended, there may be insufficient space for instruments to maneuver in treating the root.

- *Over-extension* occurs when the dentist removes unnecessary tooth structure. Consequently, a weak tooth structure could be left behind after the treatment.

Both under-extension and over-extension could compromise the subsequent treatment stages. Therefore, the access opening preparation stage is considered the most critical stage in the root canal treatment.

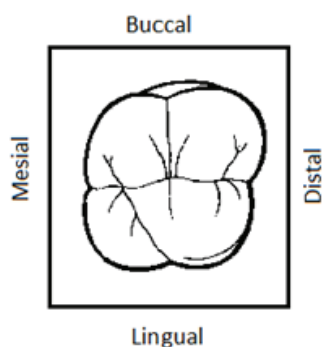


Figure 2.2: Tooth surfaces

In practice, an expert dentist provides feedback on student's work by specifying the location of errors in the student's outcome. To identify specific areas on a tooth, the crown of the tooth is divided into surfaces that are named according to the direction in which they face. As shown in Figure 2.2 the surfaces of a premolar and molar teeth from occlusal view are

- Buccal: The surface that faces the cheek
- Lingual: The surface that faces toward the tongue
- Mesial: The surface that faces the front of the mouth
- Distal: The surface that faces the back of the mouth
- Occlusal: The surface that is used for biting or chewing.

## **CHAPTER III**

# **OBJECTIVE ASSESSMENT AND FEEDBACK GENERATION FRAMEWORK**

In this chapter, we describe the framework for objective assessment and feedback generation for technical skill training using a dental VR simulator.

### **3.1 Framework Overview**

As shown in Figure 3.1, our approach to objective feedback begins with an assessment of outcome where the procedure outcome will be evaluated to identify the location, the type, and the severity of errors. Then, to determine the portions of the procedure responsible for errors, the way the procedure is executed will be assessed. In the subsequent step, the relation between procedure and outcome will be determined to provide the feedback in the following step. In providing the feedback, some modalities might be more appropriate than the other, therefore, multiple modalities will consider in this framework.



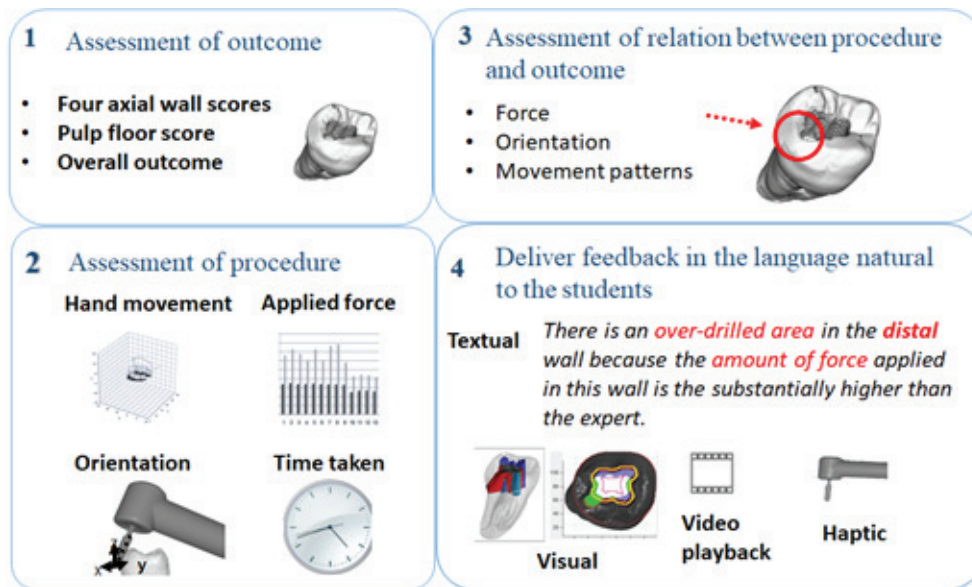


Figure 3.1: Objective assessment and formative feedback framework

The major components of the framework are: (i) Automated outcome scoring system, (ii) Correlator and (iii) Feedback system as shown in Figure 3.2. The procedure log is the kinematic variables log collected by the simulator while the student performs the procedure. The correlator component incorporates the procedure log of the student as well as that of the expert. The collided voxels log contains the timestamps and the locations of voxels that are drilled out from the tooth volume. The correlator component also references to the collided voxels log to obtain the temporal information of the error voxels. Each component of the framework will be discussed in the following sessions.

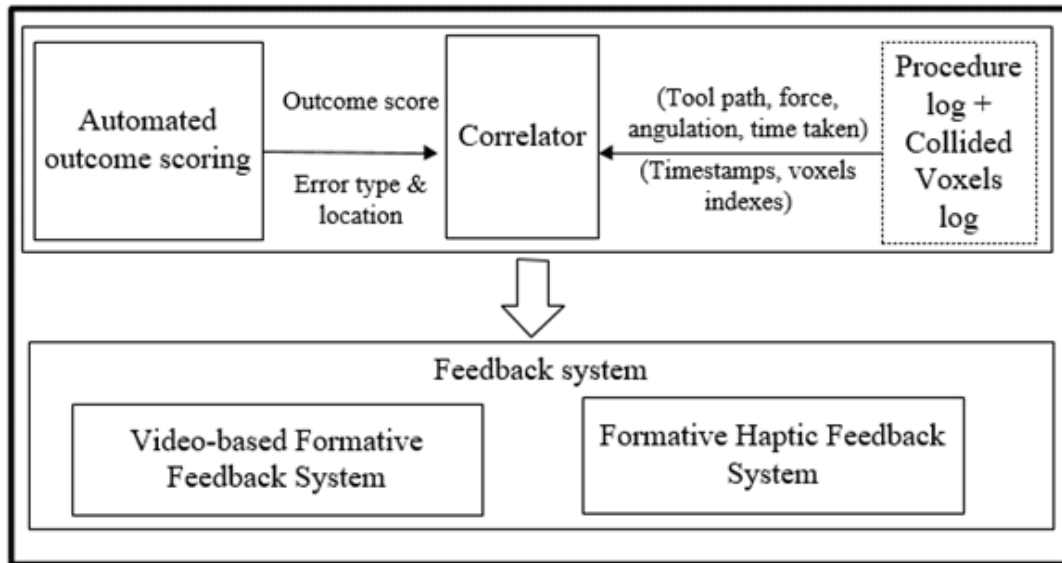


Figure 3.2: Framework components

### 3.1.1 Automated Outcome Scoring System

The main aim of the automated outcome scoring system is to evaluate and assign the scores on the outcome and to identify the types and locations of errors in the outcome. In this thesis, the student's work is compared with the standard virtual templates using the template matching technique to evaluate and determine the score for the student's outcome automatically. The system identifies the regions of the outcome that deviate from the standard outcome and the score is determined based on the degree of deviation.

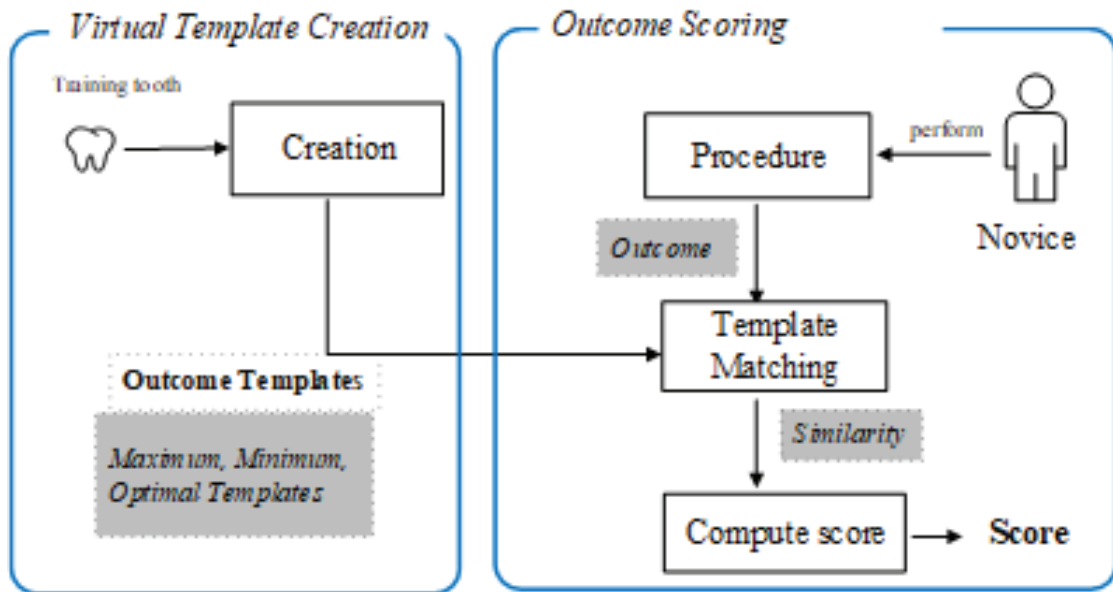


Figure 3.3: Automated outcome scoring

The overview of automated outcome scoring is shown in Figure 3.3. The correct identification of location and shape of pulp chamber is crucial for the accurate scoring result. Due to the unique nature of each tooth, there is no standard way of pinpointing the pulp chamber from each tooth. Automated outcome scoring system faces difficulties in addressing the variation in internal anatomy of the individual tooth. The presence of the external bone fragments in tooth surrounding and the pulp stones inside the tooth create challenges in locating the pulp chamber inside the training tooth. The details of the automated outcome scoring system are discussed in Chapter 4.

### 3.1.2 Correlating Procedure and Outcome

This component is responsible for correlating the errors identified in the performance outcome with the responsible portions of the procedure. The correlated information is used in generating the feedback in the subsequent feedback system. The correlator identifies when and how an error was made during procedure execution for each error diagnosed in the outcome. The problem of correlation procedure and outcome is an instance of the well-known credit assignment problem [134]. The problem is concerned with determining and distributing the success and the failure of a system's overall performance to the various contributions from different system's components

[134]. The details on the credit assignment problem are further discussed in Chapter 5.

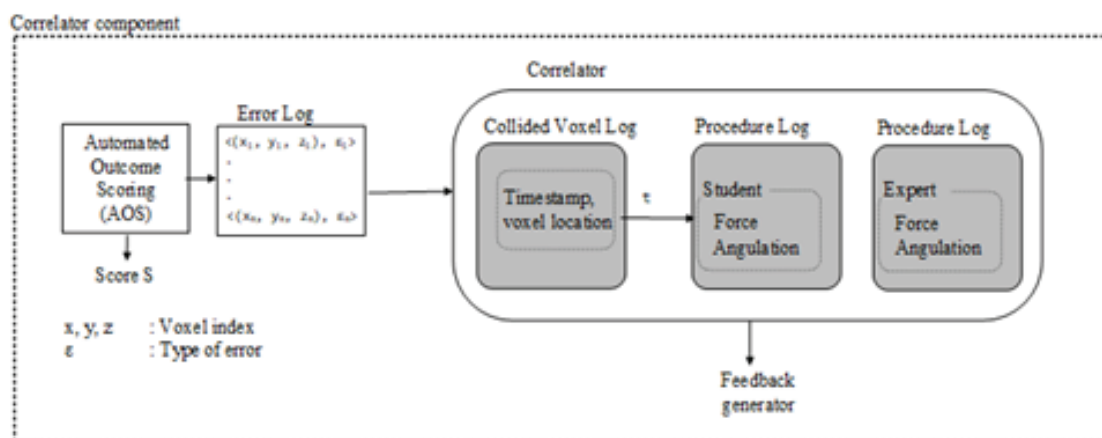


Figure 3.4: Correlator overview

The overview of correlator component is shown in Figure 3.4. The types and the locations of errors in the outcome obtained from the automated outcome scoring component serve as inputs into the correlator. Since the information on errors is at the fine-grained voxel level, the correlator component transformed them into clusters. It then employed the procedure log containing the kinematic variables of the student together with the experts procedure log which was obtained before the training. Among all the kinematic variables available in the log, the correlator component extracts the timestamps, the stage, the force (in x-, y- and z- axes) and the angulation (in x-, y- and z- axes) of the instrument, and the drilling status of the instrument from the procedure log. As the correlator component needs to decide when the error occurs, it incorporates the collided voxels log. Similar to the procedure log, the collided voxel log is collected in 1000 ms while the student executes the procedure. It contains the timestamps, the locations of voxels which were collided and drilled out with the instrument during the procedure. For each error voxels in the outcome, the spatial locations of errors are mapped to the collided voxel log to identify the portion(s) of the procedure responsible for each error. The details of the correlator component are further discussed in Chapter 5.

### 3.1.3 Feedback System

The outputs of the correlator component are information on types and locations of errors in the outcome, and portions of the procedure identified as causes responsible for them. Although information necessary to generate the feedback is obtained from the correlator component, another challenge is to decide how to effectively convey to the students. Primarily, the formative feedback will be delivered to the trainee through video-based formative feedback system to assist them in their learning process.

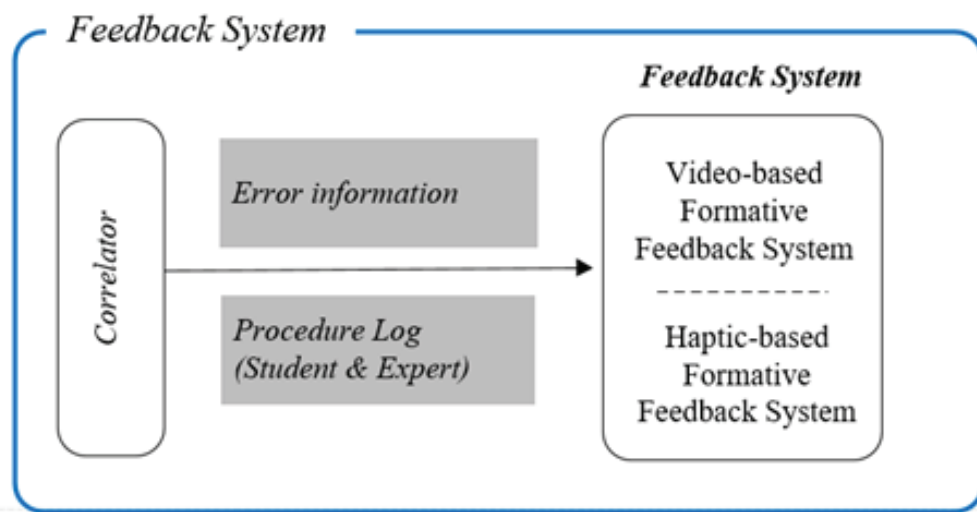


Figure 3.5: Feedback system overview

As shown in Figure 3.5, the input to the video-based formative system consists of the error information (the location and the type of errors), the procedure logs of trainee and expert, and the temporal information of the portion of the procedure that is identified as the cause leading to the error from the correlator component. The feedback information is transformed as simulator playback together with video control features. The playback replays the procedure while it interactively visualizes the simulation by highlighting the error areas within the tooth volume at the identified point of time associated with the wrong actions determined from the correlator component. The design features, and the findings from the evaluation study on effectiveness of the of video-based formative feedback system are discussed in details in Chapter 6.

Video-based feedback system is the visual feedback modality aims to communicate with the students on the errors in the outcome and the portions of the procedure

accountable for them. The kinematic variables of procedure (the applied force and the angulation of the instrument in our case) are analyzed as part of the identification and to assign the credit/blame to the portions of the procedure each error in the correlator component. Video modality alone is not sufficient to convey the feedback involving the kinematic variables such applied force information on the instrument. Therefore, the haptic modality-based force formative feedback was designed and implemented as part of the feedback component of the framework.

For each procedure stage, the students applied force on the instrument is compared with that of the expert and provided to the trainee. Together with the expert, the trainee has to decide the stage to get the force feedback training. For the stage determined to train again, the simulator is rewound, and the correct application of force is trained by using haptic feedback. The details on the formative force feedback approach and the findings of its effectiveness evaluation study are described in Chapter 7.

## **3.2 Evaluation**

We employed the VR dental simulator developed by Rhienmora et al. [55]. The simulator operates on a standard PC connected to two GeoMagic Touch haptic<sup>TM</sup> devices [127] as a dental handpiece and as a dental mirror (Figure 3.6). The stylus of the haptic device controls the position of the virtual dental bur on the display screen. The monitor was placed at eye level, and the haptic device was positioned at elbow level directly in front of the participant. A virtual high-speed handpiece with a tapered bur diameter of 1 mm and a length of 6 mm was employed. The tooth model was acquired using three-dimensional micro-CT (RmCT, Rigaku Co., Tokyo, Japan). The operator receives different force feedback depending on the density values of various tissues while cutting the tooth. A study of the construct validity of the simulator showed the haptic force feedback to the operator to be similar to working in the real situation [128]. In the simulator, the tooth is stored in the form of a three-dimensional grid of voxels representing the density of the structure at each point using a value between 0 and 255 with 0 representing a transparent voxel.



Figure 3.6: Simulator interface

We have selected access cavity preparation of the root canal treatment procedure to demonstrate and evaluate our approach. This procedure was chosen because it exclusively involves drilling, which is supported by the simulator, and because the outcome is challenging to score, being a complex function of the internal tooth anatomy. In the access cavity preparation phase, the endodontist drills a small access hole through the surface of the tooth crown to gain access to the pulp chamber and root canals for treatment. The ideal result of access opening preparation is to create an unobstructed passageway to the pulp space and the apical portion of the root canals. The perfect shape of the opening is a function of the tooth shape, tooth size, and the number and location of the root canals. The number and location of the root canals can differ in the same tooth (e.g., mandibular left second molar) across different patients. While the act of drilling may seem at first relatively simple, the access to the complicated root canal systems dictates the final treatment outcome. During data collection, data was gathered on elapsed time and kinematic variables concerning

- the position of the handpiece in x, y and z-axes\*,
- the angulations of the handpiece with respect to x, y, and z-axes\*,
- the transformation of the handpiece from the original position, drilling enabled/not enabled,
- the position of the mirror in x, y, and z-axes\*,
- the angulations of the mirror with respect to x, y, and z-axes\*,

- the transformation of the mirror from the original axes, and the force applied on the handpiece in x, y and z-axes\*. \*x, y, and z-axes (x, buccolingual direction; y, mesio-distal direction; and z, long axis of the tooth)

We considered three stages of access preparation:

- Stage 1, initial drilling to shape the outline;
- Stage 2, extend the opening to the distal canal orifice;
- Stage 3, extend the opening to all the remaining canal orifices.

Evaluation will be carried out on three key components of the system (Figure 3.7) is designed for our study.

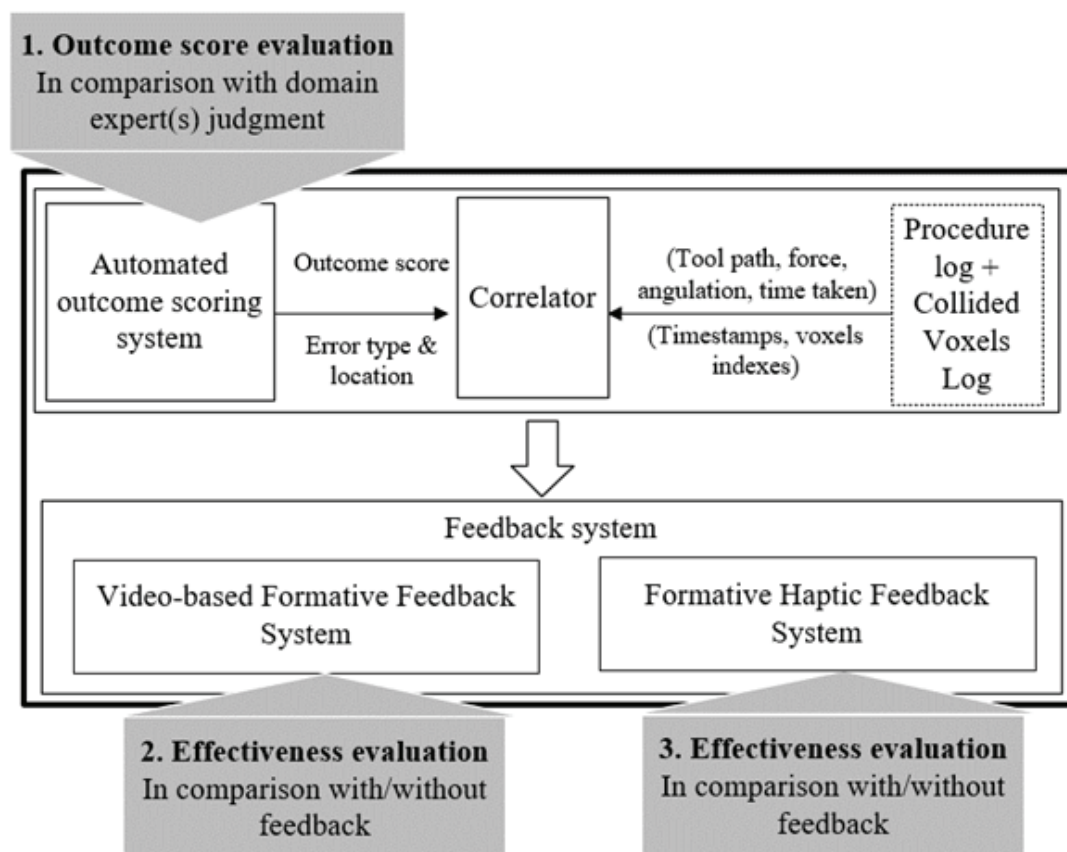


Figure 3.7: Evaluation plan

The first evaluation point is the automated outcome scoring system. The outcome score from the system is evaluated for compliance with domain expert judgment (Chapter 4). Evaluation studies are carried out to determine the effectiveness of the video-based (Chapter 6) and haptic-based feedback (Chapter 7) system in dental sur-



gical skill training. In general, the studies involve the comparison of skill achievement between three groups of students, one experimental group trained with the simulator with feedback, one experimental group trained with the simulator without feedback and a control group trained in the traditional way.

## CHAPTER IV

### AUTOMATED OUTCOME SCORING SYSTEM

Having defined the framework, in this chapter, we describe the first component of the framework, the automated outcome scoring system in details. Attempts to evaluate tooth preparation objectively include the use of pulpal floor measuring instruments [129], CAD systems [130], and the computer-based approaches using simulators discussed in Chapter 2. The need for valid objective assessment using simulators has been noted since early 2000 [131], yet the integration of efficient grading and evaluation systems using simulators into curricula is still open to research.

In a survey conducted by Wang and colleagues [58], they found that only a few VR simulators [48, 49, 51, 132–134] provided immediate feedback on errors or score performance of trainees [58]. Among existing dental skill training simulators, the Dentsim [67] and EPED [135] simulators provide training in intracoronal and extracoronal restoration by using plastic teeth and tracking kinematic data of the instruments using sensors. In practice, the look and the feel of artificial teeth are different from natural extracted teeth and usually have neither simulated anatomy nor pathology; those that do exhibit either of them are expensive [28]. In the absence of pathological insinuations, the trainee ignores merely the limits of the carious lesion, without knowing with the certainty of the distribution of the dental disease [136].

The number of voxels cut in the operating area by a novice compared to skilled dentists is a commonly used metric for outcome assessment. Examples include the percentage of caries removed, the percentage of healthy tissue removed and injuries (e.g. pulp exposed) considered in cavity preparation [26, 49, 53, 132] and how much the operator has carved into risk areas by measuring the removed amount of bone, enamel, dentin, and pulp on the virtual tooth in tooth extraction [137]. Although a significantly different number of voxels removed in each area can indicate severe problems in the procedure, a similar number of voxels removed does not always mean error-free performance. IDEA simulator [133] measures and records task time, the percentage of desired

material removed, and deviation from the assigned drilling task. The measures were further processed to get the final score. However, the authors noted that this grading system has not yet been validated [133]. For quantitative evaluation for three Periodontics procedures, iDental [54] uses performance metrics such as (a) periodontal pocket probing check task: deviation of pocket depth, maximum contact force, probing angulation; (b) Calculus detection: the number of identified calculus, reported value of position and size; (c) Calculus removal task: the number of removed calculus, the damage to the neighboring gingiva, and the operation angle of the probe.

In a construct validation study of Simodont simulator [134] provided immediate feedback for a manual dexterity exercise to the trainee. The feedback information included a percentage score for each of the following: target (task completion percentage), error scores (Leeway Bottom, Leeway Sides, container bottom and container sides) and time taken to drill (in seconds). No further discussion on the scoring system is provided further.

## **4.1 Outcome Scoring System**

To provide formative tutoring feedback, automated outcome scoring (AOS) must have the ability to identify the type, location, and severity of errors and be robust enough to account for a variety of possible outcomes. Since the optimal access opening route is an unobstructed passageway to the pulp space, we first locate the pulp chamber from the training tooth and project vertically from the base of the pulp chamber to capture the morphological information (shape, size, and location) of the pulp as shown in Figure 4.1 (a, b). Using the projection, the areas of the training tooth are virtually removed to create the optimal outcome template Figure 4.1(c).

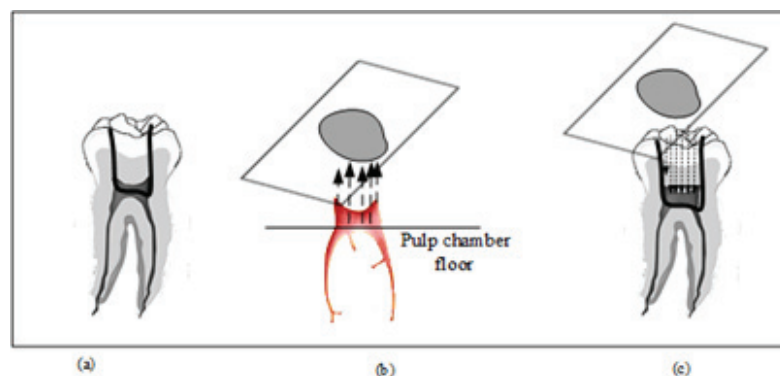


Figure 4.1: Virtual optimal template creation (a) pulp chamber location (b) projection from pulp chamber (c) removal of tooth mass for projected area

To permit a clinically acceptable amount of variation in outcome, Max (maximally acceptable) and Min (minimally acceptable) templates are defined by expanding and compressing the optimal template, respectively. Given a tooth and a procedure, a wide range of outcomes is possible. Our approach to outcome scoring is to evaluate the voxels in the tooth volume and label them with scores with respect to reference templates. In AOS, a new tooth volume is created and the voxels tagged with scores according to their locations. We call this a score cube.

