

Hierarchical Facial Data Modeling for Visual Expression Synthesis

Zhang, Y. ^{*1}, Prakash, E. C. ^{*2} and Sung, E. ^{*1}

*1 Division of Control & Instrumentation Engineering, School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798.

*2 Center for Graphics and Imaging Technology, Division of Computer Science, School of Computer Engineering, Nanyang Technological University, Singapore 639798.

Received 19 April 2002

Revised 27 February 2003

Abstract: This paper presents a new method for modeling of a personalized 3D face model that has a hierarchical structure of the skin, muscle, and skull for realistic expression synthesis. We start from an accurate facial mesh reconstructed from the individual face measurements. The deformable skin model has a multi-layer structure to approximate different types of soft tissue and takes into account the nonlinear stress-strain relationship and incompressibility of the real skin. The 3D face model incorporates the skull structure which extends the scope of facial motion and facilitates interactive facial muscle insertion. Under the muscular force, the deformation of the facial skin is evaluated by the numerical integration of the governing Lagrangian dynamics. The resulting animation system enables realistic facial expressions to be dynamically synthesized.

Keywords: Facial expression synthesis, Hierarchical model, Range data, Dynamic simulation.

1. Introduction

A good model of human face is essential for the applications such as teleconferencing, man-machine interface, virtual reality, and surgical facial planning. In the literature of 3D face modeling, several approaches have been proposed; for example, key-frame interpolation approach (Parke, 1972), parameterized models (Choi et al., 1994; DiPaola, 1991; Parke, 1982), control-point models (Kalra et al., 1992; Su et al., 2001), a performance-driven approach (Williams, 1990) and physically-based models (Hashimoto et al., 1996; Sera et al., 1996; Terzopoulos and Waters, 1990; Wu et al., 1999; Zhang et al., 2001). However, all the previous models are adapted from a template face with low resolution, and therefore could not represent a personalized face in an anatomically accurate way. Moreover, they consider the skin as an infinitesimally thin surface with no underlying structure, and nonlinear dynamics of the real skin have not been considered. Another drawback of previous methods is that the inner skull has been underemphasized despite its importance in the jaw motion.

This paper presents new methods that are developed to model an accurate personalized face for realistic expression synthesis. Our original contributions are:

- We develop an anatomy-based face model with the hierarchical structure of skin, muscle and skull based on individual facial measurements.
- Our skin model employs a multi-layer structure constituted of different spring sets to simulate the nonhomogeneity of the soft tissue. By taking into account the nonlinear stress-strain relationship of soft tissue, a kind of nonlinear springs is used to simulate the nonlinear facial skin deformation under facial muscles' contraction.

2. Data Acquisition

The facial geometry and color information is obtained by using a Minolta VIVID 700 Digitizer™. To recover the whole facial geometry, scans are taken from several different view-angles of the subject's face. When each scan is complete, the device has acquired two registered images: a laser range map (Fig. 1 (a)) and a reflectance (RGB) image (Fig. 1 (b)). The range map can be transformed into a 3D mesh representing the surface of the person. Software manipulation allows the surface mesh of different viewpoints to be automatically registered and merged into a surface model (Fig. 1 (c)). The generated mesh consists of more than 10K triangles. In order to enable interactive visualization and fast simulation, we employ Schroeder's adaptive reduction algorithm (Schroeder et al., 1992) for thinning and retiling. Figure 1 (d) shows the adaptively reduced facial mesh.

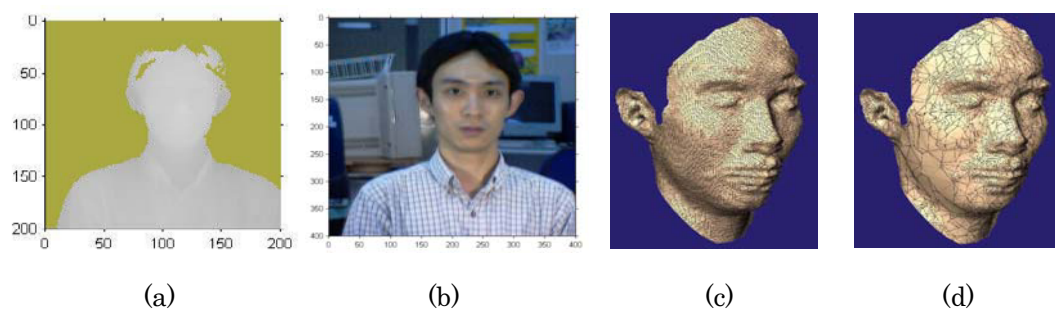


Fig. 1(a). Range image; (b) Reflectance image; (c) Original facial mesh (12,194 vertices and 23,642 triangles); (d) Adaptively reduced mesh (3,766 vertices and 7,206 triangles).

