CDA JOURNAL, VOL 37, Nº3

Three-Dimensional Surface Imaging of the Face

DAVID C. HATCHER, DDS, MSC, AND CRAIG DIAL, DRT

ABSTRACT Accuracy, precision, quality, and simplicity are goals of diagnostic 3-D surface imaging of the face. The rewards include improved diagnosis, an ability to create a patient-specific model, simulate treatment, and improve treatment outcomes. This article discusses 3-D surface imaging of the face and selected clinical applications that add value to the image data.

AUTHORS

David C. Hatcher, DDS, MSC, is an oral and maxillofacial radiologist in Sacramento.

Craig Dial, is a dental radiography technician in Sacramento

Two-dimensional photography has been the standard for facial imaging for decades but recently several methods have been developed for 3-D facial imaging. The current methodologies for creating 3-D facial images include close-range stereophotogrammetry, structured light, laser scans, CT scans, and MRI. Acquiring 2-D facial images is simple, does not require expensive equipment, and 2-D images are easy to format and print using a variety of software packages, such as Dolphin Imaging (Chatsworth, Calif.). All 2-D images have inherent limitations when the results are compared to the anatomic truth (anatomy as it exists in nature).

With a 2-D system it is difficult to control focal length, point of view, and lighting, and it is impossible to produce accurate images that could be used for anthropometric measurements. Precision is a measure of reproducibility and the inability 2-D systems to control focal length, point of view and lighting greatly reduces their precision. The inability to provide depth using a 2-D photographic system and the inability to calibrate for focal length produces an inaccurate representation of the anatomy.

All imaging technologies allow for the capture and display of anatomy. The capture and display variables of interest for this article are point of view, POV; field of view, FOV; focal length, and anatomic accuracy. These are important variables when discussing the differences in 2-D and 3-D imaging.

Point of View

Point of view refers to the visualization perspective. For example, a cephalometric projection usually refers to a lateral or posteroanterior POV. Using 2-D techniques, the visualization and capture

POVs are identical. Conversely, using 3-D techniques, the capture POV does not necessary match the display POV. For example, a 3-D CBCT capture may occur by circling a central ray around the head but the display angle can be user defined and is infinite. Therefore, there is an infinite number of viewing angles. In addition, a 3-D volume (CBCT), using software tools, can be reformatted or sliced along any plane, oblique plane, or curved plane to reveal the internal anatomy.

Field of View

Field of view refers to the dimensions and the anatomy captured by the imaging sensor. The variables related to the FOV include sensor size and spatial relationships between imaging source, anatomy, and sensor. It is often the goal of imaging to have the FOV match the region of interest (ROI) by collimating the X-ray beam and the sensor in order to minimize the radiation burden. To appropriately select the FOV requires imaging objectives that are designed to answer the clinical question being investigated. Matching the FOV with the ROI has the added advantages of controlling the radiation burden and improving the quality of the exam by reducing scatter radiation.

Focal Length

Focal length refers to the distance from the object being imaged to the recording sensor. The recording sensor includes film (X-ray or photographic) or digital sensor (CCD chip).

Anatomic Accuracy

Anatomic accuracy is an ideal imaging goal to accurately represent the anatomy as it exists in nature, i.e., the anatomic truth. The projection geometry associated with 2-D techniques does not produce accurate anatomic images. Three-D digital techniques have the opportunity to produce anatomically accurate images.

Imaging the craniofacial structures using 2-D film and digital acquisition techniques occurs with multiple POVs, multiple FOVs, multiple focal lengths, and variable resultant accuracy. This method deconstructs the anatomy into a collection of 2-D images. For the clinician to understand the anatomy using the 2-D deconstruction method requires a virtual reconstruction that attempts to reassemble disparate POVs, FOVs , focal lengths, and variables of differing accuracy. This is a difficult if not an impossible task.

> **THE IDEAL FACIAL** photographic system would be one that could accurately reproduce the 3-D geometry of the face, produce repeatable results, is low risk, is simple, fast, and produces photorealistic images.

The ideal facial photographic system would be one that could accurately reproduce the 3-D geometry of the face, produce repeatable results, is low risk, is simple, fast, and produces photorealistic images. Ideally, image data can be manipulated by software for future visualization and analysis.