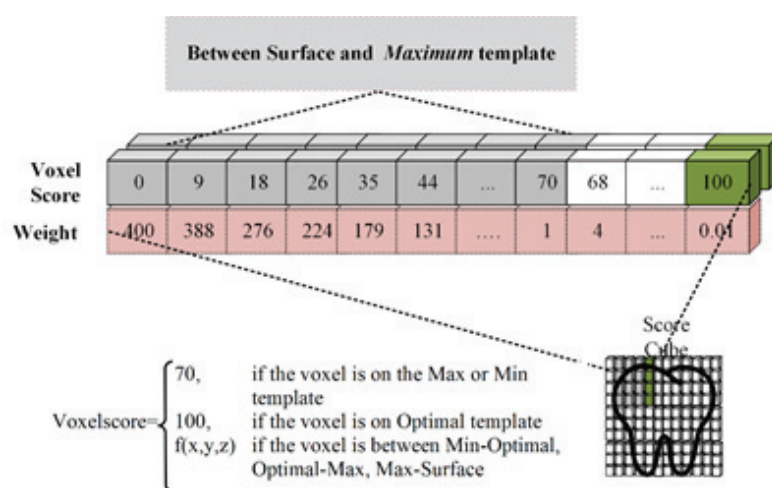


Figure 4.2: Portion of score cube between surface and Max template

The score cube (Figure 4.2) is a 3D volume of the same size as the training tooth. To fill in the cells of the score cube, we first define a voxel scoring function. In

practice, endodontists evaluate each access opening preparation on a scale between 0 and 100, with 0 representing perforation, 1-69 representing unacceptable, and 70-100 representing clinically acceptable. Following this scoring scheme, the score of 100 is assigned to the optimal template area, and the score of 70 is assigned to Max and Min templates. The values of the voxels in the regions Min to Optimal, Optimal to Max and Max to Surface are filled using linear interpolation. Figure 4.3 illustrates the z-axis cross-section view of the score cube.

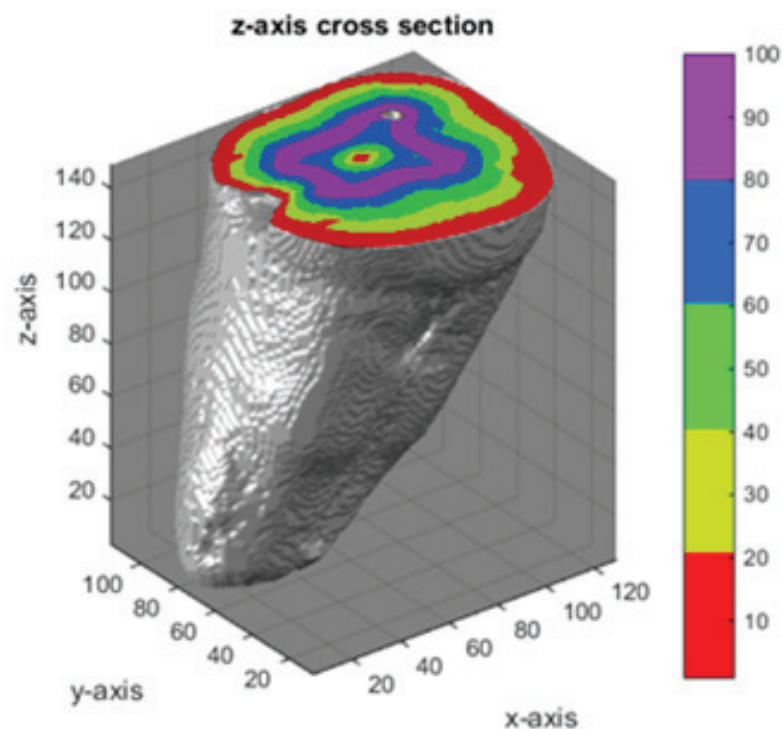


Figure 4.3: Score cube z-axis cross-sectional view

Initially, we computed the overall outcome score as the average of the scores of voxels on the drilled area surface. However, in comparison with scoring by endodontists, we found that our scores computed in this way did not correspond well with the expert scores. In subsequent interviews we found that experts assign a significantly higher weight to more severe errors than to minor errors, such that a wall with a small area that is considerably over drilled (close to perforation) is given a much lower score than a wall with numerous small amounts of over drilling, even if the average amount of over drilling in the second case is more significant than in the first case. Thus, the linear scheme was adjusted with a non-linear five-parameter logistic weight function [138]

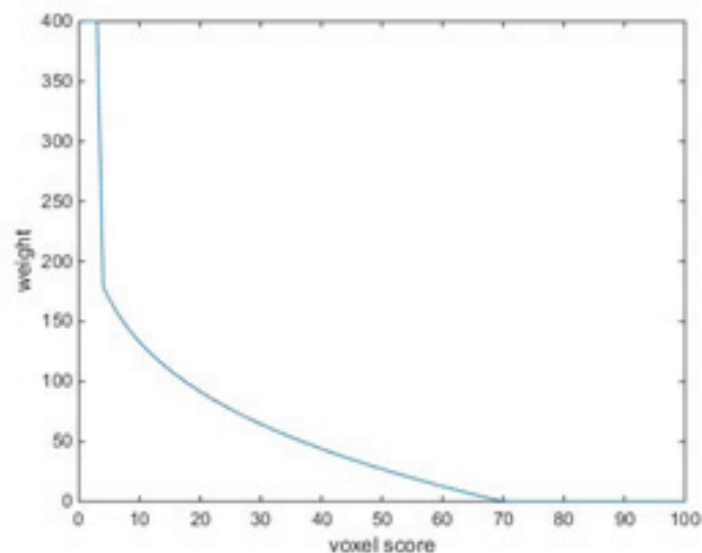
shown in Eqn. 4.1 to apply a higher weight to more severe errors at the voxel level. The parameters were estimated by curve fitting with least squares method before the experiments as shown in Figure 4.4. The range of the weight values [0,400] was determined experimentally.

$$weight = D + \frac{(A - D)}{1 + \left(\frac{voxelscore^B}{C}\right)^E} \quad (4.1)$$

where

- D is the top plateau of the curve (the highest weight value),
- A is the bottom plateau of the curve (the lowest weight value),
- C is the voxel score at which the middle weight value is attained,
- B is the slope factor, and
- E is the symmetry factor.

To effectively communicate the assessment results, feedback should be made in a language easily comprehensible to students. Endodontic surgeons evaluate and communicate about errors in terms of scores for the four axial walls (Lingual, Buccal, Mesial, Distal), the pulpal floor (Figure 4.5), and an overall score.



Note: The voxel scores between 70 and 100 are assigned near zero value weights.

Figure 4.4: Non-linear weight function

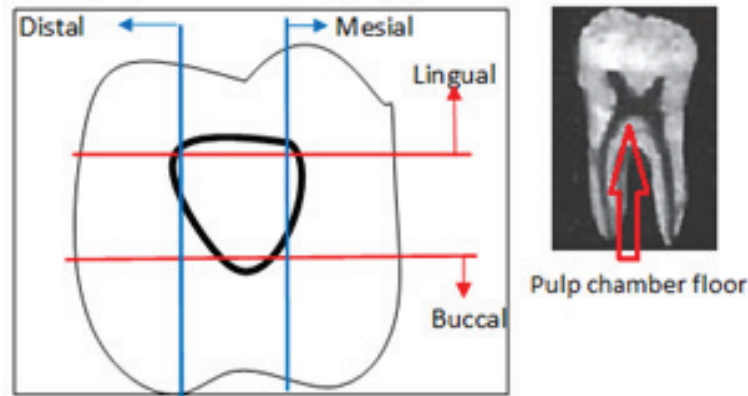


Figure 4.5: Tooth walls

To translate the weighted voxel-level scores into the wall level score, the scores of the four axial wall regions are first determined. Typically, human experts use subjective judgment in determining the axial wall regions of the tooth. To standardize the scoring, we define the four walls by projecting rays from the center of the tooth to the four diagonals as shown in Figure 4.6(b). At the center of the pulp chamber, a minimum bounding box is drawn within the pulp area. From the center through the four corners of the bounding box, the lines are projected towards the tooth surface, to create four non-overlapping regions. In this manner, AOS determines non-overlapping axial wall regions for scoring.

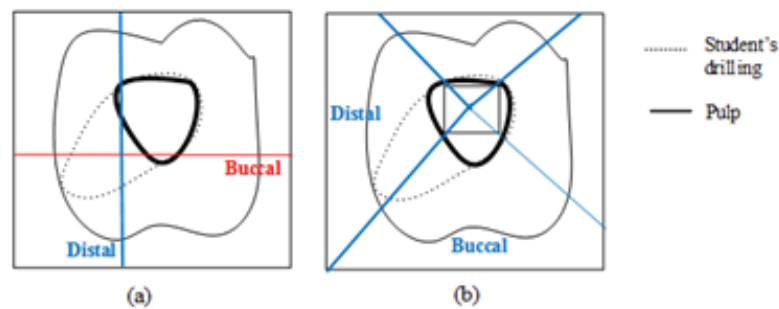


Figure 4.6: Segmentation of tooth into four axial walls (a) Original tooth walls (b) Adjusted tooth walls

After determining the walls for scoring, AOS first extracts the surface contour from the drilled outcome area. The contour of the wall is then mapped onto the score cube, and the voxels on the surface are assigned the score points from the weighted-scoring function. As shown in Figure 4.7, the wall score is computed as the average

weighted scores of the drilled area surface on the tooth slices in the wall region. The overall outcome score is determined by the average score of four axial walls and the pulp floor.

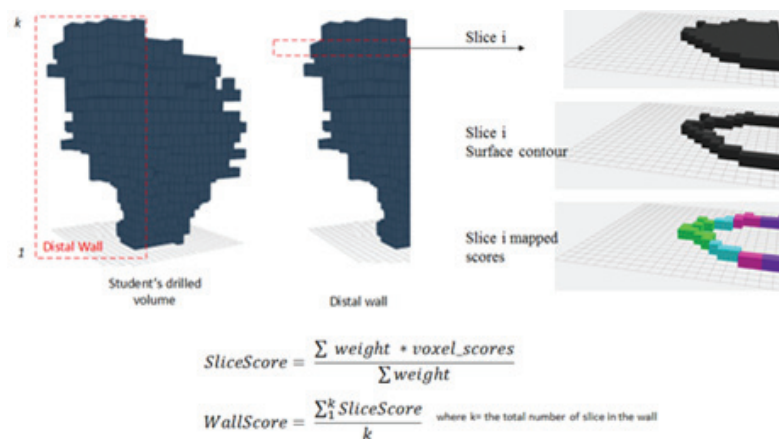


Figure 4.7: Wall score computation

## 4.2 Experimental Evaluation

We sought to evaluate the degree of agreement between AOS and human expert scores over a range of outcomes, including varying numbers of errors, types of errors, and severities. Fifteen outcome samples using the mandibular left second molar were prepared by an experienced endodontist who is familiar with the use of the VR system. During data collection, the expert deliberately committed a range of errors on the training tooth to reflect the types of errors committed by the students during the access opening procedure. The set of outcomes contained optimal results and those with errors including perforation of the walls, floor, and both, as well as various combinations of more minor over and under drilling errors. Five endodontists (R1 - R5) who had professional training and experience in root canal treatment participated as raters in the experiment. The raters were selected by expertise levels which varied from one year to more than ten years of experience. The raters received verbal instructions to score the four axial walls (Lingual, Buccal, Mesial, Distal) and the pulp floor using the standard scoring scheme to which they were accustomed. Human raters usually score outcomes using the external view of the tooth, making it difficult to see some errors. Thus, any differences in rating between AOS and the human raters could be due to limitations of



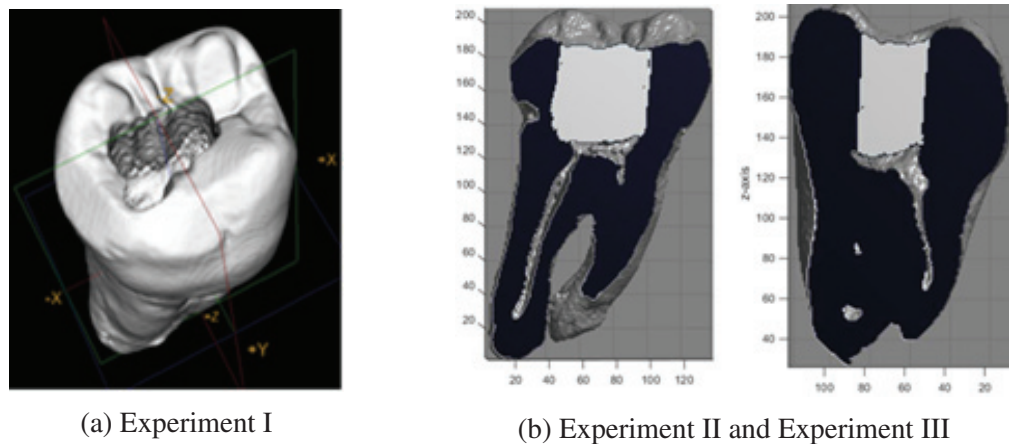


Figure 4.8: Example views the experiments

perception or to differences in judgment. To separate these two factors and determine the extent of the influence of perception on scoring, we ran three sets of experiments.

In experiment I the raters were provided with a 360 degree external view of the drilled tooth (Figure 4.8a)) just as they would have in a clinical setting with a real patient. In experiment II the level of information provided to the raters was increased by additionally providing mid cross-sectional views of the drilling area (Figure 4.8b)). This provided raters with visual information on the depth, size, and shape of the drilling at the center of the pulp.

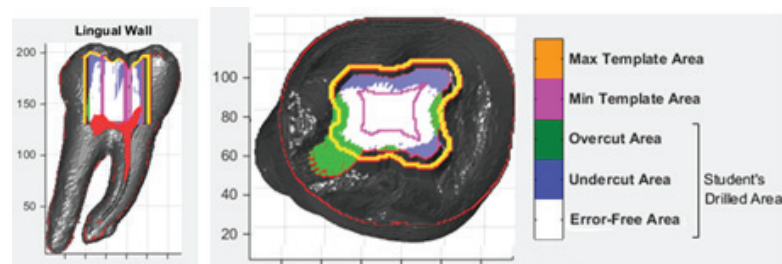


Figure 4.9: Min, Max and Optimal templates overlaid on student's drilling area from the lingual wall and occlusal view

Note: Overcut area indicates student's drilled area beyond optimal template; Undercut area suggests student's drilling needs further extension to get the optimal result; Error-Free area represents student's drilling within the optimal drilling area.

In experiment III the ideal drilling area based on the internal anatomy of the tooth along with the acceptable drilling areas were provided as visual guidelines for all

Table 4.1: Mean and standard deviation values for the scores of AOS and human raters in all experiments

	<b>R1</b>		<b>R2</b>		<b>R3</b>		<b>R4</b>		<b>R5</b>	
<b>Experiment I</b>										
Outcome	59.8	25.0	66.4	27.7	66.6	27.5	64.2	26.4	69.1	28.6
Axial	63.8	22.7	70.1	17.5	69.1	18.1	66.8	14.7	71.3	18.9
Floor	60.0	22.4	94.7	11.3	94.0	10.7	89.7	11.9	92.7	8.0
<b>Experiment II</b>										
Outcome	61.5	9.5	65.5	27.1	69.5	28.8	61.3	25.3	65.8	28.2
Axial	59.7	17.6	68.4	17.1	76.0	18.0	66.7	16.7	71.5	20.6
Floor	68.7	13.0	90.0	15.0	78.3	13.3	84.0	11.7	83.7	23.9
<b>Experiment III</b>										
Outcome	48.5	20.6	64.7	26.4	72.3	29.6	67.4	27.8	70.8	29.2
Axial	52.8	15.0	67.3	13.4	77.7	16.5	72.3	15.0	76.4	15.9
Floor	65.3	14.1	94.0	12.4	91.3	14.1	89.0	11.1	92.0	16.6
<b>AOS</b>										
Outcome	66.1	27.3								
Axial	76.4	15.9								
Floor	93.1	10.5								

axial walls and the pulp chamber floor from both lateral and top views separately (Figure 4.9). In each experiment, raters gave a score on the four axial walls and the pulp floor, with the overall outcome score then computed as the average of the five component scores.

## 4.3 Results

### 4.3.1 Bland-Altman plots

The performance data of access opening scores for human raters in three experiments are summarized in Table 4.1. The mean values for raters R2 through R5 are quite close, but those for R1 are consistently lower. The internal consistency among raters was further analyzed using a Bland-Altman plot.

The Bland-Altman plot is a scatter plot, in which the difference between the

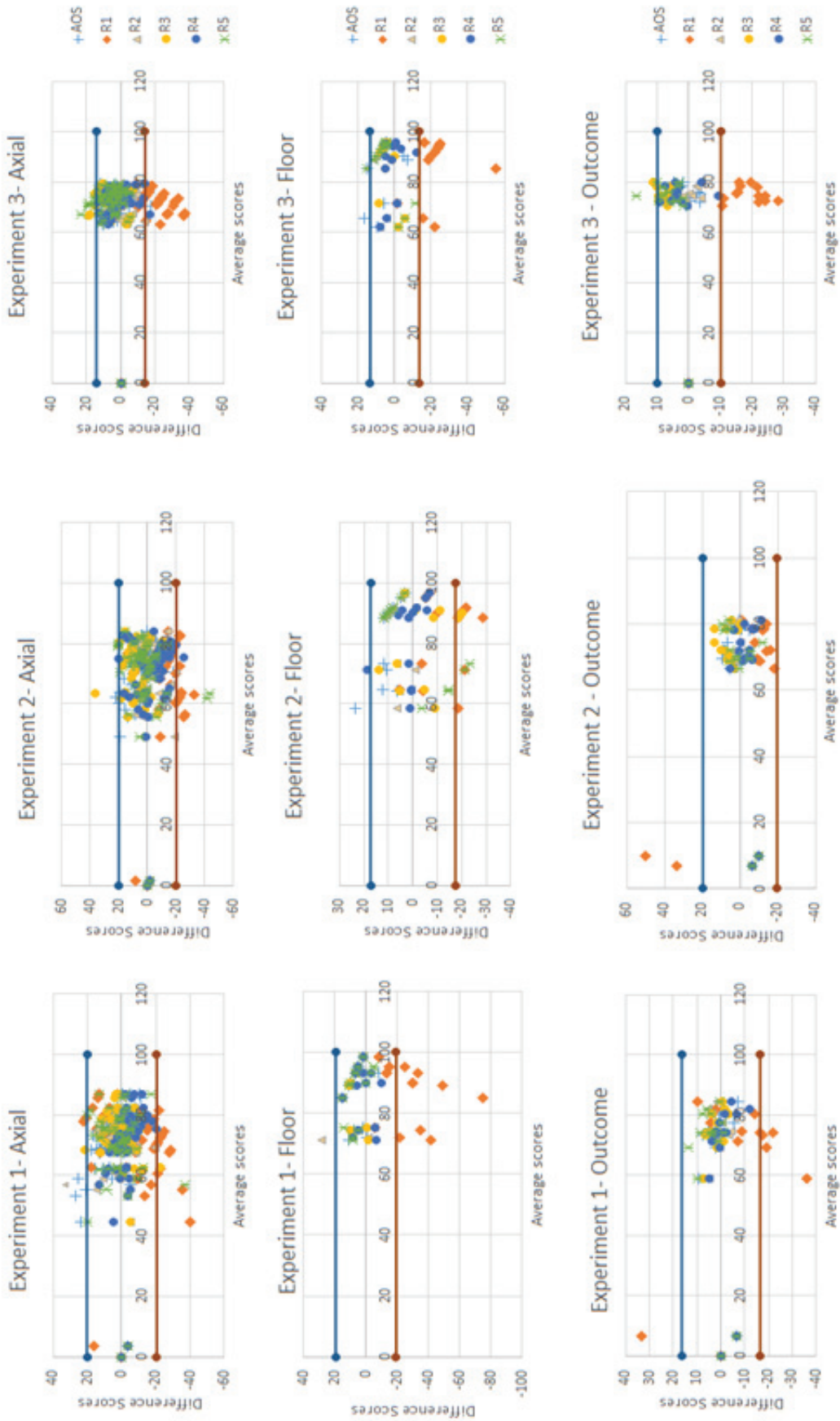


Figure 4.10: Bland-Altman plots of scores between AOS and human raters in three experiments

Table 4.2: Information-based disagreement between AOS and human raters with confidence intervals

	AOS			
	R2	R3	R4	R5
<b>Experiment I</b>				
Outcome	0.06 (0.04 , 0.01)	0.04 (0.02 , 0.06)	0.06 (0.04 , 0.10)	0.06 (0.03 , 0.09)
Axial	0.17 (0.14 , 0.21)	0.19 (0.16 , 0.22)	0.17 (0.15 , 0.21)	0.15 (0.13 , 0.19)
Floor	0.04 (0.01 , 0.09)	0.04 (0.01 , 0.09)	0.08 (0.05 , 0.15)	0.07 (0.04 , 0.12)
<b>Experiment II</b>				
Outcome	0.08 (0.05 , 0.11)	0.07 (0.04 , 0.10)	0.1 (0.07 , 0.15)	0.11 (0.07 , 0.17)
Axial	0.15 (0.12 , 0.19)	0.13 (0.11 , 0.16)	0.21 (0.18 , 0.24)	0.17 (0.14 , 0.22)
Floor	0.05 (0.01 , 0.10)	0.2 (0.15 , 0.28)	0.15 (0.12 , 0.19)	0.14 (0.06 , 0.27)
<b>Experiment III</b>				
Outcome	0.06 (0.05 , 0.09)	0.1 (0.06 , 0.12)	0.07 (0.04 , 0.11)	0.08 (0.05 , 0.12)
Axial	0.16 (0.15 , 0.19)	0.1 (0.09 , 0.12)	0.13 (0.11 , 0.16)	0.1 (0.09 , 0.13)
Floor	0.05 (0.02 , 0.11)	0.06 (0.02 , 0.12)	0.09 (0.06 , 0.13)	0.08 (0.03 , 0.15)

scores is presented in the vertical y-axis against the average between these scores on the horizontal x-axis. Two lines are drawn horizontally at the limits of agreement, which are defined as the mean difference plus and minus 1.96 times the standard deviation of the differences. With Bland-Altman plots, we can investigate the existence of any systematic difference between the raters and identify possible outliers. From Figure 4.10, it is observable that the rater R1's scores widely deviate from the other human raters. Rater R1 was thus eliminated from further analysis.

### 4.3.2 Information-based disagreement

To examine the degree of subjectivity in the scoring of outcomes, we evaluated the agreement among the five experts. Table 4.2 shows the information-based measure of disagreement (IBMD) [139] between AOS and human raters. The minimum disagreement (0.04) was found at the floor between AOS and R2 and R3 in experiment I. The maximum disagreement (0.21) between AOS and human raters was observed for rater R4 at the axial walls in experiment II.

An information-based disagreement value of 0.2 is observed for the floor score of rater R3 in experiment II as well. Post-experiment interviews indicated that this was likely due to difficulties in perception. The rater may have relied too heavily on

Table 4.3: Information-based disagreement between human raters R2 vs. R3, R4, R5

	<b>R2</b>	<b>R3</b>	<b>R4</b>	<b>R5</b>
<b>Experiment I</b>				
Outcome	0.05 (0.026 , 0.080)	0.07 (0.041 , 0.112)	0.06 (0.030 , 0.096)	
Axial	0.14 (0.113 , 0.172)	0.13 (0.108 , 0.163)	0.15 (0.121 , 0.192)	
Floor	0.04 (0.011 , 0.106)	0.08 (0.036 , 0.158)	0.08 (0.037 , 0.140)	
<b>Experiment II</b>				
Outcome	0.07 (0.041 , 0.116)	0.07 (0.042 , 0.114)	0.08 (0.061 , 0.118)	
Axial	0.16 (0.129 , 0.194)	0.18 (0.148 , 0.217)	0.17 (0.135 , 0.221)	
Floor	0.19 (0.126 , 0.253)	0.12 (0.094 , 0.155)	0.11 (0.037 , 0.214)	
<b>Experiment III</b>				
Outcome	0.12 (0.085 , 0.158)	0.07 (0.037 , 0.106)	0.09 (0.064 , 0.135)	
Axial	0.17 (0.145 , 0.194)	0.12 (0.104 , 0.150)	0.15 (0.133 , 0.185)	
Floor	0.06 (0.022 , 0.105)	0.06 (0.032 , 0.108)	0.03 ( 0 , 0.077)	

the cross-sectional view which was taken mid-volume, revealed only half of the drilling area, thus raters needed to estimate the depth of the drilling in the rest of the pulp floor.

Table 4.4: Information-based disagreement between human raters R3 vs. R4, R5, and R4 vs. R5

	<b>R3</b>	<b>R4</b>	<b>R5</b>
<b>Experiment I</b>			
Outcome	0.05 (0.035 , 0.076)	0.05 (0.028 , 0.090)	0.08 (0.049 , 0.139)
Axial	0.11 (0.085 , 0.135)	0.16 (0.121 , 0.196)	0.14 (0.116 , 0.188)
Floor	0.06 (0.036 , 0.104)	0.06 (0.029 , 0.112)	0.08 (0.031 , 0.141)
<b>Experiment II</b>			
Outcome	0.14 (0.090 , 0.184)	0.10 (0.058 , 0.160)	0.13 (0.095 , 0.171)
Axial	0.21 (0.186 , 0.257)	0.15 (0.117 , 0.201)	0.18 (0.141 , 0.225)
Floor	0.23 (0.174 , 0.328)	0.13 (0.084 , 0.188)	0.19 (0.143 , 0.272)
<b>Experiment III</b>			
Outcome	0.08 (0.047 , 0.132)	0.04 (0.025 , 0.068)	0.07 (0.036 , 0.131)
Axial	0.12 (0.097 , 0.150)	0.11 (0.094 , 0.134)	0.12 (0.091 , 0.152)
Floor	0.08 (0.050 , 0.133)	0.04 (0.009 , 0.101)	0.1 (0.071 , 0.145)

Table 4.5: Mean Absolute Error of pairs of AOS and human rater scores

	AOS					R1					R2					R3					R4				
	R1	R2	R3	R4	R5	R2	R3	R4	R5	R3	R4	R5	R4	R5	R5										
<b>Experiment I</b>																									
Outcome	17	4	3	4	4	15	16	16	17	3	4	4	3	4	5										
Axial	20	11	12	11	10	16	17	15	17	8	8	10	6	10	9										
Floor	33	3	3	6	5	35	34	30	33	3	6	6	4	6	6										
<b>Experiment II</b>																									
Outcome	18	5	5	6	7	17	21	15	18	5	4	5	9	6	8										
Axial	19	10	9	13	11	13	18	15	17	10	11	11	14	10	11										
Floor	24	3	15	10	9	22	12	16	22	14	9	6	9	18	13										
<b>Experiment III</b>																									
Outcome	18	4	6	5	5	16	24	19	22	8	4	6	5	3	5										
Axial	24	10	6	8	7	15	25	20	24	11	8	10	8	7	8										
Floor	28	4	4	7	6	29	26	24	28	8	10	4	6	3	7										

### 4.3.3 Mean Absolute Errors

Mean Absolute Error (MAE) values were computed to quantify the degree of disagreement. As shown in Table 4.11 and Table 4.5, MAE values between AOS and human raters are higher in the axial scores than the overall and floor scores in experiments I and III while large MAE values for floor scores were observed in experiment II. The minimum MAE value between AOS and human raters across all experiments was 3, and the maximum was 15. The values of MAE also are correlated with the degree of disagreement between AOS and human raters observed with IBMD. For example, Raters R2 and R3 in experiment I have the minimum IBMDs and the minimum MAEs.

