

A 3-D Reconstruction System for the Human Jaw Using a Sequence of Optical Images¹

Sameh M. Yamany, Member, IEEE, Aly A. Farag*, Senior Member, IEEE

Computer Vision and Image Processing Laboratory

Electrical and Computer Engineering Department, University of Louisville, Louisville, KY 40292, USA.

David Tasman and Allan G. Farman

School of Dentistry, University of Louisville, Louisville, KY 40292, USA.

*Contact author: farag@cairo.uofl.edu

Abstract

This paper presents a model-based vision system for dentistry that will assist in diagnosis, treatment planning and surgical simulation. Dentistry requires an accurate 3-D representation of the teeth and jaws for diagnostic and treatment purposes. The proposed integrated computer vision system constructs a 3-D model of the patient's dental occlusion using an intra-oral video camera. A modified shape from shading (SFS) technique, using perspective projection and camera calibration, extracts the 3-D information from a sequence of 2-D images of the jaw. Data fusion of range data and 3-D registration techniques develop the complete jaw model. Triangulation is then performed, and a solid 3-D model is reconstructed. The system performance is investigated using ground truth data, and the results show acceptable reconstruction accuracy.

Keywords

Image Sequence Analysis, Camera Calibration, Shape Representation, Data Fusion, Registration, Dental Occlusion.

I. INTRODUCTION

Dentistry requires accurate 3-D representation of the teeth and jaw for diagnostic and treatment purposes. For example, orthodontic treatment involves the application, over time, of force systems to teeth to correct malocclusion. In order to evaluate tooth movement progress, the orthodontist monitors this movement by means of visual inspection, intra-oral measurements, fabrication of casts, photographs, and radiographs; this process is both costly and time consuming. Moreover, repeated acquisition of radiographs

may result in untoward effects. Obtaining a cast of the jaw is a complex operation for the dentist, an unpleasant experience for the patient, and also may not provide all the necessary details of the jaw.

Oral and maxillofacial radiology can provide the dentist with abundant 3-D information of the jaw. Current and evolving methods include computed tomography (CT), tomosynthesis [1], tuned-aperture CT (TACT) [2], and localized, or “cone-beam,” computed tomography [3]. While oral and maxillofacial radiology is now widely accepted as a routine technique for dental examinations, the equipment is rather expensive and the resolution is frequently too low for 3-D modeling of dental structures. Furthermore, the radiation dose required to enhance both contrast and spatial resolution can be unacceptably high.

Much efforts have focused recently on computerized diagnosis in dentistry[4]. Bernard et. al. [5] developed an expert system where cephalometric measurements are acquired manually from the analysis of radiographs and plaster models. Laurendeau et al. [4] presented a computer-vision technique for the acquisition of jaw data from inexpensive dental wafers. That system was capable of obtaining imprints of the teeth. Usually, most of the 3-D systems for dental applications found in the literature rely on obtaining an intermediate solid model of the jaw (cast or teeth imprints) and then capturing the 3-D information from that model. User interaction is needed in such systems to determine the 3-D coordinates of fiducial reference points on a dental cast. Other systems that can measure the 3-D coordinates have been developed using either mechanical contact [6] or a traveling light principle [7][8]. Goshtasby et al. [9] designed a range scanner based on white light to reconstruct the cast. The scanner used the subtractive light principle to create very thin shadow profiles on the cast.

The authors of this paper are developing a system for dentistry to replace traditional approaches in diagnosis, treatment planning, surgical simulation, and prosthetic replacements ([10], [11], [12]). Specific objectives of our dental imaging research (the Jaw Project) include: (i) to design a data acquisition system that can obtain sequences of calibrated video images, with respect to a common reference in 3-D space, of the upper/lower jaw using a small intra-oral camera ; (ii) to develop methods for accurate 3-D reconstruction from the acquired sequence of intra-oral images. This involves using a new algorithm for shape from shading (SFS) that incorporates the camera parameters; (iii) to develop a robust algorithm for the fusion of data acquired from multiple views, including the implementation of an accurate and fast 3-D data registration; (iv) to develop a specific object segmentation and recognition system to separate and recognize individual 3-D tooth information for further analysis and simulations; and (v) to develop algorithms to study and simulate

tooth movement based on the finite element method and deformable model approaches. This research has an immense value in various dental practices including implants, tooth alignment, and craniofacial surgery. The research has also wide applications in teledentistry, dental education and training.

This paper describes the project's first phase concerning the development of a 3-D model of the jaw, not from a cast, but from the actual human jaw. The work reported here is original and novel in the following aspects: (1) data acquisition is performed directly on the human jaw using a small off the shelf CCD camera, (2) the acquisition time is relatively short and is less discomforting to the patient compared to current practices. Also, the acquired digital model can be stored with the patient data and can be retrieved on demand. These models can also be transmitted over a communication network to different remote practitioners for further assistance in diagnosis and treatment planning. Using these models, dental measurements and virtual restoration can be performed and analyzed.

II. SYSTEM OVERVIEW

As shown in Figure 1, our approach to reconstruct the human jaw consists of the following stages. The first stage is data acquisition. A small intra-oral CCD camera, with a built-in white light, mounted on a five link 3D digitizer arm, is calibrated and then placed inside the oral cavity. The camera acquires a set of overlapping images $\{I_j \mid j = 1, 2, \dots, J\}$ for different parts of the jaw such that $\bigcup_{j=1}^J I_j$ covers the entire jaw. The images are preprocessed to reduce noise, sharpen edges, and remove specularities. Removing specularities is important because it affects the accuracy of the reconstructed surfaces using SFS. We used an approach similar to Tong and Funt method to remove specularities [13]. In this approach changes in the reflection map is removed from the luminance image by calculating the logarithmic gradient of the image and thresholding at locations of abrupt chromaticity change. In addition, we used median filtering to remove speckle noise from the images. Using a modified SFS algorithm, which accounts for the camera perspective projection, J sets of 3D points are then computed. The SFS technique is described in detail in section III. To obtain accurate metric measurements, range data is obtained using the five link digitizer arm. These data consist of reference points on the jaw. Fusion of the range data and the SFS output provides accurate metric information that can be used later for orthodontic measurements and implant planning. This stage is described in section IV. A fast registration technique is required to merge the resulting 3-D points to obtain a complete 3-D description of the jaw [14], [15]. The final stage transforms this model into patches of free

