EXAGGERATION OF EXTREMELY DETAILED 3D FACES

Andrew Soon Won-Sook Lee

School of Information Technology and Engineering, University of Ottawa
800 King Edward Avenue, Ottawa, ON, Canada K1N 6N5
E-mail: asoon@discover.uottawa.ca, wslee@uottawa.ca

ABSTRACT

Exaggeration is often used in art and entertainment to capture the interest and attention of an audience. We present an approach to automatically exaggerate the distinctive features of 3D faces which contain skin detail down to the pores. Mesh adaptation and model simplification are used to produce two low resolution approximations of the face. The detail of the original high resolution face is captured by achieving two sets of model parameterizations: the high resolution face with respect to the face obtained using model simplification (simplified model) and the simplified model with respect to the model obtained using mesh adaptation (working model). The working model is exaggerated using a vector-based algorithm that automatically identifies the prominent features with respect to an average face. The resulting model and the parameterizations drive a two-stage model reconstruction process that generates the high resolution exaggerated model which preserves the original level of detail. The results of our testing show that the proposed methodology is capable of producing exaggerated models from an initial face model comprising roughly 2,000,000 triangles.

KEY WORDS

entertainment, exaggeration, high resolution, 3D faces

1. Introduction

Exaggeration is commonly used in art and entertainment to varying degrees and for different reasons. Regardless of its purpose, exaggeration usually captures the attention and the interest of an audience. Many studies have been conducted into the use of exaggeration in digital works. However, many of these earlier studies like [1], [2] and [3] were conducted in the context of 2D images and not extremely detailed 3D models. Some of today's most advanced scanning technologies can yield extremely accurate digital representations of real-world objects. For example, the laser scanning services offered by XYZ RGB (http://www.xyzrgb.com) utilize a combination of technologies developed by the National Research Council of Canada [4] capable of capturing surface detail in the order of about 100 microns. The resulting scan data can yield polygonal models comprising several million triangles. Working directly with such models to create high resolution exaggerations is expensive, inefficient and potentially intractable from both a computational and a resource requirement standpoint.

A common approach to avoid these pitfalls is to first perform the desired changes on a low polygon approximation of the original model and then use advanced techniques like displacement mapping to achieve higher resolution renderings or models. Previous work such as [5], [6], [7] and [8] has demonstrated that the low polygon models used in this practice can be constructed in many different ways. However, methodologies such as [5] which rely on low resolution models containing arbitrary structures (i.e. connectivity between vertices) have some significant drawbacks. Without performing any further work on these models, it is not possible to determine the locations of landmarks such as the eyes and nose; consequently, automatic exaggeration of features is not possible either. Additionally, animating these arbitrary structures would either require mechanical work or the use of other techniques in the literature, which could add significant costs to the workflow.

We propose an approach to efficiently construct exaggerated versions of extremely detailed 3D faces obtained by technologies such as [4]. Low polygon approximations of the detailed model are prepared by employing mesh adaptation and model simplification techniques. A two-step procedure captures the high resolution detail of the original face using the low polygon models by performing model parameterization. Parameterization is achieved by mapping points in 3D space to a surface and allows a high resolution model to be reconstructed from the low resolution models. A vector-based caricature algorithm is applied to the low polygon model constructed using mesh adaptation. The algorithm automatically identifies and exaggerates the pronounced features of the face by comparing it to an average face. The resulting exaggerated model and the model parameterizations drive a two-stage model reconstruction process that produces a high resolution exaggerated model which preserves the level of detail in the original face.

This paper consists of eight sections. An overview of our proposed approach is presented in Section 2. Section 3 describes the derivation of the low resolution model from a generic head model using surface fitting. Section 0 explains the two-step procedure to capture high resolution detail using the low polygon models. The technique of automatically exaggerating characteristic features is discussed in Section 5 and the process of constructing the high resolution exaggeration is covered in Section 6. Sample results achieved with our methodology are presented in Section 0 and are followed by a conclusion in Section 8.