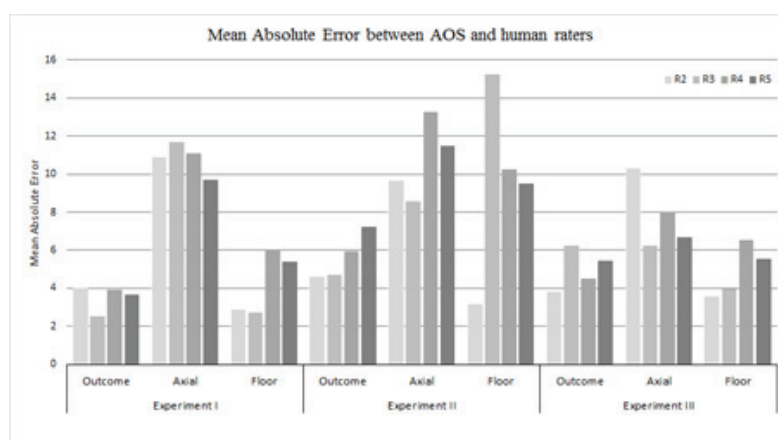


Figure 4.11: Mean Absolute Error between AOS and human raters

Table 4.6: Chi-square values from Kruskal-Wallis tests for R2, R3, R4, R5 (Significant results are indicated with the bold-faced fonts)

Raters	R2	R3	R4	R5
Outcome	0.326	5.215	6.857	1.374
Axial	1.656	<b>21.551</b>	<b>21.551</b>	<b>6.973</b>
Floor	1.228	<b>12.726</b>	4.239	1.847

#### 4.3.4 Perceptual effects

In evaluating the AOS, we are aware of the fact that any differences in rating between AOS and the human raters could be due to limitations of perception or to differences in judgment. To confirm the perceptual effects of additional information provided in each experiment on the scores, the non-parametric Kruskal-Wallis tests (two-sided) were conducted. Kruskal-Wallis test determines if there are statistically significant differences between the assessment scores across three experiments for each rater. Dunn (Dunn-Bonferroni) post hoc method was applied following a significant Kruskal-Wallis test. All statistical analyses were performed using SPSS 18 (SPSS, Chicago, IL, USA) and statistical significance was defined as a p-value less than 0.05.

Table 4.6 shows the Chi-square (2) values of Kruskal-Wallis tests. For the overall outcome scores, there were no statistically significant differences among raters across the experiments. No statistically significantly different scores across all experiments were found for rater R2 meaning that R2 scored consistently throughout the three experiment regardless of the level of information provided. On the other hand, significant differences were found between experiments for raters R3, R4 and R5 in axial walls scores. The post hoc analysis revealed that experiment I axial wall scores of R3 were significantly different from experiment II. For the rater R4, axial wall scores from experiment III were statistically different from experiment I and II. For the rater R5, significant differences were found for the scores between experiment I and III. From the post-hoc results, there is no clear pattern of increase in response to increasing amount of information provided to the raters, and we can conclude that the differences in scores between AOS and the human rater would less likely be due to the limitations of perception.

### 4.3.5 Discriminatory Power of the AOS Scores

Our final analysis examines the extent to which the AOS scores can discriminate among the various grades assigned by human raters. AOS scores can be viewed as a measure that indicates to which grade the drilled tooth should be assigned. If we view one tier of a grade as the positive class and one as the negative class, then the sensitivity and specificity of the measure can be adjusted based on the threshold one chooses, with sensitivity typically traded off against specificity. ROC analysis allows the discriminatory power of a test or measure to be evaluated independent of any arbitrary threshold [140]. The basic concept is to plot the sensitivity in function of 1-specificity for all possible thresholds. The area under the resulting curve (AUC) is an indication of the discriminatory power of the measure. A perfect measure will have an AUC of 1 while a measure that is no better than a random guess will have an AUC of 0.5. For this analysis, the grades are assigned using the scheme: A ( $\text{score} \geq 80$ ), B ( $70 \leq \text{score} \leq 79$ ), C ( $0 \leq \text{score} \leq 69$ ).

Table 4.7 shows the Area Under Curve (AUC) values of the AOS scores for the two cutoffs among the three grades: A:BC, AB:C. An AUC value should be above 0.8 to be regarded very good performance for any test. Low AUC values of (0.5 AUC 0.8) are observed for all axial walls across three experiments for all raters for the cutoff A: BC. This is not surprising since the small AUC values coincide with large mean absolute error values in Table 4-5. The highest AUC among pairs is a perfect 1.00 for all raters for AB:C in outcome scores of experiment I. In fact, among the cutoffs, the AUC values are highest for the AB:C cutoff at the boundary between clinically acceptable (B) and clinically unacceptable (C). One possible explanation is the relatively broad range of values of scores in the C category, making it easier to differentiate between that category and the others.

## 4.4 Discussion

In root canal treatment, the internal anatomy - pulp and the roots in particular - dictate the size, shape and the location of the access opening. In root canal treatment, an access route to the cavity cannot be assumed to take a predetermined geo-



Table 4.7: ROC AUC values for AOS scores over all grades in all experiments

	Rater 2		Rater 3		Rater 4		Rater 5	
	{A:BC}	{AB:C}	{A:BC}	{AB:C}	{A:BC}	{AB:C}	{A:BC}	{AB:C}
<b>Experiment I</b>								
Outcome	0.8	1	0.9	1	0.6	1	0.8	1
Axial	0.7	0.7	0.6	0.7	0.5	0.6	0.7	0.8
Floor	1	-	0.9	-	0.9	0.8	-	-
<b>Experiment II</b>								
Outcome	0.7	0.7	0.9	1	0.5	0.8	0.7	0.6
Axial	0.7	0.7	0.8	0.8	0.6	0.6	0.7	0.7
Floor	0.9	0.9	0.9	0.9	1	1	1	1
<b>Experiment III</b>								
Outcome	-	0.9	0.8	1	0.7	0.8	0.9	1
Axial	0.6	0.8	0.7	1	0.6	0.6	0.6	1
Floor	1	-	0.9	0.9	1	-	1	1

Note: '-' indicates AUC is not computed because the rater's scores did not fall within the interval of a certain grade tier. For example, R2's floor scores are above 70 for all samples, and thus the AUC between AB: C cannot be computed.

metric shape. The anatomy of the pulp chamber of a tooth determines the access cavity shape. Many dental students find endodontic treatment challenging to learn because of root canal anatomy variation among individuals [125, 141]. Limited anatomical variation in plastic teeth potentially constrains the students exposure to multiple anatomical and pathological specimens essential for skill training [29, 30]. Extracted teeth are available as an alternative; however, with hygienic and ethical concerns, this is not a viable skill training option. In our approach, with the virtual tooth models created from CT scans of teeth and templates created on the fly, a variety of realistic training teeth can be employed.

In dental schools there is much discussion of standardization of training and assessment [21, 142–146]. Our evaluation of expert scoring highlights the significant variation that is typically found. Subjective definition of the axial wall regions of the tooth noted earlier could be one of the factors leading to such variation in scores. This is seen in the disagreement in scores at the junctions of adjacent walls among the experts and between AOS and the experts. The standardized AOS definition of axial wall regions helps to address this issue. Our AOS has the potential to help create easily communicated standards for assessment. The fact that the templates are related to the anatomical structure in a very clear way also helps to provide an objective basis for

discussion of standards for endodontic surgical outcomes.

In assessing outcomes of endodontic surgery, it is essential to identify both areas of under preparation and of over-preparation. Under preparation can lead to instrument breakage in subsequent stages or the trainee might end up with inaccessible roots. Over-preparation can result in weakened tooth structure due to excessive removal of tooth mass. The inclusion of Max/Min-templates in the score cube addresses both injuries and underprepared areas. Scoring metrics based on the identification of injuries [49, 53] do not capture regions of under preparation and do not identify near the injury. Metrics based on the amount of tissue removed [137] recognize neither areas of over nor under preparation and cannot provide nonlinear scoring. Our results suggest that if similarity with expert judgment is a desired criterion, researchers using template-based approaches should consider the use of non-linear scoring functions. The use of reference templates seems most suited to this.

The objective consistency of AOS scores and a high degree of agreement with experts make it a promising addition to existing VR simulators. The translation of detailed level scores into terminology commonly used in dental surgery supports natural communication with students and instructors. With the reference virtual templates created automatically, our scoring cube based approach is robust and is not limited to the access opening procedure in root canal treatment. It is applicable in scoring the outcome of any dental surgery procedure involving the act of drilling or milling.

One of the advantages of a scoring system to a training program is that the trainee can review specific areas of weakness in the task with reference to the scoring system. The outcome assessment can be done with or without the presence of an examiner. The use of automated outcome scoring for assessing a student task's outcome could provide increased accuracy in scoring and grading. Although we did not show in this study, the scoring system is general enough to use with a variety of training teeth. In our study, we have demonstrated the scoring approach using a molar tooth with four root canals. Other than molar teeth if the training tooth is chosen either as premolar, canine and incisor type, the templates can be created in the same manner without requiring any further changes as they have less complex root canals system consisting of a maximum of two roots. To use the scoring system with a new training tooth, the challenge lies

in the correct identification of pulp chamber inside the tooth. With one - two orifices, the identification of pulp chambers in incisors, canines and premolar is relatively simple compared to the molar tooth. Depending on the user's requirement for the scores such as the way the scores are penalized with the weight, it is optional to fit the non-linear weight function again for the new tooth. Templates can be created based on the identified pulp chamber for a new training tooth and the score cube can be created with the templates in a straight forward manner. Once the score cube is in place, scoring can be done on the student's preparation using the new training tooth.

In this chapter, we presented an approach that provides scores to the 3D voxel structure commonly used in VR dental simulators at a sufficient level of detail to allow correlation with procedure kinematic variables collected by such simulators. The fine-grained voxel level scores provide the precise error information that can be difficult to quantify in irregular 3D objects such as teeth with complex internal anatomy. To efficiently communicate outcome score results, detailed level scores are translated into the language of the coarser level standard scoring system used by dental schools. The algorithm has been implemented for the procedure of access opening to the root canals. Agreement between system scores and those of expert endodontists is evaluated on fifteen outcome samples with a range of error types and severity. Results show a high degree of agreement between system scores and those of experts, while at the same time highlighting the variability in the subjective expert judgments.

## **CHAPTER V**

### **CORRELATOR: CORRELATING PROCEDURE AND OUTCOME**

Having identified possible errors in the outcome and their causes using the automated outcome scoring system, we will outline our approach to correlating the procedure and the outcome in this chapter.

The problem of correlation procedure and outcome is an instance of the well-known credit assignment problem. It is concerned with determining how the success of a system's overall performance is due to the various contributions of the system's components [134]. In his seminal work, Minsky exemplifies the problem with chess (checker) board game. In the game of chess, the player would receive a reinforcement signal after a long sequence of moves. During play, each ultimate success (or failure) is associated with a vast number of internal decisions. If the run is successful, appropriate credits should be assigned to individual moves resulting from the multitude of decisions for their contribution to success or blames in the case of loss. In situations in which the assignment of credit is extended over time as in a chess game, the problem is called as the temporal credit assignment problem.

Formally, the temporal credit assignment problem concerns determining which of the past actions were responsible for an eventual success (or failure). In reinforcement learning, temporal credit assignment is mostly attempted with temporal difference (TD-) learning algorithms [227]. It is based on the difference between temporally successive predictions. An initial prediction is made first, and when the observation is available, the prediction is fine-tuned in accordance with the observation. In short, TD-learning estimates the value of state-action pairs. The value itself is the (discounted) expected accumulated future reward when taking a particular action in the state. The value can then serve as an immediately available alternative for the delayed reward signal. In essence, the task of correlating procedure and outcome is considered as

credit assignment problem. However, the function cannot be implemented with existing reinforcement learning based approaches straightaway due to the following reasons. Existing approaches from reinforcement learning are based on learning the optimal policy (stage-action pairs) and emphasize the temporal aspect of each action. In this regard, existing approaches are more applicable for a study that involves learning such as training a robot on how to perform dental surgery. This thesis focuses on the direct assignment the credit/blame to the associated actions while bypassing the learning stage.

The original credit assignment problem was conceptualized on apportioning credit /blame for the various actions leading to the outcome: success/failure. During execution, some activities could be explicitly recognized as wrong actions. Among a multitude of such activities, a specific portion of them may be considered as secondary actions resulting from the domino effects of the one game-changing primary action. Intuitively, the credit/blame should be assigned to the central action that necessarily leads to the secondary actions. The correct credit/blame assignment to the right action is particularly important for our approach as the assignment information will be subsequently used in providing feedback. It poses a challenging problem as well since the procedure and outcome could not be correlated instantly regardless of the availability of localized error information.

In motor skill acquisition, the correct assignment of credits to actions that are responsible for the success is essential for shaping learning. Due to its relevance, the credit assignment problem has been extensively investigated in motor control learning studies. Berniker and Kording [20] examine credit assignment regarding allocating the cause of observed errors to changes in the properties of the body versus the surrounding world. In their study, probabilistic Bayesian models are built to allocate the errors in proportion to the optimal estimate of where the errors arise.

Dam and colleagues [43] investigate how people solve the credit assignment problem while learning a motor skill with reinforcement. They quantitatively vary reward functions and examine whether and how people assign credits to different movement properties during movement reinforcement learning. Intelligent Tutoring Systems (ITS) also widely study the idea of modeling error assessment as credit/blame assignment problem as a function to provide feedback. In designing an Intelligent Tutoring

System for Introductory Physics, Liew et al. [119] describe an approach to solving a credit assignment problem to used as a function to generate feedback that is specific to the problem context. Similarly, Mengshoel and Wilkins [132], use an interactive multimedia-based crisis management ITS in identifying types of errors made by damage control assistant students and study with credit assignment problem to recognize and giving them feedback on those errors. Bayesian networks and techniques from Uncertainty subdiscipline of Artificial Intelligence are widely used to address the credit assignment problems in ITS [43].

## 5.1 Error Regions in the Outcome

The outputs of the automated outcome scoring system are numeric scores range between 0 and 100 and the error log detailing the location (x, y, z) and the type, and the location the errors in the outcome. An example of the outcome scores and the error information for the four axial walls, the pulp chamber floor and overall scores along with the error regions is shown in (Figure 5.1).

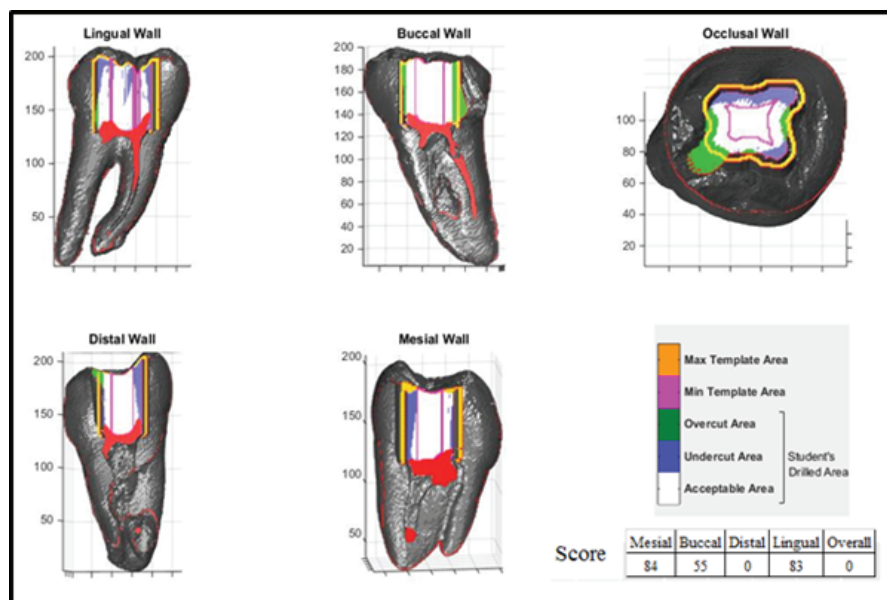


Figure 5.1: The outcome scores and the error information for an outcome

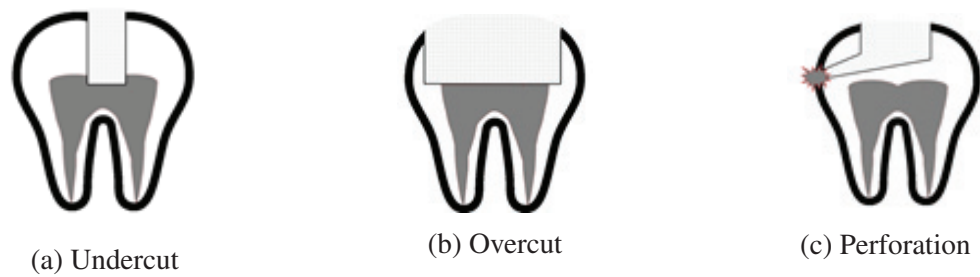


Figure 5.2: Operative errors

## 5.2 Types of Errors and Causes

As described in Chapter 3, the formative feedback from the framework is designed to contain the three aspects of error: the types (what), the location (where) and the time they are committed (when). In practice, several mishaps or errors can arise during the root canal treatment, ranging from treating the wrong tooth, missed canals to identify, damage to existing restoration, under-/over-extension and access cavity perforations to crown fractures [126]. To simplify the process of diagnosing the errors in the outcome, we define the types of errors (Figure 5.2) to be included in the assessment process as follows:

- Undercut: when the dentist drills a hole with a small diameter, the roots remain inaccessible
- Overcut: when the dentist drills a large mass of tooth unnecessarily
- Perforation: when the dentist accidentally punches a hole through the tooth surface or gum with an instrument

Once the errors in the outcome are localized, the next step is to identify the originating source actions for each error in the procedure that are accountable for them and assign the blame to them. However, there could be more than one underlying cause (action) contributed to each type of error in the outcome. To provide the formative feedback, it is necessary to understand the underlying actions that could lead to the errors. Therefore, for each operative error, we study the possible causes as follows.

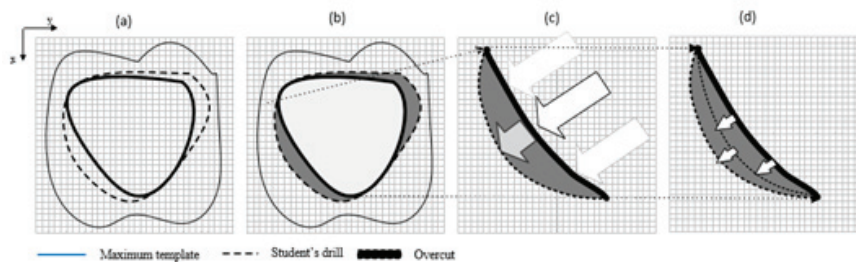


Figure 5.3: Overcut example

(a) Maximum template and the students drilling area (b) Overcut area (b) Inappropriate amount of force applied in the area where the student should stop drilling leading to the overcut error (c) Right amount of force is applied but repeated drilling in the same area

- Case 1: Overcut As shown in Figure 5.3, the overcut case occurs when the student's drilling reaches to the area beyond the maximum area defined by the standard virtual templates (Figure 5.3 (a)). In the resulting outcome, the overcut regions will be recognized as an overcut error (the filled area in Figure 5.3 (b)). Although the filled area is labeled as errors, the actions taken in that region could not be immediately labeled as wrong actions. In fact, two conditions could lead to the overcut errors.

- Case 1.1: Improper amount force was exerted on the instrument or inappropriate orientation of the instrument at the area where the student should have stopped drilling (marked with the virtual maximum template 5.3 (c)).

- Case 1.2: The right amount of force was applied, but the student did not recognize the area to stop drilling and repeatedly drilled at the same region with the uniform force and with fixed tool orientation. Repeated drilling in the same neighborhood eventually leads to an overcut error in that area 5.3 (d)).

- Case 2: Undercut In contrast to Case 1, undercut cases occur when the student did not clear the internal tooth anatomy entirely as required. The undercut regions can be determined by the optimal virtual template as shown in Figure 5.4(a). The possible condition of this error is

- Case 2.1: Inappropriate orientation of instrument or an insufficient number of passes or improper amount of force exerted at the student's drilled area prevented



him from reaching to the optimal drill area Figure 5.4 (c).

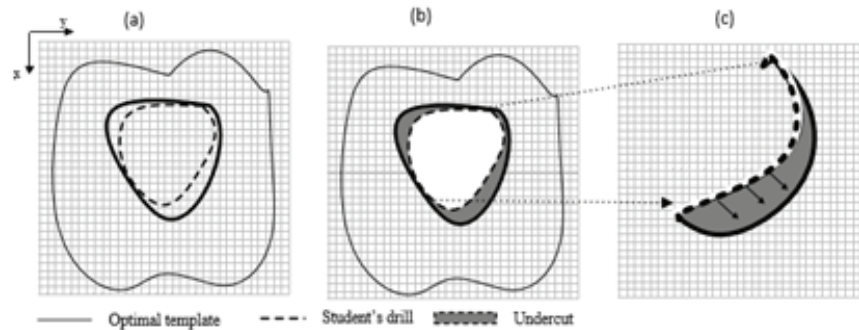


Figure 5.4: Undercut example

(a) Optimal template and the students drilling area (b) Undercut area (b) Inappropriate amount of force applied in the area where the student should stop drilling

- Case 3: Perforation The perforation cases occur when the student's drilling area reaches beyond the virtual maximum template, and the instrument punches through one of the tooth walls, resulting in an irreversible hole in the wall as shown in Figure 5.5 (a-b). Similar to over-cut cases, perforation cases can be occurred from

- Case 3.1: Improper amount force was exerted on the instrument or inappropriate orientation of the instrument at the area where they should have stopped (marked with the virtual maximum template Figure 5.5 (c)).

- Case 3.2: The right amount of force was applied, but the student didn't recognize the area to stop and drilled over and over at the same region with the uniform force with the fixed tool orientation as in Figure 5.5 (d). Repeated drilling in the same neighborhood could eventually lead to over-cut error in that area.

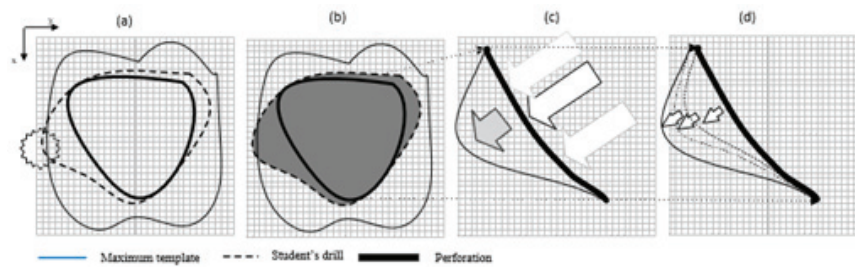


Figure 5.5: Perforation example

(a) Maximum template and the students drilling area (b) Perforation area (b) Inappropriate amount of force applied in the area where the student should stop drilling leading to the perforation error (c) Right amount of force is applied but repeated drilling in the same area causing perforation

Through cases 1-3, we can conclude that errors in the outcome are caused by either the inappropriate amount of force applied or the incorrect orientation of the instrument. Errors identified in the outcome are considered to be caused by the actions taken during the procedure. Additionally, the errors are contributed from the omitted actions as well. Consider overcut errors occur from over drilling of the tooth (drill actions were taken by the trainee), on the other hand, undercut errors occur at the area of the tooth when the trainee did not remove as required (drill actions are omitted). As omitted actions are not explicitly present in the procedure, an additional step is necessary to identify the neighboring portions of the procedure of the excluded actions before assigning the credit/blame.

For some errors, the responsible portions of the procedure can be identified based on the information obtained by correlating the outcome and the procedure. On the other hand, some errors caused by a consequence of more holistic characteristics of the procedure such as the incorrect tool angulation throughout one stage due to incorrect finger positions and misunderstanding of the sequence of stages, are more challenging to identify and do not address in this current formative feedback system design.

### 5.3 Procedure and Outcome Correlation

The details on the procedure and outcome correlation inside the correlator component are shown in Figure 5.6. The error information from the automated outcome

scoring system are at the voxel level. We assume that providing feedback at the lowest voxel level is tedious and unnecessary. Therefore, the voxel level error information is grouped into cluster levels. With a simple flood fill segmentation method, the errors in the 3D tooth volume are segmented to get the coarser cluster level representation. Grouping error voxels into regions of under and over cutting enables generating feedback in a natural manner. After discussion with the expert and through experiments, the clusters consisting of less than 50 voxels are considered as the minor errors which don't attribute to the performance and discarded from further analysis.

For each error cluster identified in the outcome, firstly spatial location is mapped to the collided voxel log to determine the relevant portion(s) of the procedure. Actions over multiple parts of a procedure may be responsible for a single error. For example, the overcut error identified at the mesial wall in Figure 5.7 is caused by repeated visits to the same region with the tool at different stages of the procedure. Some error clusters may also spread across more than one wall, and a single wall may contain more than one type of error.