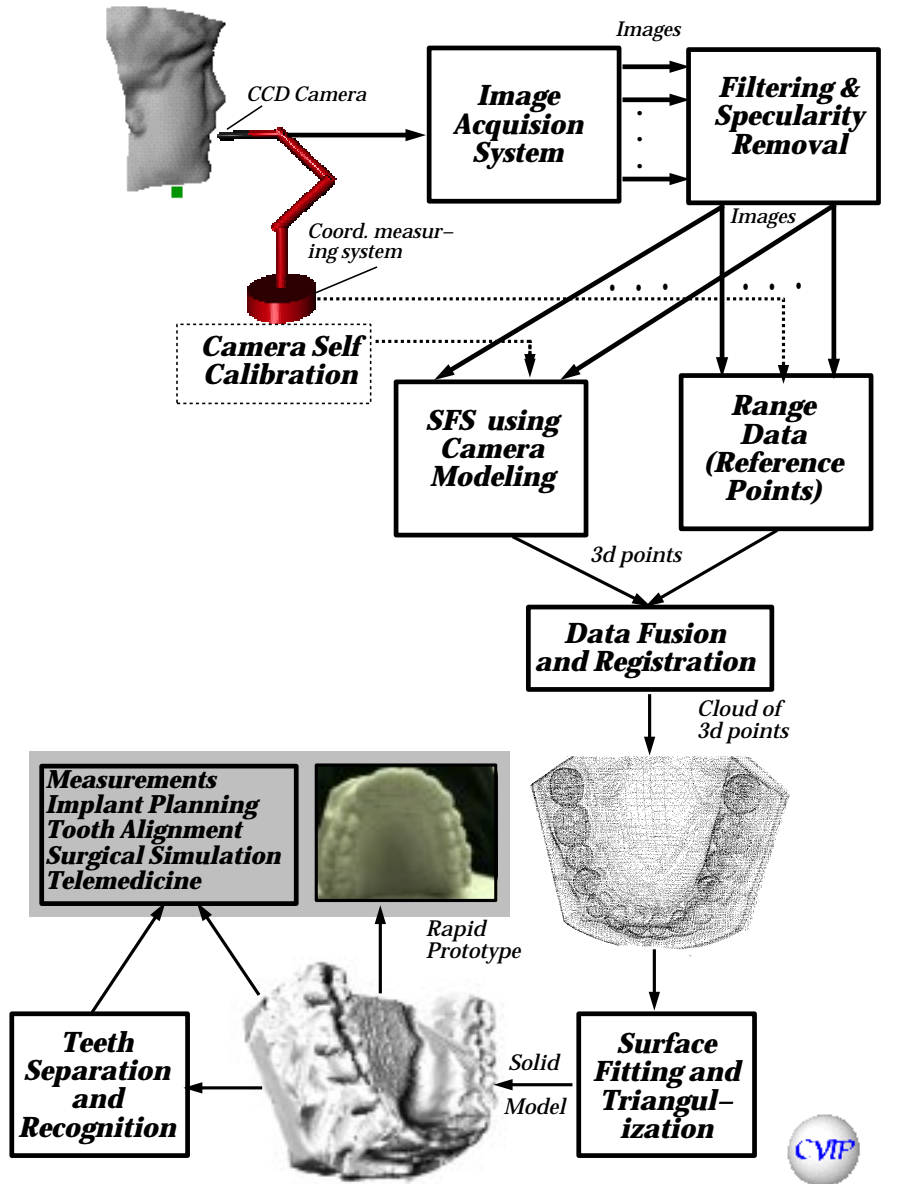


Fig. 1. Reconstructing the 3-D model of the human jaw starts by capturing a sequence of video images using a small intra-oral CCD camera. These images are preprocessed to remove specularity. Reference points are obtained using the CMM system. The range data are fused to the shape from shading (SFS) output and then registration takes place. A cloud of points representing the jaw is obtained and, by triangulation, a solid digital model is formed. This model is reproduced using a rapid prototype machine. Further analysis and orthodontics application can be performed on the digital model.

form surfaces using a triangulation technique. This step enables the development of a 3-D solid model for visualization. A cast can be fabricated from this model via rapid prototyping. Further processing that can be carried out on the digital model includes tooth separation, force analysis, implant planning, and surgical simulation.

III. SHAPE FROM SHADING USING PERSPECTIVE PROJECTION AND CAMERA CALIBRATION

Among the tools used in shape extraction from single view is the shape from shading (SFS) technique. SFS has been primarily studied by Horn [16] and his colleagues at MIT. There have been many developments in the algorithms (e.g., [17], [18], [19], [20], [21]). SFS assumes that the surface orientation at a point \mathbf{M} on a surface S is determined by the unit vector perpendicular to the plane tangent to S at \mathbf{M} . Under the assumption of orthographic projections, the elemental change in the depth Z at an image point (x, y) can be expressed as $\delta z \approx \frac{\partial Z}{\partial x} \delta x + \frac{\partial Z}{\partial y} \delta y$. The partial derivatives are called surface gradients (p, q) . The normal to a surface patch is related to the surface gradient by $\mathbf{n} = (p, q, 1)$. Assuming that surface patches are homogeneous and uniformly lit by distant light sources, the brightness $E(x, y)$ seen at the image plane often depends only on the orientation of the surface. This dependence of brightness on surface orientation can be represented as a function $R(\cdot)$ defined on the Gaussian sphere. Thus, the SFS problem is formulated as finding a solution to the brightness equation: $E(x, y) = R(p, q, \mathbf{L})$, where $R(p, q, \mathbf{L})$ is the surface reflectance map and \mathbf{L} is the illuminant direction.

A number of algorithms were developed to estimate the illuminant direction (e.g., [20]). Because the white light beam in our design is built in the CCD camera, the assumption that the illuminant direction is known beforehand is valid. However, the assumption of orthographic projection is not adequate for the dental application because the camera is very close to the object. Some SFS approaches using perspective projection were found in the literature (e.g. [22], [23], [24]). However, most of these approaches ignore the camera extrinsic parameters, hence cannot provide metric information of the depth. In our approach, the CCD camera is calibrated and the camera parameters are used in the SFS algorithm to obtain a metric representation of the teeth and gum surfaces. To calibrate the camera, the relation between the 3D point $\mathbf{M} = \{X, Y, Z\}$ and the corresponding image coordinates $\mathbf{m} = \{x, y\}$ is written as; $s\vec{\mathbf{m}} = \mathbf{P}\vec{\mathbf{M}}$ where s is a scalar, $\vec{\mathbf{m}}$ and $\vec{\mathbf{M}}$ are the extended vectors $[\mathbf{m}^T \ 1]^T$ and $[\mathbf{M}^T \ 1]^T$, and \mathbf{P} is called the camera calibration matrix. In general, $\mathbf{P} = \mathbf{A} [\mathbf{R}, \mathbf{t}]$ where \mathbf{A} is a matrix containing all the camera intrinsic parameters and

\mathbf{R} , \mathbf{t} are the rotation matrix and translation vector. The matrix \mathbf{P} has 12 elements but has only 11 degrees of freedom because it is defined up to a scale factor (e.g., [25]).

The standard method of calibration is to use an object with known size and shape and extract the reference points from the object image. It can be shown that given N points ($N \geq 6$) in general positions, the camera can be calibrated [25]. The perspective projection matrix \mathbf{P} can be decomposed as $[\mathbf{B}\mathbf{b}]$ where \mathbf{B} is a 3×3 matrix and \mathbf{b} is a 3×1 vector such that;

$$s\vec{\mathbf{m}} = \mathbf{B}\mathbf{M} + \mathbf{b} \quad \text{or,} \quad (1)$$

$$\mathbf{M} = \mathbf{B}^{-1}(s\vec{\mathbf{m}} - \mathbf{b}) = f(s(x, y)) \quad (2)$$

This last equation represents a line in the 3D space corresponding to the visual ray passing through the optical center and the projected point \mathbf{m} . By finding the scalar s , $f(s(x, y))$ will define a unique 3D point \mathbf{M} on the object. The surface normal at \mathbf{M} is defined to be the cross product of the two gradient vectors $\mathbf{p} = \frac{df(s(x, y))}{dx}$, $\mathbf{q} = \frac{df(s(x, y))}{dy}$. The surface reflectance $R(\cdot)$ becomes a function of the scalar s defined in Eq.(1), i.e.,

$$R(s) = \frac{(\mathbf{p} \times \mathbf{q}) \cdot \mathbf{L}}{|\mathbf{p} \times \mathbf{q}| |\mathbf{L}|} \quad (3)$$

