International Journal of Information Technology
Vol. 8, No. 2 September 2002

Constructing A Realistic Face Model of An Individual for Expression Animation

Yu Zhang[a] , Edmond C. Prakash[b] and Eric Sung[a]

[a] IMRL, School of Electrical and Electronic Engineering,

[b] Division of Software Systems, School of Computer Engineering,
Nanyang Technological University, Singapore 639798

ABSTRACT

This paper presents a method for creating photorealistic textured 3D face models of specific people for dynamic facial expression animation. The modeling approach reconstructs an accurate geometrical face model based on the individual face measurements containing both shape and texture information that is acquired from a laser range scanner. By using a semi-automatic registration and merging technique, a 3D dense face mesh is recovered from the partial range data obtained from arbitrary multiple different views. A model editing and adaptive meshing scheme is then used to refine the surface model. Having recovered the facial geometry, we add realism by mapping the model with high-resolution texture images. The resultant synthetic face has been shown to be visually similar to the true face. Based on the geometrically accurate surface model, a physically-based face model with a hierarchical structure of the skin, muscles and skull is developed from anatomical perspective. The dynamic displacement of nodes in the skin lattice under the influence of internal muscular forces is calculated by a numerical integration method. Using our technique, we have been able to generate highly realistic face models and flexible expressions.

Keywords: Face modeling, Facial expression animation, Personalized model, Laser range scanner, Hierarchical structure, Dynamic deformation.

1. Introduction

One of the most interesting and challenging problems in computer graphics is the effortless generation of realistic looking, animated human face models. Indeed, attempts to model and animate realistic human faces date back to the early 70's [19], with many efforts made on this topic since then. The computer generation of faces has attracted a number of investigations such as low bandwidth teleconferencing [2], speech synthesis for the deaf and hard of hearing [7], computer synthetic facial surgery [15], human face shape and motion estimation [5] and face recognition [12]. The complexity of human facial anatomy and our natural sensitivity to facial appearance increase the difficulty of modeling facial appearance and subtle expressions.

The typical processes of computer facial animation are composed by two stages: (1) design a 3D facial mesh, (2) animate the 3D mesh in a controlled fashion to simulate facial actions. In the first stage, many previous approaches were based on the face models from standard polygonal surface meshes [6, 20, 25] to parametric surfaces [18, 24]. All these face models are adapted from a generic face with low resolution instead of generated directly from the underlying data set, and therefore could not represent a personalized face in an anatomically accurate and realistic way. The simple and effective method to reconstruct a specific person's face model is the image-based technique using the shape features extracted from facial images. The most frequent method proposed is to use facial images captured from frontal and side views and detects a number of feature points either automatically or interactively on the two pictures [1, 4, 13]. Another method is the stereo reconstruction which uses the geometric/homologous relation between stereo images to recover the surface depth [10]. This approach often results in noisy data. Some researchers also construct face models from video sequences [11, 17]. It acquires a video stream with a cheap and entirely passive sensor such as an ordinary video camera. The more frames there are, the better the results will tend to be. The disadvantage of these kinds of techniques is that they are either manually intensive or computationally expensive and the reconstructed 3D shape is not completely accurate.

In the animation stage, several approaches have been proposed to produce expressive and plausible animation of the face model. The keyframing method animated a synthetic face by using a linear combination of previously modeled facial expressions [16, 19]. Although simple to implement, it requires a full database of key expressions to be modeled before hand. Using parameterized models [6, 20], the animator can create facial expressions by specifying appropriate sets of parameter value sequences, eliminating the need for a complete bank of models. Though parametric models can be applied at relatively low computational costs, realistic blending between facial expressions is problematic since each parameter affects a disjoint set of vertices in the face model. There are also approaches to deform predefined skin with a space filling function whose purpose is to deform the skin surface corresponding to desirable muscle actions [14, 23]. The purely geometric nature of these prior animation methods ignores many of the complexities of human face anatomy. They consider the skin as an infinitesimally thin surface with no underlying structure, and deformations are generated by geometrically manipulating the surface. Consequently, attempts to mimic many of the subtleties of facial tissue deformation are spurious.

This paper develops 3D physically-based face models of specific people for realistic facial animation. Our modeling approach is based on individual facial measurements including information about face shape and face texture which are used to create precise geometry. To obtain a complete model of the face, we acquire three views of the person's face by using a non-contact 3D laser range scanner, each producing separate 3D reconstructions of the visible face regions. Through semi-automatic registration, the meshes of different views are merged into a single one. A highly accurate geometric face model of an individual suitable for animation is then reconstructed after mesh editing and adaptive meshing. Based on this personalized surface model, we develop a physically-based face model by taking into account the facial anatomy and the hierarchical structure. It incorporates a physically-based approximation to facial skin tissue, a set of anatomically-motivated facial muscle actuators and underlying rigid skull. Lagrangian mechanics governs the dynamics, dictating the deformation of facial surface in response to muscular forces. By solving the governing dynamics numerically, flexible and realistic expressions can be generated.

The rest of the paper is structured as follows. Section 2 describes the reconstruction of a photorealistic personalized face model from individual facial measurements. Section 3 presents the physically-based face model with the hierarchical structure. The modeling process of multi-layer skin, skull and muscles is detailed. The motion dynamics and our numerical solution strategy are also explained. The simulation results are shown in Section 4. Section 5 gives concluding remarks and our future work.

2. Reconstruction of A Personalized 3D Face Model

A triangular mesh that describes the surface of the face is the basic facet of the proposed facial animation system. To enhance the realism, rather than shading the piecewise planar triangular surface through standard Gouraud shading, a high-resolution color image of the face will be used for photorealistic rendering of the face. The geometry and color information of the facial surface is obtained by scanning a subject using a Minolta VIVID 700 Digitizer^{TM}. The construction of a full 3D face model needed for facial animation requires a number of steps: 1) acquisition of range and color data from a number of viewpoints chosen to get complete coverage of the subject's face, 2) registration of these data into a single coordinate system, 3) representation of all the data by a surface model that agrees with all the images, 4) reduction of the size of the data set for further animation. Fig. 1 shows a flow diagram of the 3D reconstruction procedure. It takes about 0.6 seconds for capturing of all data in each scan; a few minutes will be necessary to generate the final 3D face model.

