A Grey-level Assisted Method for CBCT Alveolar Bone Height Measurements

THESIS

Presented in Partial Fulfillment of the Requirements for the Degree Master of Science in the Graduate School of The Ohio State University

By

Soumya Padala B.D.S., M.D.S.
Graduate Program in Dentistry

The Ohio State University
2014

Master's Examination Committee:

Zongyang Sun, Advisor
Do-Gyoon Kim
Michael Beck
Copyright by
Soumya Padala
2014
Abstract

Introduction: Measuring alveolar bone height from CBCT images reliably and accurately requires consistent and precise identification of landmarks. Previous studies have indicated a conventional vision-based method may lead to unreliable and inaccurate alveolar bone height measurements. This study tested a new method of landmark determination based on grey-level value changes between different tissues involving the alveolar bone in a pig model.

Methods: Twenty 5-month-old pig heads underwent CBCT scans (0.4mm voxel-size). Independently analyzed by three calibrated, blinded raters to determine labial/buccal alveolar bone height at two locations; mandibular central incisor and maxillary molars using direct-visual (in Dolphin-3D) and grey-level assisted methods (in ImageJ) entailing boundary detection based on eyes and changes in grey-level values, respectively. CBCT measurements were subsequently compared with physical truth measured after gingival removal. Inter-rater reliability and accuracy of measurements were statistically analyzed.

Results: Inter-rater reliability of measurements at the molar regions were higher with the grey level assisted method than with the vision-based method. The accuracy of alveolar
bone height measurements was generally similar between the two methods, with mild improvement with the grey-level assisted method at the left molar region. Regardless of methods, measurements from the left and right molar regions demonstrated a small but significant side-related difference.

Discussion: A grey-level assisted method was more reliable than a vision-based method. This new method is not based on rater’s experience, easy to learn and apply.

Conclusion: Compared to the vision-based method, an interactive grey level assisted method can be used to measure alveolar bone height with superior reliability and at least comparable accuracy.
Dedication

This document is dedicated to my God “PTLJC” who is my Best Friend without whose strength I wouldn’t be where I am today. I am always smiling because of the hope that comes from God alone!

I would also like to dedicate it to my pastor Kannan for being constant source of encouragement and for spending numerous hours praying for me and my family and friends. So that we may all live our lives such that it would glorify God!

My parents Ramababu & Leela, who always encouraged me to pursue my dreams and forever sacrificed their needs to support mine. They were with me when I walked first time in to the school and today its my last day in school they are with me as well!

My mentor Katherine Vig, for spending numerous hours encouraging me and helping me realize that when we pursue our dreams it’s the journey not the destination that matters the most.

My sister Shweta (Akka I call her) for her unconditional love and support that allowed me to work and study here in the USA while she took care of my parents. She is a sister that is like a mother to me.

In my residency program to my right sat my little sib Kathy that corrected me when I was wrong and patiently helped God in molding me. She will forever be a sister to me. To my left sat a good friend Jeff, who was an inspiration constantly. His smile gave me hope when I was down.

Thank you all for your unconditional support, encouragement and care. You will all forever be part of my life!
Acknowledgments

I would like to thank the following:

- My thesis committee, Zongyang Sun, Mike Beck, Do Gyoong Kim for their feedback and guidance during this project.
- Boon Ching Tee for all help with data compilation and analysis.
- Kathy Elias for assisting with data collection
- Jeffery Price for his support with formatting during the entire period of research
- Ammar Almahdi and Debbie Pack for pig head collection and scanning
- Jahanazeb Chaudhry and Katherine Vig for their constructive comments during preparation of the manuscript.
- The faculty, staff, alumni, and patients of The Ohio State University, College of Dentistry, Section of Orthodontics, who gave me the opportunity to live my dream.
- My fellow residents, who were a source of encouragement.
- My family, for their boundless love and support.
- Delta Dental for the research grant that funded this project.
Vita

2000-2005 .................Dental School, Sri Ramachandra University, Chennai, India

2006-2009 ..................Resident, Graduate Orthodontics Program, Sri Ramachandra University, Chennai, India

2009-2010 ..................Intern, Graduate Orthodontics Program, The Ohio State University

2010-2011 ..................Fellow, Graduate Orthodontics Program, University of Connecticut

2011-2014 ..................Resident, Graduate Orthodontics Program, The Ohio State University

Publications

Patients', parents', and orthodontists' perceptions of the need for and costs of additional procedures to reduce treatment time.

Cone-beam computed tomography evaluation of alveolar ridge width and height changes after orthodontic space opening in patients with congenitally missing maxillary lateral incisors.

A biomechanical approach to second-molar intrusion.
Alveolar ridge width and height changes after orthodontic space opening in patients congenitally missing maxillary lateral incisors.

Comparative evaluation of condylar position in symptomatic (TMJ dysfunction) and asymptomatic individuals.

Short-term zoledronic acid reduces trabecular bone remodeling in dogs.

**Fields of Study**

Major Field: Dentistry

Specialty: Orthodontics
# Table of Contents

Abstract................................................................................................................................. ii  
Dedication ............................................................................................................................. iv  
Acknowledgments .................................................................................................................. v  
Vita ........................................................................................................................................ vi  
Table of Contents .................................................................................................................. viii  

Chapter 1 Introduction ............................................................................................................. 1  
  Overview and History of Cone-beam Computed tomography (CBCT) ................................ 1  
  Principles of CBCT ............................................................................................................... 1  
  Applications of CBCT in Dentistry ..................................................................................... 2  
  Applications of CBCT in Orthodontics ............................................................................... 3  
  Significance of Alveolar Bone height measurements in Orthodontics ............................... 4  
  Accuracy and Inaccuracy of Morphological Alveolar Bone Analysis Using CBCT .......... 5  
  Factors affecting the accuracy of alveolar bone height measurements .............................. 5  
  Factors potentially affecting the quality of CBCT images ................................................... 6  
  Factors potentially affecting the image analysis process ..................................................... 12  

viii
List of Tables

Table 1: Interrater agreement (ICC) of physical measurements ........................................ 58
Table 2: Interrater agreement (ICC) of physical measurements with 95% confidence intervals .......................................................................................................................... 59
Table 3: Interrater agreements (ICC) of CBCT measurements ........................................ 60
Table 4: Mean differences between physical and CBCT measurements and the limits of agreement (LOA, 1.96SD) for each CBCT method categorized based on side. ............ 62
Table 5: Two-way repeated-measures ANOVA between two CBCT methods at molar and incisor locations. ........................................................................................................ 65
Table 6: Two-way repeated-measures ANOVA summary ............................................... 65
Table 7: Post hoc comparisons ...................................................................................... 67
List of Figures

Figure 1: Diagrammatic representation of cone beam CT (CBCT) acquisition: a: X-ray source; ......................................................................................................................... 2

Figure 2: Buccal alveolar bone height measured from mesiobuccal and distobuccal cupsal tips of 2 distal most fully erupted maxillary molars to alveolar crest following the long axis of the tooth at 4 locations per side in maxilla................................................................. 18

Figure 3: Labial alveolar bone height measured from centre of incisal edge of mandibular central incisor to alveolar crest following the long axis of the tooth at one location per side in mandible........................................................................................................ 19

Figure 4: Prior to application of Initial reorientation................................................................. 21

Figure 5: Initial reorientation - occlusal plane was made parallel to the floor in the coronal and sagittal views, and the head’s midline was bisected by the midsagittal plane as seen in axial view.................................................................................................................. 22

Figure 6: Measurement of maxillary molars: coronal axis adjusted to pass through the cusp tip in sagittal view and in axial view, sagittal axis is set to pass through buccolingually center of maxillary posterior teeth on the side of measurement. .......... 23

Figure 7: Measurement of mandibular incisors: Sagittal axis centered mesiodistally to pass through the mandibular central incisor in axial view. Sagittal view, displays site of measurement corresponding to physical measurement................................................. 24
Figure 8: Tiff image (left) converted into grayscale values in excel (right) using image-to-results tool in ImageJ software. ................................................................. 25

Figure 9: A macro was written to categorize gray values of pixels of an image into 20, 50 and 85th percentile groups with a 3-color scale, based on which the rough contour of the tooth and alveolar bones became visible. ................................................................. 26

Figure 10: Most incisel and second to most labial cell with highest gray value compared to the surrounding was marked as incisel edge................................................................. 27

Figure 11: Most occlusal and second to most labial cell with highest gray value compared to the surrounding was marked as cusp tip. ................................................................. 27

Figure 12: Approximately 7x7 cells in the region of interest which most likely would have the alveolar bone crest were highlighted and a 2D line diagram was developed for these cell values. ................................................................. 29

Figure 13: In the graph, a particular pattern of three consecutive rows of cells with...... 30

Figure 14: Alveolar bone height calculated using the Pythagoras principle $a^2+b^2=c^2$ considering the length of hypotenuse as the alveolar bone height measurement. Length of two sides of the triangle was calculated using row and column numbers noted for marked cells ................................................................. 31

Figure 15: Bland-Altman plots for molars: Differences against the mean of the clinical and radiological measurements................................................................. 63
Figure 16: Bland-Altman plots for incisors: Differences against the mean of the clinical and radiological measurements. .............................................................. 64

Figure 17: Repeated measures ANOVA ........................................................................................................... 66

Figure 18: Comparison of alveolar bone height measurements between the two methods at different locations. P-values were based on two-way repeated-measures ANOVA..... 66
Chapter 1 Introduction

Overview and History of Cone-beam Computed tomography (CBCT)

Computed tomography (CT) became available for 3-dimensional (3-D) imaging in the 1980’s, but due to high cost, limited access, and radiation exposure, its application in the craniofacial and dental field was limited to diagnosis of severe craniofacial anomalies, planning for complex surgeries, and other dental situations such as pre implant dental imaging. 1-2 To lower the cost and radiation associated with 3-D radiographic imaging, CBCTs were developed as an alternative to CT used for oral and maxillofacial structures. 3 Although CBCT was first described in 1978, 4 it only became commonly available in the late 1990s. 5-6 The first dedicated CBCT scanner for the oral and maxillofacial region was described in 1998, particularly for implant planning. 5 Currently, there are 20 manufacturers offering a total of 47 devices produced in 7 different countries. 3

Principles of CBCT

CBCT is a medical image acquisition technique based on a cone shaped X-ray beam centered on a two dimensional (2D) detector. The source-detector system rotates around the object one or two times and produces a series of 2D images. 7 These images are
subsequently reconstructed to 3D data set using various modification of a cone beam algorithm originally developed by Feldkamp.⁸ (Fig. 1).

![Figure 1: Diagrammatic representation of cone beam CT (CBCT) acquisition: a: X-ray source;](image_url)

Conversely, conventional CT involves multiple rotations around the target object and acquires a series of images in the axial planes, which are subsequently converted into 3D data set with individual slices stacked one after another in a straight fashion or in a continuous spiral fashion.¹

**Applications of CBCT in Dentistry**

Despite its short history, CBCT has been used in a number of areas in dentistry as follows. In dental implantology, it has been used for identification of anatomical structures,⁹ and assessing bone quality and implant stability.¹⁰ In oral pathology, it has been imperative
in differentiating solid from fluid-filled lesions (periapical granulomas from cysts) using grayscale values in the lesions. In addition, CBCT are used for identification, diagnosis, determination of the severity of diseases and surgical assessment of pathology. Prior to surgical procedures, CBCT is used to determine extent of the fracture, localization of impacted teeth, and thickness of palatal mucosa. Furthermore, CBCT is widely used for TMJ assessment to find out subtle osteoarthrosis alterations such as subchondral cysts and sclerosis, osteophyte formation, surface erosion, and bony remodeling. Caries diagnosis in unrestored teeth is also possible with CBCT due to reduced artifacts in the absence of restorations. Moreover, higher resolution of CBCT scans are employed in the field of periodontics to evaluate periodontal bony defects and in endodontics for detection of root fractures and localization and characterization of root canals.