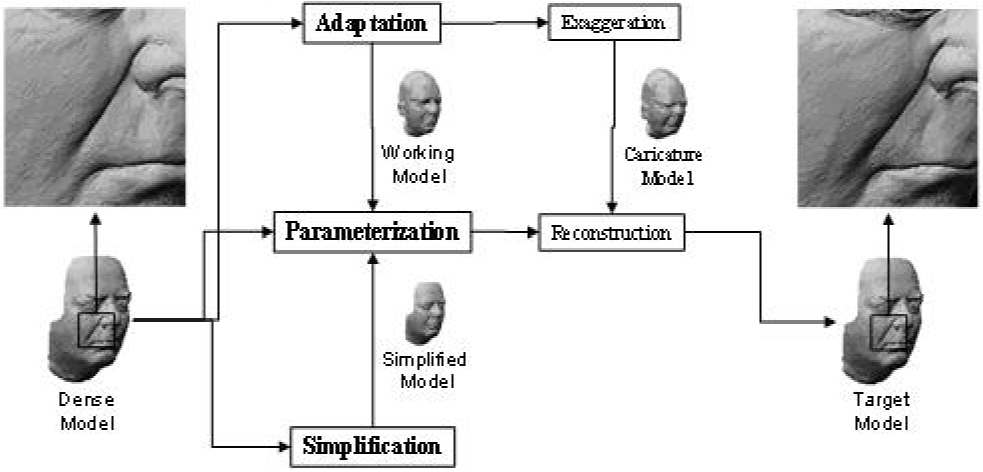


Figure 1: Flow diagram for producing extremely detailed exaggerated models

2. Overview

Figure 1 shows a flow diagram of our approach. The models and stages shown in the workflow are briefly discussed below. The names assigned to the various models are carried over into the discussions in proceeding sections.

- *Dense Model*: The high resolution head or face which is to be exaggerated.
- *Adaptation*: The mesh adaptation technique described in [9] used to fit a generic model to the dense model.
- *Working Model*: The low polygon approximation obtained using adaptation.
- *Simplification*: Any model simplification technique which produces a good low polygon approximation that preserves the shape of the dense model.
- *Simplified Model*: The result of performing simplification on the dense model.
- *Parameterization*: Captures the high resolution detail of the dense model using the point-to-surface mapping algorithm described in [7] and [8]. The dense model is parameterized with respect to the simplified model and the simplified model is parameterized with respect to the working model.
- *Exaggeration*: The vector-based caricature algorithm [1] which automatically exaggerates the most pronounced features of the working model with respect to an average face.
- *Caricature Model*: The result of performing exaggeration on the working model.
- *Reconstruction*: A two-step procedure which reconstructs the dense model using the caricature model and the model parameterizations. The first step is to reconstruct the simplified model, which is in turn used in the second step to reconstruct the dense model. Both reconstructed models are exaggerated versions of the originals.
- *Target Model*: The exaggerated dense model produced using model reconstruction. The dense model and the target model have the same level of detail.

3. Constructing the Working Model

The *working model* is constructed by using the adaptive mesh procedure described in [9] to fit a generic head model (Figure 3a) to the high resolution *dense model*. As a result, the point and polygon structures of the generic model are inherited by every working model produced. Additionally, the animation structure incorporated into the generic model makes it possible to animate each working model. The generic model has a set of 163 vertices classified as feature points which represent the most characteristic points used for human recognition.

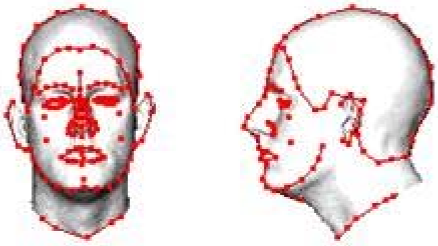
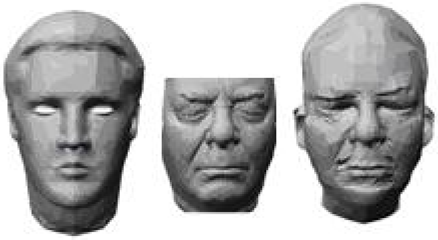


Figure 2: Feature points marked on the front and side view images of a dense model



a. b. c.