Data acquisition from different viewpoints: Unlike some scanners (e.g., Cyberware) which can rotate the full 360 degrees around the head, VIVID can only capture a partial facial region facing the scanner during each scan. To recover the whole facial geometry, scans are taken from several different angles of the front of the head relative to the scanner. The subject is asked to first look directly onto the scanner and then rotate his or her head about the vertical axis by 45 degrees. In each scan, the scanner projects a laser strip onto the head and shoulders of the subject, who sits motionless in front of the scanner. As the data acquisition speed of this scanner is high (0.6 seconds per scan), this constraint can be met effortlessly. Once each scan is complete, the device has acquired two registered images of the subject: a 200×200 range image – a topographic map that records the distance from the sensor to points on the facial surface, and a 400×400 reflectance (RGB) image – which registers the color of the surface at those points. Fig. 2 shows range maps and color images of three views.

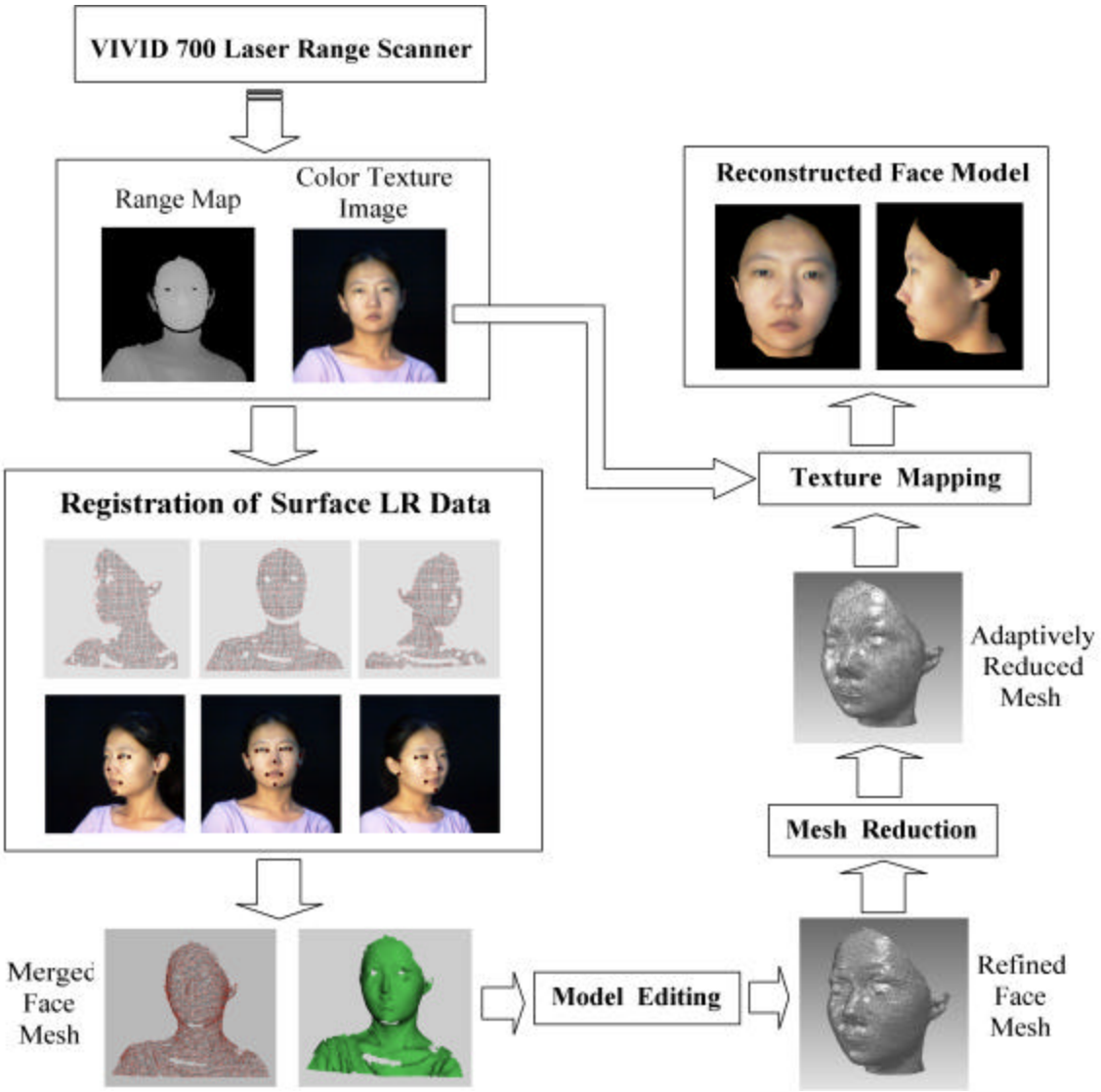
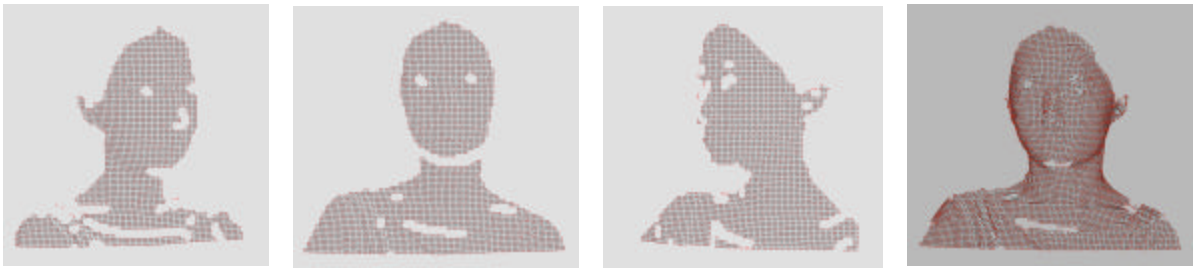


Figure 1. An overall flow diagram for face reconstruction from laser range scans.