3. Nonlinear Deformable Facial Skin Model

We have developed a multi-layer skin model based on the mass-spring-damper (MSD) system as shown in Fig. 2.

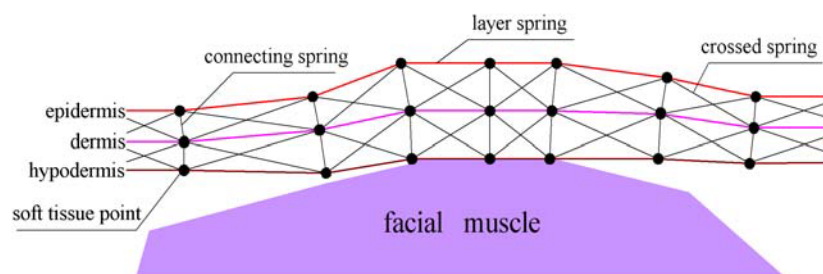


Fig. 2. Multi-layer MSD facial skin model.

In this model, each vertex corresponds to a point mass with mass density m and there are different kinds of springs connecting it to its neighbors. 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. According to their locations and function, all springs are categorized into the following three different sets:

- Layer-spring set: 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.
- Connecting-spring set: springs link a vertex on the layer with its corresponding vertex on the adjacent layer. They resist the traction stresses or compression between layers.
- Crossed-spring set: springs link a tissue layer vertex with the neighbors of its corresponding vertex on the adjacent layer. In this model, the crossed-spring is mainly used to resist shearing and twisting stresses.

In order to simulate nonlinear deformation of the skin, we use a nonlinear function to describe the stress-strain relationship directly. Suppose an arbitrary soft-tissue point \mathbf{x}_i is connected to one of its neighbors \mathbf{x}_j by a spring with the rest length d_{ij} . We introduce a function $K(\mathbf{x}_i, \mathbf{x}_j)$ to modulate a constant spring stiffness k_0 which is called *basic spring stiffness*:

$$K(\mathbf{x}_i, \mathbf{x}_j) = (1 + (|\mathbf{x}_i - \mathbf{x}_j| - d_{ij})^2)^\alpha k_0 \quad (1)$$

Total elastic force applying on node \mathbf{x}_i is:

$$\mathbf{F}_{ela}(\mathbf{x}_i) = - \sum_{j \in \Omega_j} K(\mathbf{x}_i, \mathbf{x}_j) \frac{(|\mathbf{x}_i - \mathbf{x}_j| - d_{ij})}{|\mathbf{x}_i - \mathbf{x}_j|} (\mathbf{x}_i - \mathbf{x}_j) \quad (2)$$

where Ω_i is the index set of neighboring mass points of \mathbf{x}_i . In Eq. (1) α is the *nonlinearity factor* controlling the modulation. By assigning different values to α and k_0 , function $K(\mathbf{x}_i, \mathbf{x}_j)$ can be chosen to model linear or nonlinear stress-strain relationship as shown in Fig. 3. For each spring set, the spring parameter values are set according to Van Gelder, (1998) to model differentiated elasto-mechanical properties of soft tissue layers.

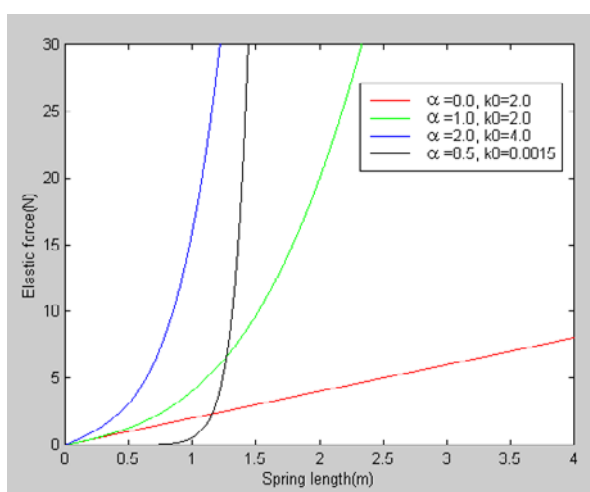


Fig. 3. The stress-strain relationship of nonlinear spring with different values of α and k_0 .