Applications of CBCT in Orthodontics

At least comparable to other disciplines of dentistry, orthodontics in the last decade has seen rapid growth of CBCT in the diagnosis, treatment planning and assessment of treatment outcomes, which has become one major driving force for the development of newer CBCT units with reduced radiation dose. Common applications of CBCT in clinical orthodontics that have been developed so far include production of 2-D cephalometric and panoramic images, localization of unerupted teeth, detection of
external root resorption, measurement of oropharyngeal airway, diagnostic aid in placement of temporary anchorage devices and for orthognathic and cleft surgeries.  

**Significance of Alveolar Bone height measurements in Orthodontics**

In orthodontics, accurate measurement of alveolar bone height is important for treatment planning and assessment of treatment outcome. The need to assess the periodontal status including bone levels before and after orthodontic treatment has been emphasized by the American Board of Orthodontics, as all examinees are required to provide a formal periodontal evaluation for patients over the age of 18 years and for those with signs of periodontal disease. Based on previous studies including clinical and radiographic examinations, it is very likely that orthodontic treatment can affect the height of the alveolar bone. One location where the alveolar bone is prone to such changes is the buccal side of the maxillary molar region. Rungcharassaeng et al performed a study on the CBCT records of 30 subjects taken before and after rapid palatal expansion (RME), and found that buccal crown tipping, reduction of buccal bone thickness, and marginal bone loss had occurred within 3 months after RME. Another area prone to such changes is the labial side of the mandibular incisors, where the bone height can be significantly reduced after excessive sagittal movements or tipping of the incisors. These studies further confirmed the importance of accurately monitoring and measuring the alveolar bone height especially in these areas.
Accuracy and Inaccuracy of Morphological Alveolar Bone Analysis Using CBCT

Reports on the accuracy of CBCT alveolar bone measurements have been somewhat inconsistent. For alveolar bone of the molars, while Timock et al found that CBCT images can produce highly accurate bone height measurements, Sun et al found that maxillary bone height measurement was systematically underestimated by 0.9 to 1.2 mm when the bone was thinned to the voxel size level. In another study, Wood et al found that using CBCT images scanned under typical orthodontic settings (0.4mm voxel size, full field of view), buccal alveolar bone height measurement from the maxillary molar region had a limit of agreement over 2 mm, suggesting a high degree of inaccuracy. Similarly, for alveolar bone of the mandibular incisors, Patcas et al found up to 2.1 mm of limits of agreement. Although the inconsistency among studies may be partly due to variations in research designs and materials such as cadaver heads, phantoms, or dry skulls without soft tissue, these studies strongly suggest that measurement of alveolar bone height in the maxillary molar and lower incisor regions from CBCT images is prone to inaccuracy.

Factors affecting the accuracy of alveolar bone height measurements

Measurement inaccuracy can be caused by numerous factors divided into two major categories. The first category includes those that can potentially affect the quality of CBCT images, such as CBCT units, scan settings, soft tissue conditions, artifacts/noises,
reconstruction algorithms, type/size/location of target structures. The second category includes those that can potentially affect the analysis process, such as software choices, raters, measurement methods, etc.

Factors potentially affecting the quality of CBCT images

CBCT units: Although there are 47 CBCT units currently available in the market, only a few of them have been compared directly. Stratemann et al \(^{29}\) determined the accuracy of measuring linear distances between landmarks commonly used in orthodontic analysis on a human skull using two CBCT systems and concluded that both CBCT systems provided highly accurate data compared with the gold standard of physical measures directly from the skulls, with less than 1\% relative error. \(^{29}\) Also, when using identical voxel sizes, the accuracy level of different CBCT devices appears hardly distinguishable. \(^{24}\) Another error is linked to a hardware limitation and the costs of large flat-panel detectors necessary to acquire large volumes. Such detectors are very costly and thus the manufacturers came up with a technical solution which is now frequently used in commercial CBCT machines: a smaller size flat-panel with the source-detector axis positioned offset to the centre of rotation. The consequence is that the central (cylindrical) part of the reconstructed volume is scanned over a complete 360° rotation, while locations at the periphery are only scanned over a 180° half rotation, this abrupt transition between the two regions may result in a ring artifact in this area in the axial
planes. Additionally, 180° rotation is preferred over 360° rotation for the purpose of reducing the effective radiation dose, although 180° rotation scans may further reduce image quality.

**Artifacts**: a CBCT artifact may be defined as a visualized structure in the reconstructed data that is not present. The following relevant artifacts have been reported: extinction artifacts; beam hardening artifacts; partial volume effect and exponential edge-gradient effect (EEGE); aliasing artifacts; ring artifacts; motion artifacts (misalignment artifacts); noise and scatter. Artifacts may not be easily depicted as specific patterns but as a more general deviation of the reconstructed density (gray) values from the ‘‘true’’ ones. Although noise is commonly not dealt with as an artifact, it is still considered as an image deteriorating factor. A smaller voxel will not detect as many photons as would a larger voxel size. A decrease in the number of photons acquired by a voxel would result in a decrease in signal leading to an increase in noise. CBCT machines for dose reduction are operated at milliamperes that are approximately one order of magnitude below those of medical CT machines. Thus, the signal-to-noise ratio is much lower than in CT. Scatter, on the other hand, is caused by those photons that are diffracted from their original path after interaction with matter. Scatter causes streak artifacts in the reconstruction that are very similar to those caused by beam hardening. Scatter is well-
known to further reduce soft-tissue contrast and it will also affect the density values of all other tissues.\textsuperscript{32, 34}

The acquisition time of state-of-the-art CBCT machines roughly ranges between 6 s and 20 s, which is a relatively long time for a human head to not perform some minor movements. Obviously, the smaller the voxel size (i.e. the higher the spatial resolution), the smaller the movement necessary to move the patient structures out of the ‘correct’ voxels. In other words, the higher the nominal resolution, the more likely motion artifacts are to appear. Typically, movement artifacts present as double contours. A sufficient fixation of the patient’s head during the scan process should help to limit the movement options for the patient.\textsuperscript{30}

**Scan settings:** Voxel size of a 3D image is similar to the pixel resolution in 2D images. Sun et al\textsuperscript{23} and Patcas et al\textsuperscript{24} reported that higher scanning resolution (smaller voxel size) produces more accurate alveolar bone. Linear measurements made with CBCTs on alveolar bone covering human mandibular anterior teeth were not statistically different between 0.4mm and 0.25mm voxel size.\textsuperscript{24} Damastra et al\textsuperscript{27} evaluated accuracy of linear measurements on dry human mandibles concluded that since there was no tangible difference in measurements on CBCTs taken at 0.25mm and 0.4mm voxel resolution, 0.4mm voxel resolution was adequate for measurements of craniofacial structures.\textsuperscript{27} Insignificant differences between CBCT and true measurements of the human temporomandibular joint and other human skull structures scanned at 0.4mm voxel size
have been reported in the past.\textsuperscript{35-36} These studies indicated that linear measurements of large craniofacial structures are acceptable for clinical use, but inaccuracy indeed occurs. In addition to voxel size, Molen\textsuperscript{37} suggested that it is important to report a CBCT scan’s spatial resolution. Although spatial resolution equal to the voxel size should be achievable in theory, this has not been the case in practice because of noise and scatter.\textsuperscript{22} Ballrick et al\textsuperscript{25} acquired i-CAT images of custom radiographic phantoms and, through varied spacing of pairs of metal wires, found that the spatial resolution was 0.84mm for scans taken at 0.3mm voxel size. Leung et al,\textsuperscript{26} confirmed this effect in his study wherein he examined the ability to distinguish alveolar bone from cementum and found that areas with bone less than 0.6mm thick were invisible on CBCT images. This phenomenon, favors improved accuracy of alveolar bone height measurements as they often exceed the magnitude of spatial resolution in comparison to alveolar bone thickness measurements. Other factors that might cause inaccuracies in measurements are the local tomography effect that occurs when the region of interest is surrounded by tissues that are not reconstructed.\textsuperscript{30,38-40} This is known to occur in CBCT scans that involve smaller field of view (FOV). Whereas, larger FOV provides less sharp reconstructions because of the greater beam angulations in the superior and inferior volume area and reduced contrast to noise ratio.\textsuperscript{24,41} Additionally, a high kilovolt technique is recommended to reduce artifacts resulting from beam hardening effects.\textsuperscript{42-43} However, this is not easy to achieve as the X-ray energy of CBCT is similar to that of panoramic radiography with a typical
operating range of 1-15 mA at 90-120 kVp, while that of medical CT is significantly higher at 120-150 mA, 220 kVp.\textsuperscript{1} Draenert et al \textsuperscript{44} compared visual spatial resolution of CBCT and multidetector CT (MDCT) in the imaging of metallic dental implants and found that beam hardening artifacts occur in dental implant scans with the NewTom\textsuperscript{®} cone beam CT but not with the MDCT.

**Type/size/location of target structure:** Less accurate measurements from maxillary molar region can be largely attributed to thin bones in this region. In pigs, the maxillary buccal alveolar bone was less than 1 mm thick, significantly thinner than the mandibular alveolar bones, which were above 1 mm.\textsuperscript{21} As a result, thin bones (near voxel size) tend to become indistinguishable from adjacent cementum on CBCT images due to partial volume averaging effect, which is a common CT artifact and occurs when a voxel lies on the borders of two objects of different densities. The voxel will then reflect the average density of both the objects rather than the true value of either object.\textsuperscript{41,45} This invisibility of some structures could also be caused by the limitations in contrast resolution related to CBCT units, which determines the ability to distinguish two objects with similar densities and in close proximity as mentioned previously.\textsuperscript{25-26,41}

**Soft tissue condition:** Among various factors affecting accuracy of CBCT measurements, the role of soft tissue has been recently revealed. Wood et al \textsuperscript{21} evaluated differences in accuracy of linear alveolar bone measurements with and without soft tissue
present. They found that measurements made on specimens with soft tissue present were more accurate as alveolus in the soft tissue absent images had greater contrast and surface brightness than did the soft tissue present images due to beam hardening effect. Beam hardening effect is the process whereby average energy level of an x-ray beam are increased by filtering out the low-energy photons. In this case presence of soft tissues allowed absorption of low energy photons from the polychromatic beam before it reached the alveolar bone and tooth crown.

**Reconstruction algorithm:** It is well known that CBCT does not satisfy the data completeness condition for exact image reconstruction and therefore allows only for approximate image reconstruction and more so when the scan is taken with 180° rather than 360° rotation. One of the first reconstruction algorithms was developed by Feldkamp. Even though it is widely used, this algorithm has many deficiencies. Maaß et al compared various algorithms used for reconstruction and found that the Feldkamp reconstruction had the lowest computational load and showed the worst cone beam artifacts. New methods such as the factorization have been developed to reduce the cone–beam artifacts. The Feldkamp algorithm, guarantees a high image quality in the central plane, whereas from a mathematical point of view, it is identical to the filtered back projection used in CT machines. Image quality, however, will degrade as a function of
distance from that plane. This additional aspect should be kept in mind, particularly when evaluating large FOV volumes.