(a) (b) (c)

Figure 2. Range map (top row) and texture image (bottom row) of a face acquired from different viewpoints: (a) view from right 45 degrees; (b) front view; (b) view from left 45 degrees.

(a)



(b) (c) (d) (e)

Figure 3. Registration process: (a) landmarks required for the registration of the laser range surfaces; (b), (c) and (d): partial face meshes of LR scan from three viewpoints; (e) merged full face mesh.

Registration: The task of registration process is to register the surface LR data scanned from different viewpoints in a common coordinate frame. The problem is to find a transformation which maps one LR surface onto another. Due to the fact that the scaling of each surface is determined by the scanning method, we are left with finding only a suitable translation and rotation in order to register the geometries. To achieve this a semi-automatic method is employed. As input to this semi-automatic set-up, we take two (or more) color images of the face from different viewpoints. Based on anthropometry of the head and face [9], a few landmark pairs, which are the most characteristic points and easy to locate facial features, as e.g. the corners of the eyes or the mouth and the tip of the nose or the chin, are selected interactively from among the face mesh vertices, and their positions in each image (where visible) are specified by hand. An illustration of typical landmarks is given in Fig. 3 (a).

The registration process can be regarded as the task of finding a transformation matrix **M** to minimize a scalar error function E defined by the square distance of n corresponding landmark points $\mathbf{u}_i$ and $\mathbf{u}_i'$. E depends on the rotation matrix **R** around three axes as well as on the translation vector **t**.

$$E(\mathbf{R},\mathbf{t}) = \sum_{i=1}^{n} (\mathrm{Mu_i} - \mathrm{u_i'})(\mathrm{Mu_i} - \mathrm{u_i'})^T \tag{1}$$

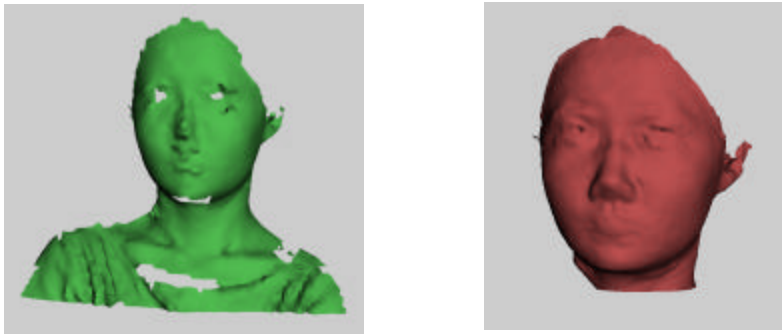
where

$$\mathbf{M} = \begin{bmatrix} \mathbf{M}_{\mathrm{rot}}(\mathbf{R}) & \mathbf{t} \\ 0 & 1 \end{bmatrix} \tag{2}$$

$$\mathbf{M}_{\mathrm{rot}}(\mathbf{R}) : \mathrm{Rotation\ \ matrix\ \ correspond\,ing\ \ to}\ \mathbf{R} \tag{3}$$

For minimizing $E(\mathbf{R},\mathbf{t})$, a method of conjugate directions presented in [21] is employed. After the registration, the partial face meshes are merged together to generate a full one (as shown in Fig. 3 (e)).

Model editing: While measuring the face with laser scanner, it is often difficult to scan the complete facial surface. Usually, range data is missing due to shading effects or because it is not possible to reach all parts of the face, especially if there are cavities and other small sized concave regions. For instance, in the nostril area, chin area, and even in the pupils, laser beams tend to disperse and the sensor observes no range value for these corresponding 3D surface points. Missing data causes gaps and holes in the reconstructed triangle mesh (see Fig. 4 (a)).



(a) **(b)**

Figure 4. Model editing process: (a) original model with errors; (b) refined facial surface after editing.

In the model editing procedure, these defects are filled with flat triangles by topological operations (see Fig. 5). Gaps between separate components of the mesh are closed by gap bridging, holes are closed by mesh growing. The mesh growth operation simply connects adjacent edges with a new one and checks that the new triangle does not overlap with any other triangle in its vicinity. Besides the error of missing data, the protruding vertices exist in the regions with high curvature such as the areas of mouth lips and eyebrows. The correction of these inaccurate vertices is done by deleting them first and then closing the hole left by the mesh growth operation. The face model after mesh correction is shown in Fig. 4 (b). As the hair of dark color absorbs all of the laser radiation, it is generally not digitized unless the hair is artificially colored white or the subject wears a white cap to cover the hair area, but that destroys the texture. For this reason, currently only the frontal part of the head is modeled. More sophisticated scanning method will be developed in the future to retrieve the complete head geometry including the hair.

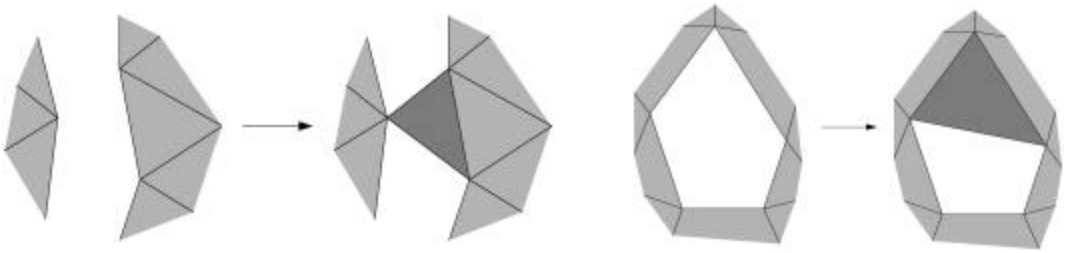


Figure 5. Topological operations to close gaps and fill holes in triangular facial mesh: (a) gap bridging; (b) mesh growth.

