Hair Shape Modeling from Video Captured Images and CT Data

Ali, Md. Haider and Toyohisa Kaneko
Toyohashi University of Technology
1-1, Hibarigaoka, Tempaku-Cho, Toyohashi City 441-8580, Japan
kaneko or haider@mmip.tutics.tut.ac.jp

Abstract

We propose a three-dimensional (3D) reconstruction method of human hair-shape from rotating head multiple video captured images and CT data. It is well known that no hair is present on the polygonal skin surface of the human head (*3D-head*) reconstructed from CT or MRI data. Our task is to reconstruct and add the hair-shape on the 3D-head to create a realistic human head model for simulating post-surgical facial expressions. Using a sculpturing technique based upon rotating head images we propose a method of reconstructing the hair-shape with the help of 3D-head. We have utilized binarized voxel data of the 3D-head (*solid-head*) in this regard. The sculpturing object in our definition is the solid-head surrounded by assumed thick hair-voxels. We sculpture the surrounding hair-voxels according to the extracted hair-region from the video captured images while keeping the internal solid-head intact. We reconstruct the concave and semi-occluded regions by digging up to the visible skin surface of the solid-head in/near the hair region. We define *complete-head* as the 3D polygonal surface obtained from solid-head including the residue-sculptured hair-voxels on it. Experimentally we have shown that our method can successfully reconstruct the concave and semi-occluded regions in the skin-hair junction regions, which is not easy to reconstruct by the conventional way.

Key words: realistic modeling, 3D head modeling, 3D hair modeling, visualization.

1. Introduction

The 3D reconstruction techniques build real world into computational models, which is urgently required in virtual reality, CAD/CAM, and other related fields[1]. In medicine, 3D modeling is very essential especially in the field of computer-integrated surgery (e.g., surgery simulation and/or image guided surgery). Our task is to reconstruct a realistic complete head model for simulating post-surgical expressive faces. We have already proposed methods for making a realistic model face by precisely pasting blended colors from three photographs to the 3D facial skin surface derived from CT data[2,3]. In those proposals we emphasized on the 3D-2D projective registration and also our interest was limited to the facial part only.

Absence of hair actually fails to provide a complete head model, especially in the simulated post-surgical expressive faces. Without hair, a human face does not look realistic and moreover is not sufficient to create a vision-convincing animation. It is known that no hair is present on the 3D-head reconstructed from CT or MRI data. Even a Cyberware Digitizer™ scanner cannot reconstruct hair well (which is usually black). That is to say, there is no commercially available instrument to reconstruct the hair shape. In this paper we propose a method of reconstructing and adding 3D hair-shape on the CT/MRI reconstructed 3D-head to make a realistic complete-model.

In this paper we use four different names of 3D data of human head. The definitions are given here for further clarification.

- **3D-Head:** The hairless polygonal-data (skin surface) of the patient's head derived by the marching cubes algorithm from the original 12 bit gray level CT slices using a skin-air threshold.
- **Solid-Head:** Binarized voxel-data of the 3D-head. This can be obtained by any of the two following ways: (1)by filling the 3D head, or (2)from the same set of CT slices using identical threshold value as used in 3D-face reconstruction. In the former case there is a possibility to yield shape-error in the multiple layer regions, e.g., ears and nostril as it is not like a simple polygon filling. The latter way on the other hand easy to implement and there is no possibility of yielding shape-error.
- **Complete-Head:** The polygonal-data obtained from the solid-head covered with residue-sculptured hair-voxels by the marching cubes algorithm.
- **Final-Head:** This is actually the complete-head but to get better surface quality, the uncovered skin surface portion is replaced by the 3D-head surface

1.1 Literature Review

There are a number of works reported on reconstructing 3D shape from a sequence of 2D views and/or silhouettes[1,4,5,6,7,8]. In almost all cases the target is to reconstruct 3D shape correctly, especially the concave or un-exposed parts. L. Zhau and W. Gu[1] used a laser range sensor in conjunction with a sequence of images in this regard. J. Zheng and F. Kishino[4] proposed a technique of detecting un-exposed regions while reconstructing 3D shape from sequential image silhouettes. They employed a filter for detecting non-smooth points in the silhouette distribution. S. Sugimoto and M. Okutomi[5] proposed a technique of estimating radii of rotating points on the object surface using spacio-temporal images. To determine the missing radius data they fitted the obtained data with a suitable sine curve. There are also some proposals of reconstructing a 3D face by modifying generic facial geometry according to the photographs[10,11].