**Factors potentially affecting the image analysis process**

**Software choice:** Azredo et al \(^{43}\) compared 5 software programs specifically designed for visualization, analysis and found no statistically significant differences between the software. Similar results were found by Weissheimer et al \(^{49}\) who compared the precision and accuracy of 6 DICOM viewers for 3D analysis of the upper airway. Wood et al \(^{21}\) compared Dolphin software with Osirix software. Dolphin is currently one of the most commonly used programs in dentistry, especially in orthodontics, and Osirix is the most widely used DICOM data viewer in the medical field. Their results demonstrated that the use of either software program resulted in an identical degree of accuracy for alveolar bone height measurements. \(^{21}\)

**Raters:** Human vision is subject to the influence of a host of factors such as lighting, fatigue, gray scale ability, and visual acuity. Images produced by CT scanners and other modalities typically contain between 12/16 bits per pixel which corresponds to 4,096 to 65,536 shades of gray. Human observers are able to discriminate between 700 and 900 simultaneous shades of gray for the available luminance range of current medical displays and in optimal conditions. Additionally, viewing angle-dependence and spatial
noise can significantly decrease the accuracy as well.\textsuperscript{50} So, there could be possible differences between the raters but most of the studies have reported high (above 0.8) inter-rater reliability negating possible differences between raters. \textsuperscript{21-22}

\textbf{Methods for analysis:} Lastly, the methods used to analyze CBCT images may also affect the accuracy of measurements. To date, two main methods have been used to analyze CBCT images taken for dentistry purposes. One is a direct-visual method which entails detection of structure boundaries or changes using human eyes. This method has been commonly used to measure alveolar bone height,\textsuperscript{21} temperomandibular joint,\textsuperscript{36} airway,\textsuperscript{49} and skull.\textsuperscript{35} Although convenient and not technique sensitive, this method is likely vulnerable to the influence of image display, scan contrast and individual variations such as lighting, fatigue, gray scale ability, and visual acuity as mentioned above, all of which may increase the bias and subjectivity of the measurements. The other method is based on image segmentation, which entails separation of different tissue types, such as bone vs. soft tissue based on their difference in gray levels. This technique has been used in analyzing upper airway by El and Palomo.\textsuperscript{51} A major advantage of using automatic segmentation method is that it eliminates operator subjectivity in boundary selection and the measurements can be highly reliable and objective.\textsuperscript{49, 52} However, automatic segmentation also has its own limitations as it has been found that gray-level values cannot be used quantitatively for segmentation in CBCTs due to
histogram shifting and excessive scatter detection. Clearly, the direct-visual method completely discards the gray-level information carried by the CBCT voxels, while the automatic segmentation relies on the gray-level information too roughly and fails to consider the variations associated with structure location, scan timing and settings.

In the past, gray levels have been used in thresholding techniques to detect boundaries in addition to histograms. Conceivably, one may use this gray-level information in a more interactive manner to determine exact boundaries of dentoalveolar structures. It is confirmed that the gray-level values of different tissue types in the same image remains proportional to their HU values measured by conventional CT. This allows one to examine the relative changes in gray-level values of neighboring voxels to determine the location of tissue transition or boundaries.

Therefore, the purpose of this study was to investigate whether a gray-level assisted method can be used to improve the accuracy and reliability of alveolar bone height measurements from CBCT images. Pigs were used in this study as they have craniofacial anatomy and function similar to humans. The tissues properties were not altered as they were fresh at the time of scan and not fixed with any solutions. The alveolar bone thickness of juvenile pigs at the ages of 3-6 months is comparable to that in adolescent humans. Based on previous studies of pigs at this age, measurements of alveolar bone height from orthodontic CBCT images obtained under common 0.4mm-voxel size full
FOV settings tend to have considerable inaccuracy at the upper molar region. In addition to the upper molars, the lower incisor region is also reported to be prone to measurement inaccuracy. These regions were therefore chosen to test this new method and address the following:

**Specific Aim**

To compare an interactive gray-level assisted method and a direct-visual method for their accuracy and reliability in measuring alveolar bone height at the maxillary molar and mandibular incisor regions from CBCT images acquired similar to Orthodontic diagnostic records setting.

**Hypothesis**

1. The interactive gray-level assisted method, results in more reliable alveolar bone height measurements than the direct-visual method among raters.

2. The interactive gray-level assisted method results in more accurate alveolar bone measurements (relative to physical measurements) than the direct-visual method.
Chapter 2 Materials and Methods

Animals and sample size

Twenty fresh domestic pig (Sus scrofa) heads, aged 5 months and equivalent to early adolescent humans, were used. Based on previous studies, sample size was calculated using G*Power 3 software. A total sample of 39 was required to achieve 80% power. With each side (right or left) treated as an individual sample, a total of 20 pig heads would be adequate to improve inaccuracy to 5% from the 25% reported by Wood et al. Under complete anesthesia, all pigs underwent an abdominal surgical procedure which was neither a component of nor a disruption to this study. Immediately after this procedure, with the pigs kept under deep anesthesia, they were euthanized by IV injection of pentobarbital through an ear vein. After confirmation of death, the pig head with all soft and hard tissues kept intact was collected from OSU university laboratory animal resources (ULAR) and frozen under -20 °C after completion of scanning.

CBCT Scan

CBCT scanning was completed within 24 hours of collection of each pig head. CBCT scans were done using iCAT 17-19 Platinum CBCT machine (120Kvp, 5mA) (Imaging Sciences International, Hatfield, PA) at 0.4mm voxel size (scanning time, 8.9 seconds). Field of view was set at height/width, 10-13/16cm similar to that used for obtaining regular Orthodontic records. The gantry bearing the x-ray source and detector rotated
180° around the pig head. Partial 180° rotation is preferred over the full 360° rotation due to decrease in effective radiation dose with the 180° rotation and, hence, recommended for clinical use. 

Scan began on the right side of the pig head and moved to the left. The orientation of the pig head in the scanner was centered in the stage with the snout pointing up on an average of 55° (ranging from 36° to 70°) to the floor as seen in Fig 4 prior to reorientation.

**Specimen dissection and physical measurements of alveolar bone height**

The pig head was thawed under room temperature for few hours prior to specimen dissection. Subsequently, after making a skin incision, superficial soft tissues composed of skin, muscles, fascia were dissected and removed without damaging the underlying bone. Alveolar labial/buccal gingiva was carefully elevated with periosteal elevators and excised with scalpel blade to prevent any damage to alveolar bone especially in the areas of interest (maxillary molars and mandibular incisors). Maxillary and mandibular bones were separated from the head. Maxillary quadrants were separated using a Ryobi 9-in band saw (Technotronic Industries North America, Andreson, SC). Mandible was not separated in to quadrants to prevent damage to alveolar bone in the mandibular incisor region during separation. As previously mentioned, buccal alveolar bone height in the region of maxillary molars and mandibular incisors has been assessed for accuracy and reliability due to its relevance in clinical orthodontics.
For each side, labial/buccal alveolar bone height was measured at 4 maxillary molar locations and 1 mandibular incisor location. Details of locations are listed below and illustrated in Fig 2,3

Figure 2: Buccal alveolar bone height measured from mesiobuccal and distobuccal cuspal tips of 2 distal most fully erupted maxillary molars to alveolar crest following the long axis of the tooth at 4 locations per side in maxilla.

M1M: Buccal alveolar bone height from mesiobuccal cusp tip of second last fully erupted maxillary molar to alveolar crest.

M1D: Buccal alveolar bone height from distobuccal cusp tip of second last fully erupted maxillary molar to alveolar crest.
M2M: Buccal alveolar bone height from mesiobuccal cusp tip of distal most fully erupted maxillary molar to alveolar crest. 

M2D: Buccal alveolar bone height from distobuccal cusp tip of distal most fully erupted maxillary molar to alveolar crest. 

CNT: Labial alveolar bone height from centre of incisal edge to alveolar crest of mandibular central incisor. 

Figure 3: Labial alveolar bone height measured from centre of incisal edge of mandibular central incisor to alveolar crest following the long axis of the tooth at one location per side in mandible 

Apical reference point for all these locations was determined as the point at the intersection of the alveolar crest and line dropped perpendicular to the center of the incisal edge/cusp tip.
Specifically, at each maxillary molar side of an animal, alveolar bone height was measured at 4 locations; M1M, M1D, M2M, M2D (Fig 2) as listed above. Linear distance between the buccal cusp tip to the buccal alveolar crest along the long axis of the tooth was measured using a digital caliper (precision 0.001 mm). For each mandibular incisor side of an animal, alveolar bone height was measured at one location; CNT (Fig 3). Specifically, the linear distance between mesiodistal center point of the incisal edge and the alveolar crest along the long axis of that tooth was measured using a digital caliper. No measurements were taken from the lateral incisor because of their increased variability in mesiodistal angulation and incisal edge form. (Fig 3).

Physical measurements of alveolar bone height were conducted by two raters (K.E & S.P). Both raters were calibrated and then independently measured each location three times, which were averaged and used as the measurement for that location.

**Measuring alveolar bone height from CBCT using a direct-visual method**

All CBCT data (DICOM files) were analyzed by 3 calibrated, independent, and blinded raters (B.T, K.E and S.P) using Dolphin-3D software (Dolphin Imaging and Management Solutions, Chatsworth, CA) at the same computer monitor setting (1600 × 1200 pixels); a standardized protocol was used (Fig 4,5) to reorient the images. More specifically, occlusal plane was set parallel to the floor in the frontal and sagittal views, and the midline of the head was bisected by the midsagittal plane in the axial view (Fig 3B).
Measurement of buccal alveolar bone height at the maxillary molar locations was performed on coronal slices in relation to the buccal cusp tips of the maxillary molars being assessed (Fig 6). The distance between the cusp tip and the alveolar crest was visually determined by the rater and was measured using a linear measurement tool of the software on three consecutive images (Fig 6). Overall, the coronal sections used for measurements corresponded to the locations subject to physical measurements.

Measurement of labial alveolar bone at the mandibular incisor locations was performed on the sagittal section, sagittal axis passing through the middle point of the incisal edge of the central incisor on axial slice (Fig 7). The distance between the labial aspect of the incisor edge to the alveolar crest was determined using the linear measurement tool. During analyses the rater was allowed to zoom in and make these measurements.

Figure 4: Prior to application of Initial reorientation
Figure 5: Initial reorientation - occlusal plane was made parallel to the floor in the coronal and sagittal views, and the head’s midline was bisected by the midsagittal plane as seen in axial view.
Figure 6: Measurement of maxillary molars: coronal axis adjusted to pass through the cusp tip in sagittal view and in axial view, sagittal axis is set to pass through buccolingually center of maxillary posterior teeth on the side of measurement.
Measuring alveolar bone height from CBCT using an interactive gray-level assisted method

The DICOM files were imported into Image J (Developed by NIH) software and converted into “tiff” format. For incisors, no additional reorientation was performed. For molars, the stacked 3-D “tiff” file was rotated using the Transform J tool to make the occlusal plane parallel to the floor in the frontal and lateral views and make the midline of head bisected by the midsagittal plane, in a similar fashion as the reorientation performed for the direct-visual method in Dolphin. Reorientation was not performed for Incisors. 3 consecutive 2-D images for each molar or incisor location were acquired.
similar to the direct-visual method mentioned above, which were used for subsequent measurement using an interactive gray-level assisted method. Specifically, the 2-D images were transformed into the gray-level values of all pixels in that image in ImageJ, and the results were subsequently exported to an Excel worksheet (Microsoft 2010) (Fig 8). 

![Figure 8: Tiff image (left) converted into grayscale values in excel (right) using image-to-results tool in ImageJ software.](image)

Then, a macro was written so that the gray values of pixels of an image were separated to 20, 50 and 85 percentile groups and labeled by 3-color scales, based on which the rough contour of the tooth and alveolar bones became visible (Fig 9).
Figure 9: A macro was written to categorize gray values of pixels of an image into 20, 50 and 85th percentile groups with a 3-color scale, based on which the rough contour of the tooth and alveolar bones became visible.