The formulation of the SFS problem becomes finding the scalar s that solves the brightness equation $g(s) = E(x, y) - R(s) = 0$. This can be solved using a Taylor's series expansion and applying the Jacoby iterative method [21]. After n iterations, for each point (x, y) in the image, $s_{x,y}^n$ is given as follows (see appendix A for more details):

$$s_{x,y}^n = s_{x,y}^{n-1} + \frac{-g(s_{x,y}^{n-1})}{\frac{d}{ds_{x,y}}g(s_{x,y}^{n-1})} \quad \text{where,} \quad (4)$$

$$\frac{d}{ds_{x,y}}g(s_{x,y}) = -\frac{d\mathbf{N}}{ds_{x,y}} \cdot \frac{\mathbf{L}}{|\mathbf{L}|} \quad (5)$$

$$\frac{d\mathbf{N}}{ds_{x,y}} = \frac{d\mathbf{v}}{ds_{x,y}} \frac{1}{\sqrt{\mathbf{v}^t \mathbf{v}}} - \frac{\mathbf{v}}{\sqrt{(\mathbf{v}^t \mathbf{v})^3}} \left(\mathbf{v}^t \frac{d\mathbf{v}}{ds_{x,y}} \right) \quad (6)$$

$$\frac{d\mathbf{v}}{ds_{x,y}} = \mathbf{B}^{-1} \vec{\mathbf{m}} \times \mathbf{B}^{-1} (0, s_{x,y-1}, 0)^t + \mathbf{B}^{-1} (s_{x-1,y}, 0, 0)^t \times \mathbf{B}^{-1} \vec{\mathbf{m}} \quad (7)$$

where $\mathbf{v} = \mathbf{p} \times \mathbf{q}$.

Even though camera parameters are used in the SFS implementation, accurate metric information cannot be deduced from the resulting shape because only one image is used. The following section describes a novel approach to perform the necessary calibration.

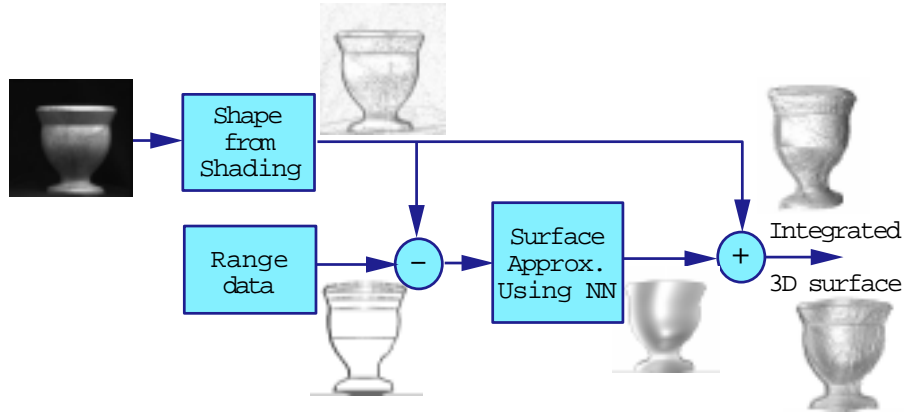


Fig. 2. *Functional block diagram for the integration process of shape from shading and range data. The surface approximation process is done using neural networks (NN). An example of using this system on the image of a vase is demonstrated.*

IV. FUSION OF SFS AND RANGE DATA

The most important information for reconstructing an accurate 3-D visible surface, which is missing in shape from shading, is the metric measurement. Shape from shading also suffers from the discontinuities due to highly textured surfaces and different albedo. The integration of the dense depth map obtained from SFS with sparse depth measurements obtained from a coordinate measurement machine (CMM) for the reconstruction of 3-D surfaces with accurate metric measurements has two advantages [26]. First, it helps to remove the ambiguity of the 3-D visible surface discontinuities produced by shape from shading. Second, it complements for the missing metric information in the shape from shading. The integration process, as depicted in Fig. 2, includes the following stages. First, we calculate the error difference in the available depth measurements between the two sets of sensory data. Then, we approximate a surface that fits this error difference. Finally, the approximated surface is used to correct the shape from shading.

A multi-layer neural network is used for the surface approximation process since neural networks was shown to be more robust in terms of complexity, speed and accuracy [26] than other computational approaches (e.g., regularization techniques). The learning algorithm applied is the extended Kalman-filter learning technique because of its fast computation of weights. The x- and y- coordinates of the data points are the input to the network, while the error in the depth value at the point (x, y) is the desired response. The error difference between the SFS and the range measurements and their x-y coordinates are used to form the training set. The input to the network is the x-y coordinates and the output is the error difference at that

coordinate. Once training is accomplished, the neural network provides the approximated smooth surface which contains information about the errors in the shape from shading at the locations with no range data. This approximated surface is then added to the SFS. The result is the 3-D surface reconstruction that contains accurate metric information about the visible surface of the sensed 3-D object. An example performed on the image of a vase is shown in Fig. 2. The output of the fusion algorithm to each image is a set of 3-D points describing the teeth surfaces in this segment. To compensate for some digitizer inaccuracy in determining the camera location in space and the occasional patient movement, we have developed a fast and accurate 3-D registration technique to link the 3-D points of all the segments to produce one set of 3-D points describing the whole jaw surface [14]. The technique is a modification of the ICP technique by Besl and McKay [27], which uses a grid transform and genetic algorithms to reduce the registration time.

V. EXPERIMENTAL SETUP

The experimental setup shown in Figure 3 consists of the following:

- (i) A MicroScribe 3D digitizer: It can digitize a working space up to 1.27m (Sphere) with a sampling rate of 1000 points/second.
- (ii) A 1/3 inch Ultra-Mirco CCD color camera(ELMO UN411E). The camera has 768Hx494V effective picture elements. We used a 5.5mm lens for this camera.
- (iii) A 150 Watt DC regulated white light source. Through a fiber optic bundle that surrounds the CCD camera, this light source illuminates the oral cavity with a stable white light. The light intensity could be adjusted manually to control the shading effect.
- (iv) An SGI Indigo2 machine that hosts the software required for the data processing, reconstruction and visualization of the 3D jaw model.

The CCD camera is mounted on the stylus of the 3D digitizer and its focal distance is adjusted such that the image will be in focus only when the stylus tip touches the tooth surface. The image plane is normal to the stylus and the stylus tip is at pixel (0,0). The CCD camera is then calibrated as shown in Figure 4. Camera calibration is done once before using the camera, and if the camera is stationary, we do not have to re-calibrate again. Yet in the proposed system, the camera will be moving; This implies the recalculation of the perspective projection matrix. However, as the camera will be mounted on a coordinates measuring system, the location of the optical center \mathbf{M}_{oc} can be tracked as the camera moves, and the