Mesh reduction: The generated triangular mesh consists of $> 10^4$ triangles. However, most of the vertices are redundant. For instance, the flat surface of forehead and cheek on the face model are usually oversampled. In order to enable dynamic simulation of the face at an interactive rate, a trade-off has to be found between computational costs and precision. Since the computing time depends mostly on the number of triangles, a good adaptive triangulation of the parameter domain is an essential step. There is a wide range of meshing and triangulation schemes available in the literature [3]. We employ an improved version of Schroder's adaptive reduction algorithm [22] for thinning and retiling. The mesh has to follow the 2D parameter domain and additional effort has to be spent at its boundaries. This compression algorithm reduces the amount of triangles used to represent complex facial structure up to 70 percent without sacrificing the visible detail of the facial surface. Fig. 6 shows two facial surface meshes of both an initial and an adaptively reduced geometry, and the texture-mapped appearance of the latter. In Fig. 6 (b) after adaptive sampling the facial surface becomes a model with finer triangular elements over the highly curved and/or highly articulate regions of face, such as the eyes and mouth, and larger triangular elements elsewhere, such as the cheek and forehead. The texture-mapped model shows that although the adaptive triangular mesh has less complexity, it preserves fine facial surface details. The resulting mesh captures the facial geometry especially including the contour line of the hair, as a polygonal surface that can be texture mapped with the reflectance image, thereby maintaining a realistic facsimile of the subject's face.

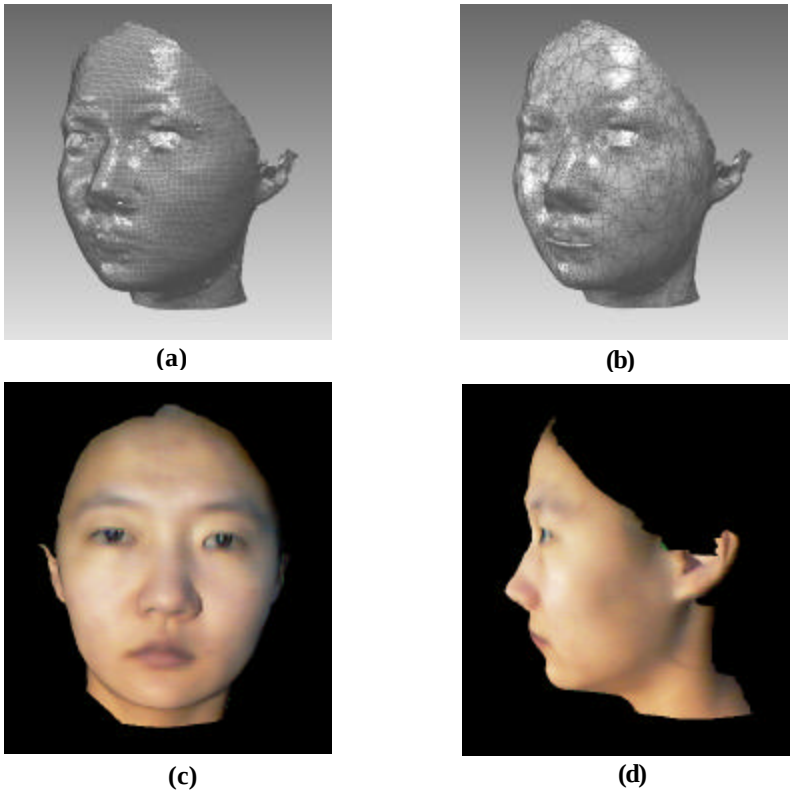


Figure 6. Adaptive triangular meshes of the acquired data: (a) original model (12194 vertices); (b) adaptively reduced model (3766 vertices); (c) and (d) front and side views of the adaptively sampled facial mesh rendered with texture mapping.

3. Physically-based Face Modeling

In accordance with the structure of the real skin, we have developed a multi-layer mass-spring-damper (MSD) tissue model. For each triangle of the facial skin one basic tissue element as shown in Fig. 7 is created. The topmost surface of the lattice represents the epidermis. It is a rather stiff layer of keratin and collagen and the spring parameters are set to make it moderately resistant to deformation. The springs in the second layer are highly deformable, reflecting the nature of dermal fatty tissue. Nodes on the bottom-most surface of the lattice represent the hypodermis to which facial muscle fibers are attached. The multi-layer structure of the model enables the soft-tissue point set V to be comprised of three subsets: $V=V_e \cup V_d \cup V_h$, where V_e, V_d and V_h are the point sets representing vertices on the epidermal, dermal and hypodermal layer respectively. In Fig. 1 every vertex $\mathbf{x}_i^e \in V_e$ in the surface mesh which represents the epidermis layer belongs to a vertex $\mathbf{x}_i^d \in V_d$ on the underlying dermis layer and a vertex $\mathbf{x}_i^h \in V_h$ on the hypodermis layer. This relation enables us to construct a spring mesh between each of the two layers.

The basic tissue elements are generated by connecting each epidermis vertex to the underlying dermis and hypodermis structure. According to their locations and function, all springs are categorized into the following three different sets:

- **Layer-spring set** (S_{la}): springs link a vertex on each layer with its neighbors on the same mesh layer. They cause the strongest internal force, which resists in-plane compression or tension. The stiffness of the layer-spring is the primary cause for the stiffness of the governing equation of the dynamic skin model.

- **Connecting-spring set** (S_{co}): springs link a vertex on the layer with its corresponding vertex on the adjacent layer. They resist the traction stresses or compression between layers. In particular, the connecting-springs are normal to the connected layers when the skin model is in its initial state.
- **Crossed-spring set** (S_{cr}): springs link a tissue layer vertex with the neighbors of its corresponding vertex on the adjacent layer. In this model, the crossed-spring is used to resist shearing and twisting stresses.

Therefore, the spring set S representing all springs of the model can be described as $S = S_{la} \cup S_{co} \cup S_{cr}$. All layer-springs have the same structure as the triangulated skin surface. The connecting-springs form prismatic volume elements.

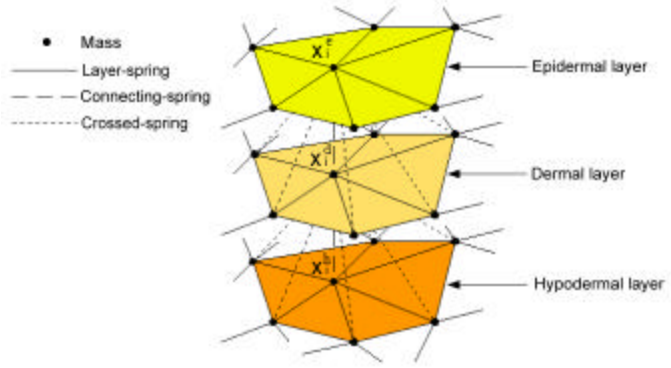


Figure 7. Multi-layer skin tissue element.