The next step involved determination of the landmarks (a particular pixel) at the tooth (cusp tip for molars and incisal edge for incisors) and at the alveolar crest for measurements of the distance between these two points. For the former, the boundary was between enamel and air, which was represented by large gray value drop moving from the enamel to the air (Fig 10). Using these characteristics, for the incisors, the most incisal and second to most labial pixel of enamel-level values were identified as the landmark representing the incisal edge. The second to most labial pixel approximated the location used for physical and direct-visual CBCT measurements. For molars, the most
occlusal cell with enamel-level gray values was identified as the landmark representing the cusp tip (Fig. 11).

Figure 10: Most incisal and second to most labial cell with highest gray value compared to the surrounding was marked as incisal edge.

Figure 11: Most occlusal and second to most labial cell with highest gray value compared to the surrounding was marked as cusp tip.
For the landmark of the alveolar bone crest, it was located at the boundary of several tissues including the enamel, periodontal ligament, alveolar crest and gingiva. To determine the exact location of the most incisal or occlusal point of the alveolar bone crest, a 7x7 cell region containing these boundaries was plotted into a line graph (Fig. 12,13). In the graph, the lines were divided into three groups based on the different patterns of value change in the horizontal direction: a bent-up pattern indicative of a strong likelihood of bone presence, a bent-down pattern indicative of a strong likelihood of bone absence, and a straight line pattern indicative of transition between the first two conditions. The straight line pattern was often shown in only one row of data and between the other two patterns. Using this characteristic, the most likely alveolar crest landmark pixel was identified (Fig 12,13). The same method of identification was used for both the incisor and molar locations.
Figure 12: Approximately 7x7 cells in the region of interest which most likely would have the alveolar bone crest were highlighted and a 2D line diagram was developed for these cell values.

After the landmark pixel locations were identified, the distance between the two landmarks were calculated using the Pythagoras principle $a^2 + b^2 = c^2$ considering the length of hypotenuse as the alveolar bone height measurement for the straightest curve cell value and the width of each pixel being 0.4 mm (Fig 14).
Figure 13: In the graph, a particular pattern of three consecutive rows of cells with bend down-straight and bend up pattern were determined and cell corresponding to the straightest curve (middle cell) was marked as possible location of alveolar crest.

CBCT measurements for this method were also conducted by the same three independent blinded calibrated raters (B.T, K.E and S.P), who measured images using Dolphin software as well.
Figure 14: Alveolar bone height calculated using the Pythagoras principle $a^2+b^2=c^2$ considering the length of hypotenuse as the alveolar bone height measurement. Length of two sides of the triangle was calculated using row and column numbers noted for marked cells.

Statistical Analysis

For quantitative diagnosis, the reliability of physical measurements between two raters (S.P and K.E) was assessed with Interclass correlation tests (ICC) using a two-way mixed effects model, single measure reliability with absolute agreement. ICC was then used to assess the reliability among the three raters, segregated based on location, sides and
The CBCT method used. The correlation coefficients (r values) from ICC were compared to assess significance between the two CBCT methods with Fisher’s Z test.

To assess the magnitude of disagreement/agreement between measurements collected directly and from 2 CBCT methods, the mean and difference between physical and CBCT methods were calculated. Any trends or outliers were demonstrated with Bland-Altman plots by using 95% limits of agreement (average differences +/- 1.96 of the standard deviation of the differences). 60

The accuracy of both methods of CBCT measurements were assessed using two approaches. In the first we looked at raw delta values using a repeated-measures ANOVA with method, location and rater (as a random effect) as the independent variables. Post hoc testing was done using the Tukey-Kramer procedure. In the second approach, the delta scores were dichotimized into clinically acceptable (≤1mm deviation from physical measurement) or unacceptable (>1mm deviation from physical measurement) and compared using multiple Cochran-Mantel-Haenszel statistics which were adjusted using the step-down Bonferroni method of Holm.
Chapter 3 Manuscript

Title Page

Localizing Alveolar Crest and Enamel Landmarks on CBCT Images Using an Interactive Grey-level Assisted Method

Soumya Padala  BDS, MDS  Graduate Orthodontic Resident
Boon Ching Tee  BS  Research Assistant
Michael Beck  DDS, PhD  Professor Emeritus
Kathy Leigh Elias  MS, DMD, PhD  Graduate Orthodontic Resident
Do-Gyoon Kim  BS, MS, PhD  Associate Professor
Zongyang Sun  DDS, MS, MSD, PhD  Associate Professor

Corresponding Author

Dr. Zongyang Sun
Division of Orthodontics- College of Dentistry - The Ohio State University
4088 Postle Hall, 305 W Twelfth Avenue, Columbus, OH 43210
Telephone : 614-247-7944  Email : sun.254@osu.edu  Fax: 614-688-3077

Abstract
Introduction: Measuring alveolar bone height from CBCT images reliably and accurately requires consistent and precise identification of landmarks. Previous studies have indicated a conventional vision-based method may lead to unreliable and inaccurate alveolar bone height measurements. This study tested a new method of landmark determination based on grey-level value changes between different tissues involving the alveolar bone in a pig model.

Methods: Twenty 5-month-old pig heads underwent CBCT scans (0.4mm voxel-size). Independently analyzed by three calibrated, blinded raters to determine labial/buccal alveolar bone height at two locations; mandibular central incisor and maxillary molars using direct-visual (in Dolphin-3D) and grey-level assisted methods (in ImageJ) entailing boundary detection based on eyes and changes in grey-level values, respectively. CBCT measurements were subsequently compared with physical truth measured after gingival removal. Inter-rater reliability and accuracy of measurements were statistically analyzed.

Results: Inter-rater reliability of measurements at the molar regions were higher with the grey level assisted method than with the vision-based method. The accuracy of alveolar bone height measurements was generally similar between the two methods, with mild improvement with the grey-level assisted method at the left molar region. Regardless of methods, measurements from the left and right molar regions demonstrated a small but significant side-related difference.
Discussion: A grey-level assisted method was more reliable than a vision-based method. This new method is not based on rater’s experience, easy to learn and apply.

Conclusion: Compared to the vision-based method, an interactive grey level assisted method can be used to measure alveolar bone height with superior reliability and at least comparable accuracy.

**Introduction**

Numerous recent studies using cone-beam computed tomography (CBCT) imaging have reported that rapid palatal expansion can significantly reduce the buccal alveolar bone height of maxillary molars and premolars, and proclination of lower incisors may significantly reduce their labial alveolar bone height. Meanwhile, studies assessing the accuracy of CBCT linear measurements continued to suggest that alveolar bone height from the maxillary molar and mandibular incisor regions tend to have considerable inaccuracy and relatively low reliability.

Measuring alveolar bone height from CBCT images can be impacted by factors in two major categories. The first category includes those affecting the quality of CBCT images such as scanning parameters, soft tissue conditions, artifacts/noises, reconstruction algorithms, and type/size/location of target structures. The second
category includes those affecting the analysis process, such as rater’s experience, measurement methods, etc. Although the accuracy and reliability can be significantly improved by controlling factors in the former category, such as using high resolution and small field of view (FOV) CBCT scans, the concern of increased radiation to patients often disallows such choices. This then leaves controlling factors in the latter category the only viable avenue to improve measurement reliability and accuracy.

To date, two main methods have been used to quantitatively analyze CBCT images taken for dentistry purposes. One is a direct-visual method which entails detection of structure boundaries using human vision, which has been commonly used to measure alveolar bone height,\textsuperscript{21} temporomandibular joint,\textsuperscript{36} airway,\textsuperscript{49} and skull dimensions.\textsuperscript{35} Specifically for the alveolar bone, to measure its height, the rater first needs to visually identify the alveolar crest and cusp tip or incisal edge landmarks. Human vision is subject to the influence of a host of factors such as lighting, fatigue, gray scale ability, and visual acuity. Images produced by CT scanners and other modalities typically contain between 12/16 bits per pixel which corresponds to 4,096 to 65,536 shades of gray, but human eyes, even under optional conditions, are only able to discriminate between 700 and 900 simultaneous shades of gray for the available luminance range of current medical displays.\textsuperscript{69} As substantiated by several recent studies,\textsuperscript{21, 23-24} these limitations of human vision can markedly increase interrater variability of the measurements. The other
method is based on image segmentation, which entails separation of different tissue types, such as bone from soft tissue based on their differences in gray levels, a method has been used recently in analyzing upper airway. While segmentation based on gray-level values effectively eliminates operator subjectivity in boundary selection and subsequent measurements, it bears a considerable risk of inaccuracy because CBCT gray-level values vary among scans and can only relatively rather than absolutely represent tissue mineral density. Therefore, for the two methods currently used for linear measurements from CBCT images, one completely disregards the information of gray-level values, thus may result in unreliable measurements among raters, while other method completely relies on the gray-level information without considering variations among CBCT scans, thus may substantially compromise measurement accuracy.

These limitations of current methods prompted us to develop a new analysis method that can result in highly accurate and reliable measurements of alveolar bone height measurements from orthodontic-grade CBCT images. As is confirmed that gray-level values of different tissue types in the same image remains proportional to their HU values reflected by conventional CT, one may examine the relative changes in gray-level values of neighboring voxels to localize tissue boundaries.
Therefore, this study was undertaken to develop a new gray-level assisted method to identify enamel and alveolar crest landmarks for alveolar bone height measurement, which was tested in the maxillary molar and mandibular incisor regions.

**Materials and Methods**

**Animals and sample size**

Twenty fresh domestic pig (Sus scrofa) heads, aged 5 months and equivalent to early adolescent humans, were used. Based on previous studies, sample size was calculated using G*Power 3 software. A total sample of 39 was required to achieve 80% power. With each side (right or left) treated as an individual sample, a total of 20 pig heads would be adequate to improve inaccuracy to 5% from the 25% reported by Wood et al. Under complete anesthesia, all pigs underwent an abdominal surgical procedure which was neither a component of nor a disruption to this study. Immediately after this procedure, with the pigs kept under deep anesthesia, they were euthanized by IV injection of pentobarbital through an ear vein. After confirmation of death, the pig head with all soft and hard tissues kept intact was collected from OSU university laboratory animal resources (ULAR) and frozen under -20 ºC after completion of scanning.

**CBCT Scan**

CBCT scanning was completed within 24 hours of collection of each pig head. CBCT scans were done using iCAT 17-19 Platinum CBCT machine (120Kvp, 5mA) (Imaging...
Sciences International, Hatfield, PA) at 0.4mm voxel size (scanning time, 8.9 seconds). Field of view was set at height/width, 17/23 cm similar to that used for obtaining regular Orthodontic records. The gantry bearing the x-ray source and detector rotated 180° around the pig head. (Fig 1) Partial 180° rotation is preferred over the full 360° rotation due to decrease in effective radiation dose with the 180° rotation and, hence, recommended for clinical use.\(^3\) Scan began on the right side of the pig head and moved to the left. The orientation of the pig head in the scanner was centered in the stage with the snout pointing up on an average of 55° (ranging from 36° to 70°) to the floor as seen in Fig 4 prior to reorientation.