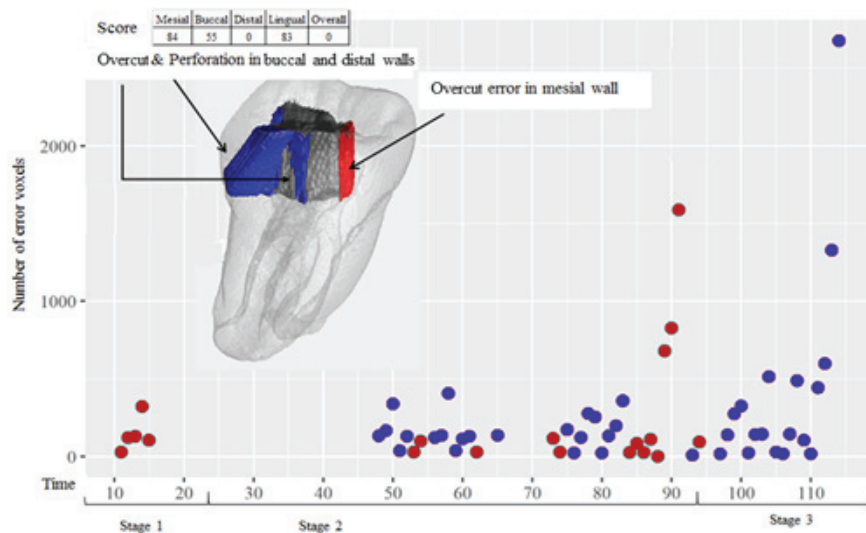


Figure 5.7: Outcome scores, errors and temporal information in procedure

In order to map the errors with the portions of the procedure, it is necessary for the correlator to obtain the timestamps at which the errors voxels are drilled out. Using the record of collided voxels over time, the temporal information of each voxel is gathered and the portions of the procedure associated with the errors are identified. Fig-

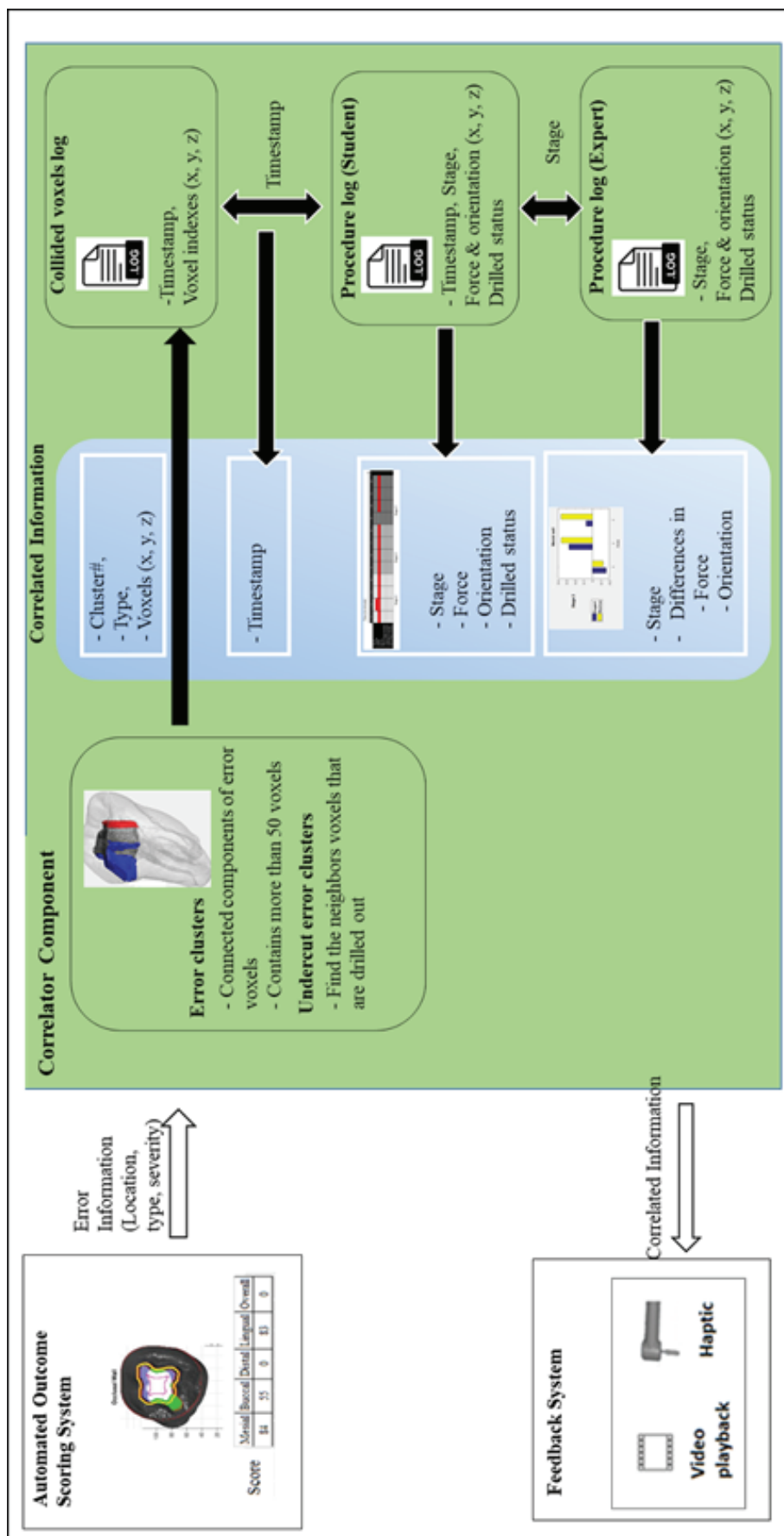
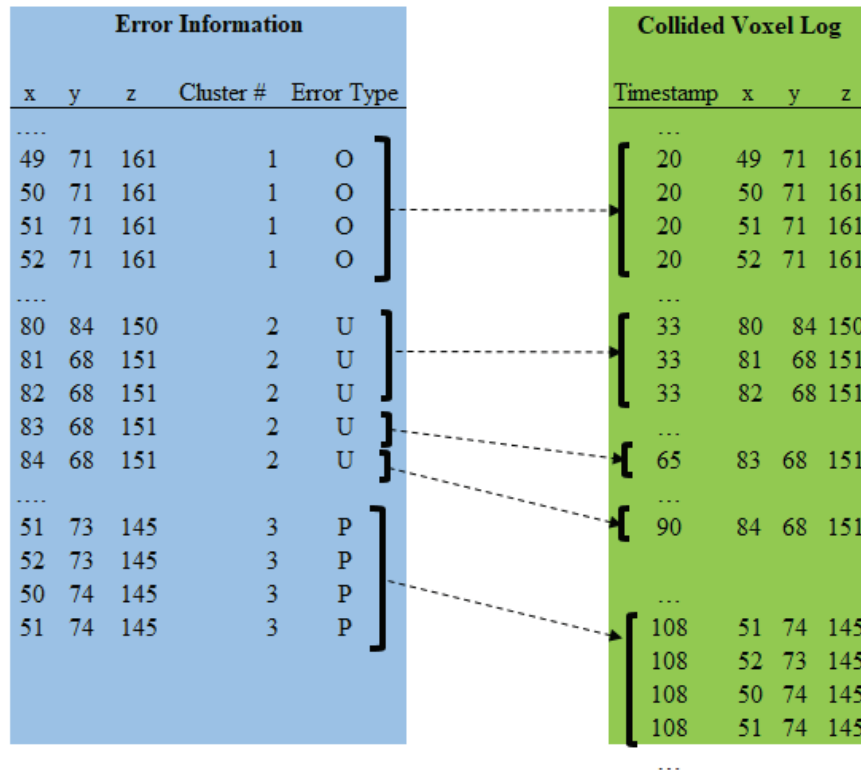


Figure 5.6: Procedure and outcome correlation flow

Figure 5.8 shows an example of mapping between the error voxels from error information from AOS Output and the collided voxel log.



\*Error type:  
 O - Overcut, U - Undercut, P: Perforati

Figure 5.8: Mapping between error information and collided voxels log

The correlator then maps the spatial information of errors (voxel indexes) into the walls. Figure 5.8 shows an example of the overcut error voxels mapped onto the respective walls and the portions of the procedure using the spatial and temporal information of errors. For the overcut and the perforation error clusters types, the error related timestamps are directly obtainable using the error voxels since the errors are caused by the drilling actions. For the undercut error clusters, the mapping cannot be done straightforwardly due to the absence of collided voxels for the undercut errors caused by the omitted actions. Therefore, the correlator firsts lookup the neighboring voxels (20 voxels in x-, y-, z- axes) for each voxels lies at the border of undercut errors with the drilled area. Using the neighboring drilled out voxels, the timestamps of the undercut error clusters.

As the length of the collided voxel log is directly proportionate to the number of voxels removed during the procedure, the larger log files take long time to carry out the mapping. Therefore, for the error clusters with more than 100 voxels (determined through experiments), error voxels of the same clusters are sampled by taking every other five voxels (determined through experiments). The worst scenario in mapping is when the error voxels are drilled out different portions of the procedure, however, since the drilling can perform in one direction from the top occlusal surface towards the pulp floor only, this issue is solved by a proper simple indexing error voxels.

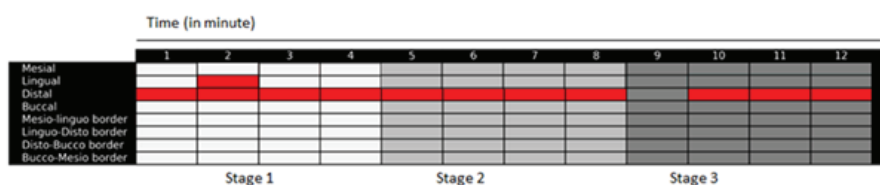


Figure 5.9: Portion of the analysis of wall, error occurrence in time (Note: Occurrences of overcut errors are marked in red color.)

After mapping the error information with the procedure, the identified portions of the procedure are extracted to analyze further with the kinematic variables on the applied force and the orientation on the instrument. In the absence of standard amount of force and orientation, the challenges come in determining whether the applied force is appropriate or the tool orientation is correct or not. The correlator tracks the trajectory of the instrument in the stages containing the timestamps related to the errors and determine the appropriate amount of force in comparison with the expert. The trainees tool orientation is analyzed in the same manner. The mapping information on each error cluster: the types of errors, their locations on the outcome, the collided timestamps during the procedure, the differences in applied force and the orientation of the instrument relative to the expert in x-, y-, z-axes are combined and forward to the Feedback system component. The pseudo code of the correlator component is shown in Figure 5.10.

```
Define Minimum_Cluster_Size
Define Neighbor_Voxels_Num

For each voxel marked as an error in the outcome
  Obtain connected components of the same error type and form error clusters
  Discard clusters with size less than Minimum_Cluster_Size

For each undercut error cluster
  Extract the cluster border voxels
  Get neighbor voxels which were drilled out (within Neighbor_Voxels_Num voxels in
  x-, y-, z- direction from each border voxel)

For each error cluster
  Get spatial information (voxel indexes) of each voxel in the cluster
  Get temporal information (timestamps) by mapping the indexes to the collided voxel
  log
  Extract kinematic variables (Force, Orientation) from procedure log using the
  timestamps
```

Figure 5.10: Pseudo code of the correlator component

## CHAPTER VI

### VIDEO-BASED FORMATIVE FEEDBACK SYSTEM

In this chapter, we present the video-based formative feedback system using a dental VR training simulator. First, we review the literature pertaining to video-based feedback in medical and dental skill training. Video recordings have been used in skill training assessment and feedback in medicine [147–150]. Muessig et al. [151] argued that video feedback could be more useful in the acquisition of technical, practical skills than in non-technical skills. Vyasa et al. [152] reported that trainees who were able to watch experienced surgeons complete endoscopic tasks benefitted more than trainees who watched and critiqued their own performance, or trainees who simply completed more repetitions. While majority report positive results on using video in the feedback process, Backstein et al. [153] failed to demonstrate an improvement in orthopedic skills using video feedback. Byrne et al. [154] also found the use of video feedback technique as effective, but not more so, as traditional feedback in a simulation of general anesthesia.

Experiential learning and reflection are the two processes through which trainees learn from training practices [155–157]. Reflection at the end of the procedure with the video playback facilitate the trainees to look at themselves from a distance and with space for reflection, thereby giving them a realistic picture of their skills or self-image [158, 159]. A study by Fierman et al. [77] demonstrated that self-observation promotes the acquisition and transfer of procedural knowledge necessary for problem solving. Strandbygaard et al. [146] concluded that that viewing ones performance, later on, might even promote reflective practice. However, evidence of the gain from observing recorded operations is still sparse.

According to a recent survey on existing VR dental simulators for skill training by Wang et al. [58], the ForssLund simulator [137], the HapTel simulator [53], and the Simodont simulator [51] have the replay feature which allow the student or instructor



to watch in full replay mode upon completion of a procedure. The existing systems provide the simple playback of the procedure carried out using the simulator without any formative feedback added. The need for recording and playback of training exercises in VR training simulators in dental education has long been recognized as the requirement [58, 160].

## **6.1 Video-based Formative Feedback System**

The formative feedback generated in this thesis contains the errors of different types located in the different regions of outcome, and the portions of the procedure, which are identified as the origins of the errors, are also temporally distributed across the procedure. To effectively communicate both spatial and temporal aspects of the performance errors, the video modality is considered as the most appropriate modality to convey the feedback. Hence, the video-based formative feedback system was implemented as a modality to provide feedback using the dental VR skill-training simulator. Video-based Formative Feedback System was implemented as the graduating project of the undergraduate students group consisting of Farin Kulapichitr, Varistha Jatuwat, Nuttanun Uthai pattanacheep from Faculty of Information and Communication Technology, Mahidol University. The formative feedback information was provided by the author, while the development team was mainly responsible to upgrade the original playback module of the simulator by integrating formative feedback with the additional video features.

### **6.1.1 User Interface Design**

The video playback interface consists of four main components: video control panel, mode control panel, the simulation panel and viewing aspect control panel (Figure 6.1).

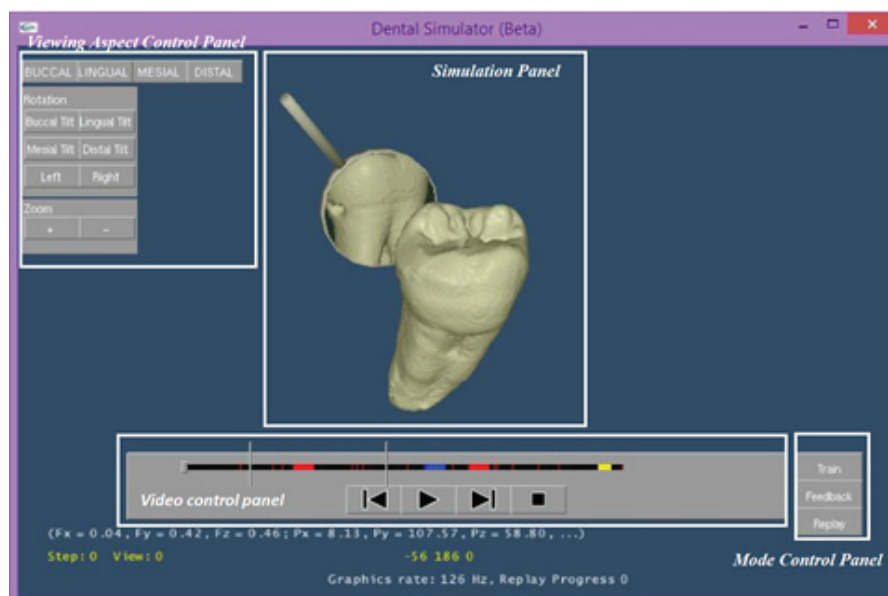


Figure 6.1: Video playback user interface

### 6.1.2 Video control panel

The video control panel offers access to several standard buttons including Play, Stop, Skip Backward and Skip Forward. The Play button toggles into Pause while the video is being played (Figure 6.2).

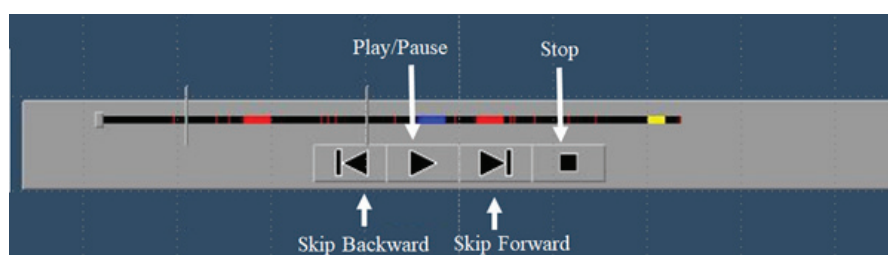


Figure 6.2: Video control panel

The skip forward and backward buttons are used to skip to the next/previous error in the video or the beginning of nearest next stage whichever comes first. They allow the user to quickly and efficiently jump to the point of the error or stage. The trainee can rewind as needed to comprehend the association between instrument movement and errors. By fast-forwarding through portions of a procedure that may not contribute to the overall assessment could reduce the time needed to complete the playback. The expert will benefit from this feature to use as a supplement in an evaluation process as well.

As the video is being played, a video progress bar (a.k.a scrubber/scrub bar/seek bar) is highlighted with red, blue, and yellow colors to denote the overcut, undercut and perforation errors respectively (Figure 6.3). The types and the temporal information of the errors are obtained from the correlator component.

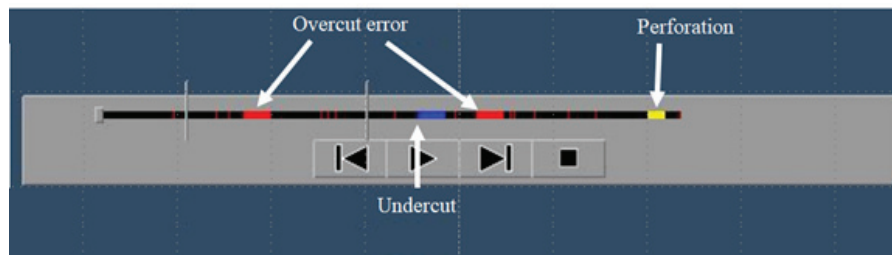


Figure 6.3: Example view of video progress with different types of error

Accordingly, the colors coded error regions serve as formative feedback that informs the users where the error was made, thus allowing them to rewind the video to the section of the error at any time.

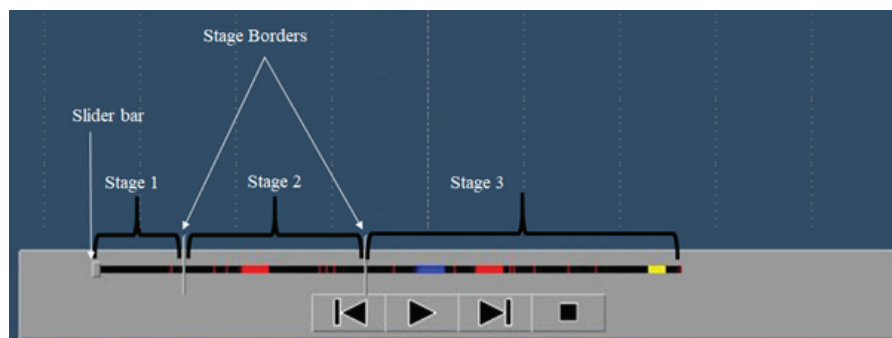


Figure 6.4: Stage Borders on the video progress bar

Additionally, as shown in Figure 6.4, two Stage Border vertical bars are drawn on the progress bar to represent the three stages involved in the access opening procedure mentioned under Simulation and Task discussed in Chapter 3. Stage Indicator(s) informs the user with

- a reference point for the current stage being played
- the time spent in each stage
- the time (when) and the type of errors (what) committed in each stage.

The last component of the progress bar is the Slider bar that provide the drag and drop functionalities of the mouse cursor to navigate to the desired region of the video playback quickly. The video progress bar succinctly presents the feedback information obtained from the correlator component. It informs the trainee with where the errors are in the outcome, what types of errors are they, when/in which stage the errors happened in the procedure.

### 6.1.3 Simulation panel

The simulation panel (Figure 6.5) hosts the video replay of the procedure consisting of the tooth, the handpiece, and the mirror. The formative feedback will feature on the graphics rendering of the tooth. As it is being drilled, with respect to the timestamps and the color coded region on the progress bar the portion of the tooth will be highlighted accordingly. Blue color represents the area of the tooth with undercut error, Red color represents the overcut area and Yellow color for the area with the perforation error. In the default video replay mode, the original tooth in opaque mode is displayed; however, for the better understanding of feedback, the tooth volume is recommended to switch to the transparent mode. This feature is supported by a shortcut key t/T.

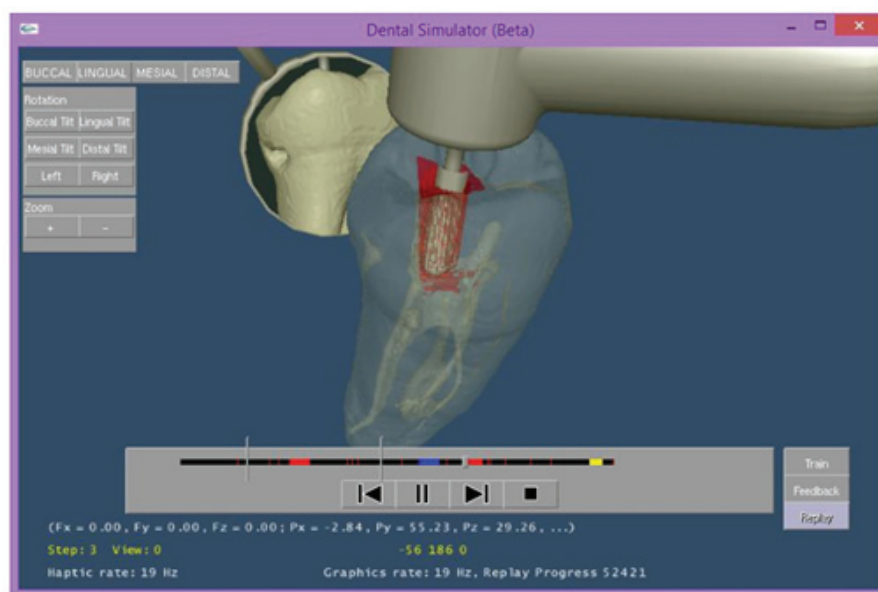


Figure 6.5: Example undercut error region on the tooth in the Simulation Panel

### 6.1.4 Mode control panel

This panel contains three more buttons to switch between two modes: Feedback, and Replay. The Feedback mode integrates the video playback system with static formative feedback according to the performance achieved by the user. This mode consists (a) video playback, and (b) the kinematics comparison graphs (Figure 6.6) with regard to the force applied to the tooth (along x, y, z-axis), the orientation of the driller (along x, y, z-axis). Kinematics graphs show the comparison between the trainee (green bar) and the expert (yellow bar). The video playback shows the tooth, and the errors of drilling are highlighted in specific color. The graph is generated by using the average values of the orientation and the force during each particular procedure stage.

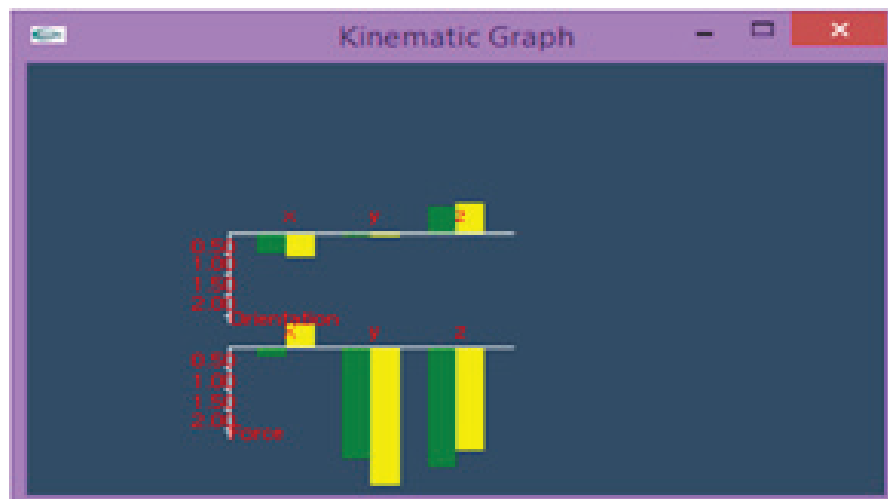


Figure 6.6: Example kinematic graphs (top: Orientation, below: Force)

We hypothesize that having the feature that displays how the expert carried out the same procedure would be a benefit to the trainees. Therefore, in the Replay mode, the system allows the user to view his/her performance in comparison with video playback the expert performance (Figure 6.7). The teeth in both windows are set to be transparent to increase the comprehension of the changes in tooth internally as it is being drilled. The expert window can be activated by pushing the button b, and all the video and camera functions work the same in student window.

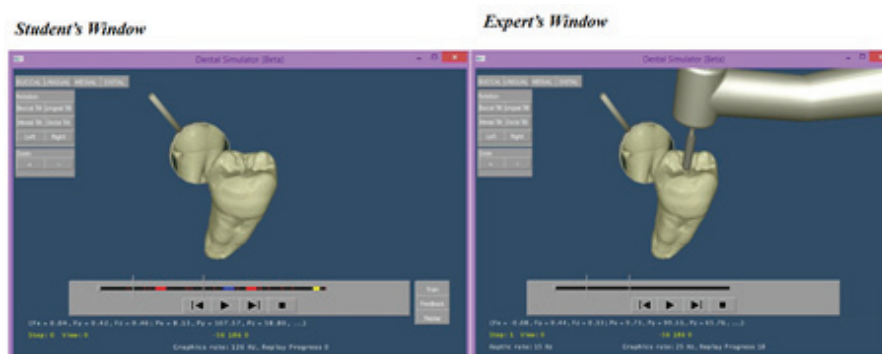


Figure 6.7: Student and expert windows in the Replay mode

### 6.1.5 Viewing aspect control panel

In the default view, the tooth is positioned with the Buccal wall facing towards the user. As the tooth is being drilled, the Lingual and Mesial walls are partially blocked from the view of the user by the driller. To provide the user with the option to view the playback from the specific wall, the axial walls are presented in Axial Wall View (Figure 6.8).

**Axial Wall View:** Dentists widely use axial walls to communicate and therefore, four axial walls (Mesial, Lingual, Distal, and Buccal) are provided to the user. Upon selection, the camera will be rotated to the selected wall, and the user can view how the tooth is being drilled from the selected wall (Figure 6.8, Figure 6.9).

