

3-D RECONSTRUCTION OF THE HUMAN JAW USING SPACE CARVING

Moumen T. Ahmed, Ahmed H. Eid and Aly A. Farag*

Computer Vision and Image Processing (CVIP) Laboratory.
University of Louisville
Louisville, KY 40292

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. The space carving algorithm is used to reconstruct the shape of the human jaw. This algorithm provides more flexibility and eliminates several constraints that imposed by other approaches like stereo. The system performance is investigated, and the results show acceptable reconstruction accuracy.

1. 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 [3], tuned-aperture CT [4], and localized, or "cone-beam," computed tomography [5]. 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.

In this paper a developed system is presented for dentistry to replace traditional approaches in diagnosis, treatment planning, surgical simulation, and prosthetic replacements ([6, 7, 8]). The objectives of this system are:

1. 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.
2. to develop methods for accurate 3-D reconstruction from the acquired sequence of intra-oral images.

The first goal is achieved by designing a setup consists of a small calibrated intra-oral CCD camera, with a built-in white light. The camera is mounted on a five link 3D digitizer arm. The arm gives the necessary transformation matrices and hence the the projection matrices required to preserve the camera calibration while moving. To achieve the second goal, several different approaches have been examined. The stereo approach has been by far the most widely used. However, One of the problems of the stereo is the correspondence problem. In addition, the stereo approach relies heavily on assumptions such as the existence of specific features in the images to produce satisfactory results. Therefore, the stereo approach is inefficient regarding objects like teeth which have no obvious geometric features. In addition, the constraints on the viewpoints make the stereo inapplicable for the jaw reconstruction. In our previous approach [1], we used the shape from shading (SFS) algorithm. However, the SFS does not provide metric information. The reconstruction results from

*This research was partially supported by grants from NSF (ECS 9505674) and The Department of Defense (USNV N00014-97-11076).

the SFS was fused with several metric measurements to enhance its performance. Here, we examined the performance of a newly published algorithm; the space carving [2]. The space carving does not impose any constraints on either the geometry of the object, the camera position, or even the texture. Therefore it is more suitable for our application.

This paper is organized as follows: Section II describes the overall system setup. Sections III presents our approach to teeth reconstruction using the Space Carving algorithm. The experimental results are shown in Section IV followed by the conclusion in Section V.

2. SYSTEM SETUP

The experimental setup shown in Figure 1 consists of the following:

1. A 3D digitizer.
2. A CCD color camera.
3. A 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.
4. A machine that hosts the software required for the data processing, reconstruction and visualization of the 3D jaw model.

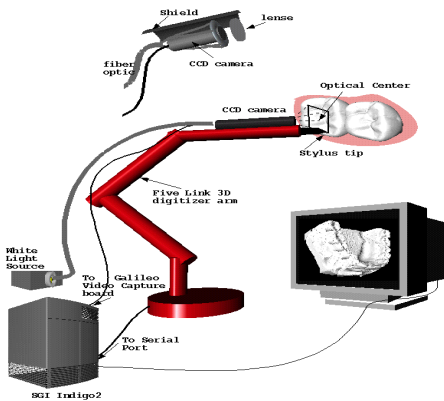


Fig. 1. The experimental setup of the proposed system consists of a CCD camera mounted on a five link 3D digitizer. A DC regulated light source is connected to a fiber optic bundle surrounding the camera. The setup is connected to an SGI Indigo2 machine.

The CCD camera is mounted on the stylus of the 3-D digitizer and is then calibrated as shown in Figure 2.

Camera calibration is done once [10] 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 M_{oc} can be tracked as the camera moves, and the 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. In addition, the arm provides the position of the initial volume enclosing the tooth or the jaw. This volume is carved by the space carving algorithm until the shape is reconstructed

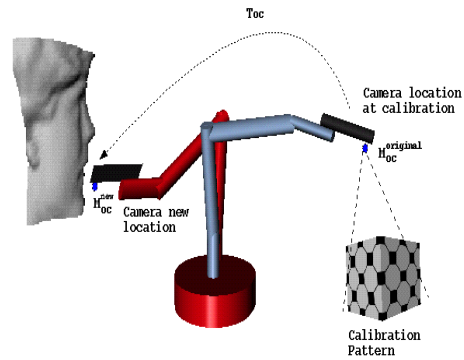


Fig. 2. Illustration of the camera self calibration. The coordinates measuring system is used to find the transformation matrix T_{oc} between the optical center M_{oc} of the camera at the calibration time and the new location while acquiring the images. This transformation matrix is used to obtain the new camera perspective projection matrix.

