FACIAL MODEL RECONSTRUCTION FOR PLASTIC SURGERY SIMULATION

Chia-Hsiang Wu, Sheng-Che Lin and Yung-Nien Sun*

Department of Computer Science and Information Engineering,
National Cheng Kung University, Tainan, Taiwan
*Department of Medicine, National Cheng Kung University, Tainan, Taiwan
E-mail:ynsun@mail.ncku.edu.tw

ABSTRACT

In facial surgery simulation, it usually requires the patient's CT (Computed Tomography) data for building the facial model. However, CT imaging is expensive, time consuming, and the patients have to be under radiation exposure as well. Thus, it is undesirable to undertake CT examination for minor surgeries; instead, we propose a new three-dimensional (3D) facial model reconstruction algorithm based on color stereovision techniques. The proposed technique provides a non-radioactive alternative for facial model reconstruction. In the proposed system, a color-coded grid pattern is first adopted to simplify the establishment of stereo correspondence. Because human facial surfaces are usually with few image features, the generated grid greatly facilitates the correspondence detection and is also robust due to its fault-tolerance capability. In addition, we setup four cameras around the patient and image the subject simultaneously to enlarge the field of view covering the whole facial areas. The resulting models from all three camera pairs were registered and integrated by the least square method. Experimental results have shown that the proposed algorithm successfully reconstructs the full facial model and recovers the 3D information with efficient computation.

1.Introduction

Since facial appearances play key roles in interpersonal activities, prediction of post-surgical human faces is a critical issue. Surgeons have to design the detailed repair plans or compensation procedures while patients are eager to know the expected post-surgery appearance before the actual surgery is carried out. Thus, it is important to develop an auxiliary system to predict the deformities for craniofacial and maxillofacial surgery. In this study, we combine the techniques of computer vision and graphics to construct the patient's 3D mesh for use in a computerized surgical simulation system. It includes the method to generate the facial model, to select features in the computer surgical simulation, and to calculate the evolution of anatomical structure in a convenient and effective manner. Thus, a practical tool for evaluating 3D facial structure for both the pre-surgery planning and post-surgery evaluation can be established for clinical usage.

In recent years, some prototype systems have been developed. Most of them utilize the CT data to generate facial model. However CT is expensive, radioactive, and time consuming, a vision-based system is a simpler, cheaper, and non-radioactive alternative that is more preferable for both physicians and patients. Thus, we propose a new system based on the structure lighting and color stereovision for facial reconstruction.

Three-dimensional reconstruction from multiple 2D images is the major research tasks in computer vision. It has been widely utilized in many applications such as reverse engineering, medical diagnosis, CAD/CAM and virtual reality. Possible approaches include shape from shading, photometric stereo, structured lighting, laser range scanner and stereovision [1]. Among them, binocular stereo is the most popular technique. We can setup a binocular vision system easily with affordable cost. Given correct image correspondences, 3D information can be recovered by triangulation. But if the depth of object surface is smoothly changed and with little surface texture, then the correspondences and resulting depths become hard to defined. Thus structured lighting system becomes a more feasible solution [2-9].

The structured lighting system is similar to a passive stereovision system with one of the cameras replaced by a projector. The projector projects a pre-designed light pattern on the measuring scene and the camera images the scene. [3,5] Similarly, binocular systems can be designed by combining stereovision and color structured patterns [10-11]. Though there are disadvantages for spatial coded patterns, e.g. decoding variation and grid resolution, the resulting system of this study provides suitable performance under reasonable expense. On the other hand, a stereo camera pair provides only partial model of the reconstructed object due to limited field of view. The proposed system acquires the object models from three views. We then register them by determining the Euclidean transformation parameters (rotation and translation) between the three resulting models and integrate them together into a single 3D representation. [19-23] The proposed system adopts the point-based method that uses corresponding points to estimate these transformation parameters. These

This work was supported by the National Science Council, Taiwan, ROC, under grant no. NSC- 89-2213-E006-034.

points can be the nature face features or artificial markers. Face feature points such as the canthi, the corners of the mouth and the nose tips are common features used in these systems. Methods that do not rely on corresponding points typically employ surface-based approaches, such as ICP (iterative closet point) algorithm [20] and its extensions [24-25]. Most of them are more complicated and time-consuming than the point-based method.

In the proposed system, the color structured lighting and robust matching algorithm are designed to obtain accurate 3D facial data. Image processing techniques are employed to detect the spot features for correspondence establishment automatically. The spatial coded patterns are adopted that is insensitive to patient motion. An adjustment algorithm is also designed to generate patterns with optimal color, shape and density. For recovering a complete facial model, we image patient's face from four different positions so that any face points can be viewed by at least two cameras. Three partial models are generated from any two consecutive cameras by triangulation. The system then registers these models and integrates them with the distance-weighted blending technique. Three slide projectors are used so that the patterns can cover the whole face. The point-based registration algorithm is then used to provide fast and simple registration.

