Title
Investigating the Accuracy and Reproducibility of 3dMDface System for Soft Tissue Analysis

Author
Kachroo, Yasir M.

Publication Date
2015

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

“Investigating the Accuracy and Reproducibility of 3dMDface System for Soft Tissue Analysis”.

A thesis submitted in partial satisfaction of the requirements

for the degree Master of Science in Oral Biology

by

Yasir Mushtaq Kachroo

2015

ABSTRACT OF THE THESIS

Investigating the Accuracy and Reproducibility of 3dMDface System for Soft Tissue Analysis

by

Yasir Mushtaq Kachroo

Master of Science in Oral Biology

University of California Los Angeles, 2015

Professor Yeumin Hong, Co-Chair

Professor Won Moon, Co-Chair

Objective: Three-dimensional (3D) soft tissue (ST) changes from orthodontic treatment can now be evaluated with the development of 3D photography. However, the accuracy and reliability of this method is still in question. This study aims to 1) evaluate the accuracy of 3dMD imaging using caliper, surface, and 3D morphometrics and 2) find standard facial expressions that can be best reproduced by 3dMD imaging for soft tissue analysis.

Methods: Three-dimensional (3D) ST facial landmarks were obtained from 40 adults not undergoing orthodontic treatment through the use of 3dMD facial photographic software. A total of 16 landmarks were used and 21 parameters were measured for surface distances (6 in vertical, 10 in anterior posterior, and 5 in transverse). 3dMD images of four different facial expressions (repose (R), maximum intercuspation (MIP), posed smile (S), and smile with lips closed (SLC)) were taken at 0hr, 1hr, 24hr, 1wk, 2wk, 3wk, and 4wk time intervals. As a feasibility test, these

measurements were taken on a mannequin at the above time intervals. Superimposition of seven 3dMD images for each facial posture per subject was performed for analysis. Error magnitude statistics: mean absolute difference (MAD), standard deviation of the error (SD), Root Mean Square Error (RMSE), relative error magnitude (REM), technical error magnitude (TEM), and intraclass correlation coefficient (ICC) were used.

3D morphometrics, seven images were taken of mannequin and subjects posing in MIP and posed smile. These images were loaded into an initial software pipeline to generate a closed mesh. Closed mesh were traced for surface curves and finally loaded into a dual pipeline to average those seven images. Color-coded displacement maps were generated from the dual pipeline images to evaluate changes across the seven time points. T-stats was performed to quantify these changes.

Results: Measurements on a mannequin confirmed the reliability of all the landmarks and parameters used in this study. For 40 subjects, 3dMD measurements between ST landmarks were reproducible with a random error lower than 1mm except for the distance from cheilion to cheilion (2.79mm). Comparing across the four facial postures, MIP showed the smallest variation with mean absolute difference (MAD) (2.03mm) and SD (0.81mm). S posture had the largest variation with MAD (2.79mm) and SD (1.28mm). For 3D morphometrics, mannequin showed that the t-stats are at near zero on the face. When applying 3D morphometrics for MIP and posed smile, areas around exocanthion, cheilion, and pronasale showed slight variation.

Conclusion: 3dMD can perform ST analysis with both accuracy and reproducibility and can be utilized for evaluating ST changes with dental treatment when using advanced technology like 3D morphometrics. Cheilion, alar base, and exocanthion might not be an appropriate landmark to

use for ST analysis. R and MIP facial postures are recommended over smile (S or SLC) for more consistent parameter measurements.

The thesis of Yasir Mushtaq Kachroo is approved by:

Neal Garrett

Carl Maida

Yeumin Hong, Committee Co-Chair

Won Moon, Committee Co-Chair

University of California, Los Angeles

2015

TABLE OF CONTENTS

Introduction

Over the past few decades, 3D imaging has progressed significantly and has found great implication in orthodontics, oral and maxillofacial surgery, and plastic surgery. The use of digital 3D photography, also known as stereophotogrammetry, is a relatively new method that allows users to acquire 3D representation of the craniofacial complex. [1, 2] Stereophotogrammetry works by photographing a 3D object from two different planes and generating a 3D reconstruction of an object. [3] There are many added benefits to this type of image acquisition. It provides us a means to measure not only linear distances but also surface distances, surface areas, and volumes. In addition, <x,y,z> coordinates can also be extrapolated to perform a variety of statistical shape analyses, which can be used for the diagnosis of craniofacial dysmorphology or changes. [4] Thus, digital 3D stereophotogrammetry has the potential to provide a more comprehensive and in-depth assessment of a patient's craniofacial morphology compared to traditional anthropometric methods. [5]

There are numerous applications of 3D imaging in orthodontics, which include pre- and post- assessment of dentoskeletal relationships, orthodontic outcomes with regard to soft and hard tissue analysis, 3D treatment planning, and 3D soft and hard tissue predictions. [2, 6] Although orthodontic treatment may primarily be targeted at modification of skeletal structure, there are often distinct soft-tissue changes to parts of the face. For example, after performing a bilateral sagittal split osteotomy (BSSO) and bimaxillary orthognathic surgery on a class III patient, the thickness of tissue in the area below the nose and upper lip was reduced by an average of 2 mm (SD 3.0 mm) of its value before surgery, while the soft tissues of the lower lip, chin, and mentolabial sulcus increased in thickness by 3 mm (SD 2.6 mm). [7] Another study

demonstrated that distraction osteogenesis of the human mandible not only lengthened deficient bone but also increased linear and circumferential dimensions of the attached soft tissue. [8] Soft tissues of the face play a major role in determining the limitations of orthodontic treatment in terms of esthetics and functionality; therefore, incorporating soft-tissue adaptations and contours is essential during treatment planning. [9] With soft tissue capture tools, orthodontists can more easily document patients electronically to evaluate, monitor and quantify outcomes. Stereophotogrammetry enables the opportunity for 3D assessment of both growth and treatment effects by superimposing images taken at different times throughout the course of treatment. A visual representation of these morphological changes can be displayed through a color-coded map. [10] Hence, rather than assessing treatment outcomes by observing soft tissue profiles using lateral cephalometric radiographs, clinicians can evaluate soft tissue changes in 3D and accurately quantify changes. [11]

Because 3D photogrammetry is a form of "indirect" anthropometry, 3D images can be manipulated in any direction without having the patient actually be there. This provides the orthodontist with considerable amounts of information without having patient recall or time restrictions during a clinical appointment. In addition, this allows for measurement of soft tissue changes that could be uncomfortable if done directly on the patient (e.g. around the eyes) or pose risk for injury. Using specific softwares, these images can be enlarged, rotated, and rendered in any way needed for better analysis. [12] More importantly, multiple 3D images taken at different time points can be superimposed along the whole surface of the face. Software tools can align 3D data sets at different time points with subvoxel accuracy and allow the clinician to use color-coded maps to measure changes with time or treatment procedures. [13] This automated method eliminates the need for observer-dependent techniques to overlap anatomic landmarks and

greatly reduces observer error. [10, 13, 14] In addition, landmarks can be assigned directly onto these 3D images and allows the clinician to collect measurements right after data acquisition. As a result, the amount of measurement time needed for both the subject and clinician is reduced. Decreasing the amount of patient interaction is particularly attractive for use with children and patients with developmental disabilities with whom behavior and cooperation could be difficult. [15]

Overall Objectives and Specific Aims

In craniofacial orthodontics and surgery, precision is absolutely critical. Recent technology, including sophisticated 3D imaging devices, has made the task of orthodontists and surgeons much easier since more accurate measurements can be made. This is a tremendous improvement over conventional two-dimensional imaging and provides an added value for patient diagnosis. [16]

The manufacturer of the technology that will be used in this study is 3dMD, Inc. (Atlanta, Georgia), a pioneer in three-dimensional imaging systems. The company's technology can capture objects with exact 1:1 proportions with an astounding speed of 1.5 milliseconds. This rapid rate ensures two things: 1) a virtual elimination of motion disturbance during the moment of capture and 2) a high degree of patient compliance as they are inconvenienced only for a moment by using completely non-invasive means. This is a vast improvement over 2D technologies whose image-capturing system is both time-consuming and uncomfortable for patients. In addition to the accuracy, speed, and comfort of 3D imaging systems, they also

provide clinicians and researchers with the ability to more easily reproduce images as the image-taking setting and protocols can be replicated with minimal variation. Furthermore, 3D systems allow for numerical data and images to be stored in a database, which can be shared and compared with other clinicians and researchers across the globe.

Perhaps the most important advantage of implementing 3D imaging systems in a clinical setting is that it permits orthodontists to evaluate and even manipulate the images in order to determine the next steps of treatment. Of course, 3D imaging also is helpful in evaluating the outcomes of patients post-treatment, and, as aforementioned, the data accumulated by clinicians and researchers creates a virtual database that can be used for anthroprometric and cephalometric studies.

Three-dimensional technology has immense potential in the field of orthodontics, and this study is an attempt to demonstrate how such technology can be used to acquire measurements, data, and surface images that 2D technology was incapable of producing. With more precise and reliable data and a proper analysis of such data, more effective and efficient treatment plans can be realized and put into practice. The goal of clinicians is to offer patients the best treatment outcome possible, and 3D imaging certainly has the potential to help the clinician deliver more optimal treatment results with less effort on his/her part and on the part of the patient. [17]

Several 3D photogrammetric devices have entered the market for 3D imaging. Although some devices have been validated independently in other studies, few attempts have been made to compare accuracy amongst different 3D photogrammetric systems. In addition, only linear measurements have been done to evaluate the accuracy of these 3D imaging systems. The first aim of our study is to evaluate the accuracy of 3D imaging (3dMDface system) and use our

methodology of quantifying accuracy to serve as a standard for future studies on various 3D imaging systems.

Certain facial expressions have been shown to be more consistent than others for the superimposition involved in 3D imaging. [11] Reproducibility of facial expression is important in assessing changes in orthodontic procedures over time. [18] The second aim of this study is to determine which facial expressions are the most reproducible. This will help to minimize the effect of changes associated with minor discrepancies in facial expression when measuring changes due to treatment. Finding the most reproducible facial expression can also have major clinical applications. In order to accurately combine cone beam computed tomography (CBCT) and 3D stereophotogrammetry, it is important that stable regions (regions of smallest variation) are registered with each other. [19] Finding the most reproducible facial expression and using that as a standard for image acquisition can greatly improve patient diagnosis.