**Specimen dissection and physical measurements of alveolar bone height**

The pig head was thawed under room temperature for few hours prior to specimen dissection. Subsequently, after making a skin incision, superficial soft tissues composed of skin, muscles, fascia were dissected and removed without damaging the underlying bone. Alveolar labial/buccal gingiva was carefully elevated with periosteal elevators and excised with scalpel blade to prevent any damage to alveolar bone especially in the areas of interest (maxillary molars and mandibular incisors). Maxillary and mandibular bones were separated from the head. Maxillary quadrants were separated using a Ryobi 9-in band saw (Technotronic Industries North America, Andreson, SC). Mandible was not separated in to quadrants to prevent damage to alveolar bone in the mandibular incisor region during separation. As previously mentioned, buccal alveolar bone height in the
region of maxillary molars\textsuperscript{21,23} and mandibular incisors\textsuperscript{24} has been assessed for accuracy and reliability due to its relevance in clinical orthodontics.

For each side, labial/buccal alveolar bone height was measured at 4 maxillary molar locations and 1 mandibular incisor location.

Details of locations are listed below and illustrated in Fig 2,3.

M1M: Buccal alveolar bone height from mesiobuccal cusp tip of second last fully erupted maxillary molar to alveolar crest.

M1D: Buccal alveolar bone height from distobuccal cusp tip of second last fully erupted maxillary molar to alveolar crest.

M2M: Buccal alveolar bone height from mesiobuccal cusp tip of distal most fully erupted maxillary molar to alveolar crest.

M2D: Buccal alveolar bone height from distobuccal cusp tip of distal most fully erupted maxillary molar to alveolar crest.

CNT: Labial alveolar bone height from centre of incisal edge to alveolar crest of mandibular central incisor.

Apical reference point for all these locations was determined as the point at the intersection of the alveolar crest and line dropped perpendicular to the center of the incisal edge/cusp tip.

Specifically, at each maxillary molar side of an animal, alveolar bone height was measured at 4 locations; M1M, M1D, M2M, M2D (Fig 2) as listed above. Linear distance
between the buccal cusp tip to the buccal alveolar crest along the long axis of the tooth was measured using a digital caliper (precision 0.001 mm). For each mandibular incisor side of an animal, alveolar bone height was measured at one location; CNT (Fig 3). Specifically, the linear distance between mesiodistally center point of the incisal edge and the alveolar crest along the long axis of that tooth was measured using a digital caliper.

No measurements were taken from the lateral incisor because of their increased variability in mesiodistal angulation and incisal edge form. (Fig 3).

Physical measurements of alveolar bone height were conducted by two raters (K.E & S.P). Both raters were calibrated and then independently measured each location three times, which were averaged and used as the measurement for that location.

Measuring alveolar bone height from CBCT using a direct-visual method

All CBCT data (DICOM files) were analyzed by 3 calibrated, independent, and blinded raters (B.T, K.E and S.P) using Dolphin-3D software (Dolphin Imaging and Management Solutions, Chatsworth, CA) at the same computer monitor setting (1600 × 1200 pixels); a standardized protocol was used (Fig 4,5) to reorient the images. More specifically, occlusal plane was set parallel to the floor in the frontal and sagittal views, and the midline of the head was bisected by the midsagittal plane in the axial view (Fig 5).

Measurement of buccal alveolar bone height at the maxillary molar locations was performed on coronal slices in relation to the buccal cusp tips of the maxillary molars
being assessed (Fig 6). The distance between the cusp tip and the alveolar crest was visually determined by the rater and was measured using a linear measurement tool of the software on three consecutive images (Fig 6). Overall, the coronal sections used for measurements corresponded to the locations subject to physical measurements.

Measurement of labial alveolar bone at the mandibular incisor locations was performed on the sagittal section, sagittal axis passing through the middle point of the incisal edge of the central incisor on axial slice (Fig 7). The distance between the labial aspect of the incisor edge to the alveolar crest was determined using the linear measurement tool. During analyses the rater was allowed to zoom in and make these measurements.

Measuring alveolar bone height from CBCT using an interactive gray-level assisted method

The DICOM files were imported into Image J (Developed by NIH) software and converted into “tiff” format. For incisors, no additional reorientation was performed. For molars, the stacked 3-D “tiff” file was rotated using the Transform J tool to make the occlusal plane parallel to the floor in the frontal and lateral views and make the midline of head bisected by the midsagittal plane, in a similar fashion as the reorientation performed for the direct-visual method in Dolphin. Reorientation was not performed for Incisors. 3 consecutive 2-D images for each molar or incisor location were acquired
similar to the direct-visual method mentioned above, which were used for subsequent measurement using an interactive gray-level assisted method. Specifically, the 2-D images were transformed into the gray-level values of all pixels in that image in ImageJ, and the results were subsequently exported to an Excel worksheet (Microsoft 2010) (Fig 8). Then, a macro was written so that the gray values of pixels of an image were separated to 20, 50 and 85 percentile groups and labeled by 3-color scales, based on which the rough contour of the tooth and alveolar bones became visible (Fig 9).

The next step involved determination of the landmarks (a particular pixel) at the tooth (cusp tip for molars and incisal edge for incisors) and at the alveolar crest for measurements of the distance between these two points. For the former, the boundary was between enamel and air, which was represented by large gray value drop moving from the enamel to the air (Fig 10). Using these characteristics, for the incisors, the most incisal and second to most labial pixel of enamel-level values were identified as the landmark representing the incisal edge. The second to most labial pixel approximated the location used for physical and direct-visual CBCT measurements. For molars, the most occlusal cell with enamel-level gray values was identified as the landmark representing the cusp tip (Fig.11).

For the landmark of the alveolar bone crest, it was located at the boundary of several tissues including the enamel, periodontal ligament, alveolar crest and gingiva. To determine the exact location of the most incisal or occlusal point of the alveolar bone
crest, a 7x7 cell region containing these boundaries was plotted into a line graph (Fig. 12,13). In the graph, the lines were divided into three groups based on the different patterns of value change in the horizontal direction: a bent-up pattern indicative of a strong likelihood of bone presence, a bent-down pattern indicative of a strong likelihood of bone absence, and a straight line pattern indicative of transition between the first two conditions. The straight line pattern was often shown in only one row of data and between the other two patterns. Using this characteristic, the most likely alveolar crest landmark pixel was identified (Fig 13). The same method of identification was used for both the incisor and molar locations.

After the landmark pixel locations were identified, the distance between the two landmarks were calculated using the Pythagoras principle $a^2 + b^2 = c^2$ considering the length of hypotenuse as the alveolar bone height measurement for the straightest curve cell value and the width of each pixel being 0.4 mm (Fig 14).

CBCT measurements for this method were also conducted by the same three independent blinded calibrated raters (B.T, K.E and S.P), who measured images using Dolphin software as well.

**Statistical Analysis**
Inter-rater reliability of physical measurements was assessed by interclass correlation tests (ICC), which was also used to assess the reliability of CBCT measurements among the three raters, stratified based on locations, sides and the analysis methods.

To assess the magnitude of disagreement/agreement between physical measurements and CBCT measurements obtained by either CBCT methods, Bland-Altman plots were used to show the mean differences and 95% limits of agreement (average differences +/- 1.96 of the standard deviation of the differences).^{60}

The accuracy of CBCT measurements obtained by the two methods were further assessed using two approaches. First, the difference between CBCT measurements and physical measurements (average of two raters) and was calculated, which was used as a dependent variable checked the variations caused by CBCT methods, alveolar bone location and rater (as a random effect) using a repeated-measures ANOVA. Post hoc testing was done using the Tukey-Kramer procedure. Then, the delta scores (absolute difference between CBCT and physical measurements) were dichotomized into clinically acceptable (≤1mm deviation from physical measurement) or unacceptable (>1mm deviation from physical measurement) and compared between CBCT methods and locations using multiple Cochran-Mantel-Haenszel statistics, which were adjusted using the step-down Bonferroni method of Holm.

**Results**

45
Physical measurements

The interrater agreement data of physical measurements are shown in Table I. Overall, the ICC between raters for physical measurements was close to 0.8 except at two locations on the right side molar region M1D and M2M, which were 0.548 and 0.682 respectively. Overall, measurements of left molar locations tended to be more reliable than right molar locations, while measurements for right incisor locations tended to be more reliable than left molar locations, although neither was significant (Table II).

Reliability of CBCT measurements

The inter-rater reliability data of CBCT measurements are presented in Table III. When molar locations were assessed individually, measurements made by the gray-level assisted method showed coefficient values in a range of 0.730-0.909, which were overall better than those made by the vision-based method (0.297-0.823). When the molar locations were analyzed together, measurements made by the gray-level assisted method remained to be more reliable than those made by the vision-based method, even though only the left side reached statistical significance. Regardless of the CBCT method, measurements from the left side were more reliable than from the right side.

For the incisor location (CNT), measurements made by the gray-level assisted method showed coefficient values of 0.940 and 0.876 for left and right side, respectively, which were roughly comparable to those obtained by the vision-based method (0.880 and 0.915, respectively). Overall, these reliability values were better than those from the
molar locations, except for that of left molar regions measured by the gray-level assisted method.

**Accuracy of CBCT measurements**

The Bland-Altman parameters and plots (mean and limits of agreement, LOA) are illustrated in Table IV and Fig. 15-16, respectively. The mean was the average variation between CBCT and physical measurements, while the limits of agreement (LOA), defined as ±1.96 SD, indicated the range of variation between CBCT and physical measurements. Overall, CBCT measurements were slightly greater than physical measurements as the mean difference between physical and CBCT measurements (Physical minus CBCT) were negative. Regardless of CBCT methods, the CBCT overestimations were greater on the right side, shown by higher negative values, which was prominent for the molar locations but not for the incisor locations. The LOA ranges of the right molars were also greater than those of the left molars, but this pattern was reversed for the incisors, where the LOA ranges were larger on the left side than on the right side. On the other hand, the mean and LOA values were not evidently different between the direct-visual and gray-level assisted methods.

Next, the results of repeated-measures ANOVA tests for the variations of methods and locations are shown in Table V and Fig 17. Overall, the location factor caused a significant impact, but the impact from the method factor alone or from interaction with
the location factor was insignificant (Table VI). Post-hoc analyses further revealed that significant variations were present between the right and left molar locations regardless of analysis methods (Table VII).

Lastly, the results about the proportions of clinically acceptable CBCT measurements are presented in Fig.18. Overall, CBCT measurements from the incisor locations were 73%-88% acceptable, while from the molar locations, CBCT measurements were 82%-95% acceptable. The highest values were present at the left molar location, showing 91% and 95% of acceptable measurements for the vision-based and gray-level assisted method, respectively. Multiple Cochran-Mantel-Haenszel Chi-square tests were performed to assess whether there were differences between the right and left locations, and between the two CBCT methods. The results showed that the only significance was between the right and left molar locations when the gray-level assisted method was used.

**Discussion**

As previous studies have indicated that alveolar bone height measurement from orthodontic CBCT images using a vision-based method can be rather unreliable and inaccurate, this study was undertaken to develop a new method based on the interactive analysis of pixel gray-level values. Specifically, two main questions were addressed: 1) whether alveolar bone height measurements obtained by using interactive gray-level assisted method was more reliable than those obtained by using a vision-based method,
and 2) whether alveolar bone height measurements obtained by using interactive gray-level assisted method resulted in more accurate alveolar bone measurements (relative to physical measurements) than the vision-based method.

Overall, our data provided an affirmative answer to the first question. That is, CBCT measurements obtained using an interactive gray-level assisted method were more reliable among raters than those obtained using the conventional vision-based method (Table II) for the maxillary molar locations. Without a set of universally accepted criteria for interpreting ICC data regarding the degree of reliability, criteria used to interpret Cohen’s kappa values have been commonly used before and were also used here. Based on this interpretation, most (6 out of 8) interrater reliability values for the molar locations were poor to fair (ICC<0.8) for the vision-based method, but were mostly good to excellent (ICC>0.8) for the grey-level assisted method (Table III). With all 4 molar locations combined for each side, interrater reliability was clearly better when the gray-level assisted method was used, although statistical significance was only reached for the left molar locations.