A 3-D facial scanner creates a point cloud (mathematical or geometrical description as x, y, and z points) of the anatomic form (size and shape) and can photo-texture this point cloud with a photograph giving the 3-D image a photorealistic appearance. Three-D cameras are similar to 2-D cameras in that they have a cone-shaped field of view, and can only sense surface information that is in their direct line of sight. Three-D scanners acquire several images, each from a unique point of view, that are combined (registered or fused) into a common 3-D coordinate system.

Laser-Based Imaging

A 3-D laser scanner uses a laser light to sequentially probe the anatomy while a CCD camera records the laser probing. The distance and angle between the CCD camera and the laser are known and form a triangle. Using triangulation mathematics the facial elevations can be computed. The laser probe is recorded on the camera x, y field of view, and x, y coordinate locations are based on the distance from the camera to the probed surface. The face is sequentially scanned by a laser light source. Digital cameras monitor the illumination and triangulation geometry allows depth information to be calculated. Laser scanning of a face can take up to 30 seconds and this relatively long scan time creates the opportunity for patient motion with loss of coherency between measurement points.

Structured Light

A structured light camera is comprised of a pattern projector and a digital camera that are spatially offset. Structured light patterns (usually white light) such as grids, dots, or stripes, are projected onto the subject while the digital camera takes an image of the subject. The reconstruction software is initially calibrated with the spatial position of the camera and the specifics of the projected light pattern. The distortion of the light pattern is then analyzed by the software and the 3-D shape is inferred from the scale of the visible distortion. A single camera technique, because of line of sight limitations, cannot

FIGURE 1. This figure shows the active stereophotogrammetry 3dMDface system created by 3dMD (Atlanta, Ga.) and installed using a wall mount strategy at DDI in Roseville, Calif. The system has the option of being portable. The 3dMDface system uses four-machine vision 2 Mpixel sensors for capturing geometry and two-machine vision sensors for capturing texture. The options for the texture sensors are either 2 or 5 Mpixels each. The six sensors simultaneously capture the facial anatomy in 1.5 milliseconds. The capture volume can be up to 0.5 x 0.6 x 0.5 meters. The accuracy has been tested within 0.2 mm depending on the lens configuration.[5]

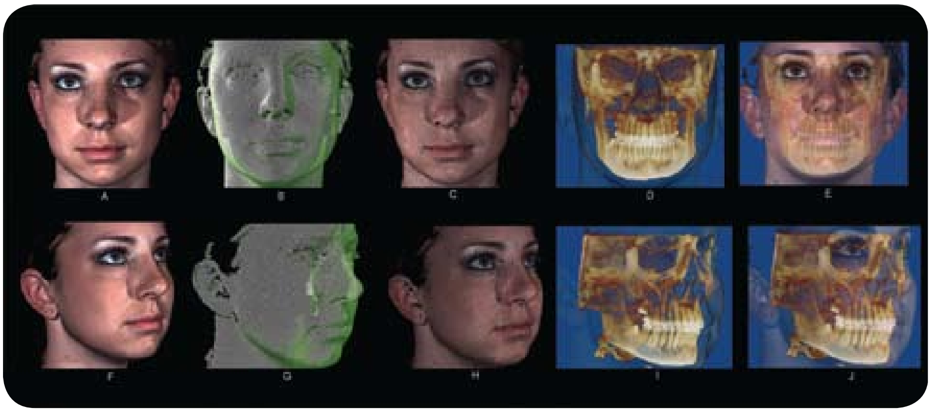

FIGURE 2. This figure shows a face image that was captured with 3dMDface using a range of 3-D visualization options. A and F show a photo-textured face from two points of view. B and G show the geometry of the face rendered as a polygon mesh. The stereophotogrammetry reconstruction algorithm creates a point cloud of approximately 100,000 surface locations. The polygon mesh was created from this point cloud using the points as vertices. C and H show the photo-textured polygon mesh using a volume render method of display. D and I show the polygon mesh of the face registered to the CBCT. The mesh and skeleton were registered using Dolphin 3-D software and displayed using a volume-rendering technique. E and J are a volume-rendering display created by Dolphin 3-D showing the 3dMD facial image registered to the CBCT generated by an Imaging Sciences (Hatfield, Penn.), iCAT machine.

acquire a full facial map of the patient with a single exposure. A full facial map requires multiple acquisitions and this reduces the accuracy and precision of the technique because the multiple images need to be stitched onto the same coordinate system and the time between exposures creates the opportunity for patient motion.[1,2]

Active Stereo Photogrammetry

Multiple calibrated cameras are spatially arranged with overlapping FOVs but different POVs that simultaneously capture the facial anatomy. This is an optical technique that does not require a structured pattern projection but uses a regular photographic flash to illuminate the face. Stereo photogrammetry uses image analysis algorithms to identify and match unique external surface features between the two photographs to generate a composite 3-D model by triangulating the points. The initial calibration procedure informs the analysis software of the precise 3-D location of each camera sensor.