Background

Since the introduction of the cephalostat, Broadbent emphasized the importance of coordinating the lateral and postero-anterior cephalometric films to achieve a distortion-free definition of skeletal craniofacial form. [6] The initial studies on this method were done on 3D analysis of maxillary growth changes in girls by Singh and Savara. [20] Since then, computer programs have been developed to collect three-dimensional coordinates directly from digital

cephalograms, eliminating the need for hand tracing and mouse-based X–Y digitizing tablets. [3, 6, 21]

Photography

Photography has been used as diagnostic tool to clinically document patients and was readily incorporated into the clinic to enhance patient records. However, there are many drawbacks with the use of photographic analysis. Several variables can cause discrepancy when comparing 2 similar photographs: distance between camera and subject, camera angle, head position (roll-pitch-yaw orientations), and photography protocol inconsistencies. [4] These problems are further complicated when integrating a patient's lateral photograph with a lateral cephalometric image for analysis. Due to all of the factors involved, achieving a repeatable and reliable alignment between the 2 images is unlikely. [6]

3D cephalometry

Despite several advancements in 3D cephalometric research, this technique is time-consuming, exposes the patient to radiation, does not clarify soft tissues, and poses challenges in accurately relating the same landmarks in the two radiographs.

Cone beam computerized tomography (CBCT) scanning

With the cone-beam systems, 3D volumetric data images can be acquired with low radiation. In addition, CBCT permits re-orientation of 2D images in coronal, sagittal, oblique, and various incline planes. Compared with conventional CT, CBCT allows gathering of patient

information with less radiation dose (15 times less). The radiation dose of CBCT is equivalent to a dose of 12 panoramic radiographs, on average. However, a major limitation of CBCT is its limited capacity in displaying soft tissues, such as the color of the skin. Furthermore, the higher price and radiation dose of CBCT devices still limit frequent use compared with conventional radiographs. [3, 6, 19, 21]

Laser scanner devices

Laser scanners were one of the first 3D surface-imaging technologies to use a laser beam across the surface area of the object to reproduce facial morphology. However, patients are scanned with their eyes closed, which may pose an issue because it may interfere with the natural facial expression. [3] If the patient were to accidentally open their eyes, laser exposure can cause harm, particularly in growing children. In addition, the slowness of the procedure in taking and developing these images may allow for distortion to occur on the scanned images. As the scanner revolves around the patient's head, the patient needs to stay motionless for one minute or longer. Due to the likelihood of patient movement related to laser scanning, extraoral laser scanning is very difficult to obtain digital models. [3]

Stereophotogrammetry

Stereophotogrammetry has been developed from old photogrammetric techniques to permit a more comprehensive and accurate evaluation of the captured subject. [3] This technique uses one or more converging pairs of views to generate a 3D model that can be viewed from any perspective. The earliest clinical use of stereophotogrammetry was reported by Thalmann-Degan in 1944 who documented changes in facial morphology produced by orthodontic treatment.

Limitations of 2D imaging and anthropometry

Craniofacial anthropometry has been the standard technique for quantifying and identifying facial dysmorphology, treatment planning, and assessing longitudinal studies. [15] Several variables can cause disparity when comparing two ostensibly similar photographs: angle of the camera, distance between the camera and the subject, and head position. [4, 15] These inconsistencies are further compounded with the difficulties in combining a patient's lateral photograph (soft tissue image) with a lateral cephalogram (hard tissue image) for analysis. Because facial images and lateral cephalograms are not taken simultaneously, any changes in facial expression, head profile, or orientation can lead to discrepancies. [19] There is software available that can calibrate the images to best fit and to account for these discrepancies, but the accuracy of those modifications can be questioned. [14, 19, 22] Overall, achieving a reproducible and reliable registration between the two images is unlikely, considering all of the variables involved.

Furthermore, finding a reliable method for measuring soft tissue changes has been a challenge. Because lateral cephalometry is mainly used for hard tissue analysis, it is difficult to accurately coordinate hard tissue landmarks with soft tissue landmarks for orthodontic applications. In addition, because lateral cephalometry is a 2D representation of 3D structures, 3D structural information of the face is lost, making it difficult to ensure the exact point being measured is the same when comparing the pre- and post-treatment soft tissue measurements. [10] Recently, 3D CT scans have been used to compensate for the drawbacks of 2D images. However, 3D CT scans still do not provide soft tissue visualizations. This is where 3D imaging plays a major role.

Future direction towards 3D morphometrics

With the advent of 3D photography systems like 3dMD, it is now possible to capture 3D facial photographs of patients for use in orthodontic practice. However, devising and integrating practical 3D analyses and utilizing the wealth of information acquired when tracking patients in three dimensions is still in its conception. Once 3D morphometrics is applied, a more rapid and detailed diagnostic imaging for orthodontics and surgical treatment planning may be possible. Visualization and 3D quantification of the human face can be accomplished with unprecedented detail and generation of normative data across various strata will be possible. Ultimately, correlation of 3D facial soft tissue photographic images to their corresponding 3D CBCT images will enable the creation of a detailed virtual patient. The ability to analyze patients or groups of patients in 3D would shift the diagnostic paradigms currently used in craniofacial analysis towards an ever-progressive direction.

Materials and Methods

Study Design

This study was conducted in the Department of Orthodontics at the UCLA School of Dentistry.

3dMD Imaging Protocol and Data Acquisition

The 3dMDface system (3dMD Inc., Atlanta, GA, USA) is an automated digital 3D stereophotogrammetric camera that is capable of taking high-resolution color surface captures of the face in approximately 1.5ms. [12][12][12][12][12][12] Multiple cameras fixed at different angles simultaneously take an image, which then get overlapped and processed by the 3dMD software to produce a 3D image. Software algorithms of 3dMDface system combine the images into a single unified 3D point cloud. To enhance visualization, points comprising the surface are linked by vertices creating a 3D polygonal mesh, and these polygons are filled in to create an "air-tight" surface. [2] Texture and color formats are then laid out over the underlying polygonal mesh to give a life-like computer generated rendering of the object.

Part 1:

A mannequin head was used because the facial contours do not change when taken images at different time points. 14 anthropometric surface landmarks were labeled directly on the mannequin using a permanent marker (Figure 1). Table 1 lists the abbreviations and definitions of landmarks used. 21 parameters (6 in the vertical dimension, 10 in the anteroposterior

dimension, and 5 in the transverse dimension) were obtained from these landmarks (Table 2). Caliper measurements were taken using a digital caliper, and surface measurements were taken using a measuring string for all 21 parameters. To analyze measurement accuracy used in this study, the determination of a reliable reference measurement method is crucial. Therefore, two observers independently measured the caliper and surface measurements 1 week apart. An average of these values was obtained and used for statistical analysis.

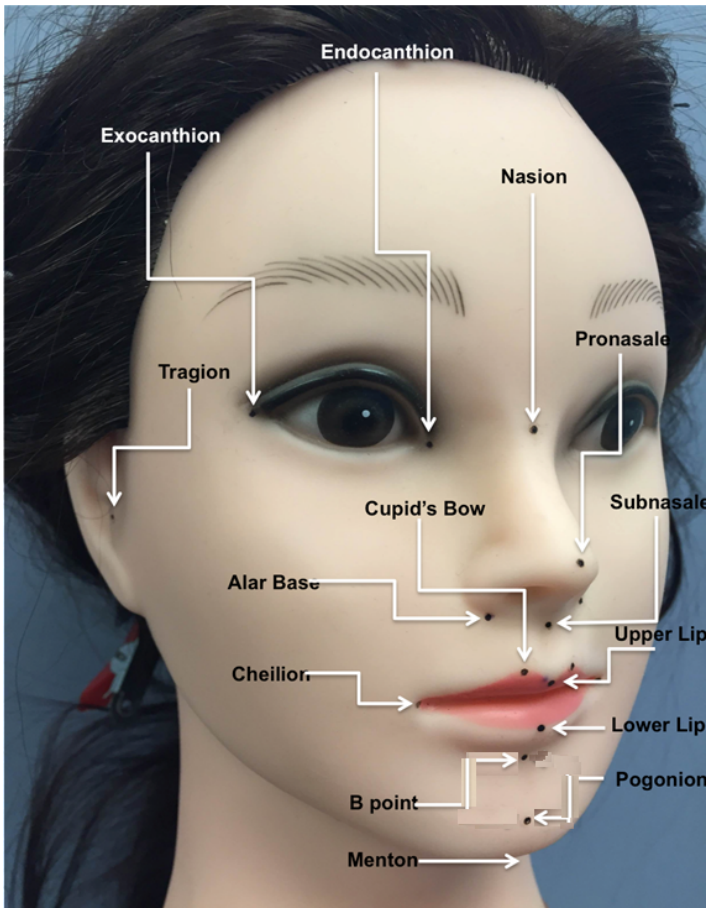

Figure 1. Mannequin with landmarks labeled.

The mannequin head was attached to a mount and secured to an adjustable chair. (Figure 2) Images using 3dMDface system were taken at seven time points (0, 1 hour, 24 hours, 1 week, 2 weeks, 3 weeks, and 4 weeks). Images were saved and imported on 3dMDvultus for 3D analysis.

Figure 2. Adjustable chair with 3dMD photogrammetric system.

Table 1. Landmark definitions

Landmark	Abbreviation	Definition
Endocanthion	En	Inner corner of eye fissure where eyelids meet
Exocanthion	Ex	Outer corner of eye fissure where eyelids meet
Nasion	N	Midpoint of nasofrontal suture
Tragion	Tr	Located at notch above the tragus of the ear
Pronasale	Prn	Most protruded point of nasal tip
Cupid's Bow	CupidB	Point on the crest of philtrum
Subnasale	Sn	Lower border of nasal septum
Alare	Alar B	Most lateral point on nasal ala
Upper lip	Ul	Junction between vermillion and soft tissue of upper lip in midline
Cheilion	Ch	Outer corner of mouth where outer edges of upper and lower vermilions meet
Lower lip	Ll	Junction between vermillion and soft tissue of lower lip in midline
Pogonion	Pg	Most projecting median point on the anterior surface of the chin
B point	B	Most concave point of lower lip between the chin and lower lip point
Menton	Me	Most inferior portion of chin

Table 2. Surface parameters

Vertical (6)	Anteroposterior (10)	Transverse (5)
N_Prn	Tr_N	Inter CD
N_Ul	Tr_Prn	Exo CD
N_Ll	Tr_Sn	Alar B
N_B	Tr_Ul	RtLt_Ch
N_Pg	Tr_R.Cub	RtLt_CupidB
N_Me	Tr_R.Ch	
	Tr_Ll	
	Tr_B	
	Tr_Pg	
	Tr_Me	

Part 2:

Forty healthy subjects with a complete dentition and who were not undergoing any orthodontic treatment (age 24-34; 21 females, 19 males) participated in this study. No selections on specific facial characteristics were made. Subjects with any craniofacial anomalies or history of facial surgery or facial paralysis were excluded from the study. Natural head posture (NHP) was adopted because it has been shown to be clinically reproducible. [23] The subject sat on an adjustable chair and was asked to look into a mirror marked with horizontal and vertical lines. Subject's head was adjusted so that their eyes were level to the horizontal line and the midline of

their face adjusted to the vertical line. This was accomplished to achieve NHP. A cross-line laser mounted on the right side of the wall was used to determine the Frankfort horizontal plane (Figure 3a).