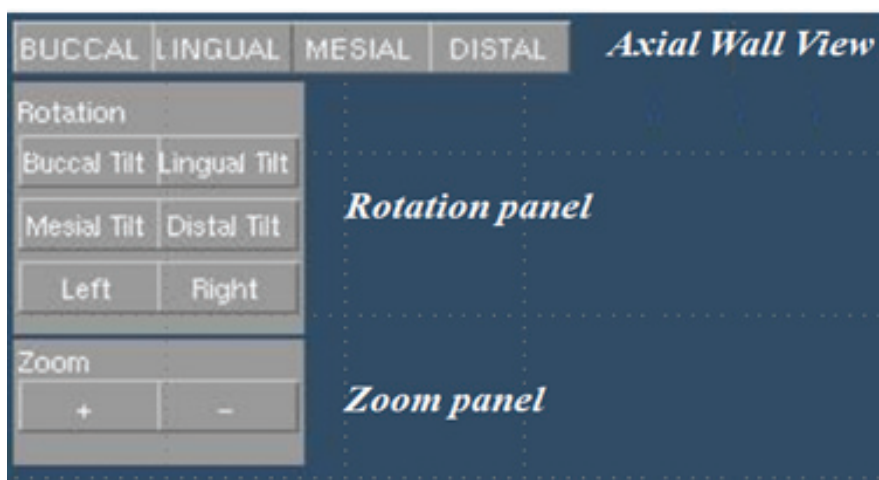


Figure 6.8: Viewing aspects control panel

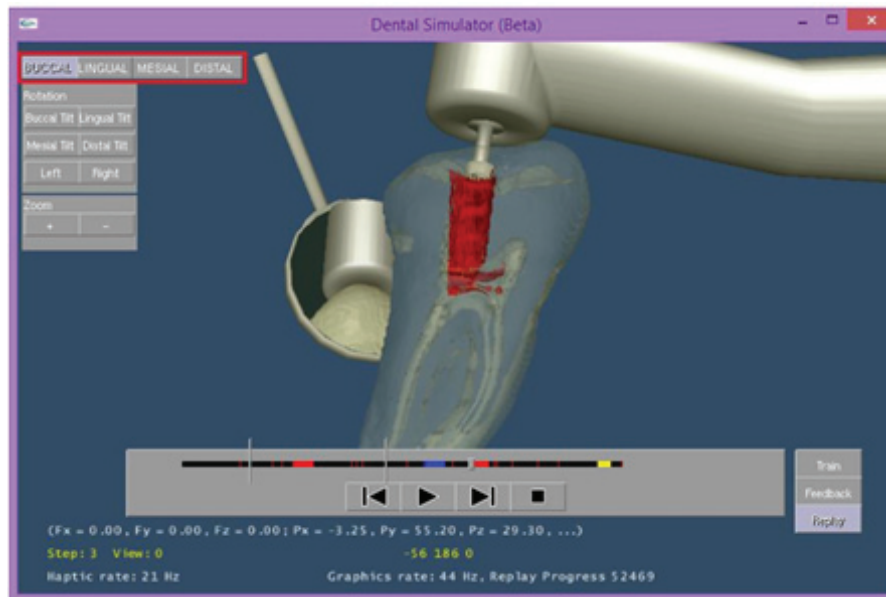


Figure 6.9: The Buccal wall view

**Rotation panel:** Using the Rotation Panel users can rotate the viewing angle step by step by tilting the walls and turning left/right. As shown in Figure 6.8, six-rotation buttons are Buccal Tilt, Lingual Tilt, Mesial Tilt, Distal Tilt, Left and Right.

**Zoom panel:** The zoom in/out functions allow the user to have a closer look at what is happening while the tooth is being drilled (Figure 6.10).

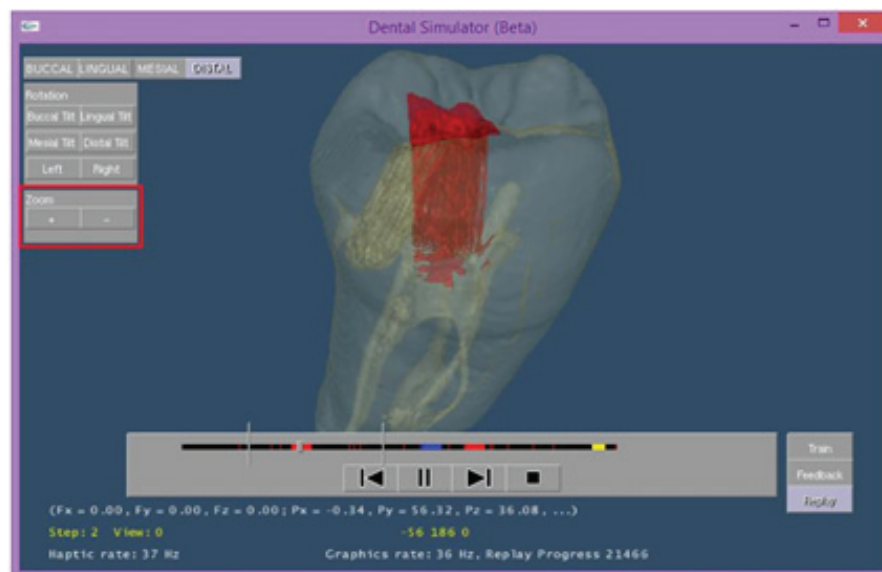


Figure 6.10: Magnified view of the simulated tooth

Auxiliary functions: The virtual handpiece and the virtual dental mirror in

the scene partially block the portion of the tooth being drilled and the highlighted error regions. Therefore, the auxiliary shortcuts are provided to remove the handpiece and the mirror from the playback scene. They are:

- Space Bar: remove/reload the handpiece and
- Space Bar + Right Mouse Button: remove/reload the mirror from the scene.

## 6.2 Experimental Evaluation

In evaluating the effectiveness, we aim to evaluate two main hypotheses as follows:

- Hypothesis I: Skill training using simulator with video-based formative feedback is better than training using simulator without feedback feature.
- Hypothesis II: Skill training using the simulator with video-based formative feedback is better than the traditional training approach.

To confirm the hypotheses the evaluation study is designed as a pre-test/post-test control group design. Experimental groups and control group are determined as:

- Experimental group I (G1): The participants in this group are trained with a VR simulator without feedback feature.
- Experimental group II (G2): The participants in this group are trained with a VR simulator with the summative score and video-based formative feedback feature.
- The control group (G3): The participants in this group are without VR simulator in the traditional laboratory.

To test the Hypothesis I, the learning gain of the training is defined as the difference between the pre- and post-training scores. The hypothesis I was tested by showing the students group trained with simulator with video-based formative feedback achieve higher learning gain than that of a control group consisting of a student group trained with the simulator without feedback. The null and alternative hypotheses are

- Null Hypothesis (H0): There will be no significant difference in learn-



ing gains between participant group trained using simulator with video-based formative feedback system (G2) and participant group trained utilizing the simulator without feedback (G1).

- Alternative Hypothesis (HA): The learning gains of the participant group trained using simulator with video-based formative feedback system (G2) will be higher (better) than that of the participant group trained utilizing the simulator without feedback (G1).

Regarding the Hypothesis II, to confirm the two training methods are equally effective, we compared whether the learning gains of the participant group trained with simulator with feedback are equivalent to outcome scores of the participant group trained in a traditional laboratory setting. The null and alternative hypotheses are:

- There will be no significant difference in learning gains between participant group trained using simulator with video-based formative feedback system (G2) and participant group trained in a traditional laboratory setting (G3).

- Alternative Hypothesis (HA): The learning gains of the participant group trained using simulator with video-based formative feedback system (G2) will be higher (better) than that of the participant group trained utilizing the simulator without feedback (G3).

Our sample size calculation is based on the similar study [161] conducted by co-investigators previously. Suebnukarn et al. [161] evaluated the effectiveness of simulator-based training using microcomputed tomography (micro-CT) tooth models on minimizing procedural errors in comparison with conventional phantom head training. According to their study, the response within each participant group is normally distributed with a standard deviation of 0.25 and the difference in the experimental and control mean 0.33. As shown by Chan [162], the required number of participants in each group is then calculated by

$$m(\text{sizepergroup}) = \frac{2c}{\delta^2} \quad (6.1)$$

where  $\delta = \frac{|\mu_1 - \mu_2|}{\sigma}$  is the standardized effect size and  $\mu_1$  and  $\mu_2$  are the means of the two treatment groups,  $\sigma$  is the common standard deviation and  $c = 7.9$

is related to a required significance level of 5%. As a result, 10 participants in each experimental group and 10 control participants are required to reject the null hypothesis that the population means of the experimental and control groups are equal with the probability (power) 0.8. The type I error probability associated with the test of this null hypothesis is 0.05. Stratified randomization is used in order to minimize the differences in gender distribution across and within groups.

Ethical approval was obtained from the Institutional Review Boards from Mahidol University and Thammasat University. A week before the first day of the experiment, the student representative was contacted regarding the purpose of the study to circulate the information on the experiments to the students for recruitment. We recruited thirty fifth-year dental students at Thammasat University School of Dentistry, Thailand. They were not admitted to the study if any of the following criteria were present: left-hand dominant individual; had prior experience with the simulation; received below 70 percent marks in knowledge assessment of the endodontic access opening. No participant dropped out from the study. The flowchart of participants through trials is illustrated in Figure 6.11.

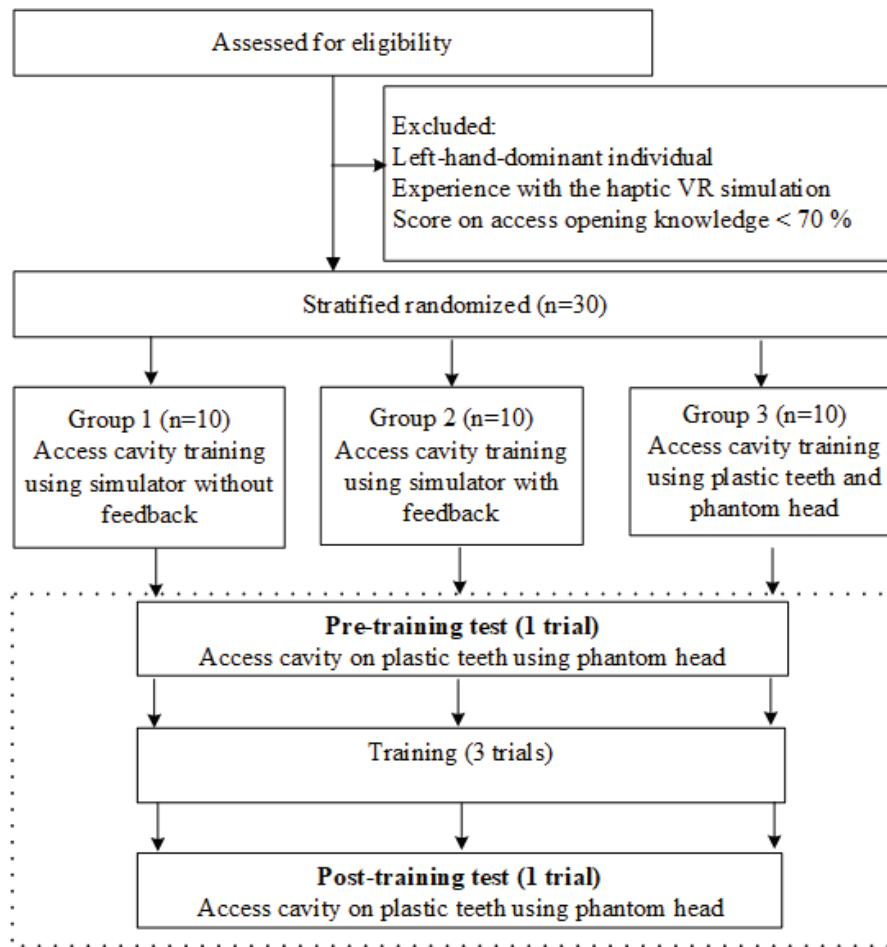


Figure 6.11: Flowchart of participants through trials

After obtaining the consent to participate in the experiment, each student was provided plastic typodont lower left permanent molar and asked to prepare the access opening stage of the root canal treatment. The artificial plastic teeth are designed for endodontic training with simulated anatomical pulp cavity and canals and have an x-ray imaging ability. Similar to working with natural teeth, trainees can experience the difference in cutting feel between the enamel and the dentin material. The teeth were acquired from Nissin Dental Products Inc (<http://www.nissin-dental.net/>), and examples of the teeth before, during and after preparation are shown in Figure 6.12. We would like to note the difference between the simulated tooth (mandibular lower right molar) and the plastic teeth (lower left molar). Lower left molar tooth is used in the evaluation study as it is the only lower molar tooth available in supply and the internal anatomy of the tooth is the most similar to the tooth used in simulation. Students were additionally

provided with a tungsten carbide bur (330), millimeter graduated periodontal probe, a mouth mirror, and a sharp straight dental probe (explorer). All teeth were coded anonymously.

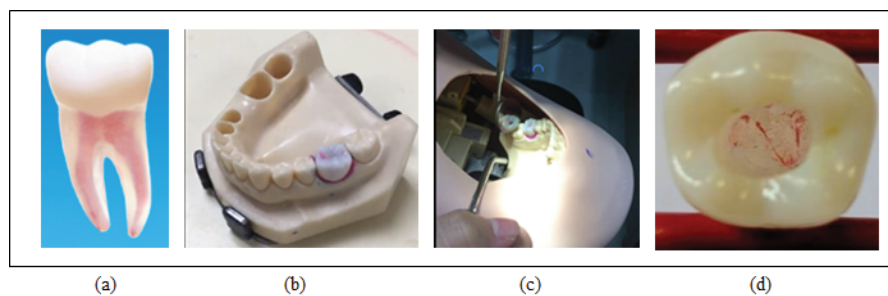


Figure 6.12: Artificial teeth (a) Before preparation (b) Mounted on the artificial jaw (c) During preparation (d) After preparation

Data were collected in separate sessions between control and experimental groups after study hours. In the pre-training session, all participants performed access opening in the laboratory using plastic teeth. During the training session, participants from G1 are trained using the simulator without feedback, participants from G2 were trained using simulator with formative feedback, and participants from G3 were trained in the traditional laboratory without the VR simulator.

Participants from G1 and G2 were briefly instructed on the use of the system, the experiment flow of assigned group and the requirements of the access opening. The participants received a verbal explanation about the use of the system from the investigators and familiarized themselves for fifteen minutes with the system interface, but not with the task. Participants from G2 were also informed that they are allowed to stop the video-feedback as they feel they understand the errors and the causes during the video-playback. During this familiarization or warm-up period, each participant was allowed to ask questions and receive further verbal explanation and suggestions from the investigators. After the familiarization, the participant continued in acquisition sessions. During the training stage, participants from G2 received scores on the outcome from the automated outcome scoring system and video-based formative feedback on the performance. They were allowed to navigate the video playback freely and exit before the replay was over (and many did) if they felt that they had understood how and what

Table 6.1: Intra-class correlation coefficient (ICC) between two evaluators (95% confidence interval)

Mesial	Distal	Buccal	Lingual	Floor	Overall
0.97 (0.955, 0.984)	0.98 (0.961, 0.986)	0.98 (0.972, 0.99)	0.98 (0.966, 0.988)	0.91 (0.848, 0.947)	0.99 (0.983, 0.994)

leads to the resulting performance score.

The primary outcome measure is the average outcome scores from each pre- and post-training retention assessed by a panel of two experts who are blinded to trainee and training status. The overall preparation score was considered as the primary dependent variable representing the success in learning outcome while the scores on axial walls and the pulp chamber floor are considered for detailed analysis of performance.

### 6.3 Results

The two experts evaluated the artificial teeth from both controlled and experimental groups in pre- and post-training steps. For all the scores normality was confirmed using the Kolmogorov and Smirnov test. Since the outcome scores in this study are normally distributed, we computed the intra-class correlation coefficient (ICC) [163] to confirm the degree of agreement between two experts in scores, ICC value ranges between 0.0 and 1.0, and the high value indicates a few variation between the scores given to each tooth by the raters. As presented in Table 6.1, high ICC values indicated the strong inter-rater agreement in all categories (the axial walls, the floor and the overall scores). All the coefficients of ICC are significant at  $p = 0.05$ . The highest ICC (0.99) was observed in the overall score while the lowest (0.91) was found in the floor scores.

The descriptive statistics of pre- and post-training scores in all groups are summarized in Table 6.2. The mean scores before training ranged between (65.80 - 63.70) while the means of after training ranged between (66.30 - 91.6). A marked decrease in the standard deviation values was observed in post-training scores compared to that of the pre-training scores in G2 (8.07 from 15.52) and G3 (7.70 from 15.71), indicating the convergence in performance of these two groups.

Table 6.2: Descriptive Statistics of Scores

	Pre-Training						Post-Training Scores					
	Mesial	Distal	Buccal	Lingual	Floor	Overall	Mesial	Distal	Buccal	Lingual	Floor	Overall
<b>G1</b>												
Mean	52.00	63.50	67.50	56.50	89.50	65.80	55.50	67.50	65.00	54.50	89.00	66.30
Median	55.00	65.00	70.00	57.50	92.50	63.50	62.50	70.00	60.00	52.50	95.00	66.50
Standard Deviation	21.50	15.64	17.52	20.69	10.92	14.98	24.32	15.86	15.09	21.53	13.08	15.91
Minimum	30.00	30.00	30.00	25.00	70.00	37.00	10.00	50.00	50.00	30.00	65.00	44.00
Maximum	90.00	85.00	100.00	100.00	100.00	95.00	85.00	90.00	90.00	90.00	100.00	91.00
<b>G2</b>												
Mean	52.50	54.50	63.00	53.00	83.50	61.30	87.00	93.00	91.00	87.00	100.00	91.60
Median	52.50	55.00	60.00	50.00	82.50	59.50	92.50	100.00	100.00	95.00	100.00	93.00
Standard Deviation	22.14	18.48	15.13	22.75	11.80	15.52	14.57	10.85	13.50	15.13	0.00	8.07
Minimum	10.00	10.00	50.00	20.00	65.00	37.00	55.00	70.00	70.00	65.00	100.00	73.00
Maximum	95.00	75.00	100.00	100.00	100.00	94.00	100.00	100.00	100.00	100.00	100.00	99.00
<b>G3</b>												
Mean	57.50	57.00	61.00	56.50	86.50	63.70	68.00	69.00	68.50	64.00	91.50	72.20
Median	65.00	62.50	65.00	55.00	92.50	67.00	70.00	70.00	70.00	65.00	97.50	74.00
Standard Deviation	22.02	17.51	16.63	16.67	17.49	15.71	10.33	13.50	8.18	9.66	11.32	7.70
Minimum	10.00	30.00	30.00	30.00	50.00	36.00	50.00	50.00	50.00	50.00	70.00	63.00
Maximum	80.00	80.00	80.00	80.00	100.00	84.00	90.00	90.00	80.00	80.00	100.00	84.00

Overall, the mean post-training scores in all categories (the axial walls, the floor and the overall scores) were higher than the mean pre-training scores in G2 and G3 compared to G1. A paired t-test was used to compare the difference between the means of pre- and post-training score. At  $p < 0.05$ , the significant differences were found G2 from G1 and G3. The student group trained with simulator with feedback achieved higher mean post-training outcome scores (G2 PostMean = 91.6) than that of the control group consisting of a student group trained with the simulator without feedback (G1 Post-Mean = 66.3), and that of participant group trained in traditional laboratory setting (G3 Post-Mean = 73).

Before further analysis, the initial differences among the three groups were determined with a one-way analysis of variance (ANOVA). The overall preparation score (S) before training was taken as the main dependent variable. Additional Tukey HSD post hoc tests were applied when significant effects were encountered. The results revealed that between-group differences were not significant ( $p = 0.809$ ):  $F(2, 29) = .214$ ,  $p < 0.05$ , and post hoc (Tukeys HSD) analyses revealed that neither of the experimental groups had significantly higher scores than the control group (G3).

The independent samples t-tests were used to compare the learning gains among all three groups. We found that the student group trained with the simulator with feedback (G2) achieved statistically significantly higher learning gain ( $30.30 \pm 17.5$ ) at

the end of the training session. The statistically significant gain compared to the control group (G1) (trained with the simulator without feedback) ( $0.5 \pm 11.375$ ),  $t(18) = -4.523$ ,  $p = 0.000$ , confirms our hypothesis I. Similarly, the student group trained with the simulator with feedback (G2) had statistically significantly higher learning gain ( $30.30 \pm 17.5$ ) at the end of the training session compared to the control group G3 (trained with traditional phantom head) ( $8.5 \pm 14.152$ ),  $t(18) = -3.068$ ,  $p = 0.007$ , confirming our hypothesis II. Figure 6.13: Plot of the change [(post-training overall score) (pre-training overall score)] vs. pre-training overall score. shows the comparison of learning gain [(Post-training score) - (Pretraining score)] with the initial pre-training scores. We can notice that low initial scores the change is high (much improved after the training period), but for high initial scores, there was little improvement.

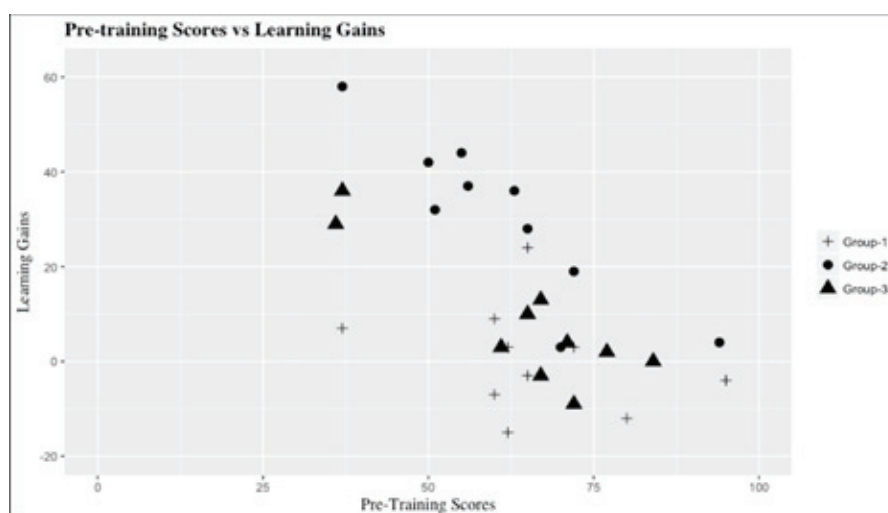


Figure 6.13: Plot of the change [(post-training overall score) (pre-training overall score)] vs. pre-training overall score.

Additionally, we analyzed the learning gain in each axial walls and the pulp floor of every group to have a better understanding of the difference in performance between groups. The negative learning gain means were observed in G1 from Buccal, and Lingual walls, indicating that scores of participants from G1 significantly dropped from pre-training scores in these two walls. In contrast, G2 and G3 have positive learning gains in Buccal, and Lingual with higher learning gains were observed in G2. From one-way analysis with Tukey post-hoc tests revealed that the mean learning gain of G2 are significantly higher than in the Mesial and Distal walls scores than that of G1 and



G3. Experts concluded the absence of the probe tool in G2 as a possible cause. The participants from G3 can use the dental probe tool to inspect the drilled area in Mesial and Distal wall while the participants from G2 were gaining the benefits of seeing the Buccal and Mesial walls from video-playback and formative feedback.

## 6.4 Discussion

Issues centered around the subjectivity on the feedback has been a topic of controversy in dental education [164, 165]. The video based-formative feedback provides purely non-evaluative objective performance feedback, specifying what was done, how it was done and how it affects the performance and thus, the system ultimately eliminates the drawbacks of face-to-face verbal feedback.

According to Weeks and Kordus [166], the most effective feedback motor skill development should create awareness of when and how an error was made which is more important than the final result itself. The video-based formative feedback system is highly suitable feedback modality to create such awareness as it provides where, when and how an error was made. This knowledge of errors can help the trainee to rectify or modify their performance on the next trial and increase the likelihood of achieving the desired outcome. This is evident in the improved post-training performance observed from participants in G2.

White et al. [9] noted that trainees knowledge is increased by making and learning from errors. The opportunity to actually see the (what) actions lead to each error instead of only being told the performance quality (grade/score) help the trainee to discover the relationship between the movement and the performance outcome. Students, especially novices, are dependent on instructors to supervise and provide feedback [167]. Performance scores observed from participants of G2 confirm without feedback training with simulator alone tells students very little about how to improve their performance.

Video-playback provides trainees with a third-person view of their performance and opportunity for reflection of own performance. Video-based feedback pro-

vides trainees with an accurate perception of their performance as well as enhancing their self-awareness. The ability to self-assess work done on a VR simulator could help students gauge the quality of their work and determine a scope to improve [168] and can contribute to the faster acquisition of skills and reduce the learning curves. The ability to self-assess is thought to be one of the essential elements for developing the motor skills [99]. The video-based formative system permits the student to conduct a self-assessment. Using the video-based formative feedback, students can have a better understanding of which movement or actions to errors and can improve their ability to recognize the error, which in turn leads generating a more accurate self-assessment that leads to improvement in performance. Also, trainees with or without their instructor can review errors at the exact time point they occurred and work out for the remedy.

Being able to compare their performance with experts performance facilitates the trainees to benchmark their performance against experts. Knowledge on the ideal procedure outcome prepared by the expert and how the expert carried out the procedure allow the trainees to know the quality of their performance and enable to plan for improvement in the next trails if needed.

As a limitation, we noted that we have considered two additional feedbacks in the experimental group: feedback specifically on outcome from automated outcome scoring system and the video-based formative feedback. In principle, it would be good to know to what extent these different aspects contribute to improvements separately. Familiarity with using plastic teeth which is as realistic as natural teeth cannot be discounted in G3 as training with phantom head with plastic/extracted teeth are the gold standard in the preclinical laboratory. Terminal feedback is known to be superior to concurrent feedback in long-term retention of the learned skill. The terminal feedback generated from our system emphasizes on the promotion of task-relevant visual aspects (internal anatomy, the errors in space and time). This evaluation study did not investigate this terminal feedback and its impact on retention of skills and identified this as the future work.

Informal interviews with instructors and the participants revealed that both students and instructors would like to have the feature to record/save the playback as a video, providing that it would enable them to keep records of their performance and

can be used to monitor their progress. The instructor can also utilize it to review each trainees skill and the performance in one place and it could provide them with the feasibility to provide more extensive feedback without requiring him to present during the training session.