The accuracy and efficiency of the analysis software can be improved by projecting a flat random pattern onto the subject to aid in the correspondence between the homologous image regions. The pattern on the surface provides the stereo localization algorithms with the base information required to build accurate geometry. The random pattern combined with the natural skin texture gives the image analysis software more detail to perform the triangulation.

Once the 3-D geometry model has been produced, the software maps the color texture information onto the model. Active stereo photogrammmetry has several advantages for clinical practice that include capture speed, a larger number of simultaneous viewpoints, and the ability to accurately compute the location of any derived point. Active stereo systems are capable of capturing the facial anatomy in less than 2 milliseconds and can compute approximately 100,000 facial coordinates[3-5] (**FIGURES 1 AND 2**).

Volumetric Imaging

Computed tomography, CT; cone beam CT, CBCT; and magnetic resonance images, MRI, are forms of volumetric imaging that produce 3-D information about the surface and subsurface anatomy. These volumetric techniques have been shown to be invaluable for diagnostic imaging but have some limitations when used for routine imaging of the surface anatomy of the face. All volumetric methods are expensive, may be associated with some risk to the patient, may not always be acquired with an optimal head position, may have distorted facial soft tissues secondary to stabilization devices, may have patient motion artifact, and may be limited to a small FOV of view, and therefore have incomplete information.

The fusion of 3-D surface imaging with volumetric imaging and associated modeling can provide the best of both technologies. This fusion occurs using multiobject software by registering or fusing the surface and volumes data onto the same coordinate system. Fusing the surface into the same 3-D coordinate space with volume data corrects volume artifacts created by patient motion and soft tissue compression, supplements missing volumetric data, and allows for correct display at a corrected head posture. The fused image data creates a patient-specific model that can be enhanced to add value to the original image data.

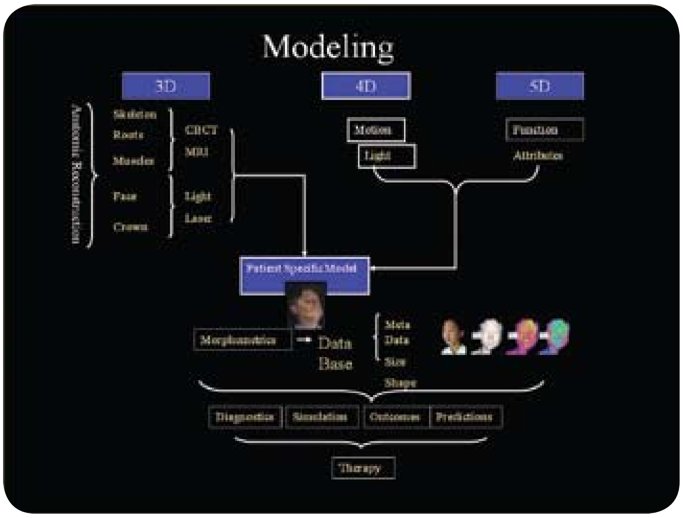

FIGURE 3. This figure illustrates the anatomic reconstruction of a 3-D patient-specific model using disparate imaging sources and fusing them on the same 3-D Cartesian coordinate system. The patient-specific model can evolve to a 4-D model by fusing a timed sequence of 3-D images onto the same 3-D coordinate system. Four-D systems can be used to evaluate change over time. Four-D modeling includes monitoring growth, development, jaw movement, facial expression, and treatment outcomes. Five-D modeling allows for the fusion of biomechanical attributes into the coordinate system and testing the biomechanical relationships between the structures. Information can be collected from the 3-D, 4-D and 5-D models, and stored in a database for retrieval and analysis. The data pool can be used for diagnostic, treatment simulation, outcomes analysis, outcome predictions and therapy.[6]

Modeling

Three-D imaging creates the opportunity for 3-D anatomic reconstruction and analysis. A 3-D image volume has a global reference or coordinate system (Cartesian coordinates) that is displayed as three orthogonal planes (axial, coronal, and sagittal). The coordinate system is often assigned to the anatomic volume by the acquisition device but can be modified later by the user with specialized software tools.