Figure 3a. Mannequin head positioned using cross-line laser as reference

Images using 3dMDface system were captured at seven time points (0, 1 hour, 24 hours, 1 week, 2 weeks, 3 weeks, and 4 weeks). For each session, the patient was asked to adopt one of the four different facial postures (Repose ©, maximum intercuspation (MIP), smile with lips closed (SLC), and Posed smile (S)). (Figure 3b) An image was taken and recorded for each facial posture.

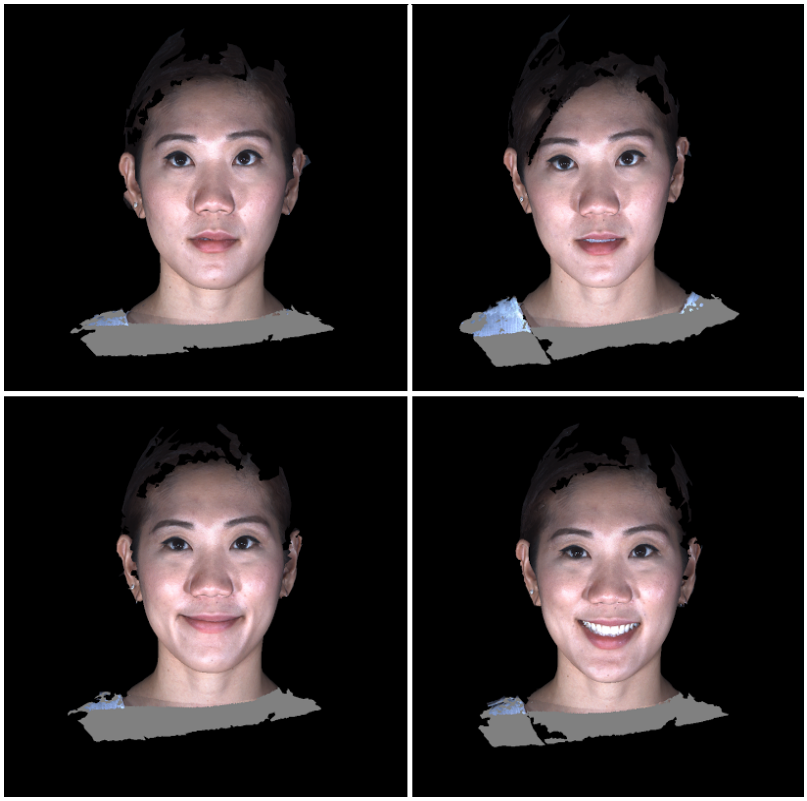

Figure 3b. Four facial postures. (a) MIP, (b) Repose, (c) SLC, (d) S

3D morphometrics

Part 3:

1. Collection of sample 3D face models

Seven images of mannequin, subject in MIP, and subject in posed smile were taken using 3dMDface system.

Figure 4: Demonstration of 3 frontal views taken from 3dMDface system

2. Surface topology correction and spherical mapping

"Topology" refers to the number of handles, islands and boundaries of the surface. When looking at the face, the natural topology is depicted as a simple sheet with a single boundary. Real 3D models of the face may exhibit additional small boundaries due to reconstruction error, but filling these does not reduce the accuracy of the representation, since no original data is altered. However, the method of 3D mesh generation, Delaunay triangulation, often leads to polygon representation of non-manifold topology, which requires a separate set of correction techniques. [24-27]

a. Non-manifold polygon correction

Removing triangles (“faces”) and vertices of non-manifold nature from polygon models is a fairly common problem in 3D modeling. We have expended considerable effort to create our own in-house non-manifold correction tool targeted specifically at the difficult mesh models created by 3dMD.

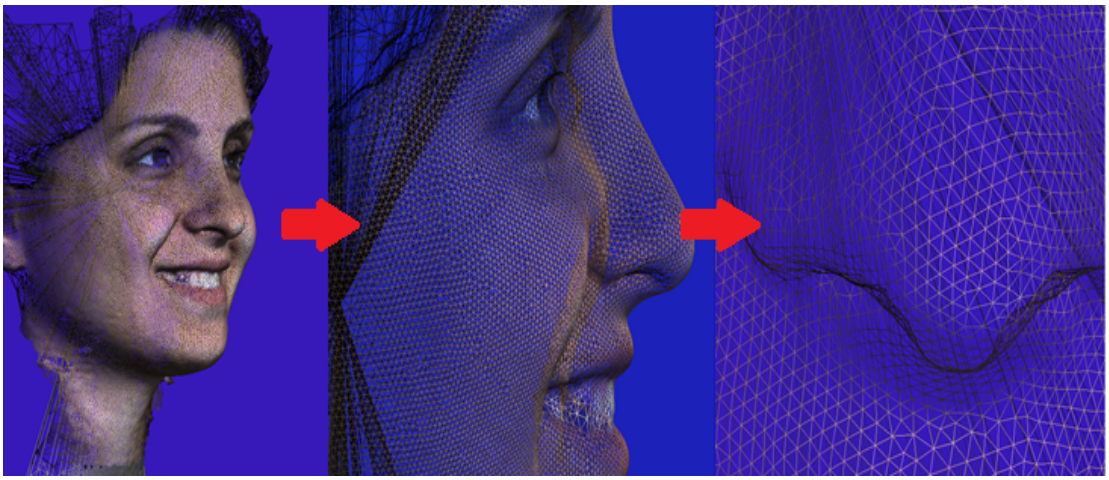

Figure 5: Illustrations of geometry created by 3dMD software zoomed in.

b. Spherical mapping

"Unconstrained spherical parameterization in ACM SIGGRAPH" was used for our spherical mapping because we were able to control the angle and area distortion. [28] Our implementation has been used on a variety of shapes, including cortical, subcortical, skull and face models.

3. Shape registration

The method we followed used the set of faces using both texture and geometry information, while maintaining sufficient flexibility to deal with non-face regions of the model.

Figure 6: (Skull image is adapted from McComb [26]) Images show the mapping process as facial 3dMD images are mapped to a sphere for shape registration using similar methodologies to past research in mapping of the human skull.

a. Initial spherical matching

A set curve-matching algorithm was used. A simple set of 21 curves was manually drawn on each model using the BrainSuite 14 software, taking roughly 5 minutes per model by a trained operator. The correspondence between landmark points on the facial surface was determined automatically via the arclength map. [29]

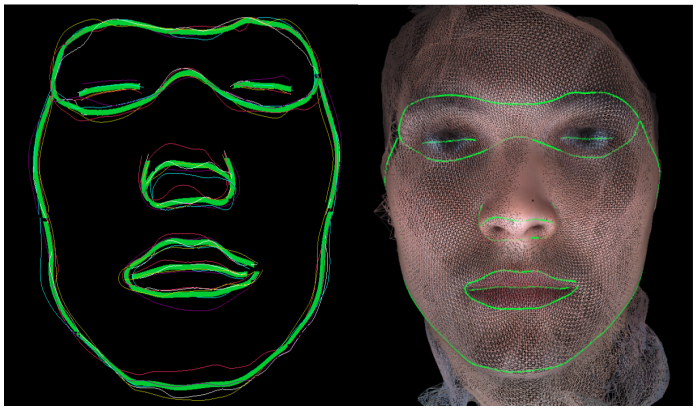

Figure 7: Illustration of the curves using curve-matching algorithm.

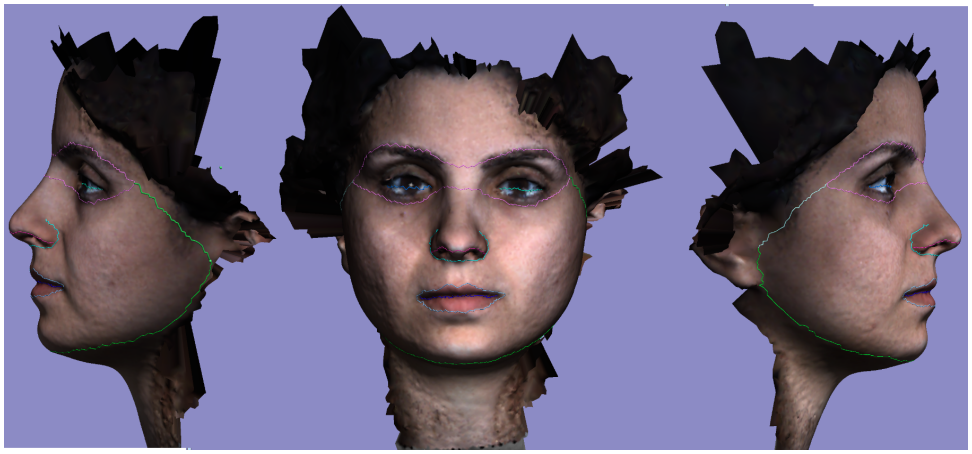

Figure 8: BrainSuite 14 surface curve manual tracing

b. Texture matching

A multi-channel Mutual Information criteria was used to minimize the mismatch often seen between texture maps. This option is employed to match 2D and 3D images from different modalities. This choice is motivated by the fact that facial texture correspondence is characterized by complex relationships between intensities of different color channels, without a straight-forward transfer function.

4. Average and distance map creation

a. Procrustes alignment/Tensor based morphometry

When computing the average face, the shape models in their original (not parametric) space have to be aligned based on the computed dense correspondence. The 7-parameter Procrustes method was used. [30]

b. Average and distance map creation, shape statistics analysis

3dMDvultus image analysis

Superimposition

Images acquired from the 3dMDface system were saved as .obj files and imported onto the 3dMDvultus software system. Before using the automatic registering feature of the software, each image was manually aligned, adjusting the rotation of the image using the XYZ rotational coordinates (Figure 9). Once the images were adequately positioned, the Registry tab was used to fine-tune and finalize the registration of the superimposed images. As a general rule, an RMS error of less than 0.5mm is acceptable. [31] The RMS error is the variation between the two surfaces.