## CHAPTER VII

### FORCE FEEDBACK

In this chapter, we present the use of haptic formative feedback to train the correct application of force in endodontic surgery.

Critical physical parameters required for precise manipulation of invasive instruments during an endodontic procedure include trajectory, orientation, and the magnitude and the direction of applied force. While visual guidance or demonstration would suffice to guide/train trajectories and orientation of instrument, it is less useful for guiding the correct application of force. At the same time, application of force is particularly challenging. The amounts of force used are on the order of tenths of Newtons, requiring a highly refined tactile sense, with incorrect force possibly causing irreversible damage to the tooth and surrounding tissue. Despite its importance, students receive very little (if any) explicit feedback on their application of force during training. Since instructors have no means to directly assess the amount of force applied by a trainee, one-to-one guided instruction/intervention on the proper application of force is very difficult in a traditional pre-clinical laboratory.

Commonly, learning by observation from expert demonstrations is at the center of dental student clinical skill training [31]. Trainee dentists typically have to rely on visual observation of experienced surgeons and verbal descriptions of proper techniques. Although demonstration takes up a significant portion of psychomotor skills training [169, 170], a typical demonstration takes a considerable time of training and requires to be conducted with care as the knowledge related to the skill underlying demonstration is often associated with tacit knowledge and not visible to the trainees [171], and needs to be clearly communicated as part of the demonstration. Specifically, in motor skill training on instrument manipulation requiring the proper application of force, verbal instructions or demonstration are not suitable to convey the magnitude or the direction the trainee should apply on the instruments. Demonstration session also limits

the precision with which an instructor can monitor a trainees drilling performance, as he/she cant feel the fine details of the trainees interaction within the confined cavity, and cant easily share the drill and the operating area for a demonstration. Teaching motor skills is considered to best explained by an instructor through physical interaction [172]. However, such approach is expensive and inefficient because it necessarily requires a low student-to-faculty ratio.

Realistic manikins in the clinical laboratories of the dental schools provide the instructors with a means to explain and improve students hand-eye coordination, and dexterity. However the verbal description of tactile sensation [168] and the necessary force and motion coordination for instrument manipulation are fundamentally difficult to explain. With less realistic soft tissues in gums and other structures around the teeth, Yamaguchi and colleagues [173] noted that learning to apply correct probing force using manikin is especially difficult in skill necessary for periodontal treatments. Numerous empirical and experimental studies have been conducted on providing realistic haptic feedback using VR dental simulators including [47, 53, 173–177]. Evidence suggests that there is a link between simulator ability and operative skills, indicating some benefit to using haptic simulation in undergraduate dental teaching [26] [134]. Studies have evaluated surgical performance and skill acquisition during training and shown that significant differences exist between experts and novices in force/torque magnitudes at the hand/tool interface [55, 76, 178]. But no work has yet explored the use of haptic feedback as formative feedback and a modality for training the correct use of force. Typically, dental students acquired motor skills by observation and practice. Alternatively, motor skills can be trained through physical guidance. Physically guiding children how to write or how to move a swimming stroke to get maximum forward extension are a few example of physical guidance. Majority of existing haptic-based motor skill training approaches are centered on the two-phase process, namely, record-and-play. In the record phase, the experts movement is recorded in terms of position, velocities, and force patterns. In the subsequent play phase, the recorded movements are haptically and visually displayed to learners during the playback mode training. Playback can be either active or playback mode. In the passive playback mode, the trainees have to grasp the end-effector of the haptic device and are physically guided through the ideal motion through

a desired trajectory to acquire the kinesthetic understanding of what is required. In contrast, the trainee moves the end-effector through a desired trajectory at his/her speed in the active playback mode. As the end-effector movements are constrained to the ideal trajectory, the learner experiences as moving along a virtual channel, which keeps the end-effector on the correct path. Yokokohji et al. [179] introduced What You See is What You Feel concept in which the experts forces are first recorded and displayed to the trainee to identify and adopt a strategy based on the force displayed. The authors also introduced various haptic training methods, including guiding a subject through a motion or restricting a subjects motion, however, any substantial results are provided. Gillespie et al. [180] presented Virtual Teacher where they demonstrated a strategy on how a teacher physically guides the movement of the trainees hand. The findings from the pilot study involving a simulated crane moving task are inconclusive, and their results show the virtual teacher has an advantage over the human teacher because of its accuracy and consistency in demonstrating motor skills.

Likewise, in the Virtual lesson system for teaching Japanese calligraphy developed by Henmi et al. [181] adopted the record-and-play strategy. In their system, the position and force trajectories of the teachers brush were recorded first, and then these trajectories are displayed to the student. Although findings from the preliminary experiment primarily focusing on force utilization reported the transfer of skill, the authors noted the need for further investigations. In a haptic-based Chinese handwriting learning study presented by Tao et al. [76], metrics such as character shape, strike smoothness, normal forces against a virtual paper, and pause-and-go motion are identified as constituents of the skills relevant in Chinese calligraphy. Post-training performance showed that most of the metrics, except the normal force pattern, were improved immediately after the training.

In addition to the calligraphy, haptic-based force feedback training is widely found in motor rehabilitation studies. Kim and Yang [182] described a rehabilitation exercise for the users with hand dysfunctions in the haptic virtual environment where the users hand movement is guided on the right track of the predefined trajectory according to the real-time guidance force. Though the results are analyzed and reported to the user during experiments, the finding details are not disclosed in their paper.

Another discipline that requires the motor skill intensively is the medicine. William and colleagues [183] demonstrated the merits of a haptic training to achieve correct hand movements with a haptic playback feature of the virtual palpatory diagnosis trainer. With active playback mode, the authors compared the hand movements between two groups of subjects, one of which received the haptic training and the other did not. Their findings indicate that the trained group performed better than their counterparts. In fact, haptic training has been studied in various areas of the medicine such as the diagnosis of prostate cancer [184], gynecologic exam procedure [185], surgical dissection training [186], and laparoscopic surgery [187]. As the person will choose a set of forces based the prior knowledge/experience, preference and environmental setting, Srimathveeravalli and Thenkurussi [188] presented the use of a unique haptic profile constructed from the haptic attributes to every task performed using the motor skill. A time series of force information is considered as haptic attributes and X, Y, Z components are regarded as the haptic profile of the person. Participants in their study are trained to reproduce an experts handwriting in terms of shape and force pattern while the reference character is visually displayed and the experts position trajectories and the forces applied were passively displayed by a haptic device. Superior results are observed in training given using force information in the form of haptic attributes as compared to training using position information in the study. Considering haptic interface as one of the modalities in training a motor skill, comparative studies with other modalities are also available in the literature. Avizzano et al. [189] compared haptic training with visual training in a task involving a simple reproduction of a predefined circle. While four critical points on the reference circle were displayed for guiding purpose in the visual training, a 2 DOF force-feedback device passively constrained participants hand close to the circular trajectory in the haptic training. As the shapes of the drawn-circles were significantly better after haptic training than after visual training, it is concluded that haptic training is more helpful for a circle drawing task.

In the study presented by Morris et al. [185], the normal force against a horizontal virtual plane was displayed haptically, visually, and visuo-haptically. Based on the results from experiments in which the participants hands were actively guided along randomly chosen paths, they concluded that force patterns could be learned through

haptic training and visuo-haptic training was shown to be the most effective method for force pattern training. Solis et al. [190] evaluated the skill-transferability of a Japanese character learning system [191] under three training methods: visual, haptic and visuo-haptic. Performance metrics considered in their study include task completion time, overall correctly applied force magnitude, and performance. Their findings indicate that haptic training can only improve task completion time while training with both visual and haptic feedback could dramatically improve participants motor skills. Feygin et al. [192] also have a similar result that haptic only training was adequate with respect to the temporal aspect of the task, while motor skill improvements were due to the training with visual information.

Kolesnikov et al. [174] first introduced the haptic playback mechanism to skill acquisition in dentistry. The passive playback feature is incorporated into the prototype of haptic-based virtual reality simulator, Periosim [174], to investigate haptic capabilities in sensorimotor skill acquisition. The realism and the effector of the simulator were evaluated with faculty members and students from a variety of clinical areas in dentistry. In the follow-up study, Kolesnikov and Zefran [193] demonstrated that there is no direct relationship between the tracking precision of a haptic playback system and its effectiveness as a tool in motor skill transfer.

## **7.1 Sensory Motor Integration**

A typical root canal preparation using hand instruments principally involves tactile sensory input, rather than visual input as occurs in routine cavity preparation tasks [123]. Endodontists use their hands almost always to explore the operating area and the objects in it. They have to recognize and distinguish the form of an object through exploration (touch) using indications about the texture, size, spatial properties and temperature of the object [194] [195].

Brain activity related to learning fine motor skills is triggered mainly by visual and tactile sensory input systems [170]. The haptic perception, the process of recognizing objects through touch [123], includes a mixture of somatosensory perceptions



of patterns on the skin surface (e.g., edges curvature, and texture) and proprioception of hand position and conformation [196]. The human hand is said to have the most substantial sensory and motor representation in the sensory cortex [197]. People with normal sensation and movement can usually identify a familiar object by touch alone, and this ability is referred as the stereognosis or hapticgnosis [198]. Two central somatosensory systems of stereognosis are (i) the kinesthetic (the proprioceptive) system which provides information about the position and movement of the body and limbs (ii) the tactile (also called cutaneous) system which provides feedback from the external world.

The haptic devices commonly used in skill training with VR systems primarily focus on stimulating the kinesthetic system and have been designed with the primary intention of generating somatosensory feedback. Sensors and actuators in the haptic devices monitor user actions and create forces which guide, resist or perturb movement, providing force feedback about the physical properties and movements of objects in the virtual environment. There is a various mechanism for providing force feedback, including mechanical levers or pneumatic actuators connected to the hand, which form an interface between the persons fingers and a computer. Haptic devices allow movement with several degrees of freedom, support and react with different amplitudes, offer manipulation in a restricted space and use various technologies. The GeoMagic touch used with our simulator is a body-based haptic device which use a connection point of the own device to provide force feedback. The device permits simulation of fingertip contact with virtual objects. The stylus tracks x, y and z Cartesian coordinates of the virtual point-probe. Its actuators communicate forces back to the users fingertips as it detects collisions with virtual 3D objects, simulating the sense of touch [199]. To train the parameters of a task beyond auditory and visual cues, haptic devices are commonly integrated into dental training systems, usually as dental mirrors [161], dental hand-pieces [161] in drilling tasks and probe explorers [175] in diagnostic tasks. The use of haptic devices allows the trainee to capture tactile sense with minimal intervention from the instructor [52], and also pave the way for haptic training.

## 7.2 Haptic Formative Feedback

We study an approach to using haptic feedback as a means to convey formative feedback on the correct application of force. We build on previous work on automated outcome assessment [200] (Chapter 4) and the correlator component (Chapter 5) to analyze the relationship between procedure and outcome. With the portions of the procedure responsible for errors classified in the correlator component, we then identify stages in which the amount and direction of applied force differ significantly from the force used by experts. Providing force feedback requires determining possible misapplications of force related to the error. We distinguish two misapplications of force: over-exertion and under-exertion of force. Interestingly, interviews with experts and detailed analyses of erroneous cinematics revealed that there is no one-to-one mapping between over-exertion/under-exertion of force with over-extension/under-extension errors. This information is conveyed to the student graphically, and the correct amount of force to apply is trained haptically.

To generate feedback, we adapt the haptic training method of Saga et al. [201, 202]. They present a haptic technique to teach hand skill tasks such as calligraphy by rendering the force applied by the expert on the haptic device but rendering it in the opposite direction. A student holding the haptic device must then apply the same amount of force in the original direction to cancel out the rendered opposite force to proceed with the operation. In this way, the student learns what the correct amount of force feels like.

This method requires knowledge of the desired (i.e., an expert's) applied force. Therefore, initial samples were obtained from an expert endodontist with more than ten years of experience. Because previous investigations (e.g., [76]) revealed that experts exert a fairly constant force along a primary axis during the procedure, the average force by the expert in each axial wall of the tooth in each stage was pre-computed as an approximately correct amount of force. To provide formative force feedback for a given stage, the expert's average force is rendered in the opposite direction to the haptic stylus. To get a sense of the magnitude and the direction of the force applied by the expert, a student holding the haptic device must then apply the same amount of force in

the original direction to cancel out the rendered opposite force as shown in Figure 7.1. While the memory is still fresh with the tactile experience, the simulator is rewound to the point before the error, and the student is asked to redo the stage where significant deviations are found. Throughout the reattempt, a hotkey feature is provided to retrieve the force training episode as required.



Figure 7.1: Opposite force training

Note:  $\vec{F}_E$  : Expert's applied force,  $\vec{F}_N$  : Student's applied force

### 7.3 Experimental Evaluation

We conducted a preliminary evaluation of the effectiveness of the described haptic force feedback for student training. The study had a pre-training/post-training control group design. Ten dental student volunteers (5 Males, 5 Females) beginning their fifth year were recruited and were randomly assigned to experimental and control groups. All participants briefly received a verbal explanation about the use of the system from the experimenter and were given 10 minutes to familiarize themselves with the system interface, but not with the task. During this familiarization period, each participant was allowed to ask questions and receive further verbal explanation and suggestions from the experimenter. After the familiarization, the participants were asked to perform the endodontic access opening procedure on the mandibular second left molar.

Data were collected in separate sessions between control and experimental groups after study hours. Participants performed one trial in each of the pre-training and post-training stages. In between the pre- and post-training trials the control group received one training trial, and the experimental group received two training trials.<sup>1</sup>

<sup>1</sup>Due to experimenter error, the control group is missing the intended second training trial. To make sure that differences between control and experimental group are not because of differences in training trials, we also analyzed performance of the experimental group after one training trial (employing perfor-

Each trial took approximately 15 minutes. During the training stage, feedback on the outcome was provided by a qualified endodontic instructor for the control group. For the experiment group, in addition to the feedback on the outcome, the instructor provided feedback on the magnitude and direction of force applied on the handpiece in each stage of the procedure using the analysis provided by the system as shown in Figure 7.2. Immediately after giving the feedback, the instructor asked the student's opinion to determine the stage to re-execute to improve the performance. The simulator was then rewound to the beginning of the selected stage, and the student was provided with three force training episodes each lasting a minute. Participants could choose to exit the force training episode before the minute was over (and many did) if they felt that they had understood how to apply force in the given situation. After the three training episodes, the student was asked to redo the stage and then continue with the remainder of the procedure. It is essential to have the student complete the procedure after redoing the stage since any changes (improvements) in the outcome of the stage will impact later steps in the procedure.

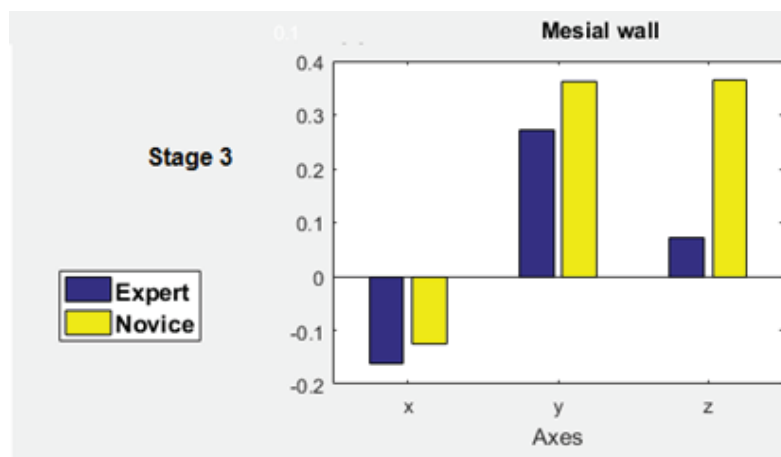


Figure 7.2: Comparison of force used in mesial wall between expert and novice in stage 3

The primary outcome measure for the evaluation was an overall preparation outcome score. The nonparametric Wilcoxon Signed-Rank test was used to examine the differences between the paired pre- and post-training outcome scores, in the same group. (The analysis of the second training trial as the post-test performance). This analyses yielded virtually identical results to the analyses reported here.

Table 7.1: Pre-training and post-training outcome scores between control and experimental groups (mean  $\pm$  SD)

	Control Group		Experiment Group	
	Pre	Post	Pre	Post
Lingual	44 $\pm$ 41	54 $\pm$ 32	40 $\pm$ 38	65 $\pm$ 9
Buccal	50 $\pm$ 29	38 $\pm$ 35	59 $\pm$ 2	54 $\pm$ 31
Mesial	46 $\pm$ 9	36 $\pm$ 33	46 $\pm$ 26	57 $\pm$ 33
Distal	80 $\pm$ 16	56 $\pm$ 34	61 $\pm$ 8	76 $\pm$ 9
Overall	55 $\pm$ 6	46 $\pm$ 17	51 $\pm$ 6	66 $\pm$ 12

The Mann-Whitney U test was used to detect any differences in training gain between student groups trained with and without haptic force feedback. A significance level of 0.05 was adopted for all reported analyses.

## 7.4 Results

None of the participants in the control or experimental group dropped out before completing the post-training assessment. There were no significant differences between the groups in terms of the outcome scores in pre-training ( $U = 9$ ,  $n_1 = n_2 = 5$ ,  $p > 0.4$ , two-tailed). Group mean outcome scores on the four axial walls (lingual, buccal, mesial and distal) and the overall summary score are shown in Table 7.1 differences were found in both detailed wall level and overall outcome scores between pre- and post-training trials in the control group. The distal wall scores are significantly improved from pre- and post-training in the experiment group ( $p = 0.043$ ). Training gain, defined as post-test score - pre-test score, was significantly larger in the experimental group ( $M = 15$ ) than in the control group ( $M = -9$ ):  $U = 3$ ,  $n_1 = n_2 = 5$ ,  $p < 0.05$ , two-tailed.

Group differences were also evident in changes of the amount of force applied: In the pre-training trial, the participants in experiment group used a more substantial amount of force than the control group and the expert across all four axial walls in y- and z-axis. The opposite pattern was observed in the post-train trials where the control group applied the more substantial amount than the experimental group and the experts. The applied force in experiment group noticeably reduced after the force training in all

walls except lingual.

## 7.5 Discussion

Studies of surgical performance and skill acquisition during training reveal that significant differences exist between expert and novices in force/torque magnitudes at the hand/tool interface during dental procedures using simulators [55, 76, 178]. Their findings are aligned with the conclusions of the study on the parameters involved in the acquisition of a practical skill in oral surgery to characterize expert and novice performance [203]. For a drilling task using an ovine jaw, Ioannou and colleagues [203], the analysis on force sensors readings reveal that regarding magnitude, the experts applied slightly more force than novice and experienced participants groups in the study while the direction of the forces applied varies widely. The authors of the studies [55, 76], also show that experts use defined force patterns while novices do not. Students tend to use significantly less main pushing force compared to experts [76] in crown preparation procedures. Less use of force in endodontic surgery may result in longer times to complete each step of the procedure, which can cause patient discomfort and damage to the tooth structure. On the other hand, too great a force can cause over-extension or in extreme cases perforation of the tooth. Even subtle differences in force can yield substantial differences in achieved results.

A known problem of conventional method in teaching skillful motion or manipulation of tools in the dental procedure is the teacher, and the student cannot hold the same tool at the same time with the right posture. To mitigate the issues arising from the spatial gap, the instructor usually holds the hand or arm of the trainee to teach the motion. This makes it difficult to show the trainees motion and force correctly. The haptic demonstration approach we presented is designed to communicate directly to the trainee's hand and there no such spatial limitation and the trainee can sense to mimic the instructors actions.

Training on the application of force while using instruments with verbal feedback such as "harder." or "2 times higher." might be insufficient in guiding novices

how to adjust their parameters [204]. Automatically moving with the haptic stylus in the student's hand along the recorded expert path will not solve the problem as it is a passive method and students still cannot learn the actual forces applied by the expert [176]. Similarly, tracking based haptic training approach shown no direct relationship between the trajectory tracking precision and the motor skill performance [193]. HapticMaster [205], a hardware component in Simodont Simulator [206] uses the same concept of rendering the countering force to achieve the realistic feedback. The fundamental difference lies in the use of formative feedback in our approach, which is generated after objectively evaluate, analyze the outcome and the procedure. Our approach still relies on the expert to determine the rewinding stage. We are investigating a threshold-based method to automate the process to determine the error-related stage and region. In this preliminary evaluation, we have considered two additional feedbacks in the experimental group: feedback specifically on the force by the instructor and the haptic feedback. In principle, it would be good to know to what extent these different aspects contribute to improvements separately.

## **CHAPTER VIII**

### **CONCLUSIONS AND FUTURE RESEARCH**

This research address the issue of objective assessment and feedback in skill training using VR simulators. Our approach to objective assessment and feedback is based on correlating the procedure and the outcome. The components are Automated Outcome Scoring System, Correlator, and Feedback System consisting of video-based and haptic-based feedback systems. While the framework is implemented for the root canal treatment using a dental simulator, our approach to assessment and formative feedback generation is general and can apply to a wide variety of surgeries as well as to other physical procedures. In this chapter, we will discuss the generality and extensibility of of each component and the framework as a whole.

#### **8.1 Automated Outcome Scoring System**

The scoring cube for outcome assessment used in the Automated Outcome Scoring System provides scores to the 3D voxel structure commonly used in VR dental simulators at a sufficient level of detail to allow correlation with procedure kinematic variables collected by such simulators. To effectively communicate outcome score results, detailed level scores are translated into the language of the coarser level standard scoring system used by dental schools. The objective consistency of the outcome scores and the high degree of agreement with experts make it a promising addition to existing VR simulators. This is the first scoring algorithm for endodontic surgery and is a scoring mechanism with a precise way of identifying errors in the outcome compared to commonly used techniques for outcome assessment.

Our scoring system relies on the availability of templates of the 3D volume objects. It is applicable to assessment of outcomes of procedures involving 3D volumne objects or regions where the desired outcome can be described in terms of templates.



Our scoring approach is robust and is not limited to the access opening procedure in root canal treatment. In addition to the access opening stage, the algorithm can apply in assessment of cavity preparation stage of the root canal treatment. In the cavity preparation stage, the canals are expanded and shaped with the small files to get access the apical portion (tip) of roots. At present there is no reliable method for evaluating the canal cleaning and shaping [207], and the systematic way of scoring of canal preparation is needed in practice. As the canals can be identified along with the pulp chamber, the cavity preparation stage can be assessed and scored using our scoring system.

Other than the root canal treatment, the scoring algorithm is also applicable in evaluating the outcome of other dental surgery procedures such as the dental implant procedure. This procedure involves replacing an extracted tooth with a metal stud or post screwed right into the jaw bone to serve as an anchor to the crown or false denture. As a part of the procedure the portion of the gum is needed to cut then drill into the jawbone in order to place the stud or post, and in this stage our outcome scoring system is applicable for assessment and scoring. Elsewhere other than dental surgery, consider hip replacement surgery. In the hip replacement surgery, the ball portion of the joint is removed by cutting the thighbone with a saw. With the same underlying concept of removing the hard bone tissue, the outcome scoring system is also applicable. It is not limited to the procedure with the hard tissue removal, drawing on the similar analogy, our scoring system is applicable to other surgery procedures involving the soft tissue removal such as surgical removal of tumors.

In addition to the dental and medical area, the scoring system can be used to evaluate the outcomes of VR simulations involving physical procedures such as gemstone faceting and the welding using VR. For instance, consider the fillet welding task where two pieces of metal are joined together either in perpendicular or at an angle. Fillet weld process outcome is evaluated from five aspects [208] known as the Root, Toe, Face, Leg and Throat as shown in Fig 8.1. The Root refers to the deepest penetrated area on the opposite angle of the hypotenuse, the Toe refers to edge or the points of the hypotenuse, the Face is the outer visual, the Legs refer to the opposite and adjacent sides to the triangular fillet weld (size of the weld) and the Throat of the weld is the distance from the center of the face to the root of the weld. Typically, the depth of the

Throat is at least as thick as the thickness of metal in process. Based on the evaluation criteria, the scoring templates for the fillet weld task on a VR simulator can be created in a straightforward manner. The non-linear scoring function can be adjusted with the desired weight function. With the templates and the weight function, the weld process outcome can be assessed and scored using a VR weld simulator and 3D volume models.

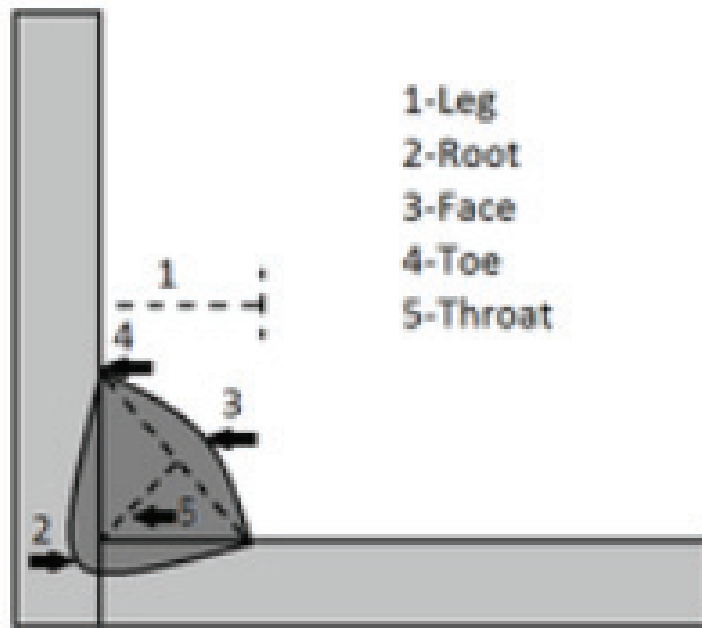


Figure 8.1: Five aspects in fillet weld [208]

## 8.2 Video-based Formative Feedback

We have shown the effectiveness of the video-based formative feedback on trainee performance through a randomized controlled trial using VR skill training simulator with our feedback system relative to training with simulator without feedback and training in a conventional setting. We found that the students trained using the simulator with video-based feedback have significant learning gains compared to the other two groups. The findings indicate that the system can potentially serve as a highly effective supplementary training tool in the skill training in dental surgery education. This is the first use of simulator replay with visualization techniques to provide feedback in dental surgery. The video-playback provides a large amount of information to the user in an easily intelligible form, and the features and functionalities are simple and general

enough to incorporate into a wide variety of computer-based simulators. Through the video-replayed at the end of procedure, the trainees are informed with the immediacy of consequences which is usually not available in a conventional training setting. The locality of the mistakes/errors manifested in their performance in terms of both spatial and temporal aspects are captured and visualized in the video-playback. The video playback with formative feedback succinctly illustrates the general feedback mechanism of "action and effect" and this simple yet powerful feedback mechanism enables the trainees to comprehend the consequences of their actions which is essential in skill training.