Figure 3: a. Generic model (5,900 triangles); b. Dense model (100,000 triangles); c. Working model (5,900 triangles)

In order to adapt the generic model, these feature points (41 major, 122 minor) must first be defined on the dense model. They are semi-automatically marked on 2D front and side view images of the dense model (Figure 2) so that corresponding points in 3D space can be calculated. The positions of the minor feature points are automatically calculated based on the manual placement of the major feature points. Radial basis function (RBF) networks, as described in [10] and [11], are used to deform the generic model at a global level using the 3D feature points as centers. The shape of the resulting model is then refined by adapting its non-feature points to the surface of the dense model to produce the working model. Figure 3 gives an example of the working model produced from a dense model containing 100,000 triangles.

4. Capturing High Resolution Detail

4.1. Point-to-Surface Mapping

The point-to-surface mapping scheme outlined in [7] and [8] is based upon simplification envelopes [12]. This scheme maps a point V in 3D space to a triangle ABC using an interpolated vertex normal (Figure 4). The ray origin P and the interpolated vertex normal $\overline{\mathbf{N}}_{\mathbf{P}}$ are given by

$$\begin{aligned} \mathrm{P} &= (1-u-v)\mathrm{A} + u\mathrm{B} + v\mathrm{C} \\ \overline{\mathbf{N}}_{\mathbf{P}} &= (1-u-v)\overline{\mathbf{N}}_{\mathbf{A}} + u\overline{\mathbf{N}}_{\mathbf{B}} + v\overline{\mathbf{N}}_{\mathbf{C}} \end{aligned}, \qquad (1)$$

where u and v are the 2D barycentric coordinates of P with respect to $\Delta\mathrm{ABC}$ and $\overline{\mathbf{N}}_{\mathbf{A}}$, $\overline{\mathbf{N}}_{\mathbf{B}}$ and $\overline{\mathbf{N}}_{\mathbf{C}}$ are the vertex normals at A, B and C respectively. For all points lying within or along the edges of $\Delta\mathrm{ABC}$, the constraints $u, v \in [0,1]$ and $u + v \leq 1$ are satisfied. The position of V can then be expressed as

$$\mathrm{V} = \mathrm{P} + d\,\frac{\overline{\mathbf{N}}_{\mathbf{P}}}{\left|\overline{\mathbf{N}}_{\mathbf{P}}\right|}, \qquad (2)$$

where d is the signed distance from P to V in the direction of $\overline{\mathbf{N}}_{\mathbf{P}}$ and $|\ |$ denotes vector magnitude.

To calculate the values of u and v in Equation (1), a new triangle $\mathrm{A}^{par}\mathrm{B}^{par}\mathrm{C}^{par}$ that is parallel to $\Delta\mathrm{ABC}$ and whose vertices are coplanar with V is defined (Figure 4). The vertices of this new triangle are obtained by finding the intersection of $\overline{\mathbf{N}}_{\mathbf{A}}$, $\overline{\mathbf{N}}_{\mathbf{B}}$ and $\overline{\mathbf{N}}_{\mathbf{C}}$ with the parallel plane. Computing the barycentric coordinates of V with respect to $\Delta\mathrm{A}^{par}\mathrm{B}^{par}\mathrm{C}^{par}$ yields the values of u and v. In other words,

$$\mathrm{V} = (1-u-v)\mathrm{A}^{par} + u\mathrm{B}^{par} + v\mathrm{C}^{par}. \qquad (3)$$

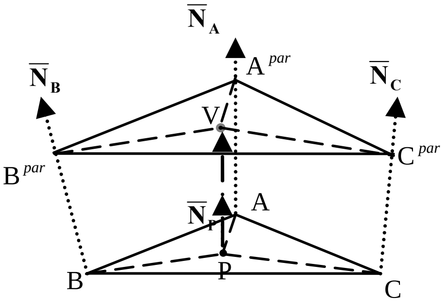


Figure 4: Example of point-to-surface mapping where point $\mathbf{V}$ is mapped to $\Delta\mathbf{ABC}$. $\Delta\mathbf{A}^{par}\mathbf{B}^{par}\mathbf{C}^{par}$ and $\Delta\mathbf{ABC}$ lie on parallel planes and $\mathbf{A}^{par}$, $\mathbf{B}^{par}$, $\mathbf{C}^{par}$ and $\mathbf{V}$ are all coplanar.