1.2 Our Method

Our method of reconstructing 3D hair-shape is simple but different from the works mentioned in the above. Insertion of solid-head into the sculpturing object and selection of hair-region silhouettes instead of complete head from the 2D image is the distinction from all other related works. Keeping the solid-head intact while sculpturing surrounded voxels to reconstruct the concave and semi-occluded regions in the resulting complete-head is a new way of 3D reconstruction.

1.3 Paper Organization

The remainder of this paper is organized as follows: Section 2 gives the outline of the proposed reconstruction algorithm. Section 3 describes the basic requirements for making a correct hair-shape. The sculpturing procedure is given in section 4. Experiments in section 5 and the conclusion and future work plan are in section 6. Finally the acknowledgement is in section 7.

2 Outline of the Reconstruction Algorithm

Fig.1 shows a flow-chart of the proposed reconstruction algorithm. A brief description of Fig.1 is given below.

The CT data provides two basic input data: (1)3D-head (polygonal-data), and (2)solid-head (voxel-data). The solid-head is the main part of the sculpturing-object. The sculpturing object is a 3D rectangular box filled with assumed hair voxels and the solid-head is placed at the center of the box. The shape and orientation of the 3D-head and the solid-head are almost identical. 3D-head can be said to be more accurate since it has subvoxel accuracy. The 3D-head provides necessary information in the form of 3D-edge to determine the camera parameters for each video captured image. Each video captured image also provides two input data for the reconstruction: (1)2D-edge, which is required for 3D-2D registration to determine the camera parameters of the video captured image, and (2)hair-region i.e., the extracted 2D hair-shape. Edge based registration in Fig.1 helps to determine the camera parameters of the video captured images by matching the projected 3D-edge with the corresponding 2D-edge.

At the hair-sculpturing stage, for each video captured image the result of edge based registration helps to position the virtual camera, which focuses towards the sculpturing-object. The technique of 2D hair-shape