Feedback from the evaluation study indicated several areas for extension and improvement. Experts and the students would like to export and store the video playback in the hardware independent multimedia file formats. Speed replay (such as x-times faster than average rate) can also be added into the system. In addition, kinematic graphs and the video-playback can be combined into one window. The system can be further extended by integrating of instrument sound while drilling and incorporating voice-over explanation of causes for each error region (such as you have over-exerted force in this region).

### **8.3 Haptic-based Feedback**

In addition to the video-based feedback, we have shown how haptic feedback can be used to train the correct application of force in dental surgery. The force-feedback represents the first time that haptics have been used to teach correct application of force in dental surgery. The mechanism enables analysis and feedback concerning the use of force which is otherwise difficult to provide in traditional dental skill training environments. Haptic sense is a unique among the senses in human beings and it is physically and functionally integrated with motor control [7]. The haptic training is one of the current research issues in human computer interaction studies involving computer-based simulators and haptic devices [4–8]. A common approach to haptic training consists of two phases. In the first phase, the expert's movement is recorded in terms of position, velocities, and force patterns. In the subsequent phase, the recorded movements are

haptically and visually displayed to learners during the playback mode training. Playback can be either active or passive mode. In the passive playback mode, the trainees have to grasp the end-effector of the haptic device and are physically guided through the ideal motion through a desired trajectory to acquire the kinesthetic understanding of what is required. In contrast, the trainee moves the end-effector through a desired trajectory at his/her speed in the active playback mode. Haptic sense can be trained through physical guidance. Physically guiding children how to write or how to move a swimming stroke to get maximum forward extension are a few example of physical guidance. Our approach deviates from the traditional haptic training approaches in utilizing the correlated information (formative feedback) from the correlator component. For the procedure stage containing portions of the procedure which are labeled as the portion responsible for a certain error region in the outcome, trainee has to undergo the haptic training. Using the expert's force in the identified procedure stage, the expert's force is rendered to the student via a haptic device and the student has to cancel it with the opposite force. Training on the proper use of force is more meaningful with the sense of touch (haptic cues) involved in the countering force. This is the first time that haptic devices have been used to teach correct application of force in endodontic surgery. From the preliminary evaluation using the randomized control trial with pre-/post- training control group, we found out that the students trained with the haptic feedback have significantly improve learning outcomes after training. The findings indicate that this assessment and feedback mechanism may be an effective addition to dental skill training. More exhaustive experiments on the effectiveness of force feedback mechanism with a larger sample of faculty and students, variety of teeth, different pathologies and retention effects should be covered out in the future.

#### **8.4 Correlator and Formative Feedback Framework**

The correlator component in our framework is responsible for correlating the procedure and outcome to generate formative feedback. Specifically, the correlator associates the error in the outcome with the procedure portions responsible for it. Cor-

relating procedure and outcome is an instance of credit assignment problem where the portions of the procedure which are responsible for the errors in performance outcome are identified. The credit-assignment problem is usually explained with the chess board game, where the win/loss status at the end of a chess game is apportioned as credit/blame to the moves taken by a player during the game. In this particular example, the actions (the moves of the player) have delayed effect since the game outcome win/loss is observable only at the end of the game. Having the action effects delayed till the end, it is not trivial to determine which of the player's moves were the most important (or detrimental) in leading to a win (or loss). Additionally, in a chess game, the individual moves are not completely independent from one another. A player's next move highly depends on the current board set reached from the previous moves. Usually a sequence of moves rather than one specific move have impact on the win or loss game outcome at the end. Specifically, the individual actions have indirect effect on the game outcome and this characteristic further underscores the hard credit assignment task. Unlike the chess game, the drilling in actions in the dental domain have immediate effects on the outcome. It simplifies the way to solve credit-assignment problem since the assignment of credit/blame only relies on the spatial information of errors (effects) to obtain the associating temporal information of actions. With the immediately observable effect of each action and the availability of localized spatial and temporal aspects of the effects, in our case we have a tractable solution to the credit assignment problem. It is important to note about the immediate effects of actions in the perforation error type. While overcut actions effects are immediately available in the outcome, perforation errors are not the immediate effects of the actions. As mentioned in the correlator component (Chapter 5), the repeated or extended drilling in the same neighborhood eventually leads to the perforation error, it is the special case of overcut error. Although, the action effect is not immediately available, immediate effects can be established from the continuous overcut actions leading to perforation.

In our prototype for the framework, the correlator relies on the locality of both temporal and spatial aspects of errors (effect of action) due the nature of error under study. The correlator works for the class of problems with the action(s) with immediate effects on the outcome. For example, consider a scenario playing a musical

instrument where the effects of an action has no associating spatial information. The correlator component can correlate the error (effects of action) in the resulting sound of a music score with how it was executed/played. In this case, the correlation is possible to establish because an action (incorrect key) or an episode of continuous actions have immediate and direct effect on the sound produced from the instrument. Specifically, the correlator component works with the class of problems which has the fundamental characteristic of having actions which have immediate effects on the outcome.

Continuing from the previous example of the welding simulator, the correlator can also be integrated into the training welding simulator to generate formative feedback as the action taken by trainee have immediate effect on the outcome of the welding process. Other simulators where the framework is applicable include calligraphy training simulators, and psychomotor rehabilitation training VR simulators to train the patients after a medical condition such as stroke or injury in hand.

Formative feedback can be further improved by providing the image-guided instructions as feedback during the procedure execution to reduce the chances of students committing the errors. Based on the reference templates, the simulator knows the ideal drilling area and shape, and guidance instructions can be generated as feedback while the students are performing the procedure. Besides the reference templates, the important anatomical landmarks can be highlighted to assist in student decisions in the drilling process to achieve a better outcome. Frequent feedback during the training could disrupt the learning process. Therefore, the feature to turn on/off the feedback can also be integrated in the future.

In determining the factors that contributed to the errors, we have focused only on the applied force and the orientation of the instrument during the procedure. However, the other variables such as the kinematics variables associated with the mirror could indicate whether the trainee properly manipulate whenever the indirect view of the operating tooth is necessary or not. Regarding the error types in correlating with the portion of the procedure, only the three most common types associated with the selected procedure stage are taken into consideration in this thesis. Errors can be further analyzed in more detailed level, for example the perforation error could be separated into either lateral or vertical perforation. Similarly, the undercut error could be sepa-

rated from incomplete error (unfinished task outcome). By analyzing at the finer level, the trainee's understanding on the task, the tooth anatomy and morphology can be determined and associated with the procedural errors. In this regard, this thesis can be extended by distinguishing the knowledge origins of the errors. Each error can be analyzed whether it is caused by the technical limitation of the trainee or the student does not have sufficient underlying knowledge to perform the task. This information will be instrumental in deriving directive feedback on correcting errors, inappropriate actions, or misconceptions.

The objective assessment and feedback generation framework in this thesis originated in dentistry. Beyond the domain of dental surgery, we have discussed examples on how each component is applicable in other procedures within dentistry, in general medicine as well as in other disciplines. Being implemented using a VR dental surgical simulator, the framework fundamentally requires a VR simulator with 3D volumetric object models. While scoring component requires the ideal outcome template, the framework requires the simulator to be equipped with the ability to capture kinematics variables along with temporal and spatial information on collided and removed voxels. Other than the simulator features, the framework is generalizable to the class of problems constitute with the actions with immediate effects on the outcome.

## **8.5 Conclusion**

To develop the required level of skills, a large portion of the curriculum is usually allocated to skill training in dental schools. Scarcity of expert time and the limitations of traditional approaches have spurred a search for alternative methods of instruction to maximize procedural skills development for the dental students. With increasing attention in the development of learner-centered curricula, our framework implemented on VR simulators could provide a valuable training resource that allows the trainee to perform independent practice with less reliance on experts. Our objective assessment and formative feedback system could be incorporated into formal skills training curricula. We would like to emphasize that virtual simulators cannot replace the

experts during training but rather complement the experts in the training process. When both the expert and the simulator actively engage in the training process, the benefits are multifold. Expert's time and workload could significantly reduce with the addition of VR simulators equipped with assessment and feedback features. Simulators can provide opportunities for students to practice independently with feedback from the simulator and their outcome automatically evaluated. Students can reflect their performance using video-based feedback from the simulators and revise the actions to improve their future performance. Expert performance on the same task can be reviewed using simulator through which trainees can learn from the expert without requiring the presence of experts during practice sessions. While simulators are the perfect platform for the deliberate practices, they can never replicate fully the clinical experience of the expert. As the simulators takes assessment and feedback for each practice session, experts can focus on qualitative feedback aspects of skill training. In a surgery, the incorrect manipulation of instruments of the dentist will be reflected as errors in the outcome, nevertheless, the correct outcome does not guarantee the correct/precise manipulation of instruments by the students. Beside the tracking sensors (if any), the experts can focus on observing students performances during practice and providing instructions to ensure the trainees adopt the correct instrument manipulations techniques and perform procedure stages in an efficient manner during the practice. Beyond instrument manipulation, experts can also observe the trainee's operating posture and the ergonomic aspects as well. Additionally, experts can confirm the students' understanding on the task, their planned actions, projected outcome and the task's ideal outcome through various tutoring intervention strategies. Additionally, the expert can provide and discuss visual cues the student should consider for the operation such as important landmarks which are not available in the textbook. Moreover, either expert or student can use the simulator with a haptic training feature to train the correct application of force during practice sessions as well.



## REFERENCES

- [1] Stefanidis, D., Arora, S., Parrack, D. M., Hamad, G. G., Capella, J., Grantcharov, T., Urbach, D. R., Scott, D. J., and Jones, D. B. (2012). Research priorities in surgical simulation for the 21st century. *The American journal of surgery*, 203(1), 49–53.
- [2] Overby, D. W. and Watson, R. A. (2014). Hand motion patterns of fundamentals of laparoscopic surgery certified and noncertified surgeons. *The American Journal of Surgery*, 207(2), 226–230.
- [3] Minsky, M. (1961). Steps toward artificial intelligence. *Proceedings of the IRE*, 49(1), 8–30.
- [4] Neges, M., Adwernat, S., and Abramovici, M. (2018). Augmented virtuality for maintenance training simulation under various stress conditions. *Procedia Manufacturing*, 19, 171–178.
- [5] Campbell, J., Hogan, T., and Fraser, M. (2018). Feeling virtual worlds: An exploration into coupling virtual and kinaesthetic experiences. *Proceedings of the Twelfth International Conference on Tangible, Embedded, and Embodied Interaction*, pp. 279–285.
- [6] Vi, C. T., Ablart, D., Gatti, E., Velasco, C., and Obrist, M. (2017). Not just seeing, but also feeling art: Mid-air haptic experiences integrated in a multisensory art exhibition. *International Journal of Human-Computer Studies*, 108, 1–14.
- [7] MacLean, K. E. (2008). Haptic interaction design for everyday interfaces. *Reviews of Human Factors and Ergonomics*, 4(1), 149–194.
- [8] Schneider, O., MacLean, K., Swindells, C., and Booth, K. (2017). Haptic experience design: What hapticians do and where they need help. *International Journal of Human-Computer Studies*, 107, 5–21.

- [9] White, C., Rodger, M. W. M., and Tang, T. (2016). Current understanding of learning psychomotor skills and the impact on teaching laparoscopic surgical skills. *Obstet. Gynecol.*, 18(1), 53–63.
- [10] Spencer, F. C. (1978). Teaching and measuring surgical Techniques-The technical evaluation of competence. *Bull. Am. Coll. Surg.*, 63, 9–12.
- [11] Müller-Tomfelde, C. (2004). Interaction sound feedback in a haptic virtual environment to improve motor skill acquisition. *International Conference on Auditory Display*.
- [12] Dreyfus, S. E. and Dreyfus, H. L. (1980). *A five-stage model of the mental activities involved in directed skill acquisition*. Operations Research Center, University of California, Berkeley.
- [13] Fitts, P. M. and Posner, M. I. (1967). *Human Performance*. Brooks/Cole.
- [14] Bloom, B. S. (1971). *Taxonomy of Educational Objectives: The Classification of Educational Goals*. David McKay Co., Inc.
- [15] Field, A. (2014). Understanding the dreyfus model of skill acquisition to improve ultrasound training for obstetrics and gynaecology trainees. *Ultrasound*, 22(2), 118–122.
- [16] Carraccio, C. L., Benson, B. J., Nixon, L. J., and Derstine, P. L. (2008). From the educational bench to the clinical bedside: translating the dreyfus developmental model to the learning of clinical skills. *Acad. Med.*, 83(8), 761–767.
- [17] Peña, A. (2010). The dreyfus model of clinical problem-solving skills acquisition: a critical perspective. *Med. Educ. Online*, 15.
- [18] Richard M. Dubinsky Mamatha Pasnoor Sonja Fabricius Richard J. Barohn Neurology residency handbook 2015-2016.
- [19] Miller, G. E. (1990). The assessment of clinical skills/competence/performance.
- [20] Anderson, L. W. and Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing : a revision of Bloom's taxonomy of educational objectives*. Longman.
- [21] Albino, J. E. N., Young, S. K., Neumann, L. M., Kramer, G. A., Andrieu, S. C., Henson, L., Horn, B., and Hendricson, W. D. (2008). Assessing dental

- students' competence: best practice recommendations in the performance assessment literature and investigation of current practices in predoctoral dental education. *J. Dent. Educ.*, 72(12), 1405–1435.
- [22] Dave, R. H. (1970). *Psychomotor levels in Developing and Writing Behavioral Objectives*, pp. 20-21. Tucson, Arizona: Educational Innovators Press.
- [23] Rogers, D. A., Regehr, G., Howdieshell, T. R., Yeh, K. A., and Palm, E. (2000). The impact of external feedback on computer-assisted learning for surgical technical skill training. *Am. J. Surg.*, 179(4), 341–343.
- [24] Ali, K., Slade, A., Kay, E., Zahra, D., and Tredwin, C. (2017). Preparedness of undergraduate dental students in the united kingdom: a national study. *Br. Dent. J.*, 222(6), 472–477.
- [25] Tanalp, J., Güven, E. P., and Oktay, I. (2013). Evaluation of dental students' perception and self-confidence levels regarding endodontic treatment. *Eur. J. Dent.*, 7(2), 218–224.
- [26] Osnes, C. and Keeling, A. (2017). Developing haptic caries simulation for dental education. *JSS*, 4, 29–34.
- [27] Nascimento, M. M., Behar-Horenstein, L. S., and others (2017). Exploring how US dental schools teach removal of carious tissues during cavity preparations. *Journal of dental*, 81(1), 5–13.
- [28] de Boer, I. R., Wesselink, P. R., and Vervoorn, J. M. (2013). The creation of virtual teeth with and without tooth pathology for a virtual learning environment in dental education. *Eur. J. Dent. Educ.*, 17(4), 191–197.
- [29] Piromchai, P., Ioannou, I., Wijewickrema, S., Kasemsiri, P., Lodge, J., Kennedy, G., and O'Leary, S. (2017). Effects of anatomical variation on trainee performance in a virtual reality temporal bone surgery simulator. *The Journal of Laryngology & Otology*, 131(S1), S29–S35.
- [30] Francis, D. M. A. (2009). Surgical decision making. *ANZ J. Surg.*, 79(12), 886–891.
- [31] Fugill, M. (2005). Teaching and learning in dental student clinical practice. *Eur. J. Dent. Educ.*, 9(3), 131–136.

- [32] Jenkins, S. M., Dummer, P. M., Gilmour, A. S., Edmunds, D. H., Hicks, R., and Ash, P. (1998). Evaluating undergraduate preclinical operative skill; use of a glance and grade marking system. *J. Dent.*, 26(8), 679–684.
- [33] Schmitt, L., Möltner, A., Rüttermann, S., and Gerhardt-Szép, S. (2016). Study on the interrater reliability of an OSPE (objective structured practical examination) - subject to the evaluation mode in the phantom course of operative dentistry. *GMS J Med Educ*, 33(4), Doc61.
- [34] Taylor, C. L., Grey, N., and Satterthwaite, J. D. (2013). Assessing the clinical skills of dental students: A review of the literature. *J. Stat. Educ.*, 2(1), 20–31.
- [35] Satterthwaite, J. D. and Grey, N. J. A. (2008). Peer-group assessment of pre-clinical operative skills in restorative dentistry and comparison with experienced assessors. *Eur. J. Dent. Educ.*, 12(2), 99–102.
- [36] Goepferd, S. J. and Kerber, P. E. (1980). A comparison of two methods for evaluating primary class II cavity preparations. *J. Dent. Educ.*, 44(9), 537–542.
- [37] Dhuru, V. B., Rypel, T. S., and Johnston, W. M. (1978). Criterion-oriented grading system for preclinical operative dentistry laboratory course. *J. Dent. Educ.*, 42(9), 528–531.
- [38] Fuller, J. L. (1972). The effects of training and criterion models on interjudge reliability. *J. Dent. Educ.*, 36(4), 19–22.
- [39] Lilley, J. D., ten Bruggen Cate, H. J., Holloway, P. J., Holt, J. K., and Start, K. B. (1968). Reliability of practical tests in operative dentistry. *Br. Dent. J.*, 125(5), 194–197.
- [40] Gaines, W. G., Bruggers, H., and Rasmussen, R. H. (1974). Reliability of ratings in preclinical fixed prosthodontics: effect of objective scaling. *J. Dent. Educ.*, 38(12), 672–675.
- [41] Houpt, I.—Kress, M., and Gerard Accuracy of measurement of clinical performance in dentistry. <http://files.eric.ed.gov/fulltext/ED080576.pdf>. Accessed: 2017-11-25.
- [42] Salvendy, G., Root, C., Cunningham, P. R., Ferguson, G. W., Hinton, W. M., Baum, S., and Khan, L. (1973). Skills analysis of cavity preparations: class 1 in mandibular right first molar. *J. Dent. Educ.*, 37(10), 11–18.

- [43] Paskins, Z., Kirkcaldy, J., Allen, M., Macdougall, C., Fraser, I., and Peile, E. (2010). Design, validation and dissemination of an undergraduate assessment tool using SimMan® in simulated medical emergencies. *Med. Teach.*, 32(1), e12–e17.
- [44] Quinn, F., Keogh, P., McDonald, A., and Hussey, D. (2003). A pilot study comparing the effectiveness of conventional training and virtual reality simulation in the skills acquisition of junior dental students. *Eur. J. Dent. Educ.*, 7(1), 13–19.
- [45] Hassanpour, N., Chen, R., Baikpour, M., and Moghimi, S. (2016). Video observation of procedural skills for assessment of trabeculectomy performed by residents. *Journal of Current Ophthalmology*, 28(2), 61–64.
- [46] Ranta, J. F. and Aviles, W. A. (1999). The virtual reality dental training system: simulating dental procedures for the purpose of training dental students using haptics. *Proceedings of the fourth PHANTOM users group*, 4, 67–71.
- [47] Thomas, G., Johnson, L., Dow, S., and Stanford, C. (2001). The design and testing of a force feedback dental simulator. *Comput. Methods Programs Biomed.*, 64(1), 53–64.
- [48] Marras, I., Papaleontiou, L., Nikolaidis, N., Lyroudia, K., and Pitas, I. (2006). Virtual dental patient: a system for virtual teeth drilling. *2006 IEEE International Conference on Multimedia and Expo*, July, pp. 665–668.
- [49] Voxel-Man. <http://voxel-man.de/>. Accessed: 2017-12-4.
- [50] Forsslund, J., Sallnas, E. L., and Palmerius, K. J. (2009). A user-centered designed foss implementation of bone surgery simulations. *Proceedings - 3rd Joint EuroHaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, World Haptics 2009*, pp. 391–392.
- [51] Bakr, M., Massey, W., and Alexander, H. (2012). Academic evaluation of simodont® haptic 3D virtual reality dental training simulator. *Gold Coast Health and Medical Research Conference*, pp. 1574–1582.
- [52] Steinberg, A. D., Bashook, P. G., Drummond, J., Ashrafi, S., and Zefran, M. (2007). Assessment of faculty perception of content validity of PerioSim,

- a haptic-3d virtual reality dental training simulator. *J. Dent. Educ.*, 71(12), 1574–1582.
- [53] Tse, B., Harwin, W., Barrow, A., Quinn, B., Diego, J. S., and Cox, M. (2010). Design and development of a haptic dental training system - hapTEL. *Eurohaptics*, pp. 101–108.
- [54] Wang, D., Zhang, Y., Hou, J., Wang, Y., Lv, P., Chen, Y., and Zhao, H. (2012). IDental: A haptic-based dental simulator and its preliminary user evaluation. *IEEE Trans. Haptics*, 5(4), 332–343.
- [55] Rhienmora, P., Haddawy, P., Suebnukarn, S., and Dailey, M. N. (2011). Intelligent dental training simulator with objective skill assessment and feedback. *Artif. Intell. Med.*, 52(2), 115–121.
- [56] Kim, L. and Park, S. H. (2006). Haptic interaction and volume modeling techniques for realistic dental simulation. *Visual Comput.*, 22(2), 90–98.
- [57] Xia, P., Lopes, A. M., and Restivo, M. T. (2012). Virtual reality and haptics for dental surgery: a personal review. *Vis. Comput.*, 29(5), 433–447.
- [58] Wang, D., Li, T., Zhang, Y., and Hou, J. (2016). Survey on multisensory feedback virtual reality dental training systems. *Eur. J. Dent. Educ.*, 20(4), 248–260.
- [59] Bogdan, C. M. (2011). Domain ontology of the VirDenT system. *International Journal of Computers Communications & Control*, 6(1), 45–52.
- [60] Bogdan, C. M. and Popovici, D. M. (2012). Information system analysis of an e-learning system used for dental restorations simulation. *Comput. Methods Programs Biomed.*, 107(3), 357–366.
- [61] Popovici, D.-M., Hamza-Lup, F. G., and Bogdan, C. M. (2013). Haptic feedback systems in education. *The International Scientific Conference eLearning and Software for Education*, p. 509.
- [62] Dalai, D., Bhaskar, D., Agali, C., Singh, N., Gupta, D., and Bumb, S. (2014). Futuristic application of nano-robots in dentistry. *Int. J. Adv. Health Sci.*, 1, 16–20.
- [63] Tavkar, A. and Pawar, A. (2017). Simulation in dentistry. *EC Dental Science*, 12, 115–121.

- [64] Tanzawa, T., Futaki, K., Tani, C., Hasegawa, T., Yamamoto, M., Miyazaki, T., and Maki, K. (2012). Introduction of a robot patient into dental education. *Eur. J. Dent. Educ.*, 16(1), e195–9.
- [65] Gallagher, A. G. and Matt Ritter, E. (2007). Virtual reality: Objective assessment, education, and training. *Emerging Technologies in Surgery*, pp. 27–33. Springer.
- [66] Buchanan, J. A. (2001). Use of simulation technology in dental education. *J. Dent. Educ.*, 65(11), 1225–1231.
- [67] DentSim is the industry leader in augmented reality dental simulation helping dental students around the world improve their skills. <http://image-navigation.com/home-page/dentsim/>. Accessed: 2017-12-4.
- [68] Szalay, D., MacRae, H., Regehr, G., and Reznick, R. (2000). Using operative outcome to assess technical skill. *Am. J. Surg.*, 180(3), 234–237.
- [69] Datta, V., Mandalia, M., Mackay, S., Chang, A., Cheshire, N., and Darzi, A. (2002). Relationship between skill and outcome in the laboratory-based model. *Surgery*, 131(3), 318–323.
- [70] Wierinck, E. R., Puttemans, V., Swinnen, S. P., and van Steenberghe, D. (2007). Expert performance on a virtual reality simulation system. *J. Dent. Educ.*, 71(6), 759–766.
- [71] Renne, W. G., McGill, S. T., Mennito, A. S., Wolf, B. J., Marlow, N. M., Shaftman, S., and Holmes, J. R. (2013). E4D compare software: an alternative to faculty grading in dental education. *J. Dent. Educ.*, 77(2), 168–175.
- [72] Callan, R. S., Palladino, C. L., Furness, A. R., Bundy, E. L., and Ange, B. L. (2014). Effectiveness and feasibility of utilizing E4D technology as a teaching tool in a preclinical dental education environment. *J. Dent. Educ.*, 78(10).
- [73] Zou, H., Jin, S., Sun, J., and Dai, Y. (2016). A cavity preparation evaluation system in the skill assessment of dental students. *J. Dent. Educ.*, 80(8), 930–937.
- [74] Cardoso, J. A., Barbosa, C., Fernandes, S., and others (2006). Reducing subjectivity in the evaluation of preclinical dental preparations for fixed prosthodont-

- tics using the kavo PrepAssistant®. *European journal of dental education*, 10(3), 149–156.
- [75] Schwindling, F. S., Deisenhofer, U. K., Porsche, M., Rammelsberg, P., Kappel, S., and Stober, T. (2015). Establishing CAD/CAM in preclinical dental education: Evaluation of a Hands-On module. *J. Dent. Educ.*, 79(10), 1215–1221.
- [76] Suebnukarn, S., Phatthanasathiankul, N., Sombatweroje, S., Rhienmora, P., and Haddawy, P. (2009). Process and outcome measures of expert/novice performance on a haptic virtual reality system. *J. Dent.*, 37(9), 658–665.
- [77] Suebnukarn, S., Haddawy, P., Rhienmora, P., and Gajananan, K. (2010). Haptic virtual reality for skill acquisition in endodontics. *J. Endod.*, 36(1), 53–55.
- [78] van de Ridder, J. M. M., Stokking, K. M., McGaghie, W. C., and ten Cate, O. T. J. (2008). What is feedback in clinical education? *Med. Educ.*, 42(2), 189–197.
- [79] Dharra A Sharma, Mohamed Faisal Chevidikunnan, Fayaz Rahman Khan, Riziq Allah Gaowgzeh (2016). Effectiveness of knowledge of result and knowledge of performance in the learning of a skilled motor activity by healthy young adults. *J. Phys. Therapy Sci.*, 28(5), 1482–1486.
- [80] Baker, C. H. and Young, P. (1960). Feedback during training and retention of motor skills. *Can. J. Psychol.*, 14, 257–264.
- [81] Gordon, N. B. (1968). Guidance versus augmented feedback and motor skill. *J. Exp. Psychol.*, 77(1), 24–30.
- [82] Hester, R., Murphy, K., Brown, F. L., and Skilleter, A. J. (2010). Punishing an error improves learning: the influence of punishment magnitude on error-related neural activity and subsequent learning. *J. Neurosci.*, 30(46), 15600–15607.
- [83] Wolpert, D. M., Diedrichsen, J., and Flanagan, J. R. (2011). Principles of sensorimotor learning. *Nat. Rev. Neurosci.*, 12(December), 1–13.
- [84] Yousif, N. and Diedrichsen, J. (2012). Structural learning in feedforward and feedback control. *J. Neurophysiol.*, 108(9), 2373–2382.