4.2. Model Parameterization

The parameterization of a high resolution model with respect to a low resolution model using point-to-surface mapping is discussed in [6]. This process is described here more generally under the name model parameterization where a *subject model* is parameterized with respect to a *control model*. Parameterization allows the control model's shape to influence the shape of the subject model since the points' positions are affected by changes to the control model's vertex normals. This parameterization is achieved by obtaining a set of mapping parameters (I, u, v, d) for each point in the subject model, where I is an identifier for the control model triangle and u, v and d are as defined in the point-to-surface mapping discussion.

Although [7] and [8] use the index into an ordered triangle list for I, an alternative is to use the texture coordinates of the triangle's vertices to add some flexibility. To improve the accuracy of the parameterization, the constraints $u, v \in [0,1]$ and $u + v \leq 1$ are enforced, smaller values of d are favored and a threshold on the value of d is used.

a b c

Figure 5: High resolution exaggerated models achieved using different parameterization techniques. a. Two-level parameterization using a simplified model; b. Single-level parameterization; c. Single-level parameterization with relaxed constraints and a liberal threshold value for displacements

4.3. Detail Capture using a Simplified Model

The high resolution detail of the dense model is captured by establishing two model parameterizations: the dense model with respect to the simplified model and the simplified model with respect to the working model. In this scheme, the simplified model acts as a middle layer to serve two functions: it transfers changes in the working model to the dense model and it also improves high resolution detail capture. This two-level parameterization scheme proved to be superior to a direct parameterization of the dense model with respect to the working model. High resolution exaggerations produced using this scheme contained fewer artefacts and possessed an overall smoother appearance (Figure 5a-b). Even with relaxed constraints and a liberal threshold value for d which lead to overall poorer results, the direct parameterization approach was unable to achieve the same level of detail as the two-level scheme (Figure 5c).

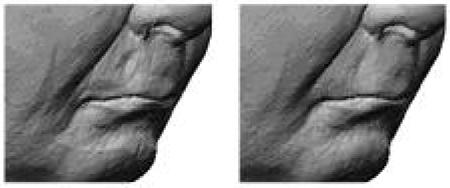

a. b.

Figure 6: High resolution exaggerated models achieved using different model parameterization strategies. a. Control models are not subdivided; b. Control models are subdivided

It is useful to state two observations about model parameterization in practice. Firstly, the suggestion made in [13] to use a smooth domain surface when deforming a control mesh applies here as well. It has been observed that subdividing the control model using Loop's algorithm [14] prior to model parameterization leads to a less faceted appearance in very high resolution target models (Figure 6). Secondly, full model parameterization is often not achievable because of the imposed threshold on displacements; that is, it is often not possible to obtain a set of mapping parameters for every vertex in the subject model. An incomplete dense model parameterization, though perhaps undesirable, has no impact on the overall workflow. On the other hand, an incomplete simplified model parameterization does present a problem since it leads to a partially reconstructed model and the dense model parameterization is dependent upon the full set of simplified model triangles. A workaround for this challenge is discussed later in Section 6.

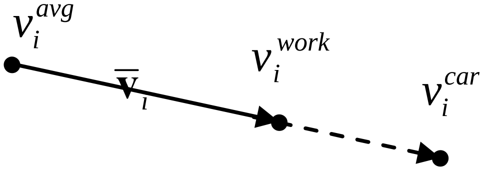


Figure 8: Vector-based exaggeration

5. Constructing the Caricature Model

5.1. Average Face Model

An average face model is constructed by taking the average of a set of working models (**Error! Reference source not found.**). Since full correspondence between working models is already achieved by way of the generic model's point structure, the average face is constructed simply by performing a point-by-point average with each working model making an equal contribution. Consequently, extraordinary characteristics present in only a small subset of faces do not typically dominate the contributions of the other faces. The average face also inherits the same point and polygon structure as the generic model.