To faithfully simulate the skin deformation, the property of volume preservation should be modeled. In the skin model, the volume variation is penalized by applying an *incompressibility constraint force* to the nodes on the epidermal layer along the normal of each node \mathbf{n}_i . The function of the constraint force is given in Eq. (3).

$$\mathbf{F}_{incom}(\mathbf{x}_i) = \rho \sum_{j \in N_i} \left(\frac{V_j - V_j^0}{V_j^0} \right)^2 \cdot \mathbf{n}_i \quad (3)$$

where V_j^0 and V_j are the initial and current volumes of the prismatic element surrounding epidermal node \mathbf{x}_i . N_i is the set of the indices of neighboring elements and ρ is the scaling factor.

4. Skull Layer

We use a generic skull model that consists of an immovable skull and a rotating jaw to map general anatomical attributes to the facial surface. In the skull model fitting process, affine transformations - rotation, translation and nonuniform scaling are applied on the skull model by the user interactively.

To distinguish between the skin part that lies over the skull or over the jaw, we have mapped the 3D facial surface and adapted skull model to a 2D image plane by using a cylindrical projection. The mapping of world coordinates $\mathbf{x} = (x, y, z)$ to 2D cylindrical coordinates (u, v) uses

$$u = \tan^{-1}\left(\frac{x}{z}\right), \quad v = y \quad (4)$$

In our previous work (Zhang et al., 2001), we have developed three kinds of muscle models to simulate the contraction of facial muscles. Based on the FACS (Facial Action Coding System) (Ekman and Friesen, 1978), we use 23 muscles to simulate facial expressions. By simple mouse clicking and dragging, one end of each fiber can be attached to the skull and the other end can be moved to a skin vertex easily. For the sphincter muscles, an elliptical geometry is scaled to the size of the mouth and the eyes and the contraction center is specified. Figure 5 illustrates the constructed model with the structure of skin, skull and muscles.

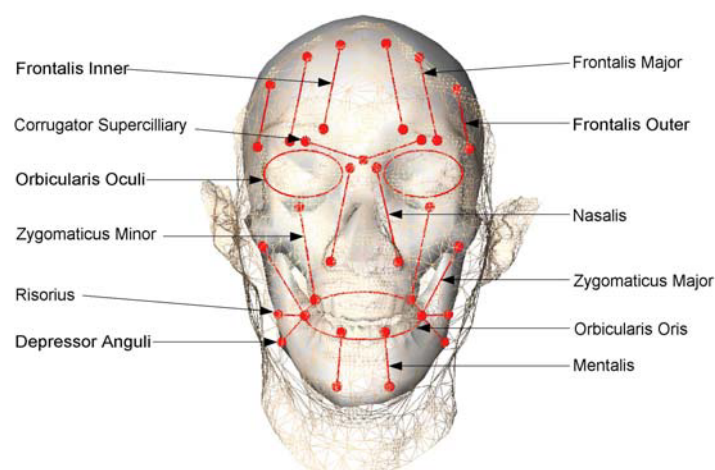


Fig. 4. Integrated face model with hierarchical structure.

5. Numerical Simulation

Facial expressions are simulated by activating a set of parameterized muscles based on the *action units* of the FACS. When facial muscles contract, the facial skin points that are in the influence area of the muscle are displaced to their new positions. 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 second-order ordinary differential equation (ODE) of type:

$$\mathbf{M} \frac{d^2 \mathbf{x}(t)}{dt^2} + \mathbf{D} \frac{d\mathbf{x}(t)}{dt} + \mathbf{K}\mathbf{x}(t) = \mathbf{F}_{mus}(\mathbf{x}(t)) + \mathbf{F}_{incom}(\mathbf{x}(t)) \quad (5)$$

We can take the elastic force expression as an external force $\mathbf{F}_{ela}(\mathbf{x}(t), \mathbf{K}) = \mathbf{K}\mathbf{x}(t)$, and take \mathbf{F}_{ela} to the right hand side of the Eq. (5). This new form of the equation will simplify the formulation procedure.

$$\mathbf{M} \frac{d^2 \mathbf{x}(t)}{dt^2} + \mathbf{D} \frac{d\mathbf{x}(t)}{dt} = \mathbf{F}_{mus}(\mathbf{x}(t)) + \mathbf{F}_{incom}(\mathbf{x}(t)) - \mathbf{F}_{ela}(\mathbf{x}(t)) \quad (6)$$