- [85] Mushtaq, F., Stoet, G., Bland, A. R., and Schaefer, A. (2013). Relative changes from prior reward contingencies can constrain brain correlates of outcome monitoring. *PLoS One*, 8(6), e66350.
- [86] Mahmood, T. and Darzi, A. (2004). The learning curve for a colonoscopy simulator in the absence of any feedback: No feedback, no learning. *Surg. Endosc.*, 18(8), 1224–1230.
- [87] Kannappan, A., Yip, D. T., Lodhia, N. A., Morton, J., and Lau, J. N. (2012). The effect of positive and negative verbal feedback on surgical skills performance and motivation. *Journal of Surgical Education*, November, pp. 798–801.
- [88] Porte, M. C., Xeroulis, G., Reznick, R. K., and Dubrowski, A. (2007). Verbal feedback from an expert is more effective than self-accessed feedback about motion efficiency in learning new surgical skills. *Am. J. Surg.*, 193(1), 105–110.
- [89] Hesketh, E. A. and Laidlaw, J. M. (2002). Developing the teaching instinct, 1: feedback. *Med. Teach.*, 24(3), 245–248.
- [90] El Boghdady, M. and Alijani, A. (2017). Feedback in surgical education.
- [91] Feil, P. H. and Gatti, J. J. (1995). Validation of a motor skills performance theory with applications for dental education. *J. Dent. Educ.*, 59, 631–639.
- [92] Sherwood, L. (2015). *Human Physiology: From Cells to Systems*. Cengage Learning.
- [93] Al-Saud, L. M., Mushtaq, F., Allsop, M. J., Culmer, P. C., Mirghani, I., Yates, E., Keeling, A., Mon-Williams, M. A., and Manogue, M. (2017). Feedback and motor skill acquisition using a haptic dental simulator. *Eur. J. Dent. Educ.*, 21(4), 240–247.
- [94] Feil, P. H., Reed, T., and Hart, J. K. (1986). Continuous knowledge of results and psychomotor skill acquisition. *J. Dent. Educ.*, 50(6), 300–303.
- [95] Sigrist, R., Rauter, G., Riener, R., and Wolf, P. (2013). Augmented visual, auditory, haptic, and multimodal feedback in motor learning: a review. *Psychon. Bull. Rev.*, 20(1), 21–53.
- [96] Shea, C. H. and Wulf, G. (1999). Enhancing motor learning through external-focus instructions and feedback. *Hum. Mov. Sci.*, 18(4), 553–571.

- [97] Wulf, G. and Shea, C. H. (2002). Principles derived from the study of simple skills do not generalize to complex skill learning. *Psychon. Bull. Rev.*, 9(2), 185–211.
- [98] Huegel, J. C. and O'Malley, M. K. (2010). Progressive haptic and visual guidance for training in a virtual dynamic task. *2010 IEEE Haptics Symposium*, March, pp. 343–350.
- [99] Dunnington, G. L., Wright, K., and Hoffman, K. (1994). A pilot experience with competency-based clinical skills assessment in a surgical clerkship. *Am. J. Surg.*, 167(6), 604–6; discussion 606–7.
- [100] Wierinck, E., Puttemans, V., Swinnen, S., and van Steenberghe, D. (2005). Effect of augmented visual feedback from a virtual reality simulation system on manual dexterity training. *Eur. J. Dent. Educ.*, 9(1), 10–16.
- [101] Suebnukarn, S., Haddawy, P., Rhienmora, P., Jittimanee, P., and Viratket, P. (2010). Augmented kinematic feedback from haptic virtual reality for dental skill acquisition. *J. Dent. Educ.*, 74(12), 1357–1366.
- [102] Rhienmora, P., Haddawy, P., Suebnukarn, S., and Dailey, M. N. (2009). Providing objective feedback on skill assessment in a dental surgical training simulator. In Combi, C., Shahar, Y., and Abu-Hanna, A. (eds.), *Proceedings of the 12th Conference on Artificial Intelligence in Medicine (AIME 2009)*, pp. 18–22.
- [103] Wierinck, E., Puttemans, V., and van Steenberghe, D. (2006). Effect of tutorial input in addition to augmented feedback on manual dexterity training and its retention. *Eur. J. Dent. Educ.*, 10, 24–31.
- [104] Buchanan, J. A. (2004). Experience with virtual reality-based technology in teaching restorative dental procedures. *J. Dent. Educ.*, 68(12), 1258–1265.
- [105] Patil, J., Saran, N., and Shetty, R. (2011). Haptic-based virtual reality dental simulator as an educational tool. *Endodontology*, 23(2), 114–119.
- [106] Bakker, D., Lagerweij, M., Wesselink, P., and others (2010). Transfer of manual dexterity skills acquired in the simodont, a dental haptic trainer with a virtual environment, to reality: a pilot study. *Bio-algorithms and Med*, 6(11), 21–24.

- [107] Shute, V. J. (2008). Focus on formative feedback. *Rev. Educ. Res.*, 78(1), 153–189.
- [108] Cannon-Bowers, J. A., Bowers, C., and Procci, K. (2010). Optimizing learning in surgical simulations: guidelines from the science of learning and human performance. *Surg. Clin. North Am.*, 90(3), 583–603.
- [109] Schroth, M. L. (1992). The effects of delay of feedback on a delayed concept formation transfer task. *Contemp. Educ. Psychol.*, 17(1), 78–82.
- [110] Eichenberg, Christiane, and Ed (2012). *Virtual Reality in Psychological, Medical and Pedagogical Applications*. InTech. 41 Madison Avenue 31st Floor, New York, NY 10010. Tel: 917-769-7640; e-mail: info@intechopen.com; Web site: <http://www.intechopen.com>.
- [111] Vanlehn, K. (2006). The behavior of tutoring systems. *Int. J. Artif. Intell. Ed.*, 16(3), 227–265.
- [112] Charles Catania, A. (2013). *Learning*, 5th edition edition. Sloan Publishing.
- [113] Gallagher, A. G. and O’Sullivan, G. C. (2012). *Fundamentals of Surgical Simulation—Principles and Practices*. Springer London, London.
- [114] Baron, R. A. (1988). Negative effects of destructive criticism: impact on conflict, self-efficacy, and task performance. *Journal of Applied Psychology*, 73, 199–207.
- [115] Sargeant, J., Mann, K., Sinclair, D., Van der Vleuten, C., and Metsemakers, J. (2008). Understanding the influence of emotions and reflection upon multi-source feedback acceptance and use. *Adv. Health Sci. Educ. Theory Pract.*, 13(3), 275–288.
- [116] Boyle, E., Al-Akash, M., Gallagher, A. G., Traynor, O., Hill, A. D. K., and Neary, P. C. (2011). Optimising surgical training: use of feedback to reduce errors during a simulated surgical procedure. *Postgrad. Med. J.*, 87(1030), 524–528.
- [117] Salmoni, a. W., Schmidt, R. a., and Walter, C. B. (1984). Knowledge of results and motor learning: a review and critical reappraisal. *Psychol. Bull.*, 95(3), 355–386.

- [118] Palter, V. N. (2011). Comprehensive training curricula for minimally invasive surgery. *Journal of graduate medical education*, 3(3), 293–298.
- [119] Stefanidis, D. (2010). Optimal acquisition and assessment of proficiency on simulators in surgery. *Surg. Clin. North Am.*, 90(3), 475–489.
- [120] Bangert-Drowns, R. L., Kulik, C.-L. C., Kulik, J. A., and Morgan, M. (1991). The instructional effect of feedback in Test-Like events. *Rev. Educ. Res.*, 61(2), 213–238.
- [121] Archer, J. C. (2010). State of the science in health professional education: Effective feedback.
- [122] Alsulaimani, R. S., Al-Manei, K., Baras, B., Alaqeely, R., El Metwally, A., and Ashri, N. (2016). Students' perception of multiple dental visits for root canal treatment: Questionnaire-based study. *Saudi Endodontic Journal*, 6(1), 21.
- [123] El-Kishawi, M. Y. (2016). Improving students' learning and performance in pre-clinical endodontics. PhD thesis School of Dentistry.
- [124] Newell, K. M. (1991). Motor skill acquisition. *Annu. Rev. Psychol.*, 42, 213–237.
- [125] Seijo, M. O., Ferreira, E. F., Ribeiro Sobrinho, A. P., Paiva, S. M., and Martins, R. C. (2013). Learning experience in endodontics: Brazilian students' perceptions. *J. Dent. Educ.*, 77(5), 648–655.
- [126] Ingle, J. I., Bakland, L. K., and Baumgartner, J. C. (2008). *Endodontics 6*. PMPH-USA.
- [127] Touch, G. Haptic force feedback device. Accessed: 2017-12-3.
- [128] Suebnukarn, S., Chaisombat, M., Kongpunwijit, T., and Rhienmora, P. (2014). Construct validity and expert benchmarking of the haptic virtual reality dental simulator. *J. Dent. Educ.*, 78(10), 1442–1450.
- [129] Schiff, A. J., Salvendy, G., Root, C. M., Ferguson, G. W., and Cunningham, P. R. (1975). Objective evaluation of quality in cavity preparations. *J. Dent. Educ.*, 39(2), 92–96.
- [130] Arnetz, G. and Dornhofer, R. (2004). PREPassistant: a system for evaluating tooth preparations. *Int. J. Comput. Dent.*, 7(2), 187–197.

- [131] Ziv, W., Small (2000). Patient safety and simulation-based medical education. *Med. Teach.*, 22(5), 489–495.
- [132] Dentistry, P. i. C. Contribution of haptic simulation for learning in odontology. <https://www.adee.org/>.
- [133] Gal, G. B., Weiss, E. I., Gafni, N., and Ziv, A. (2011). Preliminary assessment of faculty and student perception of a haptic virtual reality simulator for training dental manual dexterity. *J. Dent. Educ.*, 75(4), 496–504.
- [134] Mirghani, I., Mushtaq, F., Allsop, M. J., Al-Saud, L. M., Tickhill, N., Potter, C., Keeling, A., Mon-Williams, M. A., and Manogue, M. (2016). Capturing differences in dental training using a virtual reality simulator. *Eur. J. Dent. Educ.*, 22(1), 67–71.
- [135] EPED. <http://www.eped.com.tw/>.
- [136] Montanet, G. C., Sánchez, M. G., and García, A. S. (2017). Haptic simulators with virtual reality environments in dental education: A preliminary teaching diagnosis. *@tic. revista d'innovació educativa*, 0(18), 14.
- [137] Intro. <https://www.forsslundsystems.com/>. Accessed: 2017-12-4.
- [138] Baud, M. (1993). Data analysis, mathematical modeling. In Masseyeff, R., Albert, W., and Staines, N. A. (eds.), *Methods of Immunological Analysis, Vol .1: Fundamentals*, pp. 656–671. VCH Publishers, Inc., New Yourk.
- [139] Henriques, T., Antunes, L., Bernardes, J., Matias, M., Sato, D., and Costa-Santos, C. (2013). Information-based measure of disagreement for more than two observers: a useful tool to compare the degree of observer disagreement. *BMC Med. Res. Methodol.*, 13(1), 47.
- [140] Fawcett, T. (2014). An introduction to ROC analysis. *Pattern Recognit. Lett.*, 35(6), 299–309.
- [141] Alrahabi, M. (2017). The confidence of undergraduate dental students in saudi arabia in performing endodontic treatment. *Eur. J. Dent.*, 11(1), 17–21.
- [142] Kramer, G. a., Albino, J. E. N., Andrieu, S. C., Hendricson, W. D., Henson, L., Horn, B. D., Neumann, L. M., and Young, S. K. (2009). Dental student assessment toolbox. *J. Dent. Educ.*, 73(1), 12–35.

- [143] Manakil, J. and George, R. (2011). Reviewing competency in dental education. *International Journal of Dental Clinics*, 3(2), 33–39.
- [144] Moor, R., Hülsmann, M., Kirkevang International, L. L., and 2013 (2013). Undergraduate curriculum guidelines for endodontology. *International endodontic journal*, 46(12), 1105–1114.
- [145] Darzi, A., Smith, S., and Taffinder, N. (1999). Assessing operative skill. *BMJ*, 318(7188), 887–888.
- [146] Strandbygaard, J., Scheele, F., and Sørensen, J. L. (2017). Twelve tips for assessing surgical performance and use of technical assessment scales. *Med. Teach.*, 39(1), 32–37.
- [147] Birkmeyer, J. D., Finks, J. F., O'Reilly, A., Oerline, M., Carlin, A. M., Nunn, A. R., Dimick, J., Banerjee, M., Birkmeyer, N. J. O., and Michigan Bariatric Surgery Collaborative (2013). Surgical skill and complication rates after bariatric surgery. *N. Engl. J. Med.*, 369(15), 1434–1442.
- [148] Hamad, G. G., Brown, M. T., and Clavijo-Alvarez, J. A. (2007). Postoperative video debriefing reduces technical errors in laparoscopic surgery. *Am. J. Surg.*, 194(1), 110–114.
- [149] Nakada, S. Y., Hedican, S. P., Bishoff, J. T., Shichman, S. J., and Wolf, J. S., Jr (2004). Expert videotape analysis and critiquing benefit laparoscopic skills training of urologists. *JSLs*, 8(2), 183–186.
- [150] Jamshidi, R., LaMasters, T., Eisenberg, D., Duh, Q.-Y., and Curet, M. (2009). Video self-assessment augments development of videoscopic suturing skill. *J. Am. Coll. Surg.*, 209(5), 622–625.
- [151] Muessig, Moritz and Sterz, Jasmina and Stefanescu, Maria-Christina and Bender, Bernd and Hofer, Sebastian H and Ruessler, Miriam (2017). The impact of video feedback on acquiring competency in basic surgical skills (sterile working) in an undergraduate medical training program: A comparative effectiveness analysis. *Journal of Advances in Education Research*, 2(3).
- [152] Vyasa, P., Willis, R. E., Dunkin, B. J., and Gardner, A. K. (2017). Are general surgery residents accurate assessors of their own flexible endoscopy skills? *J. Surg. Educ.*, 74(1), 23–29.

- [153] Backstein, D., Agnidis, Z., Regehr, G., and Reznick, R. (2004). The effectiveness of video feedback in the acquisition of orthopedic technical skills. *Am. J. Surg.*, 187(3), 427–432.
- [154] Byrne, A. J., Sellen, A. J., Jones, J. G., Aitkenhead, A. R., Hussain, S., Gilder, F., Smith, H. L., and Ribes, P. (2002). Effect of videotape feedback on anaesthetists' performance while managing simulated anaesthetic crises: a multicentre study. *Anaesthesia*, 57(2), 176–179.
- [155] Kolb, D. A. (2014). *Experiential Learning: Experience as the Source of Learning and Development*. FT Press.
- [156] Boud, D. (1987). A facilitator's view of adult learning. *Appreciating Adults Learning: From the Learner's Perspective*, London: Kogan Page, 4(1), 222–239.
- [157] Schon, D. A. (1984). *The reflective practitioner: How professionals think in action*. Basic books.
- [158] Frances F. Fuller and Brad A. Manning (1973). Self-Confrontation reviewed: A conceptualization for video playback in teacher education. *Rev. Educ. Res.*, 43(4), 469–528.
- [159] Ray E. Hosford (1981). Self-as-a-Model: A cognitive social learning technique. *Couns. Psychol.*, 9(1), 45–62.
- [160] Duță, M., Amariei, C. I., Bogdan, C. M., Popovici, D. M., Ionescu, N., and Nuca, C. I. (2011). An overview of virtual and augmented reality in dental education. *Journal of Oral Health and Dental Management*, 10(1).
- [161] Suebnukarn, S., Hataidechadusadee, R., Suwannasri, N., Suprasert, N., Rhiemora, P., and Haddawy, P. (2011). Access cavity preparation training using haptic virtual reality and microcomputed tomography tooth models. *Int. Endod. J.*, 44(11), 983–989.
- [162] Chan, Y. H. (2003). Randomised controlled trials (RCTS) - sample size: The magic number? *Singapore Med. J.*, 44(4), 172–174.
- [163] PsycNET. <http://psycnet.apa.org/psycinfo/1997-08357-000>. Accessed: 2017-12-17.

- [164] Burnett, A. C. and Linden, G. J. (1988). The reproducibility of the assessment of restorations by dental students and their teachers. *J. Dent. Educ.*, 52(10), 568–570.
- [165] Williams, J. C., Baillie, S., Rhind, S. M., Warman, S., Sandy, J., and Ireland, A. (2015). A guide to assessment in dental education. <http://www.bristol.ac.uk/pure/about/abr-terms.html>. Accessed: 2015-1-17.
- [166] Weeks, D. L. and Kordus, R. N. (1998). Relative frequency of knowledge of performance and motor skill learning. *Res. Q. Exerc. Sport*, 69(3), 224–230.
- [167] Hauser, A. M. and Bowen, D. M. (2009). Primer on preclinical instruction and evaluation. *J. Dent. Educ.*, 73(3), 390–398.
- [168] Roy, E., Bakr, M. M., and George, R. (2017). The need for virtual reality simulators in dental education: A review. *Saudi Dent J*, 29(2), 41–47.
- [169] Feil, P. H., Guenzel, P. J., Knight, G. W., and Geistfeld, R. (1994). Designing preclinical instruction for psychomotor skills (i)–theoretical foundations of motor skill performance and their applications to dental education. *J. Dent. Educ.*, 58(11-12), 806–812.
- [170] Magill, R. A. and Anderson, D. (2013). Sensory components of motor control. *Motor Learning and Control: Concepts and Applications*, pp. 112–138A. McGraw-Hill Higher Education, New York.
- [171] Anning, A. (2001). Knowing who I am and what I know: developing new versions of professional knowledge in integrated service settings. <http://www.leeds.ac.uk/educol/documents/00001877.htm>.
- [172] Corno, M. (2006). Haptic playback : Modeling , controller design , and stability analysis. In Sukhatme, G. S., Schaal, S., Burgard, W., and Fox, D. (eds.), *Robotics: Science and Systems II*.
- [173] Yamaguchi, S., Yoshida, Y., Noborio, H., Murakami, S., and Imazato, S. (2013). The usefulness of a haptic virtual reality simulator with repetitive training to teach caries removal and periodontal pocket probing skills. *Dent. Mater. J.*, 32(5), 847–852.



- [174] Kolesnikov, M., Žefran, M., Steinberg, A. D., and Bashook, P. G. (2009). PerioSim: Haptic virtual reality simulator for sensorimotor skill acquisition in dentistry. *Proceedings - IEEE International Conference on Robotics and Automation*, pp. 689–694.
- [175] Luciano, C., Banerjee, P., and DeFanti, T. (2009). Haptics-based virtual reality periodontal training simulator. *Virtual Real.*, 13(2), 69–85.
- [176] Rhienmora, P., Gajananan, K., Haddawy, P., Dailey, M. N., and Suebnukarn, S. (2010). Augmented reality haptics system for dental surgical skills training. *Proceedings of the 17th ACM Symposium on Virtual Reality Software and Technology - VRST '10*, 1(212), 97–98.
- [177] Medical and dental. <http://www.novint.com/index.php/medicaldental>. Accessed: 2017-11-25.
- [178] Mm, B. (2016). Can virtual simulators replace traditional preclinical teaching methods: A students' perspective? *Int J Dent Oral Health*, 2(1).
- [179] Yokokohji, Y., Hollis, R. L., Kanade, T., Henmi, K., and Yoshikawa, T. (1996). Toward machine mediated training of motor skills. *Proceedings of the IEEE International Workshop on Robot and Human Communication*, pp. 32–37.
- [180] Gillespie, R. B., O'Modhrain, M., Tang, P., Zaretzky, D., and Pham, C. (1998). The virtual teacher. *Proceedings of the ASME Dynamic Systems and Control Division*, pp. 171–178.
- [181] Henmi, K. and Yoshikawa, T. (1998). Virtual lesson and its application to virtual calligraphy system. *IEEE Int. Conf. Robot. Autom.*, 2(May), 1275–1280.
- [182] Kim, Y. K. and Yang, X. (2007). Hand-writing rehabilitation in the haptic virtual environment. *Proceedings of the 2006 IEEE International Workshop on Haptic Audio Visual Environments and Their Applications, HAVE 2006*, pp. 161–164.
- [183] Williams, R. L., Srivastava, M., Conaster, R. R., and Howell, J. N. (2004). Implementation and evaluation of a haptic playback system. *Haptics-e*, 3(3), 160–176.

- [184] Burdea, G., Patounakis, G., Popescu, V., and Weiss, R. E. (1999). Virtual reality training for the diagnosis of prostate cancer. *IEEE Transactions on Biomedical Engineering*, 46(10), 1253–1260.
- [185] Moraes, R. M., Souza, D. F. L., Valdek, M. C. O., and Machado, L. S. (2006). A virtual reality based simulator for gynecologic exam training. *7th International Conference on Information Technology Based Higher Education and Training, ITHET*, pp. 786–791.
- [186] Chial, V. B., Greenish, S., and Okamura, A. M. (2002). On the display of haptic recordings for cutting biological tissues. *Proceedings - 10th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, HAPTICS 2002*, pp. 80–87.
- [187] Shen, Y., Devarajan, V., and Eberhart, R. (2005). Haptic herniorrhaphy simulation with robust and fast collision detection algorithm. *Volume 111: Medicine Meets Virtual Reality 13 Studies in Health Technology and Informatics series*, pp. 458–464. IOS Press.
- [188] Srimathveeravalli, G. and Thenkurussi, K. (2005). Motor skill training assistance using haptic attributes. *In Eurohaptics Conference, 2005 and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, 2005. World Haptics 2005.*, pp. 452–457.
- [189] Avizzano, C. A., Solis, J., Frisoli, A., and Bergamasco, M. (2002). Motor learning skill experiments using haptic interface capabilities. *International Workshop on Robot and Human Interactive Communication*, pp. 198–203.
- [190] Solis, J., Avizzano, C. A., and Bergamasco, M. (2003). Validating a skill transfer system based on reactive robots technology. *Proceedings - IEEE International Workshop on Robot and Human Interactive Communication*, pp. 175–180.
- [191] Solis, J., Avizzano, C. A., and Bergamasco, M. (2002). Teaching to write japanese characters using a haptic interface. *Proceedings - 10th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, HAPTICS 2002*, pp. 255–262.

- [192] Feygin, D., Keehner, M., and Tendick, F. (2002). [haptic guidance: Experimental evaluation of a haptic training method for a perceptual motor skill]. *Haptic Interfaces for Virtual Environment and Teleoperator Systems, 2002. HAPTICS 2002. Proceedings. 10th Symposium on*, pp. 40–47.
- [193] Kolesnikov, M. and Žefran, M. (2010). Haptic playback: Better trajectory tracking during training does not mean more effective motor skill transfer. *Haptics: Generating and Perceiving Tangible Sensations*, July, Lecture Notes in Computer Science, 6192 LNCS, pp. 451–456.
- [194] Boehm, A. (1941). STEREOGNOSIS AND TACTILE 'AUTOSENSATIONS'. *Br. J. Psychol.*, 4, 327–334.
- [195] Lederman, S. J. and Klatzky, R. L. (1993). Extracting object properties through haptic exploration. *Acta Psychol.*, 84(1), 29–40.
- [196] Streri, A. and Spelke, E. S. (1988). Haptic perception of objects in infancy. *Cogn. Psychol.*, 20(1), 1–23.
- [197] Jones, L. A. and Lederman, S. J. (2006). *Human Hand Function*. Oxford University Press, USA.
- [198] Connell, L. A. (2007). Sensory Impairment and recovery After Stroke. PhD thesis University of Nottingham.
- [199] Salisbury, K., Brock, D., Massie, T., Swarup, N., and Zilles, C. (1995). Haptic rendering: Programming touch interaction with virtual objects. *Proceedings of the 1995 Symposium on Interactive 3D Graphics I3D '95*, pp. 123–130.
- [200] Yin, M. S., Haddawy, P., Suebnukarn, S., and Rhienmora, P. (2016). Toward intelligent tutorial feedback in surgical simulation : Robust outcome scoring for endodontic surgery. *21st International Conference on Intelligent User Interfaces, IUI 2016*, pp. 402–406.
- [201] Saga, S., Kawakami, N., and Tachi, S. (2005). Haptic teaching using opposite force presentation. *Poster of World of Haptics Conference*.
- [202] Saga, S., Vlack, K., Kajimoto, H., and Tachi, S. (2005). Haptic video. *ACM SIGGRAPH. Emerging Technologies*.

- [203] Ioannou, I., Kazmierczak, E., Stern, L., Smith, A. C., Wise, L. Z., and Field, B. (2010). Towards defining dental drilling competence, part 1: a study of bone drilling technique. *J. Dent. Educ.*, 74(9), 931–940.
- [204] Rhienmora, P., Haddawy, P., Suebnukarn, S., and Dailey, M. N. (2009). A VR environment for assessing dental surgical expertise. *Proceedings of the 2009 Conference on Artificial Intelligence in Education (AIED)*.
- [205] Van der Linde, R. Q., Lammertse, P., Frederiksen, E., and Ruiter, B. (2002). The HapticMaster, a new high-performance haptic interface. *Proc. Eurohaptics*, pp. 1–5.
- [206] Koopman, P., Buis, J., Wesselink, P., and Vervoorn, M. (2010). Simodont, a haptic dental training simulator combined with courseware. *Bio-Algorithms and Med-Systems*, 6(11), 117–122.
- [207] Torabinejad, M. and Walton, R. E. (2009). *Endodontics: Principles and Practice*. Elsevier Health Sciences.
- [208] Lecture 14 welded connections. [http://www.magnetomechanical.com/resources/STEELandSTRUCTURAL/STEEL\\_WELD\\_DESIGN.pdf](http://www.magnetomechanical.com/resources/STEELandSTRUCTURAL/STEEL_WELD_DESIGN.pdf). Accessed: 2018-3-15.