For the vision-based method, our reliability data were consistent with those reported before. In previous studies, 2 raters have been used to evaluate reliability. In this study, we included three raters, which represents an improvement, and found similar results, further confirming that it is hard to achieve highly reliable measurements using a
vision-based method. The challenge of using a vision-based method can be attributed to several factors such as difference in rater’s visual acuity, application of measurements standards, measurement tools and rater’s experience etc. In a study evaluating diagnostic accuracy of radiologists based on experience, it was reported that greater specificity and accuracy was noted with radiologists who were more experienced and routinely used that particular technique. The three raters included in this study were not experienced radiologist and clearly had substantial variability amongst them in identifying the landmarks visually. One more factor that may add to the challenge is the variation of image display settings, such as brightness and contrast used by different raters.

In contrast to a vision-based method, our interactive gray-level assisted method minimizes the dependence on rater’s visual acuity, experience and image display, as landmark identification is conducted by observing the patterns of grey-level value change. More specifically, for the enamel landmark, the boundary pixels between enamel and air are substantially different so the identification should be minimally unambiguous. For the alveolar crest landmark, identification of the transition between bent-up and bent-down patterns is slightly harder and requires some experience from the raters in observing line graphs in Excel. While this is a potential source of inter-rater variability, learning to observe the patterns of line graphs is still much easier and more objective than learning to observe the actual CBCT images.
Surprisingly, regardless of the methods, reliability values of CBCT measurements were better at the left molar locations than the right molar locations. The exact reason for this difference is not completely clear, but it may be related to two possible factors. One is that alveolar bone at the left molar region may be anatomically different from that on the right side, which may be more difficult to measure. This was somewhat supported by the reliability data of physical measurements, which showed the same pattern of being better at the left side. (Table I). The other potential factor is that the CBCT images were not representing the physical truth equally between the two sides, thus creating an artificial systematic side-related error between the two sides. To date, no studies have reported this kind problem for the CBCT imaging technique, but based on our data significantly higher overestimation of the measurements from the right molar than the left molar locations regardless of the CBCT measurement methods (Table IV), this possibility is not completely unlikely.

Compared to the molar locations, improvement of the measurement reliability at the incisor locations was minimal. Very likely, this is because that vision-based methods already resulted in highly reliable measurements (≥ 0.88), leaving little room for improvement. These data were also comparable to those reported before for the same locations. Interesting, the side-related differences observed from the molar region were not observed from the incisor region, which is probably because the incisors are immediately adjacent and identical to each other.
For the second major question which asked about measurement accuracy, our data indicated that the grey-level assisted method slightly improved the accuracy of CBCT measurements at the left molar region, but not at the right molar region. More specifically, in the left molar region, the mean difference between physical and CBCT measurements was 0.121 mm and -0.041 mm for the vision-based and gray-level assisted method (Table IV), respectively. Some improvement was also shown in the incisor region, with the mean difference between physical and CBCT measurements being -0.305 and -0.127 mm for the vision-based and gray-level assisted method, respectively.

For the finding of no consistent improvement in accuracy for the gray-level assisted method, several possible explanations can be offered. First, the accuracy of alveolar bone measurements with a vision-based method were surprisingly good already, which makes further improvement difficult. (Fig. 17). Second, there was a significant systematic difference of measurements from the left and right molars regions, which if caused by errors of the CBCT images, cannot be overcome by an analytical method which does nothing to correct errors of the CBCT images.

It is worth further noting that there was an unexpected but significant side-related difference between the right and left molar regions, which were reflected by their mean deviations from physical measurements (Table IV) and by their proportions of acceptable measurements (Fig. 18). Our data clearly showed that this side-specific difference is
independent of the analysis methods. The reasons for these side-specific differences, however, can be rather ambiguous. As mentioned earlier, it may be because that the left molar region is anatomically easier to measure accurately than the right side (Table I), or because of errors of the CBCT scan and reconstruction algorithms that resulted in artificial differences between the left and right molar regions.

CBCT incorporates a relatively short source-to-object distance with relatively divergent x-ray beam geometry, producing a projection with marked differential magnification. CBCT images are produced only after application of a reconstruction algorithm, such as that of Feldkamp et al. Such algorithms incorporate geometric correction factors compensating for differential distortion peripheral to the central x-ray beam projection along with back projection corrections for tube position, tube, object, and detector distances. Resulting CBCT images are therefore corrected to produce orthogonal projections with no differential magnification between bilateral or midsagittal structures. It is believed that this correction accounts for the consistent and remarkable (2 mm) accuracy of CBCT images. Therefore, if the systematic trend of overestimation of CBCT measurements on the right side versus measurements on the left side, 0.12-0.4 mm and 0.041-0.14 mm respectively (Table IV), was indeed caused by errors of the CBCT technique, we speculate that it is more likely related to reconstruction algorithm rather than the scan itself. Clearly, more studies are required to test these speculations.
The data found in this study have potential clinical implications. In our study, we used five month-old pigs as the in vitro model, which have similar craniofacial anatomy and function to humans [8-9], and the age is equivalent to that of human adolescents.\textsuperscript{23} Therefore, the findings from this study is more relevant and useful for the adolescent orthodontic population.\textsuperscript{21} Although, adult human cadavers\textsuperscript{22, 24} and skulls\textsuperscript{26} have been used in the past, they have their own limitations such as decreased radiodensity\textsuperscript{22} and absence of soft tissue which is not a true representation of a clinical scenario\textsuperscript{21}. Our pigs underwent scanning within 24 hours of their time of euthanasia to prevent the effect of embalming or tissue fixation, which has been noted to alter image quality by altering tissue contrast.\textsuperscript{73} In this sense, our specimens strongly simulated the condition of CBCT scan received by clinical orthodontic patients.

In our study alveolar bone height measurements were made from the cusp tip or incisal edge to alveolar crest. Identifying cusp tip or incisal edge does not pose a problem, but identification of bone margin due to relatively similar density with adjacent cementum/PDL is difficult to accurately determine bone margin using a vision-based method. As previously found, when the alveolar bone thickness is reduced to a level below or near the voxel size, the voxels lying on the alveolar bone will reflect an average density of the alveolar bone and the periodontal ligament, rather than the true value of the alveolar bone. As a result, a thin layer of alveolar bone with a thickness near or below the
voxel size 0.4 mm of the CBCT images can become indistinguishable from the adjacent periodontal ligament and not considered bone when taking alveolar bone height measurement. More specifically, the vision-based method requires certain contrast resolution for distinguishing two objects of similar densities in close proximity. Using a high-contrast line pair phantom, Ballrick et al examined this aspect of iCAT system and found that a minimum distance of 0.86 mm is required for clear distinction between two metal plates of the same density. Putting this in perspective, PDL space of 0.5 mm can pose difficulties in differentiating alveolar bone from cementum visually, thus almost impossible to determine the location of the alveolar bone crest.

Compared to the vision-based method, our data showed that the interactive gray-level assisted method has superior reliability and comparable accuracy. Clinically, a method that leads to better reliability of CBCT measurements without losing the accuracy of measurement is of great significance and value. This is also a great improvement compared to the automatic/fixed segmentation method, which improves reliability but compromises accuracy. Recently, an in vitro study that evaluated diameter of titanium implant rods based on fixed segmentation using CBCT’s had reported inaccuracies in their CBCT measurements even though it was acceptable for their study design. When manual and fixed segmentation techniques were compared for airway evaluation using CBCT, they concluded that manual segmentation seems to be the method with the
greatest accuracy and allows the most operator control. However, manual segmentation is also significantly more time-consuming and impractical for clinical use; it takes approximately 1 hour for airway volume calculations, whereas the same procedure with automatic segmentation was possible in less than 5 minutes. All automatic segmentation programs compared with the manual segmentation program showed high correlation and poor accuracy. Therefore, considering these methods that have been discussed, our interactive grey-level assisted method have excellent reliability together with good to excellent accuracy, user-friendly and less time consuming, which would allow more objective and comparable use in clinical research.

As with many studies, this study has several limitations. The CBCT unit we used in this study may be best set and calibrated for human specimens rather than pig specimens although no study to date has found that as an issue. Additionally, the unexpected side differences between right and left may have confounded our results. Furthermore, left side which appeared to be free from errors either due to image acquisition technique or physical anatomic variation, gray-level assisted method appears to be promising especially on the left side. As discussed above, we do not have a clear explanation for this confounder at present. Furthermore, the unexpected good accuracy of alveolar bone height measurements from these pigs limited our ability to fully test this new method. Finally, even though pigs represent craniofacial anatomy similar to
adolescents, artifacts related to movement and metal braces might obscure the images in a clinical setting.

Overall, this study was to compare and assess accuracy and reliability of alveolar bone height measurements using gray-level assisted method to a conventional vision-based method. We did find improved reliability and slightly improved accuracy in certain molar locations. Nevertheless, more studies involving patient’s CBCT images are necessary to confirm the improvement of accuracy and reliability with this gray-level assisted method. For test of accuracy, human clinical trials are possible if CBCT images are compared to clinical/surgical exam after flap elevation for periodontal treatment.

**Conclusions**

Compared to the direct visual method, our data showed that an interactive grey level assisted method can be used to measure alveolar bone height with superior reliability and at least comparable accuracy to a conventional vision-based method.
Chapter 4 Results

Physical measurements

Interrater agreement values of alveolar bone measurements from the physical measurements are shown in Table 1. Overall ICC between raters for physical measurements was close to 0.8 except at two locations on the right side molar region M1D and M2M, which were 0.548 and 0.682 respectively. Overall, measurements of left molar locations tended to be more reliable than right molar locations, while measurements for right incisor locations tended to be more reliable than left molar locations, although neither was significant (Table 2).

Table 1: Interrater agreement (ICC) of physical measurements.