Given n mass points of the skin mesh the Eq. (6) establishes the equilibrium of forces for each of the m_k of the diagonal mass matrix $\mathbf{M} \in \mathbb{R}^{3n \times 3n}$ with $diag(\mathbf{M}) = [m_1, m_1, m_1, \dots, m_n, m_n, m_n]$ and describes their positional movement $\mathbf{x}(t) = [\mathbf{x}_1(t), \dots, \mathbf{x}_n(t)]$ over time t . \mathbf{D} denotes the damping matrix. It is a sparse matrix, element of $\mathbb{R}^{3n \times 3n}$ consisting of n matrices $\mathbf{D}_k \in \mathbb{R}^{3 \times 3}$. \mathbf{F}_{mus} , \mathbf{F}_{incom} and \mathbf{F}_{ela} are vectors of dimension $3n$ and represent the vectors of muscular, incompressibility constraint and elastic force respectively.

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

$$\begin{cases} \frac{d\mathbf{x}_i(t)}{dt} = \mathbf{v}_i(t) \\ \frac{d\mathbf{v}_i(t)}{dt} = \frac{\mathbf{F}_{mus}(\mathbf{x}_i(t)) + \mathbf{F}_{incom}(\mathbf{x}_i(t)) - \mathbf{F}_{ela}(\mathbf{x}_i(t)) - \mathbf{D}\mathbf{v}_i(t)}{m_i} \end{cases} \quad (7)$$

The solution of the nodal displacement, velocity and acceleration at time $t+\Delta t$ can be obtained by numerical integration of Eq. (7) using second-order Runge-Kutta method. The algorithm runs as follows:

1. Initialization:
 - Form mass matrix \mathbf{M} and damping matrix \mathbf{D} ;
 - Initialize displacement, velocity and acceleration: \mathbf{x}^0 , \mathbf{x}'^0 and \mathbf{x}''^0 ;
 - Store matrix value in memory.
2. Time loop:
 - Calculate the muscular force vector \mathbf{F}_{mus} based on the muscle models;
 - Compute the incompressibility constraint force vector \mathbf{F}_{incom} due to volume variation and elastic force vector \mathbf{F}_{ela} ;
 - Solve for acceleration at time $t+\Delta t$ from stored matrix (\mathbf{M}, \mathbf{D}) and computed force vectors;
 - Calculate new velocity and position at time $t+\Delta t$.

6. Simulation Results

Our facial animation system is programmed with C++/OpenGL and runs on an Intergraph Zx10 with dual Pentium III 730MHz, 512MB memory. Figure 5 shows the facial model in the neutral expression and five typical expressions synthesized with the given simulation parameters. Each example shows two views of the rendered face with facial texture, the wire-frame mesh representing the skin layer of epidermis, and 3D optical flow which characterizes the pattern of each expression. The result agrees well with our anticipation:

- **Happiness** : Two dominant contractions correspond to the left and right *zygomatic majors*, which matches well with the muscle actuation in real smiling.
- **Anger** : The contractions of the pairs of *corrugator supercillii* and *nasalis* are dominant, which result in frowning of the eyebrows and raising of the alae nasi.
- **Surprise** : The contractions of pairs of *frontalis major*, *frontalis inner* and *frontalis outer* raise eyebrows and jaw rotation results in mouth opening.
- **Disgust** : It is generated by the contractions of left *zygomastics major*, *zygomastics minor*, *nasalis* and pairs of *corrugator supercillii* and *orbicularis oculi*.
- **Sadness** : The contractions of left and right *orbicularis oculi* are dominant in lowering down the eyelids and contraction of *depressor anguli* matches the muscle actuation in sadness.