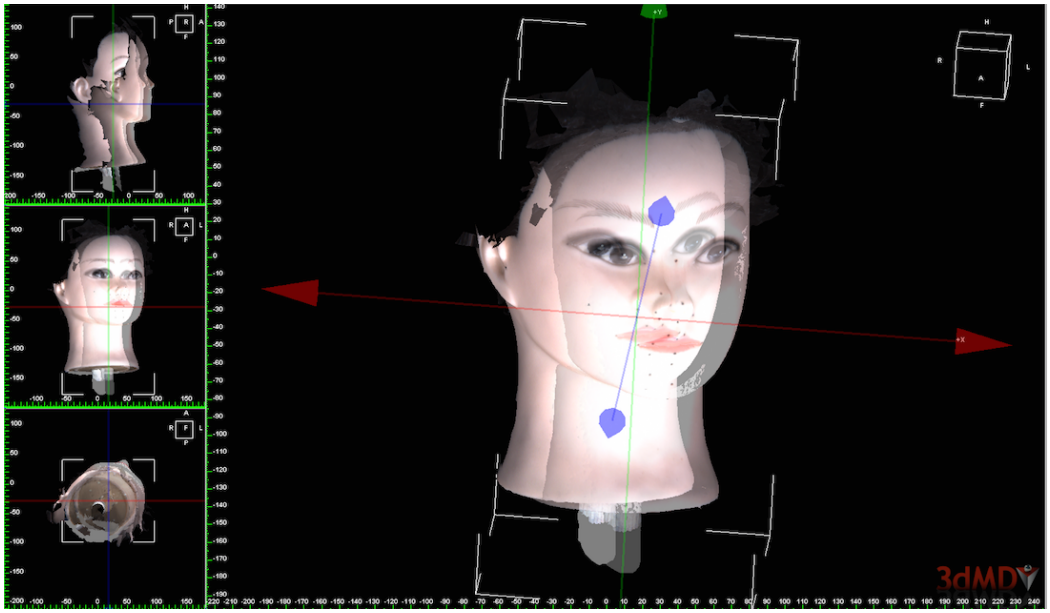

Figure 9. Superimposition of images using XYZ directionality feature.

Landmarking and analysis of multiple 3dMD images

A custom landmarking template was created that contains all 21 landmarks of interest. Each landmark was placed using a mouse on the superimposed 3D image using the landmark tab. The 3dMDvultus software automatically records the XYZ position coordinates for the placed landmark. The landmark itself displays as a green dot on the 3dMD image. (Figure 10)

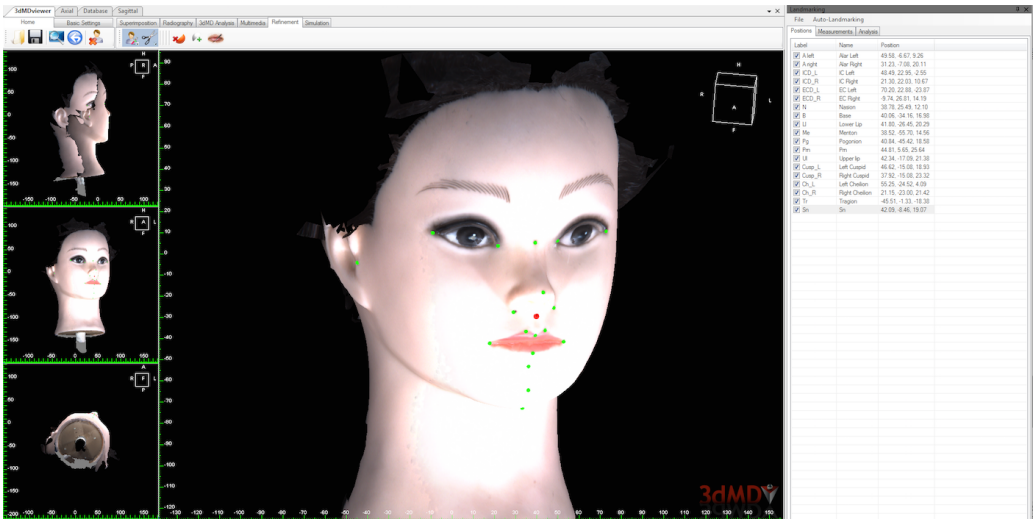

Figure 10. Surface landmarks placed on mannequin face. XYZ coordinates automatically recorded by 3dMDvultus software.

An analysis template was then generated to look at caliper and surface measurements. 3dMDvultus uses the XYZ coordinates of two points to find the straight-line caliper distance. Topographical measurements were also taken between the two points on the surface of the 3dMD image using the shortest path along the contour of the surface. (Figure 11)

Figure 11. Surface measurement of Tr_Rt.CupidB displayed

A comprehensive report including the landmark XYZ positions and surface measurements were saved as an XLS file (.xls). An intraobserver study was done with two observers measuring and recording all values individually.

Quantitative analysis

Accuracy is defined as the degree to which a measurement deviates from its "true" value. For this study, the accuracy was calculated by comparing the "true" values obtained from direct digital caliper and measuring string measurements to the measurements obtained from 3dMDvultus.

The true value for each of the 21 parameters were obtained by averaging the values recorded between observer 1 and observer 2. The measured mean was obtained averaging the values of the seven time points recorded using 3dMDvultus across all 40 subjects. Accuracy

estimates were accomplished by calculating several error magnitude statistics: mean absolute difference (MAD), standard deviation of the error (error SD), Root Mean Square Error (RMSE), relative error magnitude (REM), technical error magnitude (TEM), and intraclass correlation coefficient (ICC).

The MAD values use the same units as the original measurement data, which makes it an intuitive method of measuring error magnitude. RMSE is a good way to incorporate both MAD and error SD. REM can be calculated by dividing the MAD for a given parameter by the measured mean and multiplying the result by 100. Thus, REM is expressed as a percentage and represents an estimate of error magnitude relative to the size of the measurement. The REM is helpful to include because it allows us to normalize each value for ease of comparison. For example, an MAD value of 5 mm for a mean measurement of 100 mm gives an REM value of 5%. On the other hand, the same MAD value of 5 mm for a mean measurement of 10 mm gives an REM value of 50%.

In addition, following Weinberg et al. (2004), REM scores were divided into 5 categories: < 1% = excellent, 1% to 3.9% = very good, 4% to 6.9% = good, 7% to 9.9% = moderate and > 10% = poor. [2]

RMSE is another frequently used measure of difference between values predicted by a model and values actually observed. The RMSE is useful because the value reports both the error SD and the mean difference. Therefore, parameters that have a large RMSE but gave a small error SD means that the average difference between the true value and measured value was large. In other words, there was a lot of variation between the true and measured value, making that particular parameter unreliable.

$$RMSE=\sqrt{\text{error } SD^2+\text{mean difference}^2}$$

TEM is also commonly included to report error magnitude. TEM "provides a standard deviation-like measure of the magnitude of error, and it is in the original units of measurement." TEM was calculated by comparing measurements between two observers. [32] The formula for TEM is:

$$TEM= (\sum D^2)/2N.$$

D represents the difference in measurement between observer A and observer B and N represents the sample size of 40 individuals x 7 time points = 280. ICC values were looked at to see any error estimates between observers. This statistic is commonly used in intraobserver studies. ICC values range from 0 to 1 and represents intraobserver error over overall error. A value closer to 0 indicates more measurement error between observers while a value closer to 1 indicates fewer measurement error. [33]

Results

Part 1: Accuracy of caliper and surface measurements using mannequin

Caliper measurements

Out of the 21 inter-landmark caliper distances measured, 19 measurements showed an MAD value of less than 1mm with the exception of Tr_Prn (1.35mm) and Tr_Ul (1.59mm). The overall TEM mean was 0.495mm. The minimum TEM was 0.025mm and maximum TEM was 1.682mm, which gives us a range of 1.657mm. REM scores were categorized as excellent (error magnitude less than 1% of the mean), very good (1%-3.9% of mean), good (4%-6.9% of mean). 12 of the 21 were deemed excellent and the other 9 of 21 were deemed very good. Table 3 organizes the grand mean statistics for caliper measurements using 3dMDface system.

Surface measurements

Out of the 21 inter-landmark surface distances measured, 16 measurements showed an MAD value of less than 1mm. N_Pg (1.64mm), ExoCD (2.11mm), Tr_Pg (2.39mm), N_Me (3.44mm), and Tr_Me (4.09mm) had MAD values larger than 1mm. The overall TEM mean was 0.683mm. The minimum TEM was 0.050mm and maximum TEM was 2.205mm, which gives a range of 2.155mm. REM scores were categorized as excellent (error magnitude less than 1% of the mean), very good (1%-3.9% of mean), good (4%-6.9% of mean). 7/21 measurements were deemed excellent, 13/21 measurements were deemed very good, and 1/21 measurements were deemed good. Table 4 displays the grand mean statistics for surface measurements using

3dMDface system. Table 5 categorizes the REM values between caliper and surface measurements.

Table 3. Error statistics for caliper measurement using 3dMDface system

	Overall Mean	Error SD	Minimum	Maximum	Range
MAD (mm)	0.359	0.035	0.010	1.514	1.504
REM (%)	0.978	0.926	0	3.785	3.785
TEM (mm)	0.495	0.036	0.025	1.682	1.657

Table 4. Error statistics for surface measurement using 3dMDface system

	Overall Mean	Error SD	Minimum	Maximum	Range
MAD (mm)	0.733	0.352	.0124	2.014	2.002
REM (%)	1.709	1.344	0.260	6.088	5.828
TEM (mm)	0.683	0.352	0.050	2.205	2.155

Table 5. REM values categorized out of 21 parameters

Categories	Caliper	Surface
Excellent	12	7
Very Good	9	13
Good	0	1
Moderate	0	0
Poor	0	0

Table 6. Comparison of error statistics for caliper versus surface measurements

	Caliper	Surface
MAD (mm)	0.359	0.733
REM (%)	0.978	1.709
TEM (mm)	0.495	0.683

Root Mean Square Error (RMSE) values were obtained and organized as increasing order for both surface and caliper measurements. (Figure 12) N_Me and Tr_Me, two of the longest surface measurements of the face, had the largest variability.

