3D Reconstruction of the Human Jaw: A New Approach and Improvements

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

Computer Vision and Image Processing Laboratory
University of Louisville, KY 40292
moumen, ahmedh, farag@cvip.uofl.edu
http://www.cvip.uofl.edu

Abstract. This paper presents a new, practical approach for 3D reconstruction of the human jaw from a sequence of intra-oral images. This research has an immense value in various dental practices including implants, tooth alignment, and craniofacial surgery. Our approach is based on the recently-proposed space carving algorithm for shape recovery. This algorithm provides more flexibility to the reconstruction process and eliminates several constraints imposed by other traditional approaches such as stereo and shape from shading. Our experimental results have shown that the approach is able to reconstruct 3D models of the human jaw with sub-millimeter accuracy.

1 Introduction

Orthodontic treatment involves the application of force systems to teeth over time 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 plastic models (casts), photographs and radiographs, a process which is both costly and time consuming. Obtaining a cast of the jaw is a complex operation for the orthodontist, an unpleasant experience for the patient and may not provide all the details of the jaw. Current technology in dental radiography can provide the orthodontist with 3D information of the jaw. While dental radiology is now widely accepted as a routine technique for dental examinations, the equipment is rather expensive and the resolution, being adequate for maxillofacial imaging, is still too low for 3D dental visualization. Furthermore, the dose required to enhance the resolution is unacceptably high. Some efforts have been devoted to computerized diagnosis in orthodontics, e.g., [2,7]. Usually, most of these 3D 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 3D information from that model. User interaction is needed in such systems to determine the 3D coordinates of fiducial reference points on a dental cast. Other systems that can measure the 3D coordinates have been developed using either mechanical contact [8] or a traveling light principle [9].

Our research lab has been involved for the last five years in a project, the jaw project, to develop a system for dentistry to go beyond traditional approaches.
same setup except for the laser projector that we do not use in this paper on purpose.

The CCD camera is mounted on the stylus of the 3D digitizer and it has to be calibrated before its use. Because of the small focal lens of the camera, it suffers from some lens distortion (mainly radial distortion). The first coefficient of radial distortion is calibrated using a straight line-based technique [12] so that all acquired images can be undistorted before processing. After correcting for lens distortion, the camera can be safely modeled as an ideal pinhole camera, whose perspective projection matrix that encompasses the camera intrinsic and extrinsic parameters can be calibrated using a non-linear approach [11]. 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. Being mounted on the digitizer arm, the camera location in the 3D space can be measured. Moreover, the arm provides the transformation that relates the new position and orientation of the camera to the world coordinate system. This transformation is used to update the camera extrinsic parameters and thus the camera perspective projection matrix. As such, the camera is maintained calibrated in all positions. In addition, the arm provides the position of the initial volume enclosing the tooth or any part of the jaw, which is to be carved by the space carving algorithm until the shape is reconstructed.

The five degrees of freedom provided by the arm enable the acquisition of a sequence of intra-oral images covering different parts of the jaw. Using the space carving algorithm, sets of voxels that represent the different parts of the jaw are computed. A fast registration technique [13] is employed to merge the resulting 3D models to obtain a complete 3D description of the jaw parts. The final stage transforms this model into patches of free form surfaces using a triangulation technique. This step enables the development of a 3D 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.

3 Space Carving

In 3D object reconstruction, we attempt to achieve the reverse process of image formation by regenerating a 3D shape from various 2D projections. Space carving [3] 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 object(s) to be reconstructed. This 3D space is then discretized into a finite set of voxels \( v_1, v_2, ..., v_n \). The idea is to successively carve (remove) some voxels until the final 3D shape, \( V' \), agrees with all the input images. An outline of the algorithm is given below.

Space carving Algorithm:
Step 1: Initialize \( V \) based on arm position and discretize it.
Step 2:
- Determine the set of voxels \( \text{Vis}(V) \) on the surface of \( V \).
- Project each voxel \( v \) on \( \text{Vis}(V) \) to the different images where \( v \) is visible.
- Determine the photo-consistency of each voxel \( v \) on \( \text{Vis}(V) \).

Step 3: If no non-photoconsistent voxel is found, set \( V^* = V \) and terminate. Otherwise, set \( V = V - \{ \text{non-photoconsistent } v \text{'s} \} \) and return to Step 2.

Each voxel on the surface of the volume, i.e., in \( \text{Vis}(V) \), 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, within some tolerance to allow for some light variations and some calibration inaccuracy, 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 2. 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. This is carefully taken care of in a multi-sweep fashion [3]. One important requirement by the algorithm is segmenting the objects from the background. This does not represent a problem with objects like teeth, since teeth are brighter than the interior of the mouth and have distinctive color from the gum. Therefore, teeth can be easily segmented from the background in the image sequences.

![Fig. 2. Basic idea of space carving. Voxels are projected to the input images using their respective projection matrices. C1, C2 and C3 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.](image-url)
4 Experimental Results and Validation

After calibrating the camera, sequences of images are captured for overlapping segments of the jaw. Each segment consists of about 3-5 images. The process of taking the images was relatively fast, taking less than a minute for each segment and a total of 10-12 minutes to cover the upper/lower jaw. The patient’s jaw should not move during the acquisition time of each segment (less than a minute). However, movements between segments are permitted because the registration technique can align the individual segments. Figure 3 shows some images taken of a patient’s jaw segment. We applied the space carving algorithm to the acquired images of each segment. The initial volume was selected as a cube and discretized into $70 \times 70 \times 70$ voxels for a total of $343,000$ voxels. Each segment is reconstructed after 4-5 passes of the space carving algorithm, which took about 16 seconds on an SGI Indigo2 machine. On average, the final volume contained 4,500 voxels. We 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 calibration errors, possible light changes from one image to another and any deviation from the Lambertian assumption (e.g., presence of specularities in some images). Figure 3 shows also the reconstructed result of one segment of the lower jaw from two different views. Once each segment of the jaw is reconstructed, they are registered [13] to compose the whole jaw, of which a part is shown in Figure 4 from two different views. The part shown in the figure took about 84 seconds of processing time for reconstruction and registration. Quantitative assessment and validation 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 as shown in Figure 5. The comparison showed accuracy within 0.47 mm, which shows that the system can achieve sub-millimeter accuracy, similar to our previous approach but without the need for any range measurements.

5 Conclusions and Future Extensions

The 3D reconstruction of the human jaw has tremendous applications. To reconstruct the human jaw, we used the space carving algorithm which does not impose any constraints on the object geometry, the position of the camera, or the texture. This gives our approach a major, practical advantage over traditional shape recovery techniques such as SFS and stereo. Moreover, this algorithm successfully reconstructed 3D models of the human jaw with sub-millimeter accuracy, which is as accurate as our previous approach [5], but without using any range measurements or laser projectors. One key advantage of this new approach is the fact that it can easily exploit any available a priori information about the shape of the tooth. The algorithm can thus start with an initial volume whose shape is closer to the target object. This is expected to enhance and speed up the results of our approach. Our current research direction is directed to investigating this possibility.
Fig. 3. Reconstruction results of a jaw segment consisting mainly of two teeth: (a)-(c) images of the two teeth acquired by the CCD camera. (d) and (e) The reconstructed teeth shown from two different views.

Fig. 4. Reconstructed part of the lower jaw shown from two views.

Fig. 5. Validation of the reconstructed model: (a) measurements from the real jaw, (b) measurements from the reconstructed model.
6 Acknowledgment

This work was supported in part by grants from the Whitaker Foundation, and the NSF (ECS-9505674) institutions.

References