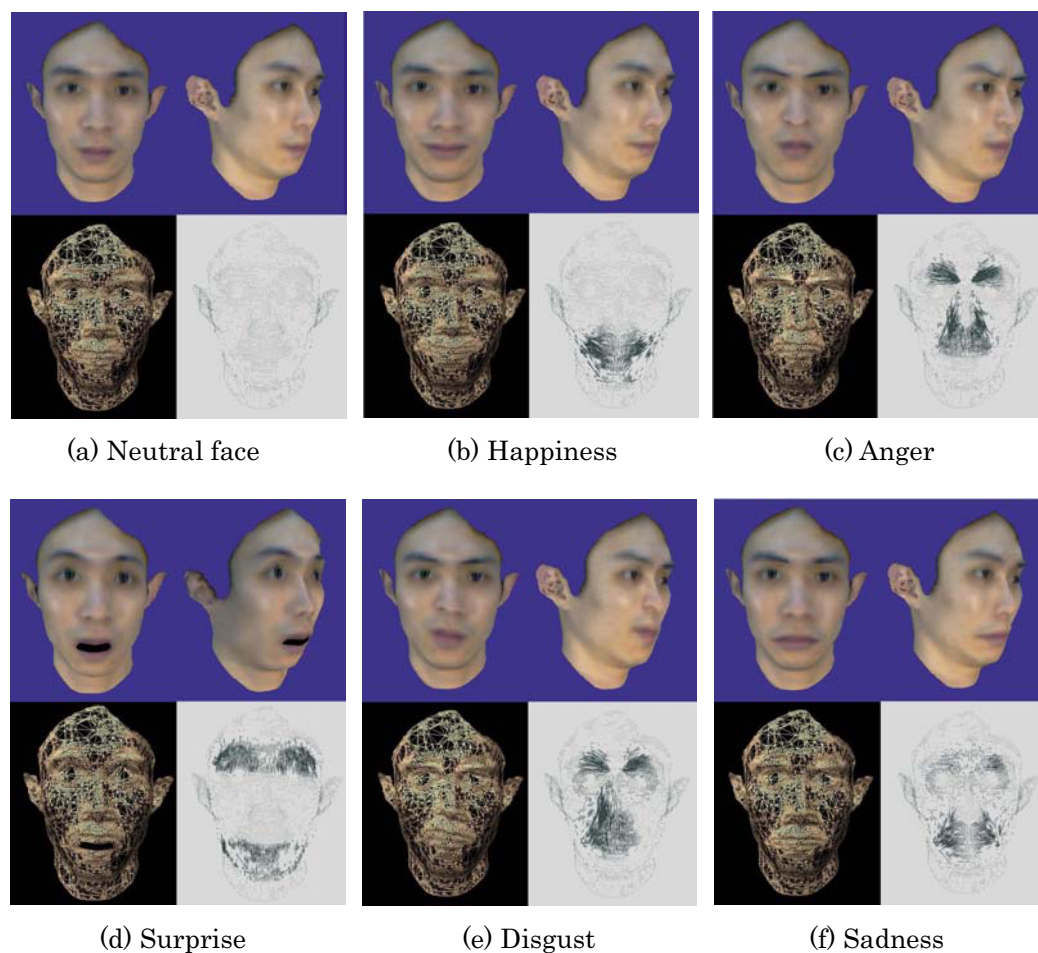
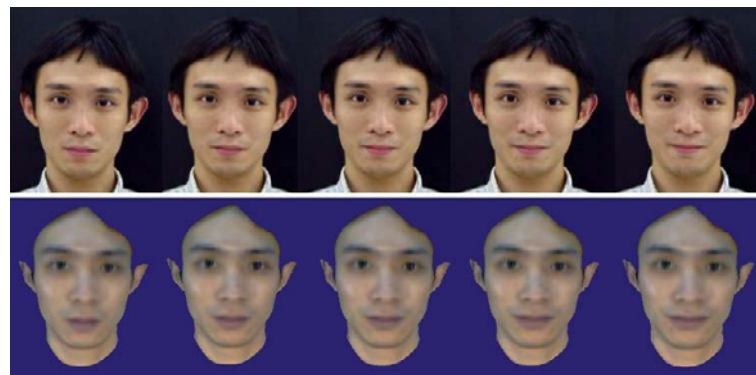
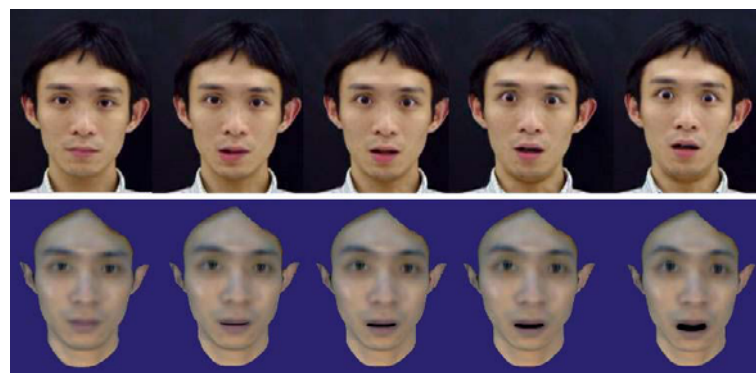


Fig. 5. Typical expressions synthesized on the 3D personalized face model.