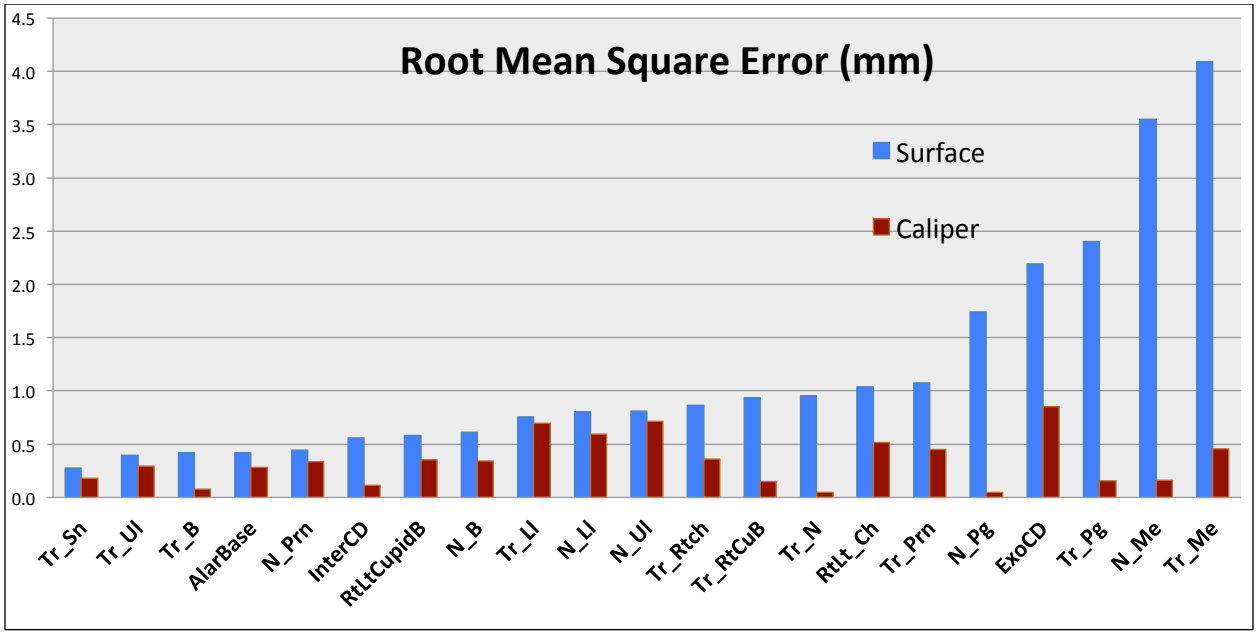

Figure 12: RMSE values showed that caliper measurements were more accurate than surface measurements across all 21 parameters. Surface measurements using parameters involving menton (N_Me and Tr_Me) resulted in the greatest deviation from the true value suggesting the least accuracy.

Part 2: Reproducibility

The overall average MAD for repose (R) was 2.32mm, maximum intercuspation (MIP) was 2.03mm, smile with lips closed (SLC) was 2.31mm, and posed smile (S) was the highest with 2.79mm. Looking at overall average REM, MIP was 0.948, R was 0.961, SLC was 0.932, S was 0.924. MIP was 1.7%, SLC was 1.9%, and S was 2.2%. For overall average ICC, R was

0.961, MIP was 0.948, SLC was 0.932, S was 0.924. Table 7 organizes each error statistic for the four different facial postures.

Table 7. Error statistics for surface measurement across 4 facial postures using 3dMDface system

	R	MIP	SLC	S
MAD (mm)	2.315	2.032	2.314	2.794
REM (%)	1.949	1.735	1.950	2.215
TEM (mm)	2.145	1.598	2.554	2.687
ICC	0.961	0.948	0.932	0.924

Although comparing error statistics across the four facial postures was the main objective of part 2 of this study, we were also interested in looking at which of the 21 facial parameters showed the highest variability. (Figure 13, 14, 15).

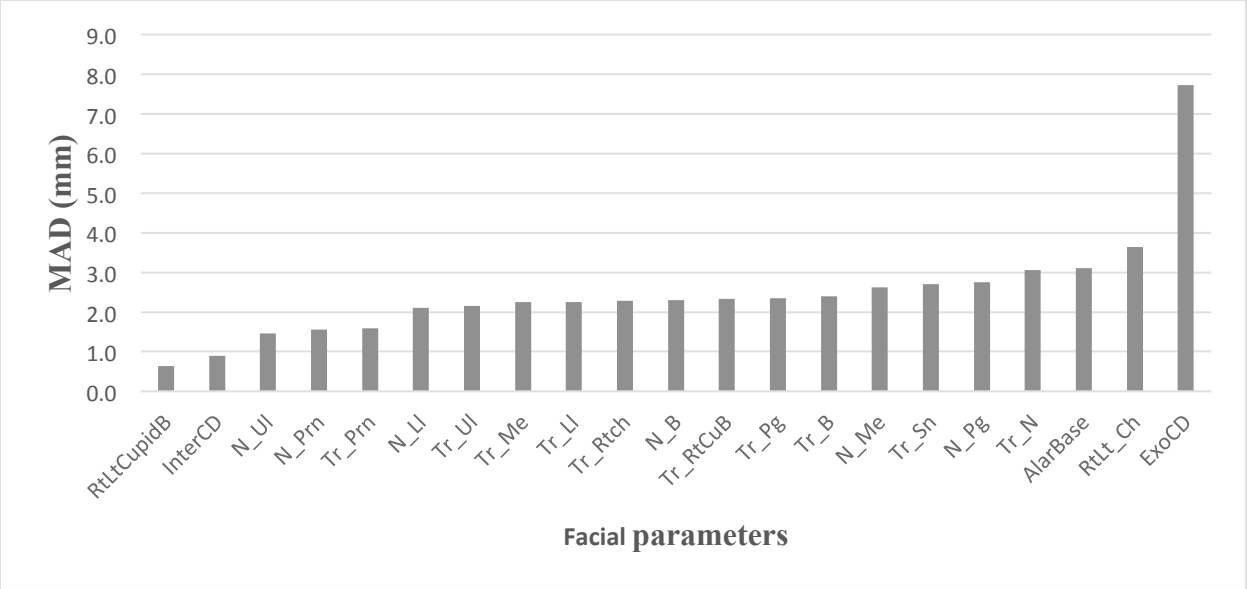

Figure 13: MAD values in increasing order across all 21 parameters. Note AlarBase, RtLt_Ch, and ExoCD showed greatest variation.

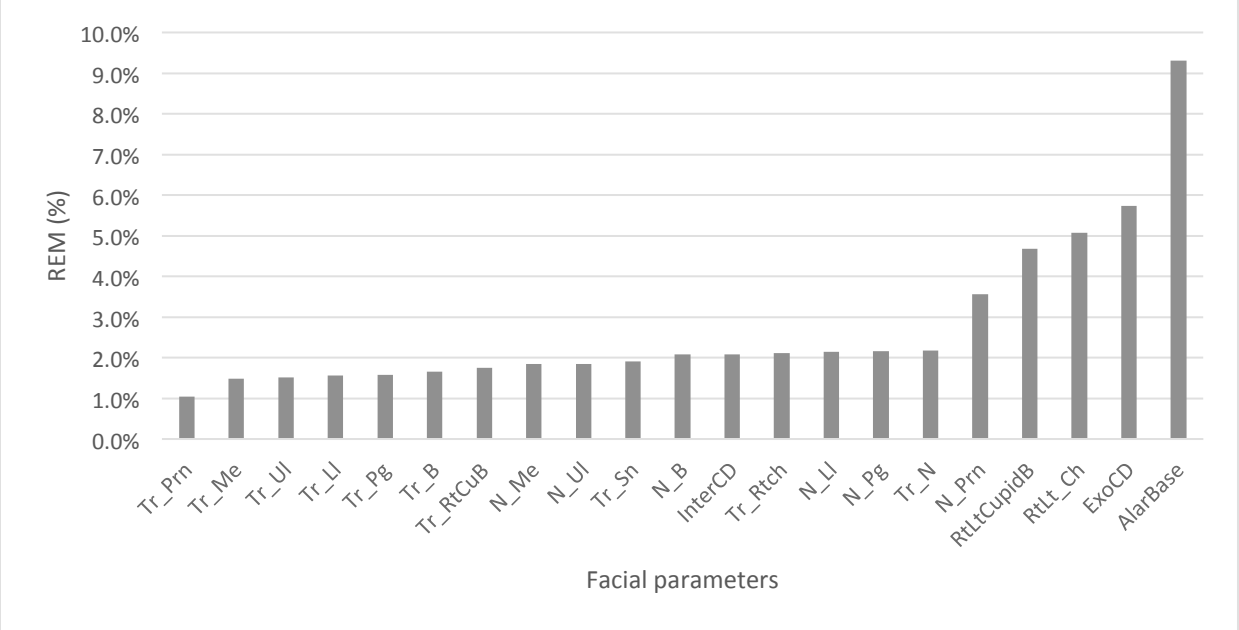

Figure 14: REM values in increasing order across all 21 parameters.

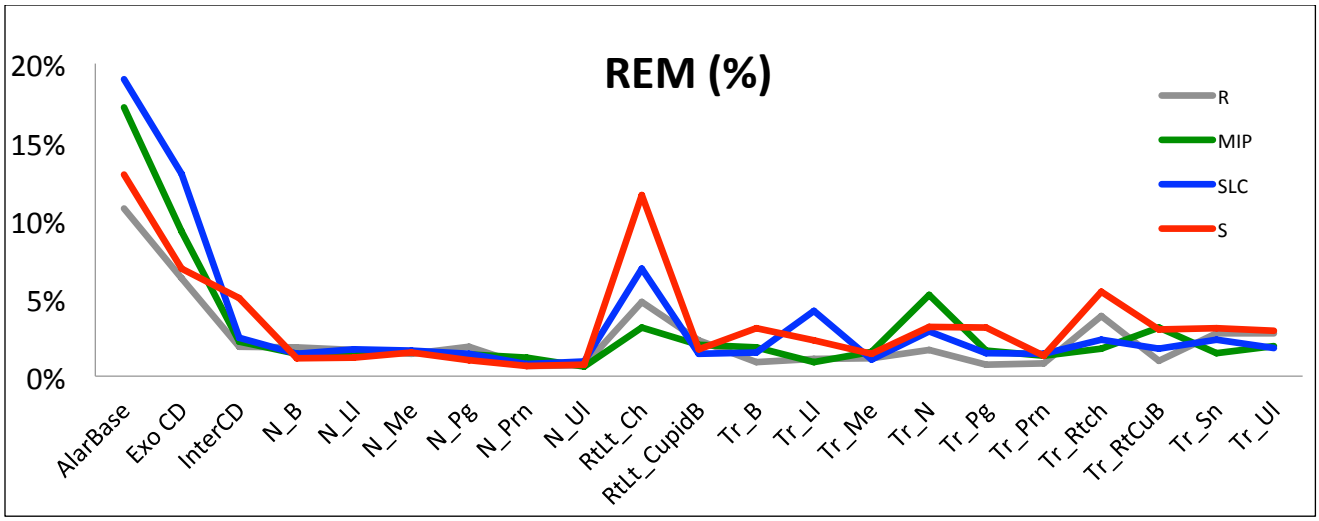

Figure 15: 3dMD surface measurements using parameters, Inter Alar Base, ExoCD, and RtLt_Ch, were least reproducible across all four facial postures.

MAD, REM, and ICC values were then compared when the three most variable surface measurements (Alar Base, RtLt_Ch, and ExoCD) were removed. Once these surface measurements were removed, MAD, REM, TEM, and ICC all improved for all facial postures.

Table 9. Error statistics for surface measurement across 4 facial postures after removing AlarBase, RtLt_Ch, and ExoCD

	R	MIP	SLC	S
MAD (mm)	2.225	1.995	2.199	2.762
REM (%)	1.856	1.663	1.877	2.125
TEM (mm)	2.145	1.889	2.087	2.519
ICC	0.961	0.960	0.940	0.948

Part 3: 3D morphometrics

The 7 (mannequin, MIP, and repose smile) 3dMD photographic images after topological correction were loaded into the intial software pipeline to generate the closed mesh forms of the images (Figure 16). These closed meshes were then traced for surface curves (Figure 17). The results of this dual pipeline process is a clean average of the 7 mannequin, 7 MIP, and 7 repose smile faces after topology correction, closed mesh creation, and shape registration.