By assembling the discrete deformable model according to histological knowledge of skin, we are able to construct an anatomically consistent facial skin surface. The reduced skin data forms the basis for the generation of the multi-layer tissue model and defines the epidermal layer. The vertices on other two layers are computed by tracing the surface normal of the epidermal layer in the direction to the underlying layer. After the automatic assembly the facial skin consisting of 7206 elements has a multi-layer structure in accordance with the real facial tissue. For clarity Fig. 8 shows a close-up view of the topology of the cheek region in the facial model. The meshes in black, blue and red color represent epidermal, dermal and hypodermal layer respectively.

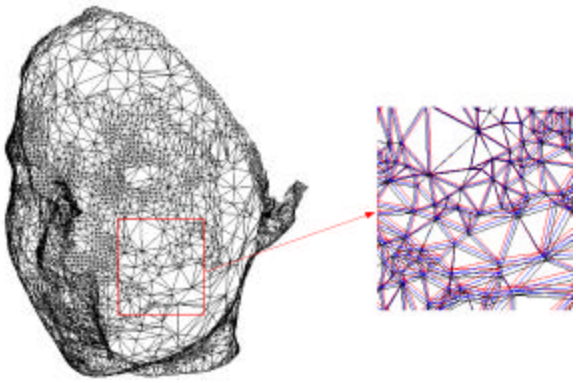


Figure 8. A close-up view of the assembled multi-layer facial skin.

Under the skin layer, there is a bone structure in our face model. We use a generic skull model that is independent from the skin mesh to map general anatomical attributes to the individual facial surface. In the skull model fitting process, affine transformations – rotation, translation and non-proportional scaling are applied on the skull model by the user interactively. Fig. 9 shows both the shape of the skull with the facial skin mesh, which identifies the nature of the skull with a texture-fitting, and the aligned mapped representation of facial surface, which shows the actual appearance. For clarity only the hypodermal layer is depicted in Fig 9 (a) and Fig. 10.

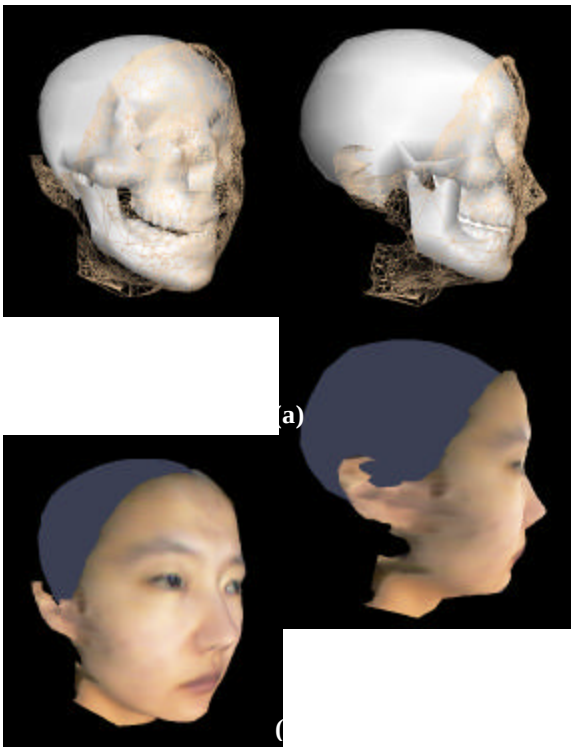


Figure 9: Generic skull model fitted to the facial skin mesh. (a) Fitting skull with skin mesh; (b) Fitting skull with texture-mapped facial surface.

Facial muscles are the major components in anatomy that are responsible for the movement of the face. In our previous work [25], we have developed three kinds of muscle models which can efficiently simulate the contraction of linear, sheet and sphincter facial muscles. Based on FACS (Facial Action Coding System) [8], we select 23 major functional facial muscles to simulate facial expressions (see Fig. 10). We use 3 sphincter muscles to represent the orbicularis oris and orbicularis oculi. The other muscles are represented as pairs of muscles which have left and right components.

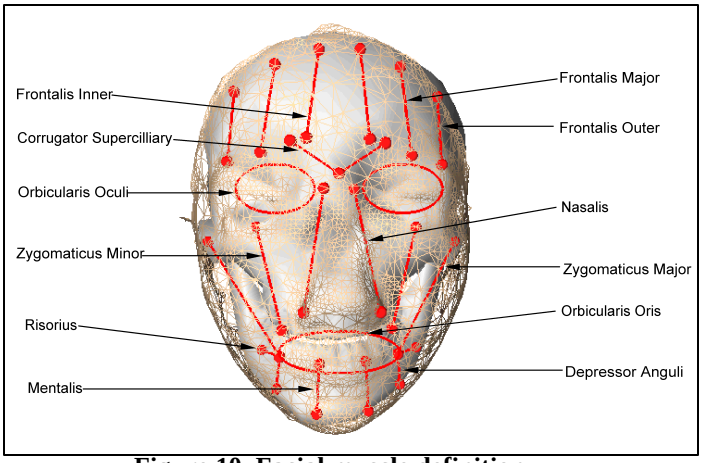


Figure 10. Facial muscle definition.