Multiple image sets can be combined into the same 3-D matrix. The process of combining these images into the same coordinate matrix is called fusion or registration. For example, a 3-D surface acquired using a visible light or laser scan can be fused onto a common coordinate matrix with a 3-D volume acquired using CBCT. Following the fusion of the two objects (surface and volume) they can be displayed, analyzed, and visualized together. The fusion of objects onto a common coordinate system can improve the accuracy and completeness of the anatomic representation. A process called anatomic segmentation can also add value to the image set. Segmentation creates an anatomic object that can be used for morphometric analysis, simulation, and biomechanical testing. For example, segmented objects may include individual teeth, mandible, maxilla, skin, and airway.

The objects are displayed and managed as rendered iso-surfaces. Each object may be assigned a local coordinate system. The global or original coordinate system monitors the position of each object using 6 degrees of freedom, DOF. The six DOF for each object are x, y, z, yaw, pitch, and roll. Fusion can also occur in 4-D. Fusion in 4-D occurs by spatially managing the object coordinate systems in a timed sequence of 3-D images. For example, the position of mandible, including the TMJs, relative to the maxilla and temporal bone, can be tracked and displayed by managing the local and global coordinate systems over time. This would allow for the creation of a virtual articulator (**FIGURE 3**).

Anatomic reconstructions can be used to create patient-specific models that can be analyzed. These models provide a visual representation of the patient and can also be used to measure size and shape of the selected attributes. The analysis data can be stored in a database for future reference. The database can later be analyzed to determine outcome values and develop prediction tables. Multiobject models can be used for treatment simulations. Treatment simulations allow the operator to iterate treatment options or rehearse a treatment (**FIGURE 4**).

Anatomic reconstructions create the opportunity for a systems or integrated diagnostic approach. This approach allows for the analysis and consideration of anatomically related structures. A developmental disturbance of the TMJ (i.e., arthritis, fracture) may have a local effect on the ipsilateral joint and a regional effect on that side of the face. The joint pathology may limit or stop growth of the affected condyle. In addition, there may be a growth reduction in the vertical dimensions of the neck, ramus, and body of the mandible. The occlusal plane may be elevated on the affected side. The lateral development of the mandible may be reduced and the cranial base (fossa) may be depressed on that side. The limited growth of the mandible may alter the occlusion, maxillomandibular spatial relationships, facial profile, facial growth pattern, and the airway shape and size.