This paper is organized as follows: Section 2 describes the procedure for pattern generation. The proposed corner detection algorithm is addressed in Section 3. The algorithm for registering and integrating incomplete models are presented in Section 4. Experimental results are given in Section 5, followed by conclusions.

2. PATTERN DESIGN

There are several coded patterns reported in the recent literature [12]. According to the temporal dependence, the patterns can be divided into two classes: temporal and spatial codes. To capture 3D model of living or moving object, we choose the spatial code design. There are two essential requirements for these patterns: 1) code-uniqueness, and 2) fault-tolerance. In our development, the code is a matrix M of size H ×W. Each element m_{ij} of matrix M is assigned with one of the possible L letters. Considering N neighboring elements P_1, P_2, P_3,… P_N of arbitrary m_{ij}, the position of m_{ij} can be computed (or decoded) from the code letters of m_{ij} and its N neighboring elements. We denote the index code word of m_{ij} by F_{ij}=(f(P_1), f(P_2), f(P_3),…..f(P_N), f(m_{ij})), where f(P_i) is the code letter of position P_i for i=1,2 ...N. We choose (3,3) de Bruijn sequences [17] for pattern encoding. The encoded matrix is established according to the method proposed by Y.-C. Hsieh[13], which ensures that each m_{ij} has three possible letters, i.e. L=3, and is uniquely indexed by its 4-neighboring system and itself. That is, m_{ij} is indexed by F_{ij}=(f(P_1), f(P_2), f(P_3), f(P_4), f(m_{ij}))) and f(P_1),f(P_2),f(P_3),f(P_4) are the letters of m_{ij}'s 4-neighbors. Since F_{ij} appears exactly only once in the matrix, the Hamming distance between each distinct F_{ij} pair will be equal to or exceed 1. If misidentified code word or loss of pattern elements should happen, this special property provides preliminary error detection and correction ability.

When the coded matrix is determined, we have to assign the color and shape for each m_{ij}. In [4,6-7], red, green and blue colors (hue=0, 85 and 170 for hue [0,254]) are first chosen to represent the three code letters. However, we have tested several different color sets and found that h=42.5, 127.5 and 212.5 are better than RGB colors. Hence, the code colors are assigned by (42.5,255,128), (127.5,255,128) and (212.5,255,128) on HSI color model. The element shape can be circles [4], slits [5,9] or other geometric appearances [6-7]. We choose squares as the element shape, since squares are effective in increasing mesh density. Fig. 1 shows the object illuminated by our patterns.

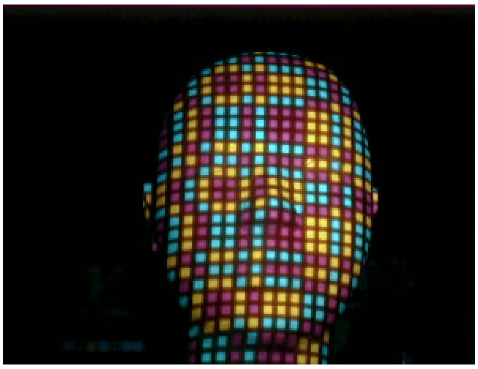
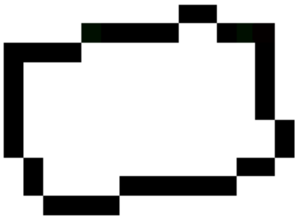


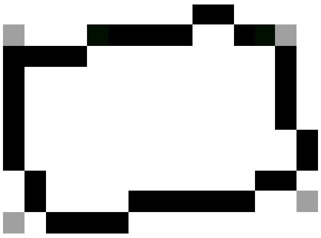
Figure 1. Object after pattern projection