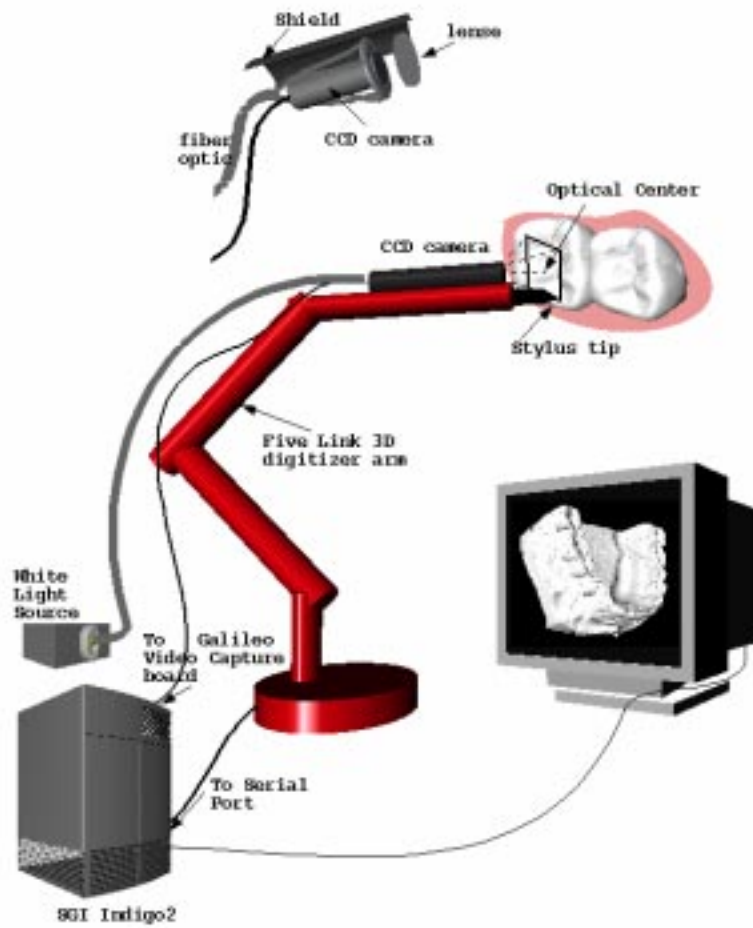


Fig. 3. *The experimental setup of the proposed system consists of a CCD camera mounted on a five link 3D digitizer. A DC regulated light source is connected to a fiber optic bundle surrounding the camera. The setup is connected to an SGI Indigo2 machine.*

camera perspective projection can be recalculated.

The five degrees of freedom provided by the arm enable the acquisition of a sequence of intra-oral images covering the upper/lower jaw. Also, with each image, the camera location in the 3-D space is measured. The perspective projection matrix is re-adjusted and the location and direction of the first pixel in the image are included. This information is used in the data fusion and registration phase to reference the image plane in the workspace.

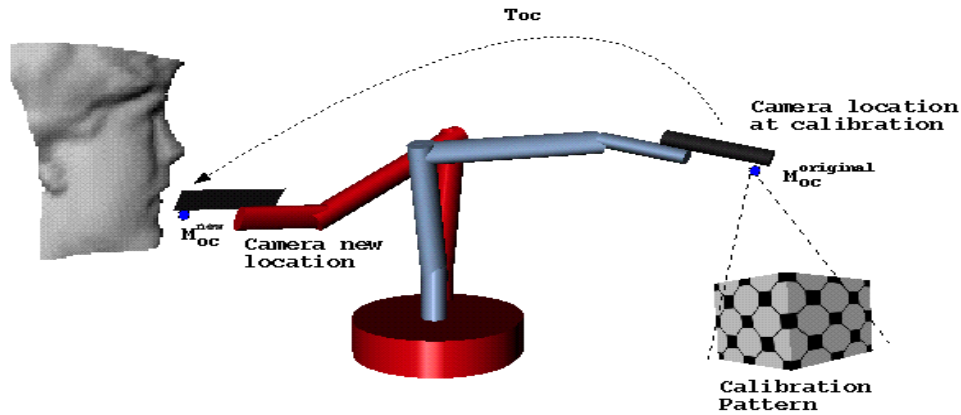


Fig. 4. *Illustration of the camera self calibration. The coordinates measuring system is used to find the transformation matrix T_{oc} between the optical center M_{oc} of the camera at the calibration time and the new location while acquiring the images. This transformation matrix is used to obtain the new camera perspective projection matrix.*

VI. RESULTS

We designed a software interface to provide the dentist with real time visualization of the reconstructed segments before registration and fusion as shown in Fig. 5. The 3D reconstructed segments are displayed with full color texture reflecting the colors of the teeth and gum. This texture is saved with the scanned model and can be used later for diagnosis especially for gum diseases. This interface also enables the dentist to know which parts of the jaw are scanned and which parts still need to be scanned. The interface then prompts the dentist to digitize some range points for each image segment. These range data are used in the fusion stage. Finally, 3D registration is performed and a complete metric model is obtained. This model is converted into a format suitable for rapid-prototyping such that a solid cast can be constructed.