(a) Happiness



(b) Surprise

Fig. 6. Comparison of the simulated expressions and actual ones. In each example, the top row shows video images of the subject generating a real expression. The bottom row is the output from the simulation (left to right).

In Fig. 6, several frames of the dynamic deformation of the face model in the synthesis of different expressions are shown. The snapshots from video sequences of the corresponding actual expressions generated by the subject are also given for comparison with our imitation of the real ones. Since our model is based on the anatomical analysis, the movement of the muscles in synthesizing each expression is similar to that of the real face. The skin is deformed along the direction provided by the muscles. As can be seen, in both examples the model produces highly realistic expression movement and natural skin motion. For the performance, we achieve the frame rate of 18 fps rendering speed on current experimental platform.

7. Conclusion and Future Work

This paper presents a new method for realistic facial expression synthesis based on an individualized 3D face model. The skin model has a multi-layer MSD structure constituted of different spring sets to simulate the nonhomogeneity of the soft tissue by taking into account the nonlinear stress-strain relationship and the incompressibility of the skin. By incorporating the underlying skull structure, the scope of facial motion is extended and facial muscles are defined at anatomically correct positions. The dynamic facial deformation is computed by solving the governing

motion equation numerically. A more sophisticated approach will be developed for an automatic and more precise fit. We also intend to quantitatively evaluate the simulation by using kinematic data which records the 3D positions of some facial feature points during the generation of real expressions.

References

- Choi, C. S., Aizawa, K., Harashima, H. and Takebe, T., Analysis and Synthesis of Facial Image Sequences in Model-based Image Coding, *IEEE Trans. Circuits and Systems for Video Technology*, 4-3 (1994), 257-275.
- DiPaola, S., Extending the Range of Facial Types, *Journal of Visualization and Computer Animation*, 2-4 (1991), 129-131.
- Ekman, P. and Friesen, W. V., *Facial Action Coding System*, (1978), Consulting Psychologists Press Inc., 577 College Avenue, Palo Alto, California.
- Hashimoto, S., Sato Y. and Oda, H., Modification of Facial Expression Using Spring Frame Model, *Proceedings of IASTED*, (1996), 9-15.
- Kalra, P., Mangili, A., Thalmann, N. M. and Thalmann D., Simulation of Facial Muscle Actions Based on Rational Free Form Deformations, *Proceedings of Eurographics'92*, (1992), 59-69.
- Parke, F. I., *Computer Generated Animation of Faces*, (1972), Master's Thesis, University of Utah, Salt Lake City.
- Parke, F. I., Parameterized Models for Facial Animation, *IEEE Computer Graphics and Application*, 2-9 (1982), 61-68.
- Schroeder, W. J., Zarge, J. A. and Lorensen, W. E., Decimation of Triangle Meshes, *Proceedings of SIGGRAPH'92*, (1992), 65-70.
- Sera, H., Morishima, S. and Terzopoulos, D., Physics-based Muscle Model for Mouth Shape Control, *Proceedings of IEEE Interantional Workshop on Robot and Human Communication*, (1996), 207-212.
- Su, M. S., Ko, M. T. and Cheng, K. Y., Feature-point-driven Facial Animation using A Hypothetical Face, *Computer Graphics Forum*, 20-4 (2001), 179-188.
- Terzopoulos, D. and Waters, K., Physically-based Facial Modeling, Analysis and Animation, *Journal of Visualization and Computer Animation*, 1 (1990), 73-80.
- Van Gelder, A., Approximate Simulation of Elastic Membranes by Triangulated Spring Meshes, *Journal of Graphics Tools*, 3-2 (1998), 21-41.
- Williams, L., Performance-driven Facial Animation, *Proceedings of SIGGRAPH'90*, (1990), 235-242.
- Wu, Y., Kalra, P., Moccozet, L. and Thalmann, N. M., Simulating Wrinkles and Skin Aging, *The Visual Computer*, 15-4 (1999), 183-198.
- Zhang, Y., Prakash, E. C. and Sung, E., A Physically-based Model with Adaptive Refinement for Facial Animation, *Proceedings of IEEE Computer Animation 2001*, (2001), 28-39.

Author Profile



Yu Zhang: He received his BE in 1997 and ME in 1999 from Northwestern Polytechnical University, 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 Computer Animation, Physically-based Modeling, 3D Visualization, and Virtual Reality.



Edmond Cyril Prakash: He 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: He 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. 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 interest are in computer vision, intelligent automation and image processing concentrating on stereovision, structure from motion and face recognition problems.