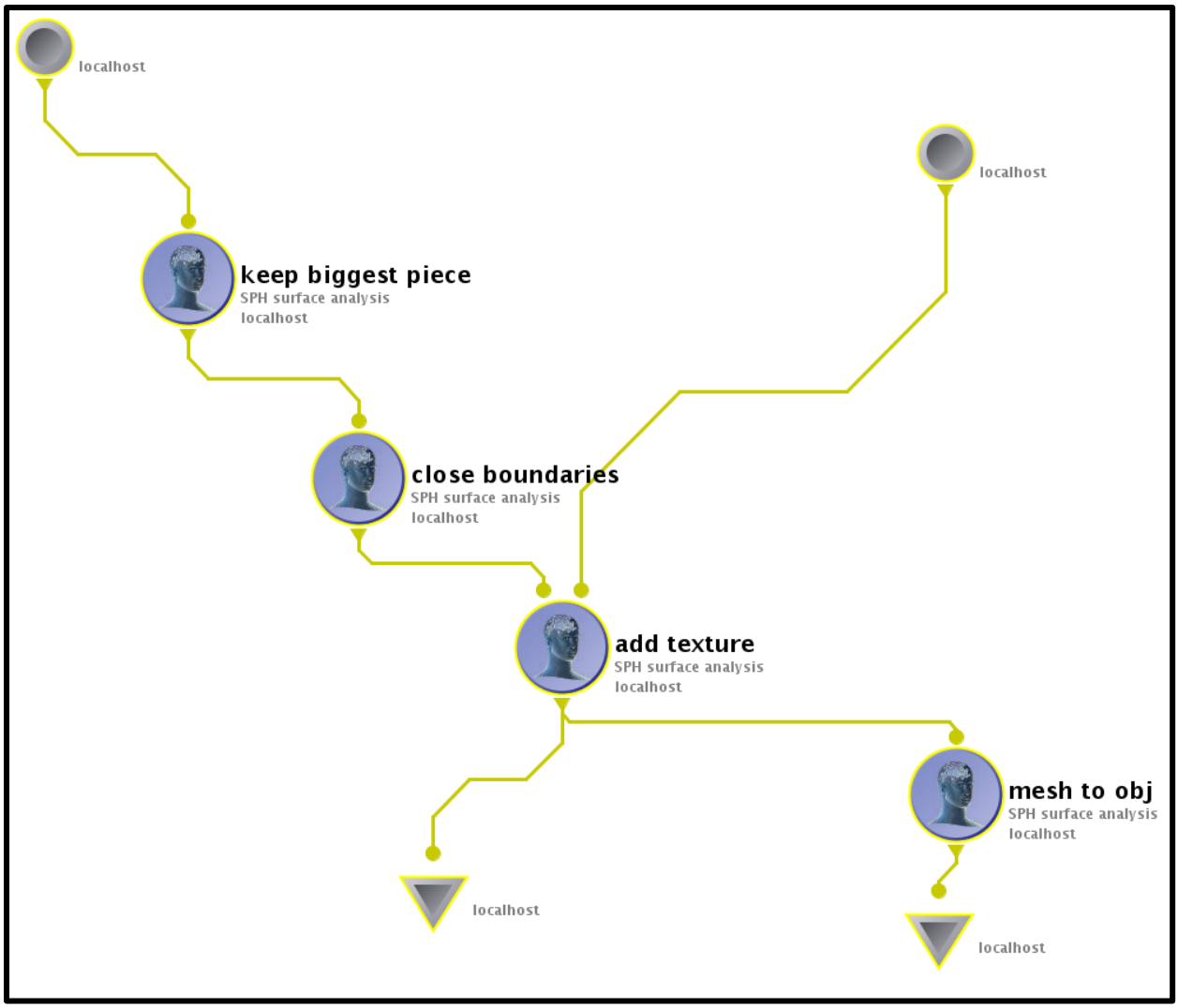

Figure 16: Snapshot of the pipeline used to create the closed mesh versions of each sample. Each surface shape raw 3dMD image was input into the pipeline to generate closed mesh geometries.

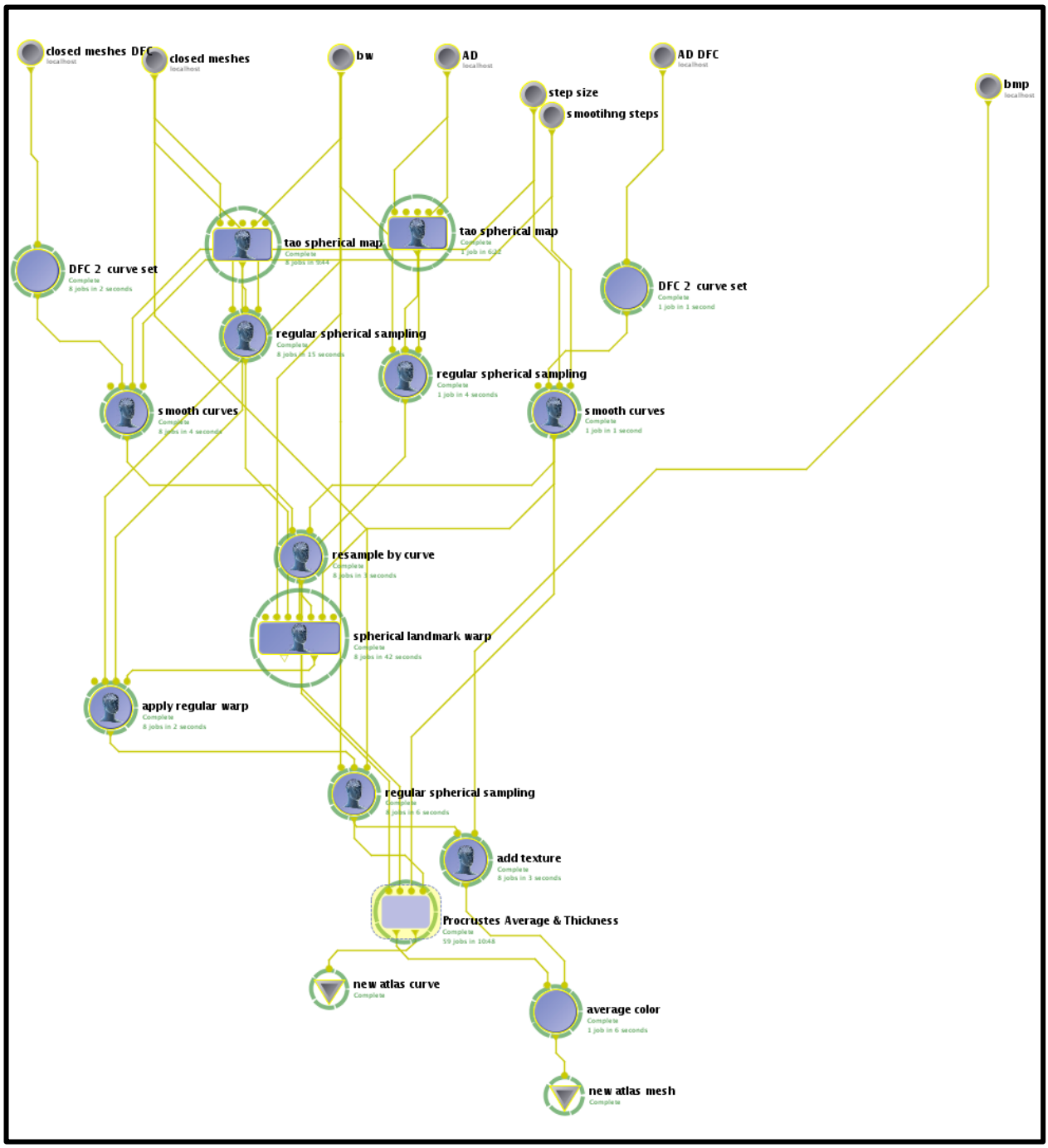

Figure 17: Snapshot of the pipeline used to create the normative average from each the closed mesh sample.

The mannequin figure below (Figure 18) illustrates the bias and reliability of the displacement measure based on seven independent acquisitions of the mannequin 3D model. The seven samples of the mannequin are brought into point-wise correspondence using landmark-based spherical registration. An average 3D model and landmark set are then constructed, and all seven models are registered once again to the unbiased average. The statistical distribution of the displacement measurements from each of the samples to the average is illustrated on the right, as a color-map mapped onto the mean shape. The color-map on the left shows the t-statistic testing the null hypothesis that the mean displacement among the seven samples is different than zero. The figure on the right shows the standard deviation, which indicates how reliable an individual measurement is expected to be.

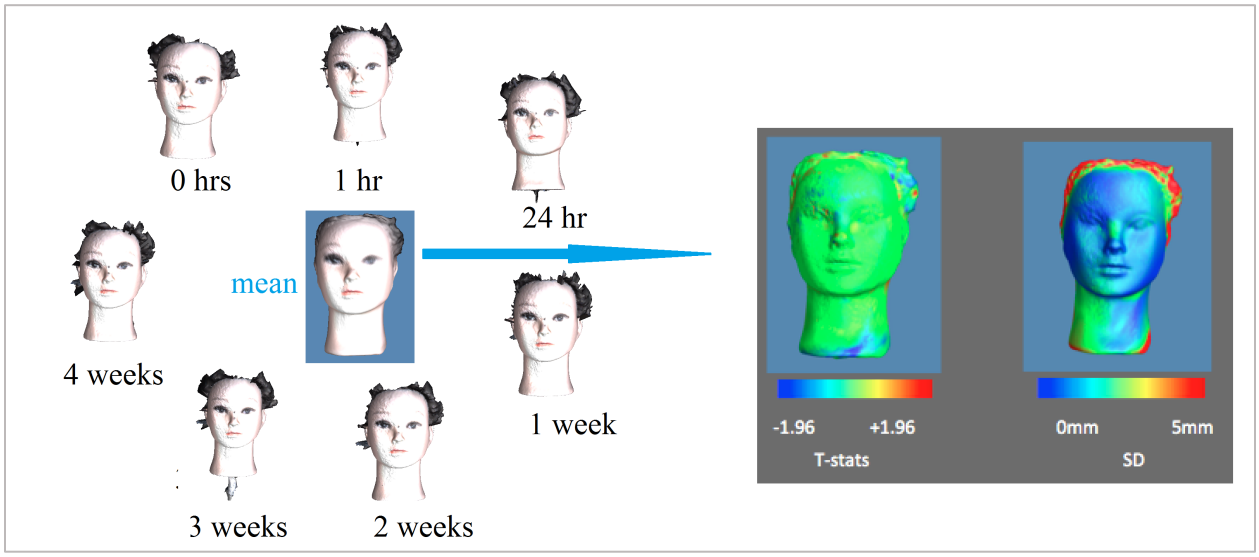

Figure 18: Average of 7 time points (T0, T1, T2, T3, T4, T5, T6) for mannequin head superimposed over the average. Demonstration of colorized distance mapping used to display results of averaging pipeline.