When facial muscles contract, the facial skin points that are in the influence area of the muscle are displaced to their new positions. As a result, the skin points that are not influenced by the muscle contraction are in an unstable state, and unbalanced elastic forces propagate through the MSD system to establish a new equilibrium state. Therefore the new position of each skin point is obtained by calculating the energy equilibrium state of the entire system. Based on the Lagrangian dynamics, the deformable facial model equations of motion can be expressed in 3D vector form by the ODE of type:

$$\mathbf{M}\frac{d^2\mathbf{x}(t)}{dt^2}+\mathbf{D}\frac{d\mathbf{x}(t)}{dt}+F_{ela}(\mathbf{K},\mathbf{x}(t))=\mathbf{F}_{mus}(\mathbf{x}(t)) \qquad (4)$$

Given n mass points of the skin mesh the Eq. 4 establishes the equilibrium of forces for each of the m_k of the diagonal mass matrix $\mathbf{M}\hat{I}\ R^{3n\times 3n}$ with $diag(\mathbf{M})=[m_1,m_1,m_1,...,m_n,m_n,m_n]$ and describes their positional movement $\mathbf{x}(t)=[\mathbf{x}_1(t),...,\mathbf{x}_n(t)]$ over time t. $\mathbf{D}$ denotes the damping matrix. It is a sparse matrix, element of $R^{3n\times 3n}$ consisting of n matrices $\mathbf{D}_k\hat{I}\ R^{3\times 3}$. $\mathbf{F}_{mus}$ is vector of dimension $3n$ and represents the vector of muscular force. It can be divided into n 3-dimensional vectors. The internal force $\mathbf{F}_{ela}$ that is basically required to be computed by the motion dynamics of the skin model is the resultant of the tension of the springs. We use a nonlinear function to describe the stress-strain relationship. Suppose an arbitrary skin point $\mathbf{x}_i$ is connected to one of its neighbors $\mathbf{x}_j$ by a spring. We introduce a function $K(\mathbf{x}_i,\mathbf{x}_j)$ to modulate a constant spring stiffness k_0:

$$K(\mathbf{x}_i,\mathbf{x}_j)=(1+|\mathbf{x}_i-\mathbf{x}_j|-d_{ij})^a\ k_0 \qquad (5)$$

where d_{ij} is the rest length of the spring, a is the nonlinearity factor controlling the modulation. The spring force function is:

$$f_s(\mathbf{x}_i,\mathbf{x}_j)=-K(\mathbf{x}_i,\mathbf{x}_j)\cdot\frac{(|\mathbf{x}_i-\mathbf{x}_j|-d_{ij})}{|\mathbf{x}_i-\mathbf{x}_j|}\cdot(\mathbf{x}_i-\mathbf{x}_j) \qquad (6)$$

Total elastic force applied on the node $\mathbf{x}_i$ is:

$$F_{ela}(\mathbf{x}_i)=\sum_{j\in\Omega_i} f_s(\mathbf{x}_i,\mathbf{x}_j) \qquad (7)$$

where Ω_i is the index set of neighboring mass points of $\mathbf{x}_i$.

We divide the second order differential equation for a single mass m_i into a system of first order differential equations by introducing the velocity function $\mathbf{v}_i(t)$.

$$\begin{cases}\dfrac{d\mathbf{x}_i(t)}{dt}=\mathbf{v}_i(t)\\[2ex]\dfrac{d\mathbf{v}_i(t)}{dt}=\dfrac{\mathbf{F}_{mus}(\mathbf{x}_i(t))-\mathbf{F}_{ela}(\mathbf{x}_i(t),\mathbf{K})-\mathbf{D}\mathbf{v}_i(t)}{m_i}\end{cases} \qquad (8)$$

The solution of the nodal displacement, velocity and acceleration at each time step is obtained by using a semi-implicit method to integrate Eq. 8 numerically.

4. Results

To demonstrate the hierarchical facial modeling based on range data and advantages of physically-based deformation, we have developed a facial animation system. It is programmed in C++/OpenGL and runs on a SGI540 Visual Workstation with dual 550MHz PIII Xeon processors, 512MB RAM. Fig. 11 depicts the windows GUI of our system which gives the user the flexibility to easily tailor the facial parameters to create facial expressions derived from facial actions.

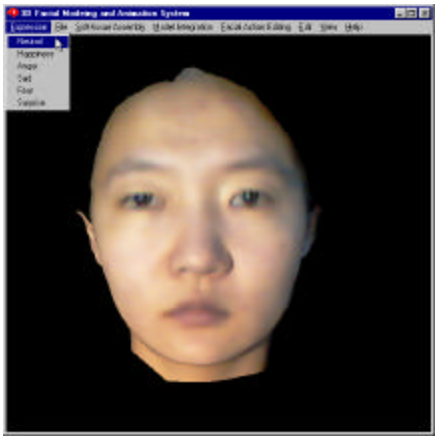


Figure 11. The GUI of our facial animation system.

In the simulation, the anatomy-based face model with physical attributes is deformed to create various expressions. Fig. 12 shows the results obtained with the described prototype. For the clarity, different views of the face model in each expression are shown. Fig. 12 (a) illustrates the texture-mapped face in the neutral state. The remaining images in Fig. 12 show the animated faces with given expression parameters. Note that all the synthesized expressions are generated dynamically. The sequences of images presented in Fig. 13 (a) shows several frames of the dynamic deformation of the face model in the simulation of "Happiness", which is created by an activation of the muscles zygomatics major, zygomatics minor and mentalis. Fig. 13 (b) shows the dynamic simulation of "Anger", which results from the contraction of the corrugator supercilliary, zygomatics minor and nasalis. Considering approximate 7,200 basic soft-tissue elements, the average simulation speed of different expressions is 12 frames per second on the current experimental platform.

We have used our system to construct face models and to simulate facial expressions for various people. The examples are shown in Fig. 14. Females and males in different ethnic groups with various ages are reconstructed and animated by our system. In all these examples, after the semi-automatic modeling stage, the expression is regenerated dynamically, thus enabling the range data of any individuals to come alive in a short time.