3. SQUARE CORNER DETECTION

To increase the resolution of the recovered model, we will detect corner points as features for establishing more correspondences. Possible solutions to locating corner points include general corner detection algorithms, polygonal approximations, or dominant points detection on planar curves [18]. However, the acquired squares are corrupted and distorted (Fig.2 (a)) that makes corner detection a very difficult task. Therefore, we develop a new algorithm to estimate corners for the corrupted or distorted squares. First, it computes the number of pixels on each line segments in both vertical and horizontal directions, which yields two force field maps Fh(i,j) and Fv(i,j), respectively. The values of Fh(i,j) and Fv(i,j) are in fact their corresponding segment lengths. For example, if (i,j) is contained in a horizontal line segment of length 3 and a vertical line segment of length 5, then Fh(i,j)=3 and Fv(i,j)=5. It is assumed that the length of line segments represents the possibility of actual square location, and the pixels near the intersections of horizontal and vertical lines may be the actual square corner positions. Thus, we define the total force by Ftotal(i,j) =Fh(i,j)+Fv(i,j). The total force value near the segment end points should become the maximum. Dividing the bounding box into four squares evenly, we then choose the pixels with the greatest force in the square to be the corner points. In case of multiple maximum, the optimal position is the interpolation among all candidates. The detected corners are shown on Fig.2 (b)



(a)



(b)

Figure 2. (a) A corrupted and distorted square. (b) Gray pixels are the detected corners.

4. REGISTRATION AND INTEGRATION

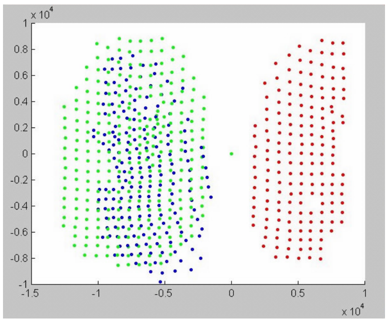
There are three 3D models reconstructed from any two consecutive cameras. We register these models by evaluating their mutually rotation and translation. Registration algorithms usually can be divided into two categories: feature-based and surface-based. In our experiment, the face grids provide very robust and convenient landmarks for feature-based registration. Utilizing the consistency of grid color code, we can establish the correspondences between these models. Then, we can make use of the feature-based registration to estimate the rotation and translation parameters by minimizing the total registration error D given in the following equation:

$$D = \frac{1}{n}\sum_{i=1}^{n}\left\| p_i - (Rq_i + T) \right\|^2$$

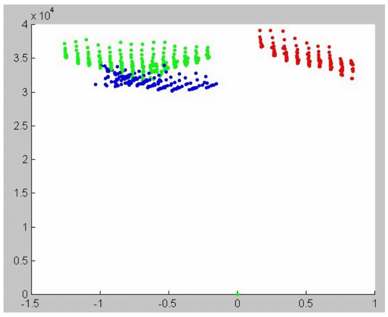
where p_i and q_i are the corresponding feature point pair of the two models, and n is the total number of correspondences. We fix p_i and transform q_i by choosing appropriate rotation R and translation T so as to best match these two point sets.

We estimate the two unknown parameters R and T by the least-square optimization. Figures 3 and 4 show the relative projection positions of these point sets before and after registration, respectively, from the front and top views.

Although these models were already aligned, some points may not exactly match due to the calibration or quantization errors. To compensate for this drawback, we calculate the 3D position by inverse-distance weighted averaging for points with two or more correspondences. The distance measure is defined as the length between a vertex and its nearest border.

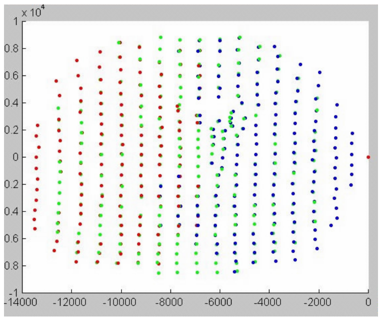


(a)

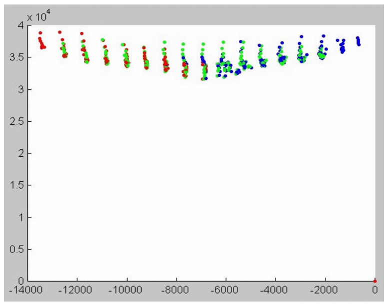


(b)

Figure 3. 3D positions of the point sets before registration. (a) Front view; (b) Top view



(a)

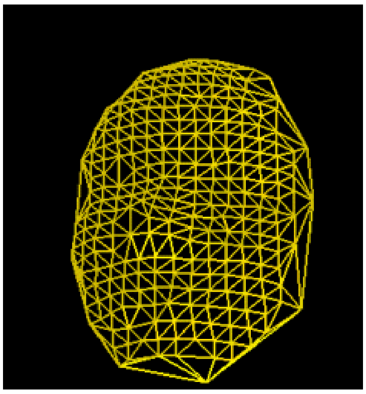


(b)