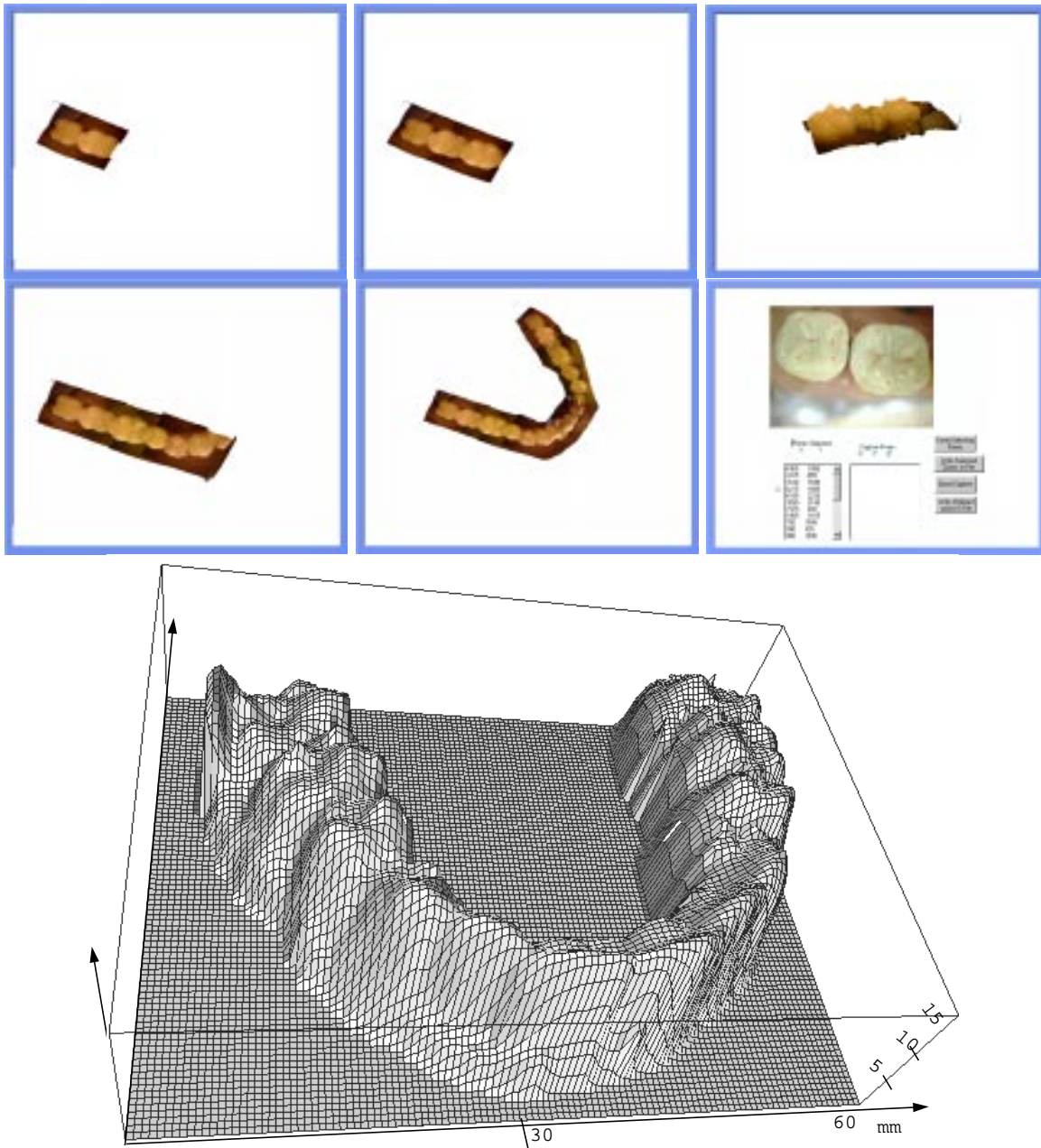


Fig. 5. *Snapshots of the software used. The reconstructed segments are displayed in real time with full color texture. The dentist can rotate and manipulate the displayed 3D model to check what parts still need to be scanned or re-scanned. Fusion with range data from the digitizer is performed next and a complete jaw model with metric information is obtained. This model can be used to measure any orthodontic parameter and can be reproduced via rapid prototyping with the same original scale.*

Although data acquisition is relatively fast, referencing the location and orientation of patient's head is very crucial, especially to register different acquisitions taken at different times, or during the same session when the patient is allowed to move the head. For this reason, four points (see Figure 6) on the patient head are captured before starting the data acquisition stage and between the different reconstruction stages. The points define the positions of the upper and lower jaw relative to a common coordinate system.

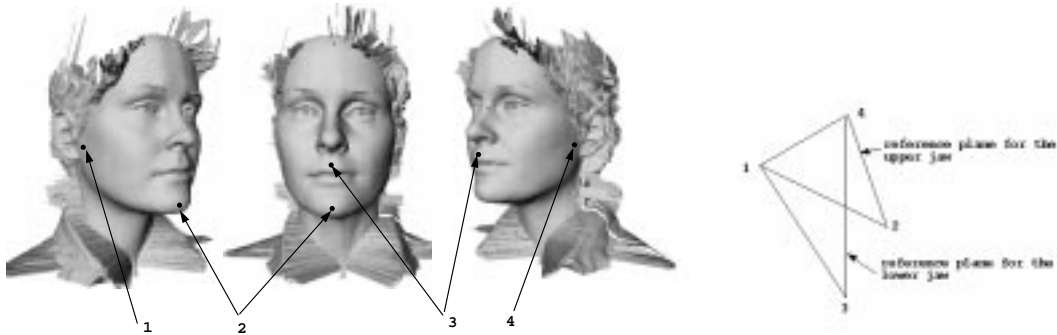


Fig. 6. *Four points are used to reference the upper and lower jaws to the same 3D coordinates system. This allows for patient head movements between acquisition stages. Also, these points are used to register different acquisition at different times for the same patient.*

Experiments were conducted on different pediatric subjects at the Orthodontics Department, University of Louisville, Kentucky. The system was used to reconstruct an accurate model of individual tooth or teeth segments. Figure 7(a) shows an example of two images taken for a patient's tooth. The complete tooth surface is covered in these two images. Using the reference data, we applied the fusion algorithm. This resulted in the 3-D visible surfaces shown in Fig. 7(b). Using the registration procedure on these two data sets produced the complete surface model of the tooth as shown in Fig. 7(c).

More results of applying the reconstruction algorithms on another subject are shown in Fig. 8. The resulting jaw models have acceptable accuracy and are faithful enough to show all the information about the patient's actual jaw in a metric space. Both the time and convenience for the patient must be considered when comparing these results with the results from scanning a cast. The reconstructed digital model can be of enormous help to the clinician in addition to numerous other advantages. A few of the suggested advantages include: helping the clinician inspect subtle problems, assisting in indirect operative procedures, improving the communication with patients because they can now see a 3-D realization of their jaw, and achieving more thorough examinations, which in turn leads to more complete diagnosis, documentation of

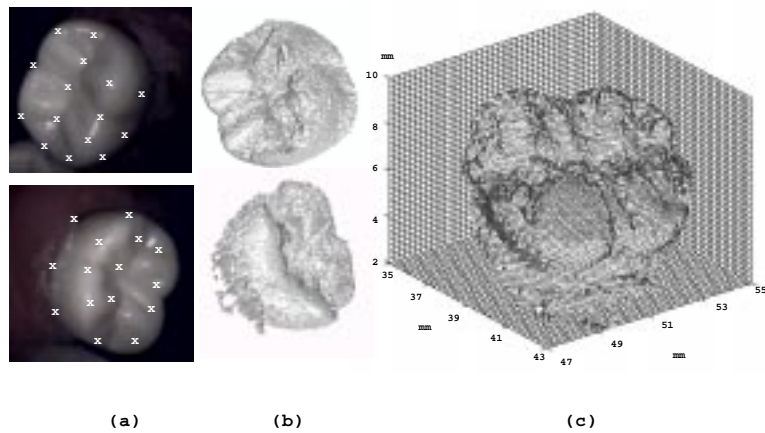


Fig. 7. (a) Intra-oral intensity images with range data marked as cross signs. (b) 3-D visible surfaces obtained from the shape from shading and the fusion process. (c) A visible surface mesh obtained from registering the two views.

dental diseases and treatment. Yamany et al. [28] performed further processing on the digital jaw model. Fig. 8(c) shows the result of the teeth segmentation and identification stage. The color code is used to visualize individual teeth segments.

Experimental Validation

The accuracy of the reconstructed surfaces is of major importance in the dental research. There exist four sources of inaccuracy in the developed approach. (1) The Camera Calibration Process: In the proposed system we used the camera calibration technique developed in our lab [29]. The approach is based on the neurocalibration technique, in which the calibration problem is mapped into a learning problem of a multi-layered feedforward neural network. The key advantage of this technique, as opposed to others, is that it starts the non-linear minimization process from a random initial point without sacrificing the calibration accuracy. The calibration accuracy, measured in terms of 3D reconstruction error, is found to be $< 0.1mm$. (2) The 3D digitizer used has a reported accuracy $< 0.23mm$. (3) The registration technique used has a reported accuracy $< 0.19mm$ [14]. It should be noticed that the registration process is used to compensate for the digitizer accuracy and the occasional patient's head movements. (4) The SFS and Data Fusion Process: To validate the SFS and data fusion methods and to find the accuracy of the proposed system, we have used a ground truth dense depth map obtained from a laser range scanner. The range scanner also provided intensity images registered with the depth map. Although laser scanners have accuracy limitation, their