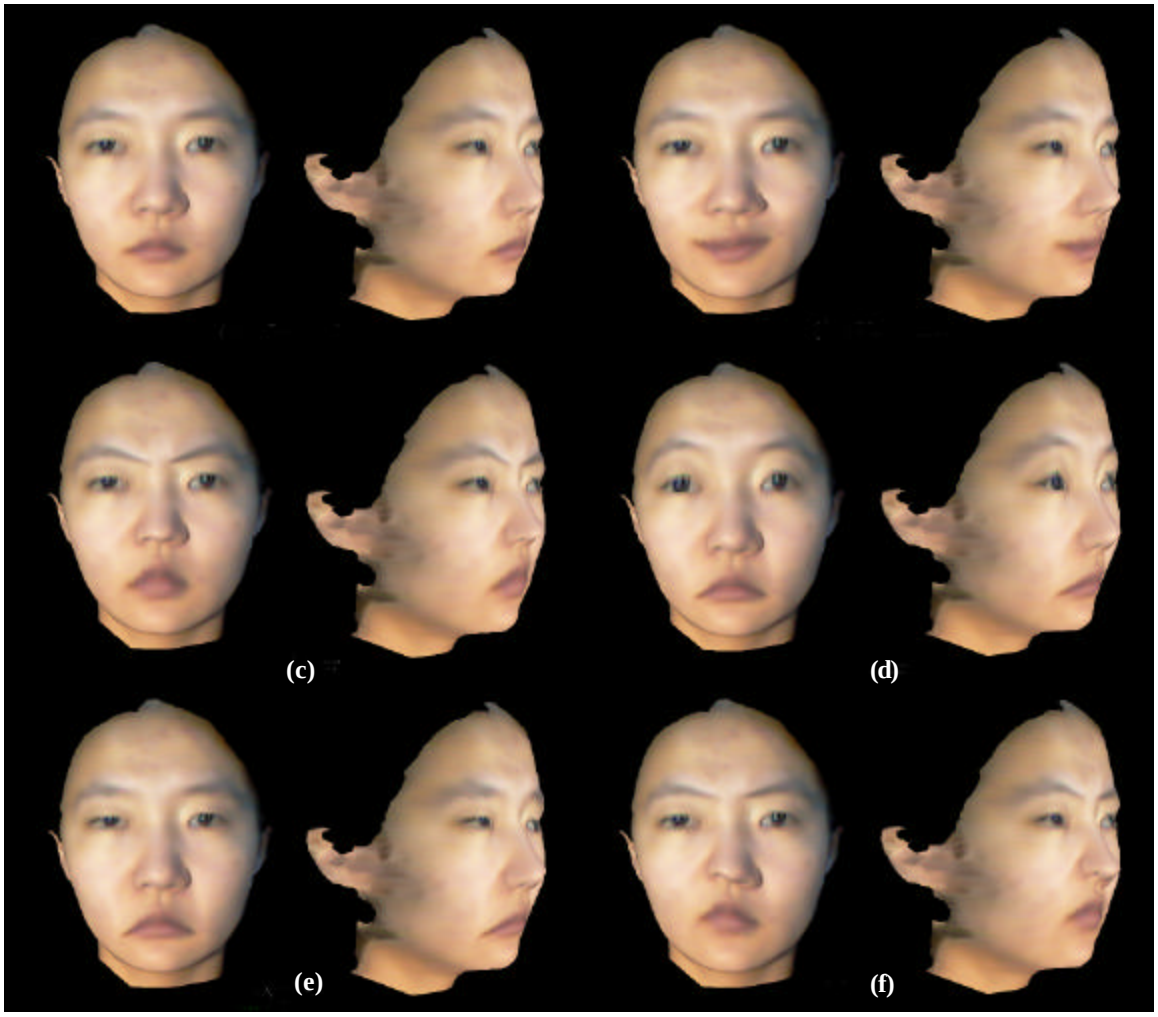
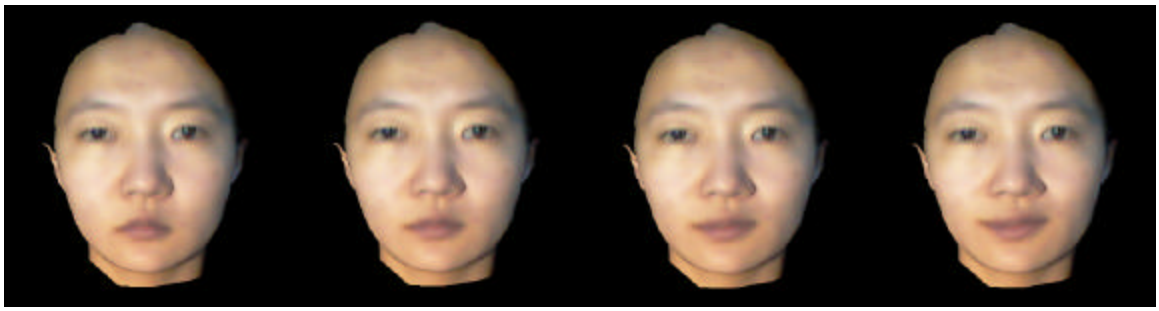


Figure 12: Expressions synthesized on the face model: (a) neutral face; (b) happiness; (c) anger; (d) fear; (e) sadness; and (f) disgust. Each expression shows the front view and view from 60 degrees.



(a)

(b)

Figure 13. Dynamic deformation of the face in the generation of “Happiness" (a) and “Anger" (b).

Figure 14: Examples of synthesized expressions on a wide variety of personalized face models.

5. Conclusion and Future Work

This paper has presented a technique for constructing a realistic and complete animation model that portrays a specific person’s facial geometry and texture. Starting with the range and reflectance data acquired by a laser range scanner from different views of a human subject, we first recover a 3D dense reconstruction of the face by a semi-automatic registration and merging of each part of face data. After model editing and mesh reduction processes, the resulting smooth and computational efficient face mesh facilitates further anatomy-based face modeling and physically-based animation. A multi-layer MSD skin model based on reconstructed facial surface is developed to approximate different types of soft tissue and simulate nonlinear behavior exhibited in the skin deformation. The underlying muscle and skull layers are also modeled to achieve an authentic and functional face model with hierarchical structure. By numerically solving the governing differential equations, various expressions can be realistically animated and the subject’s face comes to life.

Currently our model only has the face part, the back of the head, ears and hair have not been incorporated. In fact, these components are also crucial to the overall realism of the model. In the future the expressions will be animated on the reconstructed full head model. The extension is straightforward. Special care should be taken on the computation optimization to achieve real-time performance. The reconstructed face model is texture mapped by the color image of the front view. It is not sufficient for a full face mapping since the texture data of two sides of the face is lost. To achieve more photorealistic textured face model, we will blend multiple texture maps into a single, comprehensive one.