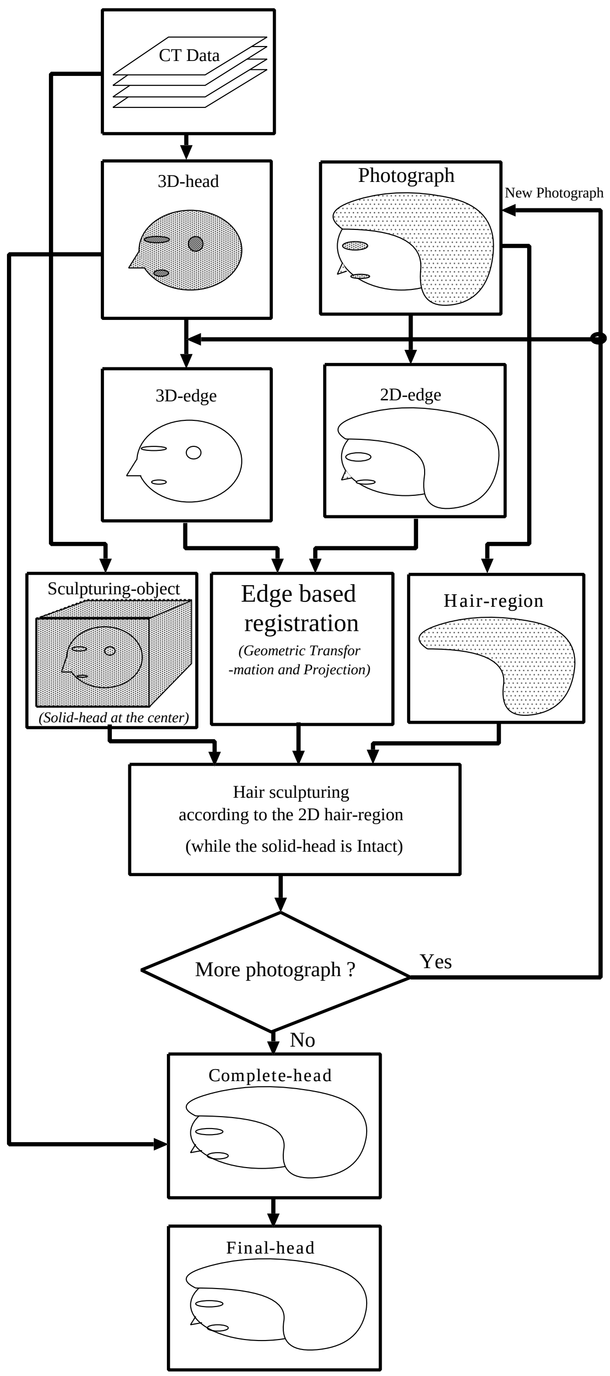


Fig.1 Flow-chart of the reconstruction method

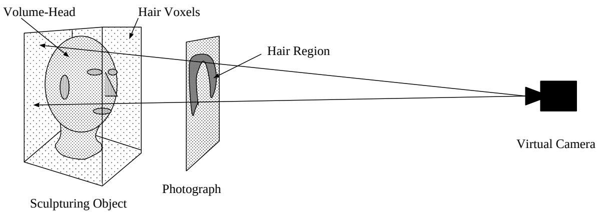


Fig.2 The technique of selecting hair-voxels in the sculpturing object.

re-projection on the sculpturing-object is shown in Fig.2.

Fig.3 The 2D virtual camera image of the solid-head (the white part and the gray region visible under hair-region). The superimposed black portion is the extracted hair-region from the original video captured image. The hair-voxels in the sculpturing-object has been divided in different groups according to this virtual camera image.

All the hair-voxels outside the re-projected ray-lines of the hair-region are cutout or removed. The hair-voxels, which are outside the hair-region but on the ray-lines to the solid-head (white region in Fig.3), are removed up to the solid-head surface. Fig.4 shows the sculpturing stage as a trans-axial view. The removed hair-voxels on the ray-lines to the uncovered solid-head actually leads to create concave and semi-occluded regions.

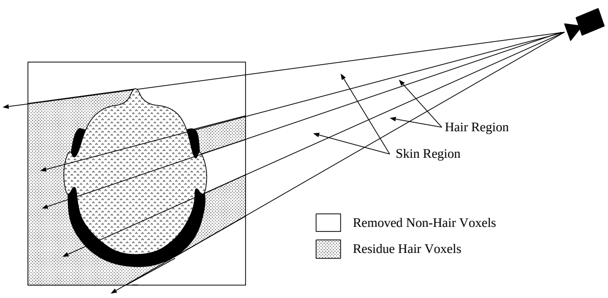


Fig. 4 Sculpturing technique from the top

The resulting hair-shape on the solid-head can be obtained by repeating the whole procedure for the rest of the video captured images. The complete-model is the polygonal surface derived from the solid-head including sculptured residue hair-voxels on the solid head, and the final-head is obtained by replacing the uncovered skin with that of the 3D-head.

3 Basic Requirements for Reconstruction

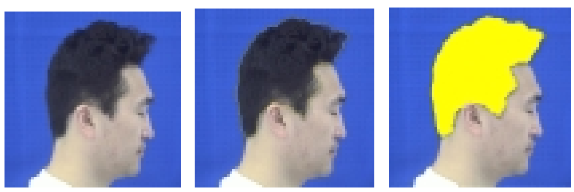
It should be noted that in this paper we emphasis on reconstructing the hair region as accurately as possible instead of the entire head. In our definition the target head (i.e., complete-head) is the combination of the hair-shape and the 3D-head (which is already available). To perform this task it needs two basic things: (1) to know the position and direction of the camera for each image, and (2) to extract the hair region correctly from each video captured image.

3.1 Camera Position Estimation

For each video captured image, we estimate the camera position by determining seven unknown parameters (six transformations and a projection function) of the virtual camera. The virtual camera is modeled as a simple pinhole camera. In the experiment we register a video captured image with the computer-generated image of the 3D-head in order to determine the virtual camera parameters. We perform the registration task automatically by our already reported *edge featured based 3D-2D projective registration technique*[3]. To obtain fast registration for the in-between images (i.e., images taken from the positions more than ten degree (10^0) far from the front, left or right), we assume an initial angle of rotation based upon the total number of in-between images and the angular span between left-to-front or front-to-right images. The rest of the camera parameters from the current image are assumed as the initial value for the next image.

3.2 Hair Region Segmentation

A semi-automatic tool called *intelligent scissors*[9] segments the hair-region from each video captured image. Fully automatic segmentation of 2D image is an unsolved problem, while intelligent scissors allow hair region to be extracted quickly and correctly using simple gesture motion with a mouse. When the mouse pointer comes in proximity to an edge, a dynamic programming based live-wire boundary wraps around the region. Finally the hair boundary is extracted by the using our already reported filling algorithm[3]. Fig.4 shows the hair-boundary and extracted hair-region, respectively, for a video captured image.