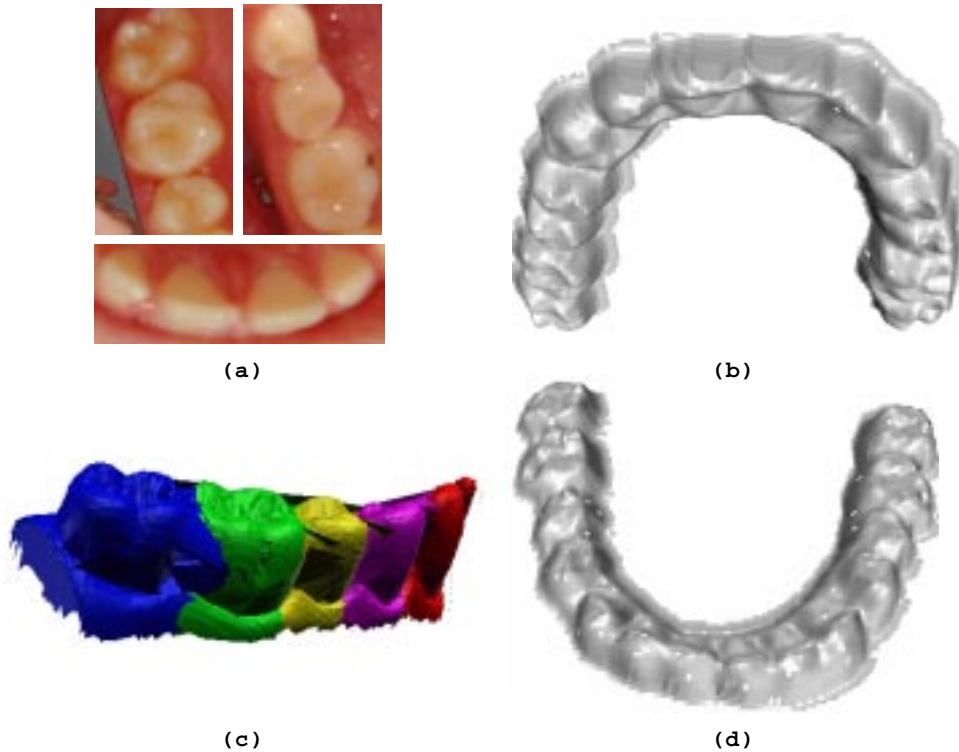


Fig. 8. (a) Some intensity images of a patient's teeth. (b and d) The upper and lower digital jaw models of this patient. (c) The result of the teeth segmentation and identification stage - different shades are used to show teeth separation.

output can be considered as ground truth since they are now widely accepted in the dentistry practices. The root mean square (RMS) error between the integrated surface and the ground truth is used as a measure of the system performance. Two different types of surfaces, a smooth and a free-form surfaces, were investigated. The neural network topology that best suited the data fusion was analyzed. The best performance (RMS of $0.24mm$ for the free-form surface and $0.132mm$ for the smooth surface) has been obtained with the network topology 2-7-3-1, that is two neurons in the input layer (the pixel coordinates in the image), two hidden layers with 7 and 3 neurons, respectively, and one neuron in the output layer (the depth measurement). Figure 9 shows the estimated RMS error with different sampling rates. In general, the RMS error for the smooth surface is smaller compared to the free-form surface. This is expected because the surface approximation process tends to smooth the surface where range data are not available. This produces a large RMS error in the case of free-form surfaces. The results show that higher sampling rate increases the accuracy in the case of free-form surfaces, yet this increases the time and complexity to acquire the data. However, the sampling

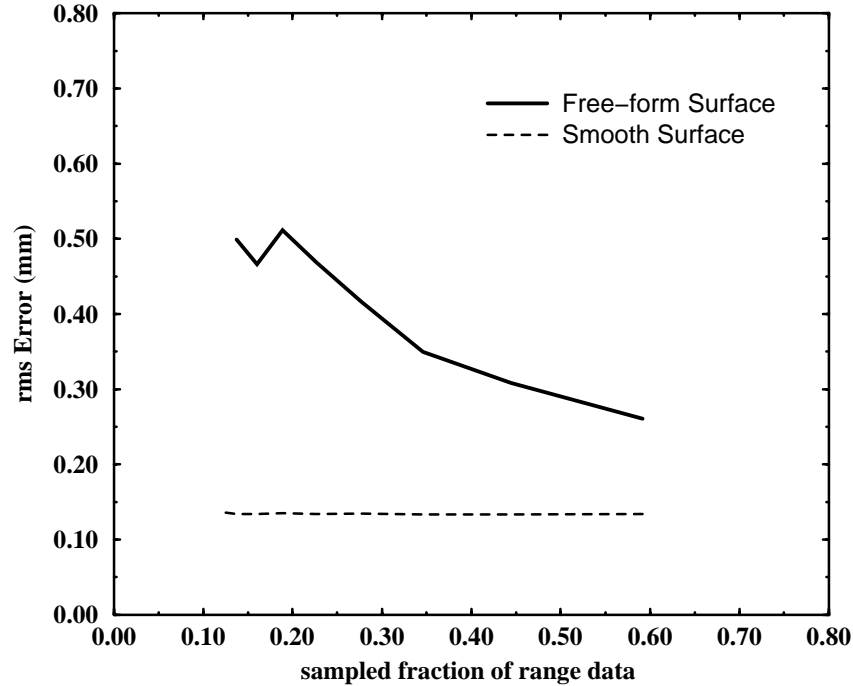


Fig. 9. The RMS error in mm between the integrated surface and the ground truth is used as a measure of the system performance. Two different types of surfaces, a sphere as a smooth surface and a scanned tooth (free-form) surface were investigated.

rate has minimal effect in the case of smooth surfaces. In the dental application, most of the teeth surfaces can be considered smooth. Rough edges are only found in the top portion of the teeth surface. In this case, most of the reference points are obtained in this area. The above analysis shows that the system can achieve acceptable accuracy with only a small number of reference points.

From the above analysis of the four sources of accuracy limitation, it can be concluded that the system can achieve an overall sub-millimeter accuracy in reconstructing the jaw model. Figure 10 shows the superposition of a reconstructed jaw model using the proposed techniques and the laser scanned model of the cast of the same subject. The above analysis shows that the system can approach the accuracy of the laser range scanner. Another validation experiment was conducted to compare the reconstructed model to the actual subject's jaw. This was done by comparing orthodontics measurements done on both the actual jaw and the reconstructed one. Forty measurements were collected and compared to their corresponding distance

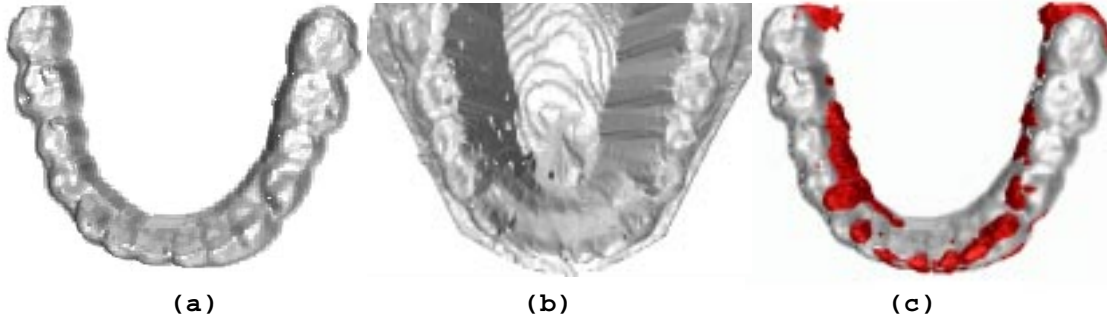


Fig. 10. (a) The jaw model obtained using our proposed approach. (b) The laser scanned model of the cast of the same jaw. (c) The two models superimposed (the dark represents the laser scanned model).