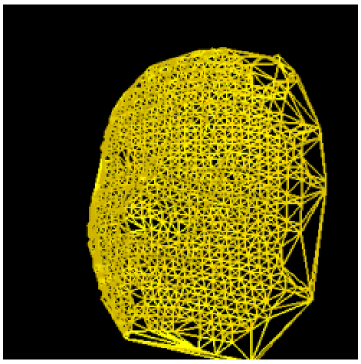
Figure 4. 3D positions of the point sets after registration. (a) Front view; (b) Top view

5.EXPERIMENT RESULTS

In our experiments, the encoded color pattern was generated on a 35-mm slide. This slide was projected onto the object surface by a slide projector. Two parallel digital cameras imaged the scene simultaneously and the 3D information was measured by the stereo triangulation. The digital cameras are AGFA e-Photo 1680 with spatial resolution 640x480. We tested our method by a plastic head model. The captured images were first transformed into the HSI color model [14], which yielded the corresponding hue, intensity and saturation images. The intensity image was used to detect the positions of squares by Canny edge detector [15]. The hue image was applied to identify the color patterns. Based on these color codes, we could determine their code configurations for each grid and then resolve the grid correspondence accordingly. In this process, the grid centroids and corners were used as the actual corresponding points. After triangulation, we first obtain a set of scattered 3D points. A 3D surface model can then be built by Delaunay triangulation [16] for generating triangles on the given 3D points. Fig.5 (a) and Fig.5 (b) are the results before and after increasing resolution. The model with corner detection is five times the number of vertices and triangles as that before the corner detection.



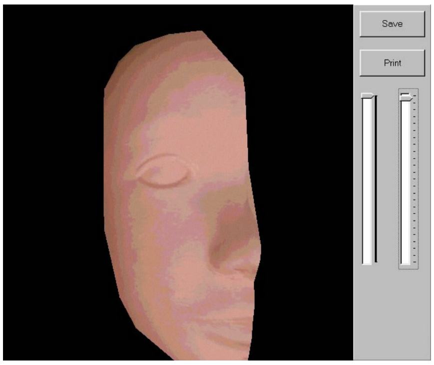
(a)



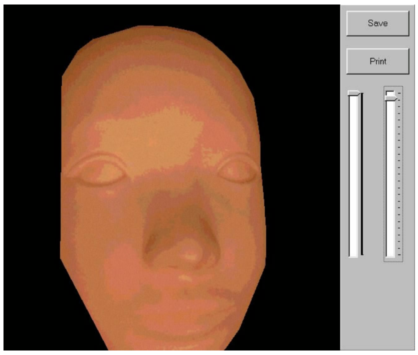
(b)

Figure 5. Model wireframe (a)before and (b)after corner detection.

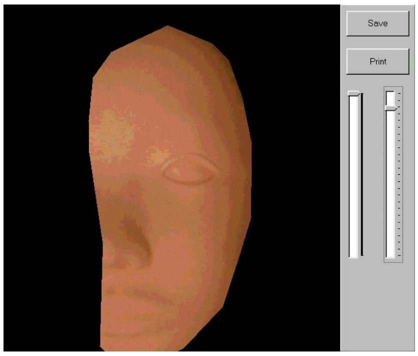
The partially reconstructed results are shown on Fig 6. Three partial models were combined to a facial model. The reconstructed model can be visualized from arbitrary viewpoints. Fig. 7 shows the texture-mapped 3D model from different views. And Fig. 8 is a textured-mapped image sequence of a real case from different viewpoints.



(a)



(b)



(c)

Figure 6. Three partial face model from different camera sets.

(a) (b) (c)

Figure 7. The complete facial model obtained by integration. (a) wireframe ; (b) texture-mapped model ; (c) Texture-mapped model from another viewpoint.

Fig. 8. An image sequence of a real subject with texture-mapping.

6. CONCLUSIONS

In order to recover 3D information from smooth facial surfaces, we propose a new system to reconstruct 3D surface from stereo images with color encoded pattern projection. The 3D object surface is efficiently computed from the color encoded stereo patterns. The de Bruijn's sequence-encoding scheme is adopted to generate the color codes. It considerably reduces the correspondence computation and provides the fault-tolerant capability. The designed corner detection algorithm is capable of estimating corners from corrupted or distorted squares. As a result, the number of vertexes and triangles are multiplied, and a dense facial model can be reconstructed with fewer squares. The proposed method is effective for instant face imaging thus the system is hardly affected by patient motion. Moreover, we use four cameras simultaneously to resolve the view range limitation problem of standard stereovision. After registration and integration, a complete facial model was successfully reconstructed. Thus, the proposed method is simple, inexpensive, non-radioactive and convenient for facial surgical planning. It provides a tempting alternative for physicians and patients in facial surgical applications.

7.REFERENCES

[1] R. Klette, K. Schluns and A. Koschan, *Computer Vision : Three-dimensional Data from Images*, Springer, Singapore, 1998.