3. SPACE CARVING

Image formation is typically achieved by projecting a 3D shape onto a two-dimensional plane. In 3D object reconstruction, we attempt to achieve the reverse process of image formation by regenerating a 3D shape from various 2D projections. However, there is no one-to-one relation between 3D shapes and their projections. Several 3D shapes can have the same 2D projections. As an illustration, a sphere and a cylinder can all have the same 2D projections from certain view-

points. This ambiguity can be resolved by using more 2D projections to estimate a 3D shape. Space carving attempts to produce the maximal 3D shape that is consistent with all the images.

Space carving starts with an initial volume, V , that includes the scene to be reconstructed. This 3D space is then discretized into a set of voxels. The idea is to successively carve (remove) some voxels until the final 3D shape, V^* , agrees with all the input images.

Each voxel in the initial volume is projected back to the different images using their respective projection matrices. To decide whether a voxel should be carved or not, the idea of color-consistency is used. The Lambertian model for the surface of the object is assumed. Under this model, light reflected from a single point on the surface of the object has the same intensity in all directions. Therefore, for a voxel to belong to the surface of the object, it must have the same color intensity for all its projections to the different images provided. Voxels that are inconsistent with a single color, are viewed as free space in which different light rays intersect. By removing all color-inconsistent voxels, we are able to approximate a maximal photo-consistent shape that is defined by all the input images. The basic idea of space carving is illustrated in Figure 3. Three input images are used to generate the 3D model of the shape shown in the images. Voxels that project on the input images to pixels of similar color are kept and assigned that color. Voxels that project on the input images to pixels of different colors are removed. Although the general idea in space carving is straightforward, modeling an algorithm to provide the desired results is not an easy task as the problem of occlusion must be treated. Given N input images and their respective projection matrices, the algorithm must be able to guarantee convergence to the maximal photo-consistent shape. Eventhough, the space carving approach requires segmentation of the object and the background, this not seems a problem with objects like teeth. Actually, teeth are brighter than the interior of the mouth. Therefore, they can be segmented easily from the interior of the mouth and the gum.

4. EXPERIMENTAL RESULTS

We applied the space carving algorithm to a sequence of images of the human jaw. About 3-5 images are used for each tooth. This number of images is showed to be enough. We also used a 15% standard deviation threshold of the grayscale values to determine whether or not the voxels should be declared photo-inconsistent and consequently carved. This relatively high threshold was chosen in order to compensate for illumination

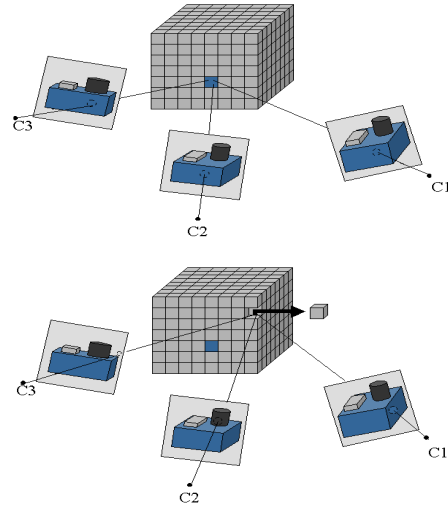


Fig. 3. Basic idea of space carving. Voxels are projected to the input images using their respective projection matrices. C_1 , C_2 and C_3 represent the optical centers of the three cameras. (a) Consistent voxels are assigned the color of their projections. (b) Inconsistent voxels are removed from the volume.

effects and errors in calibration. The initial volume was discretized into $70 \times 70 \times 70$ voxels for a total of 343,000 voxels. The object is reconstructed after 4-5 passes of the space carving algorithm, which took about 12 seconds on SGI Indigo2 machine. On the average, the final volume contained 4,500 voxels. Figure 4 shows the input images and the reconstruction results. Figure 4(f) shows a group of teeth that are registered to form a jaw segment. This segment took about 42 second of processing time for reconstruction and registration. Quantitative assessment of the reconstruction is obtained by comparing some tooth measurements (e.g., height and width) from the reconstructed model to those of the real tooth. The comparison showed accuracy within 0.47 mm, which shows that the system can achieve sub-millimeter accuracy, similar to our previous approach, without the need of any range measurements. The interested reader is referred to [11] for more details.