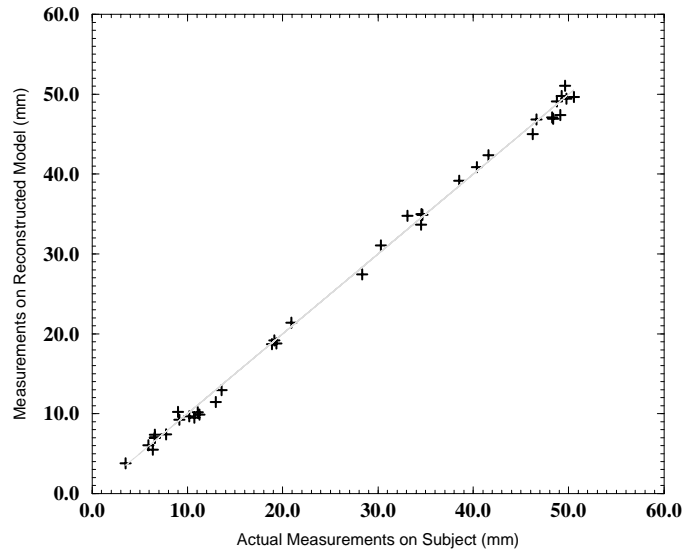


Fig. 11. Corresponding distance measurements on both the actual subject's jaw and their reconstructed model.

measurements on the digital model as shown in Figure 11. The average error in distance calculation was found to be 0.74mm which is an acceptable resolution for many orthodontics and maxillofacial applications.

VII. CONCLUSIONS AND FUTURE EXTENSIONS

The 3-D reconstruction of the human jaw has tremendous applications. The model can be stored along with the patient data and can be retrieved on demand. The model can also be used in telemedicine, that

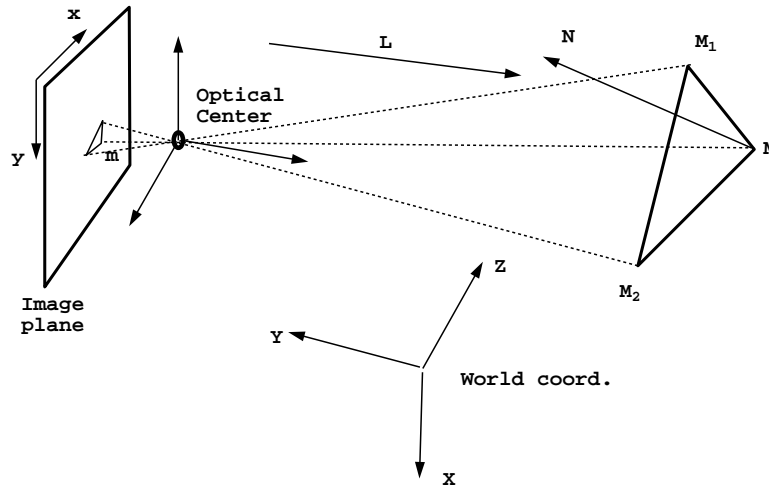


Fig. 12. Considering the surface to be a set of triangular patches each with a normal N , the dot product of the light direction L and the normal N is called the reflectance which determines the image intensity at the corresponding pixel location m .

is it can be transmitted over a communication network to different remote dentists for further assistance in diagnosis. Dental measurements and virtual restoration could be performed and analyzed. This work enables many orthodontics and dental imaging researches, applied directly to a digital jaw model, not to a cast, using computer vision and medical imaging tools.

The paper describes the result of the first phase in a project aimed to replace current dental modeling practices. The next phase includes the analysis and simulation of dental operations including tooth alignment, implant planing, restoration, and measurement of distances and orientation of teeth with respect to each other. We have some preliminary results in this second phase [28].

Acknowledgment: This work was supported by grants from the Whitaker Foundation (Research Grant 94-0228) and the NSF (Grant ECS-9505674).

From Fig. 12 and using equation[1] we can define the following;

$$\mathbf{M} = \mathbf{B}^{-1}(s_{x,y} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} - \mathbf{b}) \quad (8)$$

$$\mathbf{M}_1 = \mathbf{B}^{-1}(s_{x-1,y} \begin{pmatrix} x-1 \\ y \\ 1 \end{pmatrix} - \mathbf{b})$$

$$\mathbf{M}_2 = \mathbf{B}^{-1}(s_{x,y-1} \begin{pmatrix} x \\ y-1 \\ 1 \end{pmatrix} - \mathbf{b})$$

$$\mathbf{p} = \mathbf{M} - \mathbf{M}_1$$

$$= \mathbf{B}^{-1}(s_{x,y} - s_{x-1,y}) \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} + \mathbf{B}^{-1}s_{x-1,y} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad (9)$$

$$\mathbf{q} = \mathbf{M} - \mathbf{M}_2$$

$$= \mathbf{B}^{-1}(s_{x,y} - s_{x,y-1}) \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} + \mathbf{B}^{-1}s_{x,y-1} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad (10)$$

The unit normal \mathbf{N} to the patch formed by \mathbf{M} , \mathbf{M}_1 , and \mathbf{M}_2 can be calculated as follows:

$$\mathbf{N} = \frac{\mathbf{p} \times \mathbf{q}}{\|\mathbf{p} \times \mathbf{q}\|} \quad (11)$$

The reflection function defined by the SFS is written as,

$$R(.) = \frac{\mathbf{N} \cdot \mathbf{L}}{\|\mathbf{L}\|}$$

Thus, the SFS brightness equation becomes

$$g(s_{x,y}) = E(x,y) - \frac{\mathbf{N} \cdot \mathbf{L}}{\|\mathbf{L}\|} \quad (12)$$

The solution to the SFS problem is to find $s_{x,y}$ such that $g(\cdot)$ is minimized. Using Taylor expansion and Jacoby iterative methods, $s(x, y)$ can be found by iteration as follows,

$$s_{x,y}^n = s_{x,y}^{n-1} + \frac{-g(s_{x,y}^{n-1})}{\frac{d}{ds_{x,y}}g(s_{x,y}^{n-1})} \quad (13)$$

$$\frac{d}{ds_{x,y}}g(s_{x,y}) = -\frac{d\mathbf{N}}{ds_{x,y}} \cdot \frac{\mathbf{L}}{|\mathbf{L}|} \quad (14)$$

Let $\mathbf{N} = \frac{\mathbf{v}(s)}{\sqrt{\mathbf{v}(s)^t \mathbf{v}(s)}}$ where $\mathbf{v}(s) = \mathbf{p} \times \mathbf{q}$

$$\frac{d\mathbf{N}}{ds_{x,y}} = \frac{d\mathbf{v}}{ds_{x,y}} \frac{1}{\sqrt{\mathbf{v}^t \mathbf{v}}} - \frac{\mathbf{v}}{\sqrt{(\mathbf{v}^t \mathbf{v})^3}} \left(\mathbf{v}^t \frac{d\mathbf{v}}{ds_{x,y}} \right) \quad (15)$$

$$\begin{aligned} \frac{d\mathbf{v}}{ds_{x,y}} &= \frac{d}{ds_{x,y}}(\mathbf{p} \times \mathbf{q}) \\ &= \frac{d\mathbf{p}}{ds_{x,y}} \times \mathbf{q} + \mathbf{p} \times \frac{d\mathbf{q}}{ds_{x,y}} \end{aligned}$$

$$\frac{d\mathbf{p}}{ds_{x,y}} = \frac{d\mathbf{q}}{ds_{x,y}} = \mathbf{B}^{-1} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

$$\frac{d\mathbf{v}}{ds_{x,y}} = \mathbf{B}^{-1} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \times \mathbf{B}^{-1} \begin{pmatrix} 0 \\ s_{x,y-1} \\ 0 \end{pmatrix} + \mathbf{B}^{-1} \begin{pmatrix} s_{x-1,y} \\ 0 \\ 0 \end{pmatrix} \times \mathbf{B}^{-1} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \quad (16)$$

Algorithm: The steps involved in the implementation are as follows:

Step 1: Read image $E(x, y)$, light direction \mathbf{L} , and camera parameters (\mathbf{B}, \mathbf{b}) .