<table>
<thead>
<tr>
<th></th>
<th>Rater1 (mean±SD)</th>
<th>Rater2 (mean±SD)</th>
<th>ICC between raters</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>L</td>
<td>R</td>
<td>L</td>
</tr>
<tr>
<td>All Molars</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L</td>
<td>8.047 ± 0.853</td>
<td>8.036±0.917</td>
<td>8.140 ± 0.795</td>
</tr>
<tr>
<td>R</td>
<td>8.036±0.917</td>
<td>8.047±0.853</td>
<td>7.927 ± 1.003</td>
</tr>
<tr>
<td>M1M</td>
<td>7.646 ± 0.633</td>
<td>7.550 ± 0.773</td>
<td>7.760 ± 0.646</td>
</tr>
<tr>
<td>M1D</td>
<td>7.326 ± 0.459</td>
<td>7.339 ± 0.434</td>
<td>7.518 ± 0.495</td>
</tr>
<tr>
<td>M2M</td>
<td>9.062 ± 0.606</td>
<td>9.110 ± 0.735</td>
<td>9.050 ± 0.578</td>
</tr>
<tr>
<td>M2D</td>
<td>8.154 ± 0.486</td>
<td>8.146 ± 0.430</td>
<td>8.234 ± 0.437</td>
</tr>
<tr>
<td>Cnt</td>
<td>12.297± 0.892</td>
<td>12.298±0.984</td>
<td>12.253±0.767</td>
</tr>
</tbody>
</table>
Table 2: Interrater agreement (ICC) of physical measurements with 95% confidence intervals

<table>
<thead>
<tr>
<th></th>
<th>L</th>
<th>95% CI</th>
<th>R</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>L</td>
<td></td>
<td>R</td>
<td>Lower bound</td>
</tr>
<tr>
<td>All Molars</td>
<td>0.925</td>
<td>0.881</td>
<td>0.953</td>
<td>0.888</td>
</tr>
<tr>
<td>M1M</td>
<td>0.796</td>
<td>0.512</td>
<td>0.901</td>
<td>0.843</td>
</tr>
<tr>
<td>M1D</td>
<td>0.838</td>
<td>0.343</td>
<td>0.948</td>
<td>0.548</td>
</tr>
<tr>
<td>M2M</td>
<td>0.877</td>
<td>0.715</td>
<td>0.949</td>
<td>0.682</td>
</tr>
<tr>
<td>M2D</td>
<td>0.875</td>
<td>0.712</td>
<td>0.949</td>
<td>0.861</td>
</tr>
<tr>
<td>Cnt</td>
<td>0.868</td>
<td>0.698</td>
<td>0.946</td>
<td>0.937</td>
</tr>
</tbody>
</table>

Reliability of CBCT measurements

Inter-rater reliability of the CBCT measurements was assessed by intra-class correlation tests. The correlation coefficient values are presented in Table 3. When molar locations were assessed individually, measurements made by the gray-level assisted method showed coefficient values in a range of 0.730-0.909, which was overall better than those made by the direct-visual method (0.297-0.823). When the molar locations were analyzed together, measurements made by the gray-level assisted method remained to be more reliable than those made by the direct-visual method, even though only the left side was statistically significant. Regardless of the method, measurements from the left side were more reliable than from the right side.
For the incisor location (CNT), measurements made by the gray-level assisted method showed coefficient values of 0.940 and 0.876 for left and right side, respectively. ICC values for direct-visual method was 0.880 and 0.915 for right and left side, respectively. Although these values indicate a side/method interaction, overall they were similar and there was not a statistical difference. In addition, these reliability values were better than those from the molar locations, except for that of left molar regions measured by the gray-level assisted method.

Table 3: Interater agreements (ICC) of CBCT measurements

<table>
<thead>
<tr>
<th>Locations</th>
<th>Direct Visual</th>
<th>Gray-level Assisted</th>
<th>P-values between methods</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Left</td>
<td>Right</td>
<td>Left</td>
</tr>
<tr>
<td>All molars</td>
<td>0.830</td>
<td>0.710</td>
<td>0.932</td>
</tr>
<tr>
<td>M1M</td>
<td>0.496</td>
<td>0.819</td>
<td>0.829</td>
</tr>
<tr>
<td>M1D</td>
<td>0.747</td>
<td>0.782</td>
<td>0.902</td>
</tr>
<tr>
<td>M2M</td>
<td>0.645</td>
<td>0.297</td>
<td>0.856</td>
</tr>
<tr>
<td>M2D</td>
<td>0.823</td>
<td>0.448</td>
<td>0.750</td>
</tr>
<tr>
<td>CNT</td>
<td>0.880</td>
<td>0.915</td>
<td>0.940</td>
</tr>
</tbody>
</table>

M1M: Mesial buccal cusp for the second last fully erupted maxillary molar; M1D: Distal buccal cusp for the second last fully erupted maxillary molar; M2M: Mesial buccal cusp
for the last fully erupted maxillary molar; M2D: Distal buccal cusp for the last fully erupted maxillary molar; CNT: Mandibular central incisor.

Accuracy of CBCT measurements

All the CBCT measurements from three raters were pooled separately for both the CBCT methods based on location. The accuracy of CBCT measurements was assessed and compared between the gray-level assisted and direct-visual methods in three different ways. First, the differences between the CBCT and physical measurements (gold standard) were depicted by Bland-Altman plots. The Bland-Altman measurements (mean and limits of agreement, LOA) and graphs are illustrated in Table 4 and Fig.15-16, respectively. The mean was the average variation between CBCT and physical measurements, while the limits of agreement (LOA), defined as +/- 1.96 SD, indicated the range of variation between CBCT and physical measurements. Overall, CBCT measurements were slightly greater than physical measurements as the mean difference between physical and CBCT measurements (Physical minus CBCT) were negative. Regardless of CBCT methods, the CBCT overestimations were greater on the right side, shown by higher negative values, which was prominent for the molar locations but not for the incisor locations. The LOA ranges of the right molars were also greater than those of
the left molars, but this pattern was reversed for the incisors, where the LOA ranges were larger on the left side than on the right side. On the other hand, the mean and LOA values were not evidently different between the direct-visual and gray-level assisted methods.

Table 4: Mean differences between physical and CBCT measurements and the limits of agreement (LOA, 1.96SD) for each CBCT method categorized based on side.

<table>
<thead>
<tr>
<th>Difference between Physical and CBCT measurements (Physical-CBCT)</th>
<th>Left</th>
<th>Right</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Direct</td>
<td>Gray-level Assisted</td>
</tr>
<tr>
<td>Molars Mean Difference</td>
<td>-0.121</td>
<td>-0.041</td>
</tr>
<tr>
<td>LOA</td>
<td>0.993</td>
<td>1.069</td>
</tr>
<tr>
<td>Incisors Mean Difference</td>
<td>-0.084</td>
<td>-0.145</td>
</tr>
<tr>
<td>LOA</td>
<td>2.900</td>
<td>2.834</td>
</tr>
</tbody>
</table>
Figure 15: Bland-Altman plots for molars: Differences against the mean of the clinical and radiological measurements.
Next, to systematically analyze the possible effects of bone locations (right molar, left molar, right incisor and left incisor) and methods (direct-visual and gray-level assisted) on the accuracy of CBCT measurements, repeated-measures ANOVA test was conducted. More specifically, the difference between physical and CBCT measurements was treated as the dependent variable, while bone location and methods were treated as independent variables, and rater was treated as a random effect. The results of this test are shown in Table 5. Overall, the location factor caused a significant impact, but the impact from the method factor alone or from interaction with the location factor was
insignificant (Table 6,7). Post-hoc analyses further revealed that significant variations were present between the right and left molar locations regardless of analysis methods (Table 7).

Table 5: Two-way repeated-measures ANOVA between two CBCT methods at molar and incisor locations.

<table>
<thead>
<tr>
<th>Effect</th>
<th>Direct-visual</th>
<th>Gray-level assisted</th>
<th>P_{value methods}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Molars</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>-0.121±0.506</td>
<td>-0.041±0.545</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>-0.349±0.705</td>
<td>-0.368±1.017</td>
</tr>
<tr>
<td></td>
<td>P_{value side}</td>
<td>0.035</td>
<td>0.000</td>
</tr>
<tr>
<td>Cnt</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>-0.084±1.480</td>
<td>-0.145±1.446</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>-0.305±0.627</td>
<td>-0.127±0.647</td>
</tr>
<tr>
<td></td>
<td>P_{value side}</td>
<td>0.794</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Table 6: Two-way repeated-measures ANOVA summary

<table>
<thead>
<tr>
<th>Effect</th>
<th>Num</th>
<th>Den</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>METHOD</td>
<td>1</td>
<td>1133</td>
<td>0.61</td>
<td>0.4364</td>
</tr>
<tr>
<td>LOCATION</td>
<td>3</td>
<td>1133</td>
<td>10.43</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>METHOD*LOCATION</td>
<td>3</td>
<td>1133</td>
<td>0.78</td>
<td>0.5043</td>
</tr>
</tbody>
</table>
Figure 17: Repeated measures ANOVA

Figure 18: Comparison of alveolar bone height measurements between the two methods at different locations. P-values were based on two-way repeated-measures ANOVA.
Table 7: Post hoc comparisons

| METHOD | GROUP | _METHOD | _GROUP | Estimate | Error | DF  | t Value | Pr > |t| | Adj P |
|--------|-------|---------|--------|----------|-------|-----|---------|-------|-------|--------|
| 1:IMJ  | LCNT  | 1:IMJ   | LMOL   | -0.1038  | 0.1146 | 1133| -0.91  | 0.3654 | 0.9856 |
| 1:IMJ  | LCNT  | 1:IMJ   | RCNT   | -0.01817 | 0.1449 | 1133| -0.13  | 0.9003 | 1.0000 |
| 1:IMJ  | LCNT  | 1:IMJ   | RMOL   | 0.2233   | 0.1146 | 1133| 1.95   | 0.0516 | 0.5177 |
| 1:IMJ  | LCNT  | 2:DOL   | LCNT   | -0.00602 | 0.1449 | 1133| -0.42  | 0.6759 | 0.9999 |
| 1:IMJ  | LCNT  | 2:DOL   | LMOL   | -0.02435 | 0.1146 | 1133| -0.21  | 0.8317 | 1.0000 |
| 1:IMJ  | LCNT  | 2:DOL   | RCNT   | 0.1604   | 0.1449 | 1133| 1.11   | 0.2688 | 0.9555 |
| 1:IMJ  | LCNT  | 2:DOL   | RMOL   | 0.2043   | 0.1146 | 1133| 1.78   | 0.0748 | 0.6316 |
| 1:IMJ  | LMOL  | 1:IMJ   | RCNT   | 0.08559  | 0.1146 | 1133| 0.75   | 0.4552 | 0.9955 |
| 1:IMJ  | LMOL  | 1:IMJ   | RMOL   | 0.327    | 0.07247| 1133| 4.51   | <.0001 | 0.0002 |
| 1:IMJ  | LMOL  | 2:DOL   | LCNT   | 0.04314  | 0.1146 | 1133| 0.38   | 0.7066 | 0.9999 |
| 1:IMJ  | LMOL  | 2:DOL   | LMOL   | 0.0794   | 0.07247| 1133| 1.1    | 0.2734 | 0.9578 |
| 1:IMJ  | LMOL  | 2:DOL   | RCNT   | 0.2641   | 0.1146 | 1133| 2.31   | 0.0213 | 0.2916 |
| 1:IMJ  | LMOL  | 2:DOL   | RMOL   | 0.3081   | 0.07247| 1133| 4.25   | <.0001 | 0.0006 |
| 1:IMJ  | RCNT  | 1:IMJ   | RMOL   | 0.2414   | 0.1146 | 1133| 2.11   | 0.0353 | 0.4109 |
| 1:IMJ  | RCNT  | 2:DOL   | LCNT   | -0.04245 | 0.1449 | 1133| -0.29  | 0.7697 | 1.0000 |
| 1:IMJ  | RCNT  | 2:DOL   | LMOL   | -0.00618 | 0.1146 | 1133| -0.05  | 0.9577 | 1.0000 |
| 1:IMJ  | RCNT  | 2:DOL   | RCNT   | 0.1785   | 0.1449 | 1133| 1.23   | 0.2183 | 0.9224 |
| 1:IMJ  | RCNT  | 2:DOL   | RMOL   | 0.2225   | 0.1146 | 1133| 1.94   | 0.0524 | 0.5222 |
| 1:IMJ  | RMOL  | 2:DOL   | LCNT   | -0.2839  | 0.1146 | 1133| -2.48  | 0.0134 | 0.2061 |
| 1:IMJ  | RMOL  | 2:DOL   | LMOL   | -0.2476  | 0.07247| 1133| -3.42  | 0.0007 | 0.0151 |
| 1:IMJ  | RMOL  | 2:DOL   | RCNT   | -0.06288 | 0.1146 | 1133| -0.55  | 0.5832 | 0.9994 |
| 1:IMJ  | RMOL  | 2:DOL   | RMOL   | -0.01891 | 0.07247| 1133| -0.26  | 0.7942 | 1.0000 |
| 2:DOL  | LCNT  | 2:DOL   | LMOL   | 0.03627  | 0.1146 | 1133| 0.32   | 0.7517 | 1.0000 |
| 2:DOL  | LCNT  | 2:DOL   | RCNT   | 0.221    | 0.1449 | 1133| 1.52   | 0.1276 | 0.7941 |
| 2:DOL  | LCNT  | 2:DOL   | RMOL   | 0.265    | 0.1146 | 1133| 2.31   | 0.0209 | 0.2876 |
| 2:DOL  | LMOL  | 2:DOL   | RCNT   | 0.1847   | 0.1146 | 1133| 1.61   | 0.1072 | 0.7431 |
| 2:DOL  | LMOL  | 2:DOL   | RMOL   | 0.2287   | 0.07247| 1133| 3.16   | 0.0016 | 0.0350 |
| 2:DOL  | RCNT  | 2:DOL   | RMOL   | 0.04397  | 0.1146 | 1133| 0.38   | 0.7012 | 0.9999 |
Chapter 5 Discussion

As previous studies have indicated that alveolar bone height measurement from orthodontic CBCT images using a vision-based method can be rather unreliable and inaccurate,\textsuperscript{21, 23} this study was undertaken to develop a new method based on the interactive analysis of pixel gray-level values. Specifically, two main questions were addressed: 1) whether alveolar bone height measurements obtained by using interactive gray-level assisted method was more reliable than those obtained by using a vision-based method, and 2) whether alveolar bone height measurements obtained by using interactive gray-level assisted method resulted in more accurate alveolar bone measurements (relative to physical measurements) than the vision-based method.