The mannequin results showed that the t-stats are near zero on the face (green face = 0, range bluest = -1.96, reddest = 1.96), and the standard deviations are sub-millimeter (blue face, range is bluest = 0, reddest = 5mm)

The results of averaging 7 time points for one individual in MIP showed that there are positive change and negative changes around the left exocanthal region. Regions in green shows no change from the average. The standard deviation was in sub-millimeters in areas around the face, which is represented in blue. (Figure 19)

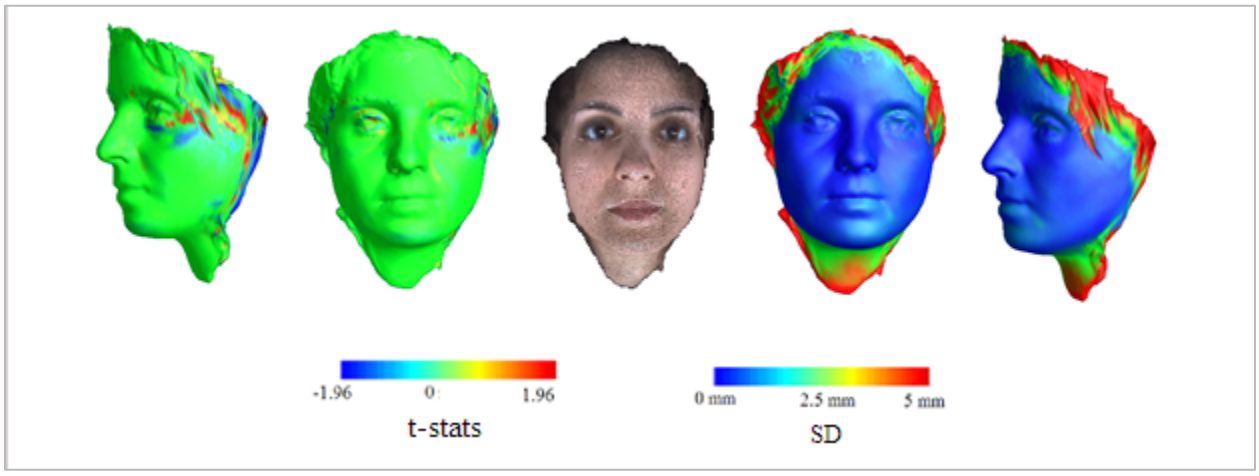

Figure 19: Average of 7 time points (T0, T1, T2, T3, T4, T5, T6) for one individual in MIP superimposed over the average. Demonstration of colorized distance mapping used to display results of averaging pipeline.

The results of averaging 7 time points for one individual in posed smile showed that there are positive changes around the right cheilion, left alar base, left endocanthion, which is represented in red. Pronasale and left side of the face near the tragion showed negative changes, which are represented in blue. Regions in green shows no change from the average. The standard deviation was in sub-millimeters around the face (represented in blue) except the regions around pronasale, cheilion, and the eyes (represented in green), which shows slight variation. (Figure 20)

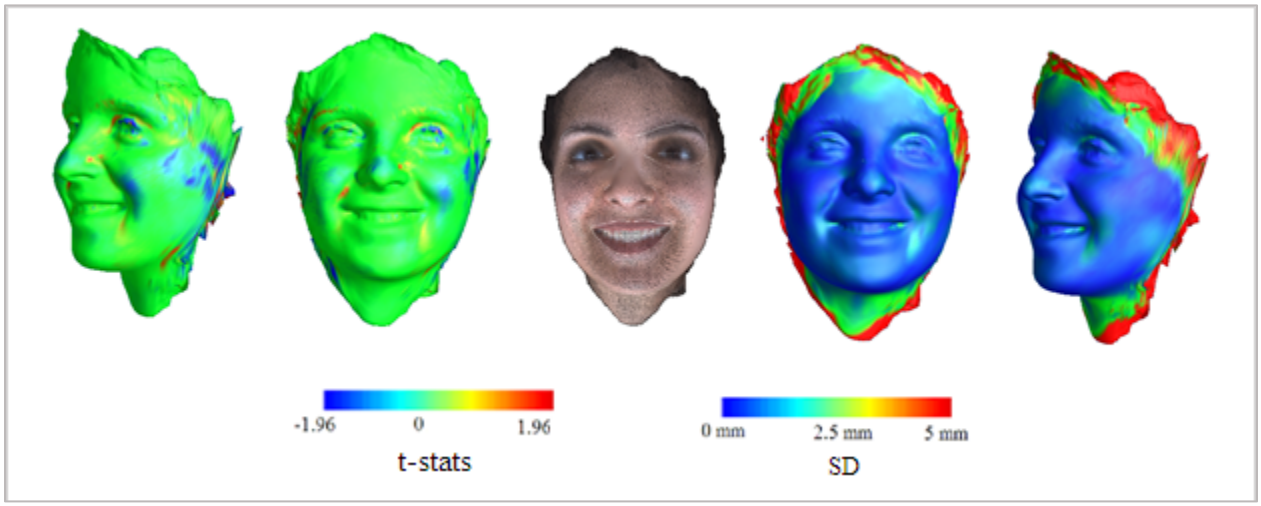

Figure 20: Average of 7 time points (T0, T1, T2, T3, T4, T5, T6) for one individual in posed smile superimposed over the average. Demonstration of colorized distance mapping used to display results of averaging pipeline.

Discussion

The objectives of this study consisted of three parts: (1) investigate the accuracy of 3dMDface system (2) find standard facial expressions that can be best reproduced on 3dMD system (3) verify accuracy using 3D morphometrics.

With regards to part 1 of this study, the 3dMDface stereophotogrammetric system was shown to be highly accurate. Similar results were found in previous studies by Gornick, de Menezes et al, and Ort et al; however, these studies only looked at caliper measurements. Our study decided to take a step further and look at not only caliper measurements but also surface measurements. Paired t-tests between the true values and 3dMD measured values for all parameters showed there was no statistical difference for caliper and surface measurements. However, caliper measurements had an overall MAD mean and overall REM mean that was lower than that of surface measurements, which signifies that the 3dMDface system provided greater accuracy with caliper measurements. All 21 caliper measurements recorded MAD values of <1 mm while 16 surface measurements recorded MAD values of <1 mm. There are certain possibilities that can affect the surface trajectory of the face when doing surface measurements. For example, the three longest parameter distances (Tr_Pg, N_Me, and Tr_Me) showed the largest variation compared to shorter distance parameters. Areas of great curvature like areas around the eyes (ExoCD) showed greater variation compared to areas of the face that are flatter like in Tr_Sn, Tr_Ul, Tr_B, N_Prn. Therefore, precaution needs to be made when quantifying long distances or curved surfaces. In order to overcome this problem, 3D morphometrics was utilized to further investigate the accuracy of 3dMDface system. Our results using 3D morphometrics showed that there was no difference when average mannequin sample was

compared with seven different time points. This means that when we quantify a 3D structure, it is more appropriate to use true 3D analysis rather than using 2D on a 3D surface, especially for long or curved distances.

In part 2 of this study, MIP (MAD=2.01mm) was the most reproducible facial posture, followed by R (MAD=2.28mm), SLC (2.31mm), and S (2.79mm). Previous studies by Sawyer et al and Popat et al [10, 14] verified reproducibility of facial expressions; however, their methodology had several limitations. Both studies had small sample sizes and their methodology of quantification was unclear. To overcome these problems associated with previous studies, we included a larger sample size of 40 human subjects and took 3dMD pictures at seven different time points. To verify our results further, two observers performed all of the measurements separately.

The advantage of MIP compared to other facial expressions is that subjects can be instructed to swallow and keep their teeth in full occlusion and lips together so that when taking 3dMD images, it is easier for subjects to reproduce this expression every time. Other facial expressions like posed smile showed less reproducibility. One explanation can be that subjects are in different emotional states when taking a picture, which can alter their natural expression. [34] When we applied 3D morphometrics to compare MIP and posed smile, similar areas like regions of the pronasale, exocanthion, and cheilion showed variation. This verifies our findings when comparing 3dMD surface measurements and 3D morphometrics.

When looking at each of the 21 parameters, Alar base (MAD=3.11mm), RtLt_Ch (MAD=3.64mm), and ExoCD (MAD-7.73mm) showed the greatest amount of variability. (Figure 16)

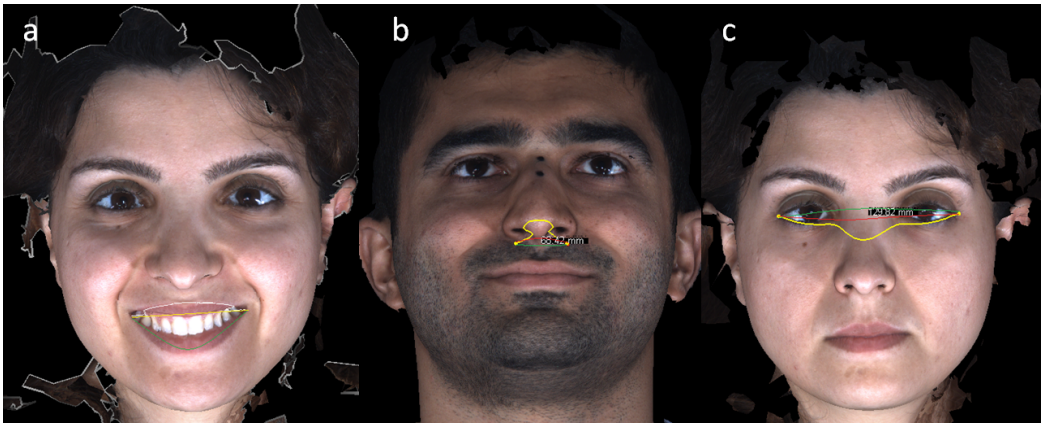

Figure 16: Note the differences in surface trajectory taken for each of the 3 parameters. (a) RtLt_Ch parameter showed variations in defining contours of teeth in S posture. (b) Alar Base parameter followed curvature of nose in SLC and S instead of surface measurement below the nose. (c) Exo CD parameter produced discrepancies when trajectory crossed over the eye in all postures (R, MIP, SLC, and S).