(a) Original photograph. (b) Hair-boundary. (c)Extracted hair-region

Fig. 5 Hair region segmentation

4 Hair Sculpturing

Our target in this work is to reconstruct and add 3D hair shape on the CT/MRI reconstructed 3D-head only. To do this, we need a good sculpturing object, which results a complete-head model after cutting out the non-hear-regions. We assume 3D sculpturing object as follows:

Initially we prepare a hairless solid-head filled with binarized voxels. As we mentioned before, there are two ways of making this solid-head: (1)by filling the 3D-head, and (2)by accumulating the binary converted CT slices. The latter way is preferable as it is easy to implement and there is no possibility of obtaining shape-error. Then a rectangular 3D box covers it. The size of the box is assumed as 30% more than that of the solid-head in each side. Except solid-head voxels rest of the box is filled with (which we call) hair-voxels. This 3D box as a whole is called the sculpturing-object. Both the solid-head-vexel and hair-voxel size in the sculpturing object are assumed as approximately one cubic *mm* ($0.908mm \times 0.908mm \times 1mm$).

Let us consider a camera position as shown in Fig.2. Suppose we see the head and the hair as in Fig.3. After identifying the hair-region shown in Fig.5(c), we sculpture the hair-voxels by removing all the voxels outside the hair-region up to the solid-head, as shown in Fig.4. After a number of camera positions around the sculpturing-object are tried, the remaining voxels on the solid-head is the resulting hair-shape.

It is true that we are able to obtain hair-shape close to the actual as the number of camera positions increases. However, from a practical viewpoint, we consider a dozen of camera positions from the left, front and right sides, while maintaining the camera height at the ear and nose level.

The sculpturing method has an special characteristic that the concave and semi-occluded parts near the skin-hair junction (especially in the forehead region) can be reconstructed reasonably well. Usual sculpturing methods using silhouettes can deal only with convex shape but fails to provide concave parts. Primarily due to the presence of hairless solid-head (3D-head), our method can deal with the concave parts in/near the forehead.

5 Experiments

The CT data we employed was a size of $512 \times 512 \times 225$ with a resolution of $0.454mm \times 0.454mm \times 1mm$. The 480×480 pixels video captured images were taken with a SONY digital video camera of 640×480 resolution. A $1m \times 1m$ blue sheet was used as the background. The person sat on a normal revolving chair. The video image was taken by keeping the camera at a fixed position while he himself rotates the chair by his leg. His head level waved slightly. Because of our superior registration scheme, there is no need to hold the head position very tightly.

(a) 3D-head (b) Textured mapped 3D-head

Fig.6 Hair-less 3D-head and its textured mapped images

Fig.6 shows the hair-less 3D-face and its textured mapped image. All the hairstyles shown in this paper are added on this 3D-head. So far we performed the reconstruction task on the same individual with three different hairstyles. One of those is the original hairstyle and the two others are wigged. One of the video captured images of original hairstyle is shown in Fig.7a, the reconstructed complete-head and final-head are in Fig.7b and Fig.7c respectively, and the textured mapped final-head in Fig.7e. Fig.8 shows the same type of images for one of the wigged hairstyles. Artificial color is added on the final-head as shown in Fig.7d and Fig.8d.

The final-heads in Fig.7 and Fig.8 were reconstructed from ten images. In this paper we emphasis only on the reconstruction of the concave and semi-occluded regions. The surface quality of the hair-region can be improved by increasing the number of images. Whereas the rest of the 3D-head (uncovered) remains unchanged as this is from the CT data.

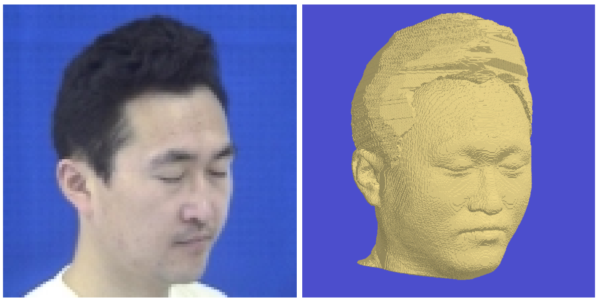
In Fig.7e and Fig.8e it is seen that for texture mapped image hair-shapes from ten images are acceptable.