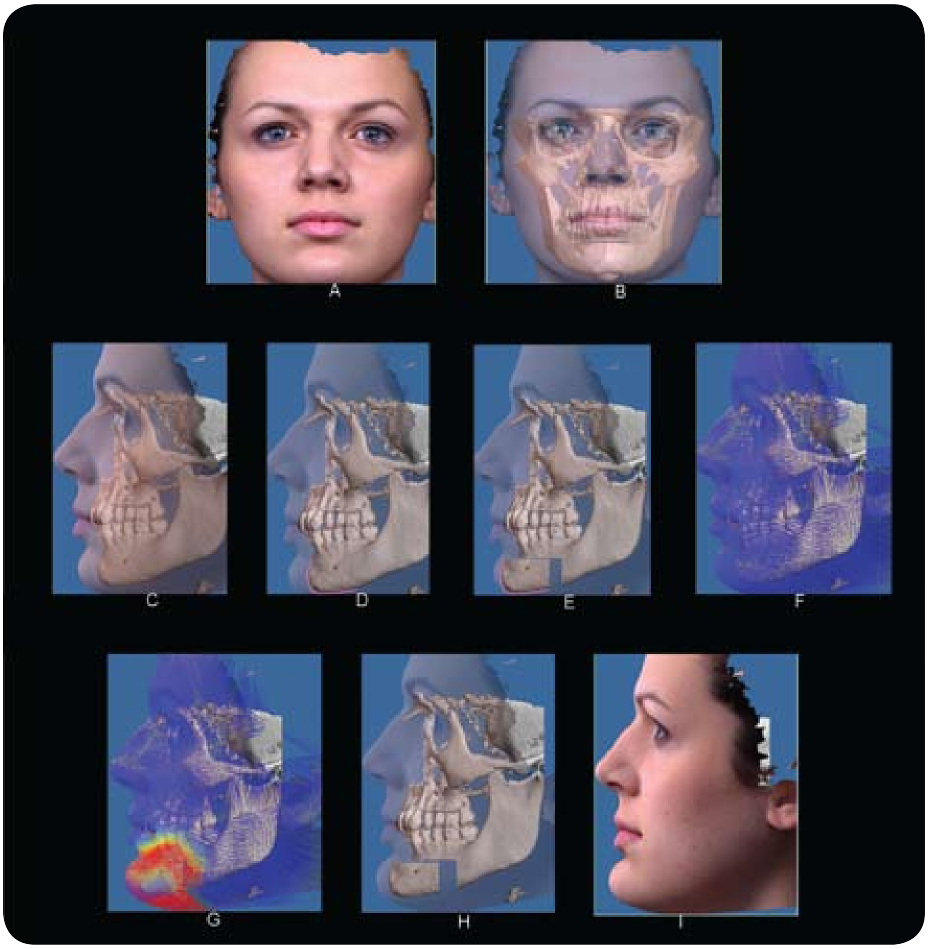

FIGURE 4. This figure illustrates functional (5-D) modeling on a patient-specific model. The software environment for generating the patient-specific model and the simulation was Vultus by 3dMD. The 3-D facial image was created by 3dMD's active stereophotogrammetry system (**FIGURE 1**). The skeletal data was produced by Imaging Sciences Int., iCAT. The skeleton was segmented in Vultus and fused with an imported facial image of the same patient to create a patient-specific model (A-B). Elastic properties through the use of a mass spring model (approximately 400,000 programmable springs and coils attached the skin to the skeleton) were applied to the model to help simulate the biomechanical relationships between the skeleton and overlying facial soft tissues (F-G). This simulation model was created by 3dMD. Using an osteotomy simulation tool, a region of the mandible was selected and outlined with a pink line (D-E). Note, the osteotomy is not anatomically correct and is only being shown as a simulation concept. The selected osteotomy segment was advanced (E). The advanced segment perturbed the springs and coils and they applied a force to the surface mesh (G). The surface mesh was deformed by the model to simulate the effect of the mandibular osteotomy (H-I). (Courtesy of Dr. William Harrell, Jr., Alexander City, Ala.)

Conclusions

The ultimate goal of imaging has been to accurately represent the anatomy, and this can be achieved for surface imaging of the face using 3-D methods, such as stereophotogrammetry systems. Software provides a multiobject environment where disparate 3-D image data can be combined into the same 3-D coordinate system and create a patient-specific model that is comprised of surface and subsurface structures. The accuracy and precision of 3-D facial images provides the opportunity for valuable clinical applications. These 3-D facial photography systems are currently being used for a variety of applications including monitoring patients through the process of growth, development and treatment, cleft lip and palate, orthodontics, orthognathic surgery, and craniofacial anomalies. ■■■■

REFERENCES

1. Enciso, R, Shaw A, et al, Three-D head antropometric analysis, in proceedings of the International Society for Optical Engineering SPIE Medical Imaging, 5029:590-7, 2003.
2. Carnicky J, Chorvat D, Three-dimensional measurement of human face with structured light illumination, *Measurement Science Rev* 6 section 2(1):1-4, 2006.
3. Adams GL, Gansky SA, et al, Comparison between traditional 2-D cephalometry and a 3-D approach to human dry skulls. *Am J Orthod Dentofacial Orthop* 126(4):397-409, October 2004.
4. Lane C, Harrell W, Completing the 3-D picture. *Am J Orthod Dentofacial Orthop* 133(4):612-20, April 2008.
5. Aldridge K, Boyadjiev SA, et al, Precision and error of 3-D phenotypic measures acquired from 3dMD photogrammetric images. *Am J Med Genet A* 138A(3):247-53, Oct. 15, 2005.
6. Usui T, Maki K, et al, Measurement of mechanical strain on mandibular surface with mastication robot: Influence of muscle loading direction and magnitude. *Orthod Craniofac Res* 6suppl 1:163-7, 2003.

TO REQUEST A PRINTED COPY OF THIS ARTICLE, PLEASE CONTACT David C. Hatcher, DDS, MSc, Diagnostic Digital Imaging, 1 Scripps Drive, Suite 101, Sacramento, Calif., 95825.