References

[1] Takaaki Akimoto, Yasuhito Suenaga, and Richard S. Wallace, "Automatic creation of 3D facial models", *IEEE Computer Graphics and Applications* 13(5): 16–22, September 1993.
[2] Chang Seek Choi, Kiyoharu, Hiroshi Harashima, and Tsuyoshi Takebe, "Analysis and synthesis of facial image sequences in model-based image coding", *IEEE Transactions on Circuits and Systems for Video Technology*, vol.4, pp. 257-275. June 1994.
[3] P. Cignoni, C. Montani, and R. Scopigno, "A comparison of mesh simplification algorithms," *Computer and Graphics*, 22(1):37–54, 1998.
[4] B. Dariush, S. B. Kang, and K. Waters, "Spatiotemporal analysis of face profiles: detection, segmentation, and registration," *Proc. of the 3rd IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 248-253, April 1998.
[5] D. DeCarlo, D. Metaxas "The integration of optical flow and deformable models with applications to human face shape and motion estimation," *Proc. CVPR'96*, pp. 231-238, June 1996.
[6] S. DiPaola, "Extending the range of facial types,"*Journal of Visualization and Computer Animation*, 2(4): 129-131, 1991.
[7] P. Eisert, S. Chaudhuri and B. Girod, "Speech driven synthesis of talking head sequences", *3D Image Analysis and Synthesis*, pp. 51-56, Erlangen, November 1997.
[8] P. Ekman and W. V. Friesen, *Facial Action Coding System*, Consulting Psychologists Press Inc., 577 College Avenue, Palo Alto, California 94306, 1978.
[9] L. G. Farkas, *Anthropometry of the Head and Face*, 2nd ed., Raven Press, 1994.
[10] P. Fua, "From multiple stereo views to multiple 3d surfaces", *International Journal of Computer Vision*, 24(1): 19-35, August 1997.
[11] P. Fua, "Using model-Driven bundle-adjustment to model heads from raw video sequences", *ICCV'99*, pp.46-53, Sept. 1999.
[12] Rien-Lin Hsu and A.K. Jain, "Face modeling for recognition", *Proc. ICIP'2001*, , Oct. 2001.
[13] Horace H. S. Ip, Lijun Yin, "Constructing a 3D individualized head model from two orthogonal views", *The Visual Computer*, (12): 254-266, 1996.
[14] P. Kalra, A. Mangili, N. Magnenat-Thalmann, D. Thalmann, "Simulation of facial muscle actions based on rational free form deformations," *Proc. Eurographics'92*, pp. 59-69, 1992.
[15] Koch R., Gross M., Carls F., Buren D., Fankhauser G. and Parish Y. , "Simulating facial surgery using finite element models," *Proc. SIGGRAPH'96*, vol.30, pp. 421-428, August 1996.
[16] Cyriaque Kouadio, Pierre Poulin, Pierre Lachapelle, "Real-time facial animation based upon a bank of 3D facial expression," *Proc. Computer Animation'98*, pp. 128-136, 1998.
[17] Zicheng Liu, Zhengyou Zhang, Chuck Jacobs, Michael Cohen, "Rapid Modeling of Animated Faces From Video," *Microsoft TechnicalReport, MSR-TR-2000-11*, 2000.
[18] M. Nahas, H. Huitric, and M. Saintourens, "Animation of a B-spline figure," *The Visual Computer*, 3(5): 272-276, March 1988.
[19] F. I. Parke, *Computer generated animation of faces*, Master's thesis, University of Utah, Salt Lake City, June 1972.
[20] F. I. Parke, "Parameterized models for facial animation," *IEEE Computer Graphics and Application*, 2(9): 61-68, November 1982.
[21] W. H. Press, S. A. Teukolsky, W. T. Verrerling, and B. P. Fannery, *Numerical Recipes in C: The Art of Scientific Computing (2nd ed.)*, Cambridge University Press, Cambridge, UK, 1992.
[22] William J. Schroeder, Jonathan A. Zarge, and William E. Lorensen, "Decimation of triangle meshes," *Proc. SIGGRAPH'92*, vol.26, pp. 65-70, July 1992.

[23] Hai Tao and Thomas S. Huang, "Facial animation and video tracking," *CAPTECH'98*, pp. 242-253, 1998.
[24] C. L. Y. Wang, D. R. Forsey, "Langwidere: A new facial animation system," *Proc. Computer Animation'94*, pp. 59-68, 1994.
[25] Y. Zhang, E. C. Prakash and E. Sung, "A physically-based model with adaptive refinement for facial animation", *Proc. Computer Animation 2001*, pp. 28-39, Nov. 2001.

Biography

Yu Zhang received his BE in 1997 and ME in 1999 from Northwestern Polytechnical University, Xi'an, China. He is currently a Ph. D student in the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore. He is a student member of IEEE Computer Society and ACM SIGGRAPH. His research interests are in the fields of computer animation, physically-based modeling, and virtual reality.

Edmond Cyril Prakash is an Assistant Professor of Computer Science at the Nanyang Technological University, Singapore. He received his BE, ME and Ph.D degrees from Annamalai University, Anna University and Indian Institute of Science respectively. Edmond's research focus is on exploring the applications of virtual reality in engineering, science, medicine and finance. He also serves as a Deputy Director for the Financial Engineering program at NTU, Singapore.

Eric Sung graduated from the University of Singapore in 1971 with BE (Hons), and an MSEE from the University of Wisconsin in 1973. He taught at the Singapore Polytechnic from 1973 to 1978. After that, he worked in the Video Development Lab of Philips Singapore for three years as a development engineer. Subsequent to that, he headed the video development laboratory of Luxor Singapore for two years. Then he was a group leader in the R&D department of King Radio Corporation, Singapore for a year. In 1985, Eric Sung joined Nanyang Technological University as a senior lecturer. He obtained his PhD from Nanyang Technological University in 1999 and is currently an associate professor. His main research interests are in computer vision, intelligent automation and image processing concentrating on stereovision, structure from motion and face recognition problems.