[2] J. L. Moigne and A. M. Waxman, “Structured Light Patterns for robot mobility,” *IEEE J. Robotics and Automation*, vol. 4, no. 5, pp. 541-548, 1988.

[3] D. Caspi, N. Kiryati, and J. Shamir, “Range Imaging with adaptive color structured light,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 20, no. 5, pp. 470-480, May 1998.

[4] R. A. Morano, C. Ozturk, R. Conn, S. Dubin, S. Zietz, and J. Nissanov, “Structured Light using Pseudorandom Codes,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 20, no. 3, pp. 322-325, Mar. 1998.

[5] J. Salvi, J. Batlle, and E. Mouaddib, “A Robust-coded Pattern Projection for Dynamic 3D scene measurement,” *Pattern Recognition Letters*, vol. 19, pp. 1055-1065, 1998.

[6] P. M. Griffin, L. S. Narasimhan and S. R. Yee, ”Generation of Uniquely Encoded Light Patterns for range data acquisition,” *Pattern Recognition*, vol. 25, no. 6, pp. 609-616, 1992.

[7] S. R. Yee and P. M. Griffin, ” Three-dimensional imaging system,” *Optical Engineering*, vol. 33, pp. 2070-2075, 1994.

[8] R. J. Valkenburg and A. M. McIvor, “Accurate 3D Measurement using a Structured Light System,” *Image and Vision Computing*, vol. 16, pp. 99-110, 1998.

[9] K. L. Boyer and A. C. Kak, “Color-encoded Structured Light for Rapid Active Ranging,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 9, no. 1, pp. 14-28, 1987.

[10] C.-S. Chen, Y.-P. Hung, C.-C. Chiang, and J.-L. Wu, “Range Data Acquisition using Color Structured Lighting and Stereo Vision,” *Image and Vision Computing*, vol. 15, pp. 445-456, 1997.

[11] A. Koschan, V. Rodehorst, and K. Spiller, “Color Stereo Vision using Hierarchical Block Matching and Active Color Illumination,” *Proc. Int. Conference on Pattern Recognition*, pp. 835-839, 1996.

[12] J. Batlle, E. Mouddib and J. Salvi, “Recent Progress in Coded Structured Light as A Technique to Solve the Correspondence Problem: A Survey,” *Pattern Recognition*, vol. 31, no. 7, pp. 963-982, 1998.

[13] Y.-C. Hsieh,” A Note on the Structured Light of Three-dimensional Imaging System,” *Pattern Recognition Letters*, vol. 19, pp. 315-318, 1998.

[14] R. C. Gonzales and R. E. Woods, *Digital Image Processing*, Addison-Wesley, Reading, MA, 1992.

[15] J. Canny, “ A Computational approach to edge detection,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 8, pp. 679-698, 1986.

[16] M. de Berg, M. van Kreveld, M. Overmars and O. Schwarzkopf, *Computational Geometry*, Springer, Berlin Heidelberg, 1997.

[17] G. Chartrand, O. R. Oellermann, *Applied and algorithmic graph theory*, McGraw-Hill, Singapore, 1993.

[18] E. R. Davis, *Machine Vision*, Academic Press, 1997.

[19]R. M. Bolle and D. B. Copper, “On optimally combining pieces of information, with application to estimating 3-D complex-object position from range data,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 8, no. 5, pp. 619-638, 1986.

[20]P. J. Besl and N. D. McKay, ”A method for registration of 3D shapes,” *IEEE Tran. Pattern Analysis and Machine Intelligence*, vol. 12, no. 2, pp. 239-256, 1992.

[21]C. Dorai, G. Wang and A. K. Jain, “Registration and integration of multiple object views for 3D model construction,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 20, no. 1, pp. 83-89, Jan. 1998.

[22]J. Tarel, “A Coarse to fine 3D registration method based on robust fuzzy clustering,” *Computer vision and image understanding*, vol. 73, no. 1, pp. 14-28, Jan. 1999.

[23]L. G. Brown, “A Survey of Image Registration Techniques,” *ACM Comput. Surv.*, vol. 24, no. 4, pp.325-375, Dec. 1992.

[24]J. Feldmar, “Extension of the ICP Algorithm to Nonrigid Intensity-Based Registration of 3D volumes,” *Computer Vision and Image Understanding*, vol. 66, no. 2, pp. 193-206, May 1997.

[25]Z. Zhang, “Iterative Point Matching for Registration of

Free-form Curves and Surfaces," *Int. J. Comput. Vis.*, vol. 13, no. 2, pp.119-152, 1994.