Inter rater reliability of CBCT methods

Overall, our data provided an affirmative answer to the first question. That is, CBCT measurements obtained using an interactive gray-level assisted method were more reliable among raters than those obtained using the conventional vision-based method (Table II) for the maxillary molar locations. Without a set of universally accepted criteria
for interpreting ICC data regarding the degree of reliability, criteria used to interpret Cohen’s kappa values have been commonly used before and were also used here. Based on this interpretation\textsuperscript{70}, most (6 out of 8) interrater reliability values for the molar locations were poor to fair (ICC<0.8) for the vision-based method, but were mostly good to excellent (ICC>0.8) for the grey-level assisted method (Table III). With all 4 molar locations combined for each side, interrater reliability was clearly better when the gray-level assisted method was used, although statistical significance was only reached for the left molar locations.

For the vision-based method, our reliability data were consistent with those reported before.\textsuperscript{23} In previous studies, 2 raters have been used to evaluate reliability. In this study, we included three raters, which represents an improvement, and found similar results, further confirming that it is hard to achieve highly reliable measurements using a vision-based method. The challenge of using a vision-based method can be attributed to several factors such as difference in rater’s visual acuity, application of measurements standards, measurement tools and rater’s experience etc. In a study evaluating diagnostic accuracy of radiologists based on experience, it was reported that greater specificity and accuracy was noted with radiologists who were more experienced and routinely used that particular technique\textsuperscript{71}. The three raters included in this study were not experienced radiologist and clearly had substantial variability amongst them in identifying the
landmarks visually. One more factor that may add to the challenge is the variation of image display settings, such as brightness and contrast used by different raters.

In contrast to a vision-based method, our interactive gray-level assisted method minimizes the dependence on rater’s visual acuity, experience and image display, as landmark identification is conducted by observing the patterns of grey-level value change. More specifically, for the enamel landmark, the boundary pixels between enamel and air are substantially different so the identification should be minimally unambiguous. For the alveolar crest landmark, identification of the transition between bent-up and bent-down patterns is slightly harder and requires some experience from the raters in observing line graphs in Excel. While this is a potential source of inter-rater variability, learning to observe the patterns of line graphs is still much easier and more objective than learning to observe the actual CBCT images.

Surprisingly, regardless of the methods, reliability values of CBCT measurements were better at the left molar locations than the right molar locations. The exact reason for this difference is not completely clear, but it may be related to two possible factors. One is that alveolar bone at the left molar region may be anatomically different from that on the right side, which may be more difficult to measure. This was somewhat supported by the reliability data of physical measurements, which showed the same pattern of being better at the left side. (Table I). The other potential factor is that the CBCT images were not representing the physical truth equally between the two sides, thus creating an
artificial systematic side-related error between the two sides. To date, no studies have reported this kind problem for the CBCT imaging technique, but based on our data significantly higher overestimation of the measurements from the right molar than the left molar locations regardless of the CBCT measurement methods (Table IV), this possibility is not completely unlikely.

Compared to the molar locations, improvement of the measurement reliability at the incisor locations was minimal. Very likely, this is because that vision-based methods already resulted in highly reliable measurements (≥ 0.88), leaving little room for improvement. These data were also comparable to those reported before for the same locations.\textsuperscript{24} Interesting, the side-related differences observed from the molar region were not observed from the incisor region, which is probably because the incisors are immediately adjacent and identical to each other.

\textbf{Accuracy of CBCT measurements}

For the second major question which asked about measurement accuracy, our data indicated that the grey-level assisted method slightly improved the accuracy of CBCT measurements at the left molar region, but not at the right molar region. More specifically, in the left molar region, the mean difference between physical and CBCT measurements was 0.121 mm and -0.041 mm for the vision-based and gray-level assisted method (Table IV), respectively. Some improvement was also shown in the incisor
region, with the mean difference between physical and CBCT measurements being -0.305 and -0.127 mm for the vision-based and gray-level assisted method, respectively.

For the finding of no consistent improvement in accuracy for the gray-level assisted method, several possible explanations can be offered. First, the accuracy of alveolar bone measurements with a vision-based method were surprisingly good already, which makes further improvement difficult. (Fig. 17). Second, there was a significant systematic difference of measurements from the left and right molars regions, which if caused by errors of the CBCT images, cannot be overcome by an analytical method which does nothing to correct errors of the CBCT images.

It is worth further noting that there was an unexpected but significant side-related difference between the right and left molar regions, which were reflected by their mean deviations from physical measurements (Table IV) and by their proportions of acceptable measurements (Fig. 18). Our data clearly showed that this side-specific difference is independent of the analysis methods. The reasons for these side-specific differences, however, can be rather ambiguous. As mentioned earlier, it may be because that the left molar region is anatomically easier to measure accurately than the right side (Table I), or because of errors of the CBCT scan and reconstruction algorithms that resulted in artificial differences between the left and right molar regions.

CBCT incorporates a relatively short source-to-object distance with relatively divergent x-ray beam geometry, producing a projection with marked differential
magnification. CBCT images are produced only after application of a reconstruction algorithm, such as that of Feldkamp et al. Such algorithms incorporate geometric correction factors compensating for differential distortion peripheral to the central x-ray beam projection along with back projection corrections for tube position, tube, object, and detector distances. Resulting CBCT images are therefore corrected to produce orthogonal projections with no differential magnification between bilateral or midsagittal structures. It is believed that this correction accounts for the consistent and remarkable (2 mm) accuracy of CBCT images. Therefore, if the systematic trend of overestimation of CBCT measurements on the right side versus measurements on the left side, 0.12-0.4 mm and 0.041-0.14 mm respectively (Table IV), was indeed caused by errors of the CBCT technique, we speculate that it is more likely related to reconstruction algorithm rather than then scan itself. Clearly, more studies are required to test these speculations.

**Implications of this study**

The data found in this study have potential clinical implications. In our study, we used five month-old pigs as the *in vitro* model, which have similar craniofacial anatomy and function to humans [8-9], and the age is equivalent to that of human adolescents. Therefore, the findings from this study is more relevant and useful for the adolescent orthodontic population. Although, adult human cadavers and skulls have been used in the past, they have their own limitations such as decreased radiodensity and
absence of soft tissue which is not a true representation of a clinical scenario\textsuperscript{21}. Our pigs underwent scanning within 24 hours of their time of euthanasia to prevent the effect of embalming or tissue fixation, which has been noted to alter image quality by altering tissue contrast. \textsuperscript{73} In this sense, our specimens strongly simulated the condition of CBCT scan received by clinical orthodontic patients.

In our study alveolar bone height measurements were made from the cusp tip or incisal edge to alveolar crest. Identifying cusp tip or incisal edge does not pose a problem, but identification of bone margin due to relatively similar density with adjacent cementum/PDL is difficult to accurately determine bone margin using a vision-based method. As previously found, when the alveolar bone thickness is reduced to a level below or near the voxel size, the voxels lying on the alveolar bone will reflect an average density of the alveolar bone and the periodontal ligament, rather than the true value of the alveolar bone. As a result, a thin layer of alveolar bone with a thickness near or below the voxel size 0.4 mm of the CBCT images can become indistinguishable from the adjacent periodontal ligament and not considered bone when taking alveolar bone height measurement. \textsuperscript{23} More specifically, the vision-based method requires certain contrast resolution for distinguishing two objects of similar densities in close proximity. Using a high-contrast line pair phantom, Ballrick et al \textsuperscript{25} examined this aspect of iCAT system and found that a minimum distance of 0.86 mm is required for clear distinction between two metal plates of the same density. Putting this in to perspective, PDL space of 0.5 mm
can pose difficulties in differentiating alveolar bone from cementum visually, thus almost impossible to determine the location of the alveolar bone crest.

Compared to the vision-based method, our data showed that the interactive gray-level assisted method has superior reliability and comparable accuracy. Clinically, a method that leads to better reliability of CBCT measurements without losing the accuracy of measurement is of great significance and value. This is also a great improvement compared to the automatic/fixed segmentation method, which improves reliability but compromises accuracy. Recently, an in vitro study that evaluated diameter of titanium implant rods based on fixed segmentation using CBCT’s had reported inaccuracies in their CBCT measurements even though it was acceptable for their study design. When manual and fixed segmentation techniques were compared for airway evaluation using CBCT, they concluded that manual segmentation seems to be the method with the greatest accuracy and allows the most operator control. However, manual segmentation is also significantly more time-consuming and impractical for clinical use; it takes approximately 1 hour for airway volume calculations, whereas the same procedure with automatic segmentation was possible in less than 5 minutes. All automatic segmentation programs compared with the manual segmentation program showed high correlation and poor accuracy. Therefore, considering these methods that have been discussed, our interactive grey-level assisted method have excellent reliability together with good to
excellent accuracy, user-friendly and less time consuming, which would allow more objective and comparable use in clinical research.

**Limitations of our study**

As with many studies, this study has several limitations. The CBCT unit we used in this study may be best set and calibrated for human specimens rather than pig specimens although no study to date has found that as an issue. Additionally, the unexpected side differences between right and left may have confounded our results. Furthermore, left side which appeared to be free from errors either due to image acquisition technique or physical anatomic variation, gray-level assisted method appears to be promising especially on the left side. As discussed above, we do not have a clear explanation for this confounder at present. Furthermore, the unexpected good accuracy of alveolar bone height measurements from these pigs limited our ability to fully test this new method. Finally, even though pigs represent craniofacial anatomy similar to adolescents, artifacts related to movement and metal braces might obscure the images in a clinical setting.

**Future Directions**

Overall, this study was to compare and assess accuracy and reliability of alveolar bone height measurements using gray-level assisted method to a conventional vision-based method. We did find improved reliability and slightly improved accuracy in certain molar locations. Nevertheless, more studies involving patient’s CBCT images are
necessary to confirm the improvement of accuracy and reliability with this gray-level assisted method. For test of accuracy, human clinical trials are possible if CBCT images are compared to clinical/surgical exam after flap elevation for periodontal treatment.

**Conclusions**

Compared to the direct visual method, our data showed that the interactive Grey level assisted method has superior reliability and comparable accuracy.
References


80


70. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159-74.