Step 2: Initialize $s_{x,y} = 0.01$.

Step 3: $\forall x, y$ get \mathbf{p} and \mathbf{q} as shown in equations[9] and [10].

Step 4: Calculate \mathbf{N} as shown in equation[11].

Step 5: Obtain the error using the brightness equation[12].

Step 6: Estimate the new $s_{x,y}$ using equations[13], [14], [15], and [16].

Step 7: Repeat steps 3 to 7 until $\max_{x,y} |g(s_{x,y})| < \epsilon$ where ϵ is a predefined positive threshold.

Step 8: Recover the surface 3-D points using equation[8].

Step 9: Construct triangular patches as shown in Fig. 12.

REFERENCES

- [1] P. F. van der Stelt and S. M. Dunn, "3D-imaging in dental radiography," in *Advances in Maxillofacial Imaging*, A. G. Farman, ed., pp. 367–372, Elsevier Science B. V., 1997.
- [2] R. L. Webber, R. A. Horton, D. A. Tyndall, and J. B. Ludlow, "Tuned-aperture computed tomography (TACT). theory and application for three-dimensional; dento-alveolar imaging," *Dentomaxillofac. Radiol.* **26**, pp. 51–62, 1997.
- [3] P. F. van der Stelt, S. M. Dunn, O. Tokuota, L. Jurgens, and J. H. Slater, "Comparison of two localized CT-techniques for detecting periodontal angular bone defects," *J. Dent. Res.* **75**(Special Issue), p. 128 (abstract 887), 1996.
- [4] D. Laurendeau and D. Possart, "A computer-vision technique for the acquisition and processing of 3D profiles of dental imprints: An application in orthodontics," *IEEE Transactions on Medical Imaging* **10**, pp. 453–461, Sept 1991.
- [5] C. Bernard, A. Fournier, J. M. Brodeur, H. Naccache, and R. Guay, "Computerized diagnosis in orthodontics," *Proc. 66th Gen. Session Int. Assoc. Dental Res. Montreal*, 1988.
- [6] F. P. van der Linden, H. Boersma, T. Zelders, K. A. Peters, and J. H. Raben, "Three-dimensional analysis of dental casts by means of optocom," *J. Dent. Res.* **51**(4), p. 1100, 1972.
- [7] S. N. Bhatia and V. E. Harrison, "Operational performance of the traveling microscope in the measurement of dental casts," *Br. J. Orthod.* **14**, pp. 147–153, 1987.
- [8] K. Tanaka, A. A. Lowe, and R. DeCou, "Operational performance of the reflex monograph and its applicability to the three-dimensional analysis of dental casts," *Amer. J. Orthod.* **83**, pp. 304–305, 1983.
- [9] A. A. Goshtasby, S. Nambala, W. G. deRijk, and S. D. Campbell, "A system for digital reconstruction of gypsum dental casts," *IEEE Transactions on Medical Imaging* **16**, pp. 664–674, Oct 1997.
- [10] M. Ahmed, S. M. Yamany, E. E. Hemayed, S. Roberts, S. Ahmed, and A. A. Farag, "3d reconstruction of the human jaw from a sequence of images," *Proc. IEEE Computer Vision and Pattern Recognition Conf. (CVPR), Puerto Rico*, pp. 646–653, June 1997.
- [11] S. M. Yamany, A. A. Farag, D. Tasman, and A. G. Farman, "A robust 3-d reconstruction system for human jaw modeling," *2nd International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI'99), Cambridge, England*, pp. 778–787, Sept. 1999.
- [12] S. M. Yamany and A. A. Farag, "A system for human jaw modeling using intra-oral images," *Proc. IEEE Engineering in Medicine and Biology Society (EMBS) conference, Hong Kong* **20**, pp. 563–566, Oct. 1998.

- [13] F. Tong and B. V. Funt, "Removing specularities from color images for shape from shading," in *Computer Vision and Shape Recognition*, A. Krzyzak, T. Kasvand, and C. Y. Suen, eds., vol. 14 of *Computer Science*, pp. 275–290, World Scientific, 1989.
- [14] S. M. Yamany, M. N. Ahmed, and A. A. Farag, "A new genetic-based technique for matching 3d curves and surfaces," *Pattern Recognition* **32**(10), p. 1817, 1999.
- [15] S. M. Yamany and A. A. Farag, "Free-form surface registration using surface signatures," *Proc. IEEE International Conference on Computer Vision (ICCV), Greece* **2**, pp. 1098–1104, Sept 1999.
- [16] B. K. P. Horn and M. J. Brooks, *Shape from Shading*, Cambridge, Mass.: McGraw Hill, 1989.
- [17] K. Kimmel and A. Bruckstein, "Tracking level sets by level sets: A method for solving the shape from shading problem," *Computer Vision and Image Understanding* **62**(2), pp. 47–58, 1995.
- [18] G. Q. W. and G. Hirzinger, "Learning shape from shading by a multilayer network," *IEEE Transactions Neural Networks* **17**, pp. 985–995, July 1996.
- [19] R. Zhang, P. Tsai, J. Cryer, and M. Shah, "Analysis of shape from shading techniques," *Proc. IEEE Computer Vision and Pattern Recognition Conf. (CVPR), Seattle*, pp. 377–384, June 1994.
- [20] A. Pentland, *Extract Shape From Shading*, Academic Press, MIT Media Lab, 2nd ed., 1988.
- [21] P. S. Tsai and M. Shah, "A fast linear shape from shading," *Proc. IEEE Computer Vision and Pattern Recognition Conf. (CVPR), Urbana, Illinois*, pp. 734–736, June 1992.
- [22] K. M. Lee and J. Kuo, "Shape from shading with perspective projection," *Computer Vision, Graphics, and Image Processing: Image Understanding* **59**, pp. 202–212, 1994.
- [23] J. K. Hasegawa and C. L. Tozzi, "Shape from shading with perspective projection and camera calibration," *Computer and Graphics* **20**(3), pp. 351–364, 1996.
- [24] K. M. Lee and J. Kuo, "Shape from shading with generalized reflectance map model," *Computer Vision and Image Understanding* **67**, pp. 143–160, August 1997.
- [25] G. Xu and Z. Zhang, *Epipolar Geometry in Stereo, Motion and Object Recognition*, Kluwer Academic, The Netherlands, 1996.
- [26] M. G.-H. Mostafa, S. M. Yamany, and A. A. Farag, "Integrating shape from shading and range data using neural networks," *Proc. IEEE Computer Vision and Pattern Recognition Conf. (CVPR), Fort Collins, Colorado*, pp. 15–20 June 1999.
- [27] B. J. Besl and N. D. McKay, "A method for registration of 3-d shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **14**(2), pp. 239–256, 1992.
- [28] S. M. Yamany, A. A. Farag, and N. A. Mohamed, "Orthodontics measurements using computer vision," *Proc. IEEE Engineering in Medicine and Biology Society (EMBS) conference, Hong Kong* **20**, pp. 536–539, Oct. 1998.

- [29] M.T. Ahmed, E.E. Hemayed, and A.A. Farag, "Neurocalibration: A neural network that can tell camera calibration parameters," *Proc. IEEE International Conference on Computer Vision (ICCV), Greece* **1**, pp. 463–468, Sept. 1999.