5. CONCLUSIONS AND FUTURE EXTENSIONS

The 3-D reconstruction of the human jaw has tremendous applications. To reconstruct the human jaw, we used the space carving algorithm. Our choice of space carving algorithm is guided by several concerns. Reconstruction using space carving does not rely on assumptions of the presence of specific image features, or on

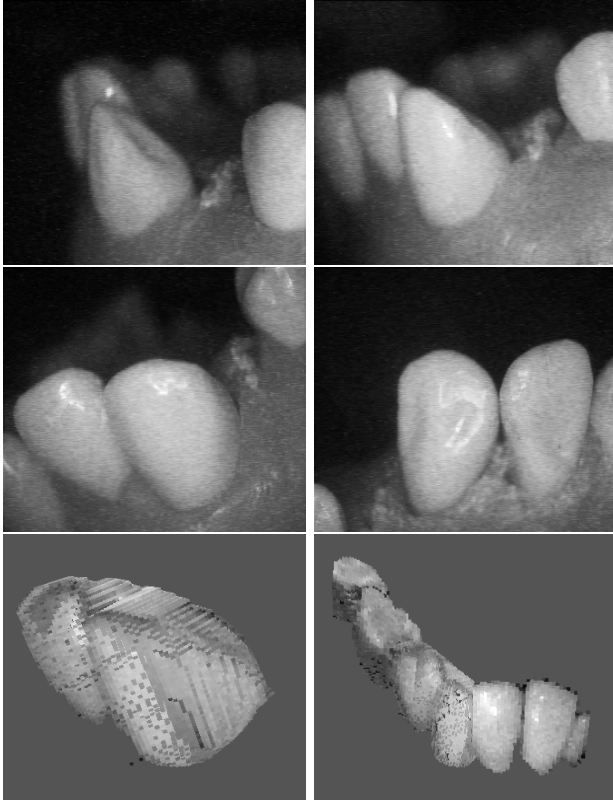


Fig. 4. (a)-(d) images of two teeth acquired by the CCD camera. (e) The reconstructed teeth. (f) A group of teeth

the position of the camera, or the object geometry and texture. This technique allows us to reconstruct accurately 3D models that agree with the input images without need to solve the difficult correspondence problem. We currently do more experiments on selecting a proper initial volume for the tooth, since we know a priori information about the shape of the tooth. This is expected to enhance the results of the space carving algorithm. Additional experiments are being conducted to compare this approach with the SFS approach.

6. REFERENCES

- [1] Sameh M. Yamany, Aly A. Farag, David Tasman and Allan G. Farman "A 3-D Reconstruction System for Human Jaw Using a Sequence of Optical Images" *IEEE Trans. on Medical Imaging*, vol. 19(5), pp. 538-547, May 2000.
- [2] K. Kutulakos and S. Seitz. "Theory of Shape by Space Carving." *Proceedings of IEEE International Conference On Computer Vision*, Corfu, Greece, pp. 307-314, Sept. 1999.
- [3] P. F. van der Stelt and S. M. Dunn, "3D-imaging in dental radiography," in *Advances in Maxillo-facial Imaging*, A. G. Farman, ed., pp. 367-372, Elsevier Science B. V., 1997.
- [4] 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.
- [5] 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.
- [6] 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.
- [7] 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.
- [8] 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.
- [9] 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.
- [10] 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.
- [11] Moumen Ahmed, Ahmed Eid, and Aly Farag, "3D Reconstruction of Human Jaw: New Approach and Improvements," *Technical Report*, CVIP Lab, University of Louisville, <http://www.cvip.uofl.edu>, March 2001.