With AlarBase, the 3dMDface system had issues following the same trajectory when the patient is in SLC or S. With patients who have wide smiles or have long noses, the 3dMDface system failed to generate the shortest distance and would often follow the outer contour of the nose. (Figure 16b) Therefore, AlarBase is an unreliable parameter to provide reproducible surface measurements. RtLt_Ch is known to have high variability since lip posture can vary greatly. [10, 14] With ExoCD, the 3dMDface system had the biggest issue providing consistent trajectories. A couple theories can explain this discrepancy. The orientation of the pupil can cause deviations of the surface trajectory. [35] Another explanation might be that the 3dMDface system has difficulty going over surfaces that reflect light. When an image is captured of the patient, light can get deflected off the eyes when the camera flashes. When AlarBase, RtLt_Ch, and ExoCD were removed, overall MAD mean and overall REM mean improved greatly across all four facial postures.

Conclusion

All 14 landmarks and 21 parameters produced accurate and reproducible measurements for caliper distance calculations using 3dMDface system and 3dMDvultus software. In human subjects, surface distance measurements by 3dMDface system produced reproducible soft tissue measurements across all four postures, excluding the following parameters: alar base, right to left cheilion, and exocanthal distance. Overall, posed smile facial posture (S) was found to be least reproducible for 3dMD surface measurements.

Limitations

In this study we used true 3D images of the mannequin and human subjects; however, the quantification in this study was done by using 2D linear distances (caliper and surface). Another limitation in this study is that facial expressions have been closely linked to emotion, it could be considered that any facial expression could vary depending on the subject's frame of mind at a particular point in time. This has important implications in assessing facial movement after a clinical intervention because the facial expression could vary depending on whether the intervention was successful or not.

Further investigation

As completion of this project is realized, the next step is to aim our focus towards 3D morphometrics and increase the sample size to further validate the most reproducible facial expression. Since posed smile facial posture was found to be least reproducible, a standardized protocol (like providing them a list of questionnaires) to assess a patient's mood can be beneficial. Verbal or nonverbal expressions can also be investigated to see whether saying a

certain word affects reproducibility of facial expression. In a previous study, clinicians found that when patients used verbal expressions like “puppy” and “baby” before taking 3dMD imaging showed greater reproducibility than those who did not use any verbal gestures. [14]

Males on average showed fewer reproducibility of facial expression than females. [14] One explanation for that discrepancy was differences in soft tissue thickness. Therefore, further investigation can be done that quantifies soft tissue thickness between genders and see whether that plays a role in reproducibility of facial expressions.

Timing of when a subject takes a 3dMD image might also play a role in reproducibility. In our study, we spread out each of the seven time points for a month. However, according to previous literature, facial expressions were shown to be highly reproducible within a 10 to 20 minute period. [10] In a future study, shortening the seven time points to 10 minutes apart might show better reproducibility among the different facial postures.

Based on the results of our study, more focus should be concentrated on the nose and forehead rather than lips and eyes as those are areas prone to change. As a last step, the whole concept should be transferred into four-dimensional representation, which means the three-dimensional capturing not only of a still image but also of a moving object using a video camera.

Clinical Implications:

When doing a clinical examination on a patient, it is imperative that accurate profile and frontal views are recorded for future references. One of the advantages of soft tissue cephalometrics is that it provides the ability to make objective measurements of important structures and relationships like quantifying facial disharmony and identifying its underlying causes. This is

exceedingly important because, as a rule, better facial esthetics are achieved if the underlying problems are identified and treated at the source [36].

There are numerous clinical implications for accurately predicting soft tissue responses to hard tissue changes. The orthodontist must understand soft tissue behavior in relation to orthopedic, orthodontic, and orthognathic changes and must also take into consideration growth and development of soft tissue traits. When treating a malocclusion nonextraction, the nose, lower face height, lip length, lip protrusion, upper incisor, interlabial gap, mandibular sulcus can all be affected. If lip posture is pushed too far forward, the result may be a masking of the chin, an increase in interlabial gap, or reduced lower face height. Treatment has to be balanced with the position of the teeth in the bone to support periodontal health and long term stability. Extraction of teeth can cause several changes. It can flatten profile, increase nasolabial angle, increase lip length, decrease lip protrusion, decrease upper incisor exposure, and increase mandibular sulcus. When there is a large nose or chin, caution should be used with respect to retracting the lips. In cases where surgery is out of the question, greater facial compensations may be necessary. Any of these changes can greater alter facial traits and may compromise facial esthetics that a patient may desire after treatment. Accurate representation of soft tissue changes is essential. Therefore, soft tissue analysis using 3D imaging using 3dMDface system can have major clinical implications towards predictable treatment outcome and patient satisfaction [37].

References:

1. Aldridge, K., et al., *Precision and error of three-dimensional phenotypic measures acquired from 3dMD photogrammetric images.* Am J Med Genet A, 2005. **138A**(3): p. 247-53.

2. <*Gornick.pdf*>.

3. Karatas, O.H. and E. Toy, *Three-dimensional imaging techniques: A literature review.* Eur J Dent, 2014. **8**(1): p. 132-40.

4. Lane, C. and W. Harrell, Jr., *Completing the 3-dimensional picture.* Am J Orthod Dentofacial Orthop, 2008. **133**(4): p. 612-20.

5. <*anthropometric measurement error.pdf*>.

6. Hajeer, M.Y., et al., *Applications of 3D imaging in orthodontics: part I.* J Orthod, 2004. **31**(1): p. 62-70.

7. Jokic, D., et al., *Soft tissue changes after mandibular setback and bimaxillary surgery in Class III patients.* Angle Orthod, 2013. **83**(5): p. 817-23.

8. Altug-Atac AT, G.B., McCarthy JG, *Comparison of skeletal and soft tissue changes following unilateral mandibular distraction osteogenesis.* Plastic Reconstruction Surgery, 2008(121:1751).

9. <*Nguyen_et_al-2000-Clinical_Orthodontics_and_Research.pdf*>.

10. Sawyer, A.R., M. See, and C. Nduka, *Assessment of the reproducibility of facial expressions with 3-D stereophotogrammetry.* Otolaryngol Head Neck Surg, 2009. **140**(1): p. 76-81.

11. Gor, T., et al., *Three-dimensional comparison of facial morphology in white populations in Budapest, Hungary, and Houston, Texas.* Am J Orthod Dentofacial Orthop, 2010. **137**(3): p. 424-32.

12. Fourie, Z., et al., *Evaluation of anthropometric accuracy and reliability using different three-dimensional scanning systems.* Forensic Sci Int, 2011. **207**(1-3): p. 127-34.

13. Solem, R.C., et al., *Three-dimensional soft-tissue and hard-tissue changes in the treatment of bimaxillary protrusion.* Am J Orthod Dentofacial Orthop, 2013. **144**(2): p. 218-28.

14. Popat, H., et al., *A comparison of the reproducibility of verbal and nonverbal facial gestures using three-dimensional motion analysis.* Otolaryngol Head Neck Surg, 2010. **142**(6): p. 867-72.

15. *<validty and reliability of craniofacial anthropometric.pdf>*.

16. Shah, N., N. Bansal, and A. Logani, *Recent advances in imaging technologies in dentistry.* World J Radiol, 2014. **6**(10): p. 794-807.

17. Papadopoulos, M.A., et al., *Three-dimensional craniofacial reconstruction imaging.* Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontology, 2002. **93**(4): p. 382-393.

18. Ort, R., et al., *The Reliability of a Three-Dimensional Photo System- (3dMDface-) Based Evaluation of the Face in Cleft Lip Infants.* Plast Surg Int, 2012. **2012**: p. 138090.

19. Maal, T.J., et al., *Variation of the face in rest using 3D stereophotogrammetry.* Int J Oral Maxillofac Surg, 2011. **40**(11): p. 1252-7.

20. Honrado C, L.S., Bloomquist D, Larrabee W., *Quantitative assessment of nasal changes after maxillomandibular surgery using 3-dimensional digital imaging system.* Arch Facial PlastSurg, 2006: p. 26-35.

21. Heike, C.L., et al., *3D digital stereophotogrammetry: a practical guide to facial image acquisition.* Head Face Med, 2010. **6**: p. 18.

22. de Menezes, M., et al., *Accuracy and reproducibility of a 3-dimensional stereophotogrammetric imaging system.* J Oral Maxillofac Surg, 2010. **68**(9): p. 2129-35.

23. Ovsenik, M., et al., *Three-dimensional assessment of facial asymmetry among pre-pubertal class III subjects: a controlled study.* Eur J Orthod, 2014. **36**(4): p. 431-5.

24. al, P.T.a.e., *"Mapping cortical change in Alzheimer's disease, brain development, and schizophrenia.,"* Neuroimage, 2004. **23**(Suppl 1): p. S2-18.

25. al, Y.S.a.e., *Inverse-consistent surface mapping with Lapace-Beltrami eigen-features.* Inf Process Med Imaging, 2009. **21**: p. 476-478.

26. al, M.S.a.e., *Framework for the statistical shape analysis of brain structures using SHARM-PDM.* Insight J, 2006. **1071**: p. 242-250.

27. al, B.F.a.e., *High resolution intersubject averaging and a coordinate system for the cortical surface.* Hum Brain Mapp, 1999. **8**: p. 272-284.

28. I.P. Friedel, P.S.a.M.D., *Unconstrained spherical parameterization in ACM SIGGRAPH.* ACM, 2005(Los Angeles).

29. al, E.D.A.a.e., *A viscous fluid model for multimodal non-rigid image registration using mutual information.* Medical Image Analysis, 2003. **7**(4): p. 565-575.

30. Everson, R., *Orthogonal, but not Orthonormal Procrustes Problems.* Advances in Computations Mathematics, 1997.

31. *<3dMDvultus Analysis User Manual 2.3 (8.JUL.14).pdf>*.

32. Mutsvangwa, T. and T.S. Douglas, *Morphometric analysis of facial landmark data to characterize the facial phenotype associated with fetal alcohol syndrome.* Journal of Anatomy, 2007. **210**(2): p. 209-220.

33. Gstoettner, M., et al., *Inter- and intraobserver reliability assessment of the Cobb angle: manual versus digital measurement tools.* Eur Spine J, 2007. **16**(10): p. 1587-92.

34. Croker, V. and S. McDonald, *Recognition of emotion from facial expression following traumatic brain injury.* Brain Injury, 2005. **19**(10): p. 787-799.

35. Eder, M., et al., *Evaluation of precision and accuracy assessment of different 3-D surface imaging systems for biomedical purposes.* J Digit Imaging, 2013. **26**(2): p. 163-72.

36. Arnett, G. W. et al. *Facial planning for orthodontists and oral surgeons*, American Journal of Orthodontics and Dentofacial Orthopedics, 2004. **126**(3): p.290-295.

37. Bergman, R., *Cephalometric soft tissue facial analysis,* American Journal of Orthodontics and Dentofacial orthopedics, 1999. **116**(4): p.373-389.