6 Conclusion and Future Works

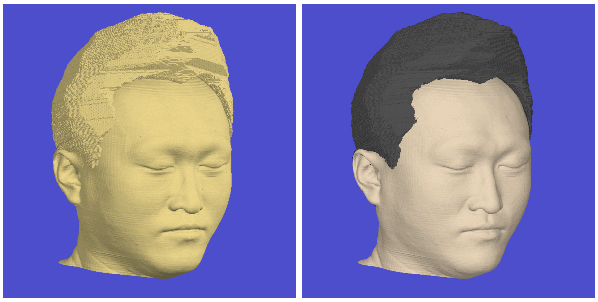
Our method is to wrap-up the facial image obtained from CT with hair-voxels and to remove the non-hair regions obtained from a sequence of images. The novelty of this research is to deal with the concave parts in/near the forehead (hair-skin junction regions). Whereas the usual sculpturing methods using silhouettes can deal with the convex shapes only. The reconstruction method is simple and easy to implement on hospital environment where a CT scanner is readily available. The additional requirements are only a computer and a digital video camera.

To obtain fully automatic registration, we discourage to use hair-shape, which covers the ears completely. This is because the ear edge is one of the landmarks for our edge-based registration.

Our next target is to simulate dynamically the post-surgical facial expressions (e.g., laughing, jaw-movement etc.) for the cancer patient having facial tumor, especially after replacing facial soft-tissue and/or removing a part of facial bone.



(a) Video Captured Image (b) Complete-head



(c) Final-head (d) Final Head with artificial color

(e) Textured mapped Final-head

Fig. 7 Reconstructed head with original hairstyle

(a) Video Captured Image (b) Complete-head

(c) Final-head (d) Final Head with artificial color

(e) Textured mapped final-head

Fig.8 Reconstructed head with a wigged hairstyle

7 Acknowledgment

The authors would like to thank Dr. R. Sekiguchi and other staffs of the National Cancer Center, East Hospital, Kashiwa, Chiba, Japan for providing medical guidance and necessary CT data.

References

1. Ling-Xiang Zhau and Wei-Kang Gu: "3D Model Reconstruction by Fusing Multiple Visual Cues," *Proceedings of ICPR'98*, Brisbane, Australia, August 16-20, pp.640-642 (1998).
2. Ali, Md. Haider, Eiji Takahashi and Toyohisa Kaneko: "A 3D Face Reconstruction Method from CT Image and Color Photographs," *IEICE Trans. on Information and Systems*, E81-D(10), pp1095-1102 (1998).
3. Ali, Md. Haider and Toyohisa Kaneko: "Automatic Reconstruction of 3D Human Face from CT and Color Photographs," *IEICE Trans. on Information and Systems*, E82-D(9) pp1287-1293 (1999).
4. Jiang Yu Zheng and Fumino Kishino: "Recovering 3D Models from Silhouette Sequence and Detecting Unexposed Regions," *IEICE Trans. on Information and Systems*, J76-D-II(6), pp1114-1122. (1993). (in Japanese)
5. Shigeki Sugimoto and Masatoshi Okutomi: "Shape Estimation of Rotating Object Using Spacio-temporal Images," *Transaction of Information Processing Society of Japan*, 40(6) pp2717-2723. (1999). (in Japanese)
6. Richard Szeliski and Richard Weiss: "Robust Shape Recovery from Occluding Contours Using a Linear Smoother," *Proc. ARPA Image Understanding Workshop*, Washington, D.C., pp939-948 (1993).

7. P. Giblin and Richard Weiss: "Reconstruction of surfaces from profiles," *First International Conference of Computer Vision (ICCV'87)*, London, England, pp136-144 (1987).

8. Y. Matsumoto, H. Terasaki, K. Sugimoto and T. Arakawa: "A Portable Three-dimensional Digitizer," *Inter-national Conference on Advances in 3d Digital Imaging and Modeling*, pp197-204 (1997).

9. Eric N. Mortensen and William A. Barrett: "Intelligent Scissors for Image Composition," *Computer Graphics Proceedings, SIGGRAPH 95*, Los Angeles, California, August 6-11, pp191-198 (1995).

10. T. Akimoto and Y. Suenaga: "Automatic Creation of 3D Facial Model," *IEEE Computer Graphics and Applications*, 13(4) pp16-22 (1993).

11. Ferderic Pighin, Jamie Hecker, Dani Lischinski, Richard Szeliski and David H. Salesin: "Synthesizing Realistic Facial Expressions from Photographs," *Computer Graphics Proceedings SIGGRAPH'98*, Orlando, Florida, pp75-84 (1998).