5.2. Vector-based Exaggeration

The caricature procedure originally proposed by Brennan [1] is used to exaggerate the working model to produce the *caricature model*. Each feature vector $\overline{\mathbf{v}}_i$ defined as

$$\overline{\mathbf{v}}_i = v_i^{work} - v_i^{avg} , \qquad (4)$$

where v_i^{work} is point i in the working model and v_i^{avg} is point i in the average face, is scaled by a constant exaggeration factor c to derive the point v_i^{car} in the caricature model, namely

$$v_i^{car} = v_i^{avg} + (1+c)\overline{\mathbf{v}}_i \qquad (5)$$

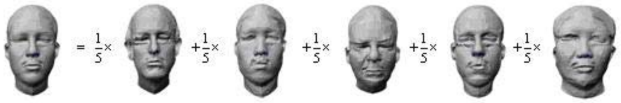


Figure 7: Average face model constructed from five working models

As shown in **Error! Reference source not found.**. Since every feature vector is scaled by the same factor, longer vectors are made even longer, achieving the effect of automatically exaggerating the most prominent facial features (with respect to the average face).

6. Constructing the Target Model

6.1. Model Reconstruction

The premise of the model reconstruction algorithm is to use the model parameterizations (i.e. sets of mapping parameters) to reproduce the vertices of the subject model. As with model parameterization, the control model is subdivided and then its vertex normals are recalculated to reflect changes to its shape. Each set of mapping parameters (I,u,v,d), where I is the control model triangle identifier, u and v are the barycentric coordinates of the ray origin and d is the displacement along the interpolated vertex normal, is applied to the new state of the control model to reconstruct each subject model point (Figure 9). Equations (1) and (2) can then be restated using the new labels in Figure 9 as

$$\begin{aligned} \mathrm{P}' &= (1-u-v)\mathrm{A}' + u\mathrm{B}' + v\mathrm{C}' \\ \overline{\mathbf{N}}'_{\mathbf{P}'} &= (1-u-v)\overline{\mathbf{N}}'_{\mathbf{A}} + u\overline{\mathbf{N}}'_{\mathbf{B}} + v\overline{\mathbf{N}}'_{\mathbf{C}} \end{aligned} \tag{6}$$

and

$$\mathrm{V}' = \mathrm{P}' + d\,\frac{\overline{\mathbf{N}}'_{\mathbf{P}'}}{\left|\overline{\mathbf{N}}'_{\mathbf{P}'}\right|}, \tag{7}$$

respectively.

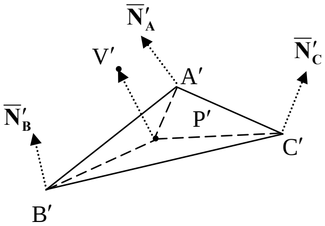
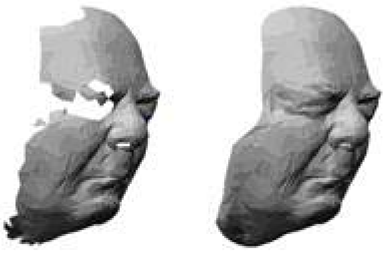


Figure 9: Reconstruction of a point $\mathbf{V}$. Its new position is denoted by $\mathbf{V}'$

6.2. Detail Reconstruction

The high resolution detail of the dense model is reconstructed in a two-step procedure to yield the *target model*. The simplified model's mapping parameters are used with the caricature model in the model reconstruction algorithm to produce an exaggerated version of the simplified model. As mentioned previously, however, this model cannot be used directly to produce the target model if the simplified model parameterization is incomplete (Figure 10a). In such an instance, the holes in the model's surface must be filled in appropriately. The method used to accomplish this is to approximate the surface in the affected areas using RBF networks. First, the correction algorithm determines the vertices that are missing in the reconstructed model and finds the corresponding points in the original simplified model. It then identifies the non-missing vertices in the vicinity of the missing vertices so that they can be used to guide the surface approximation. The RBF networks are initialized with the original positions (i.e. positions in the original model) of these neighbouring vertices and then updated with the new positions of the same points. The networks can then evaluate good approximations of the positions of the missing vertices, thus filling in the missing areas (Figure 10b). Once corrected, the reconstructed simplified model and the dense model parameterization are used by the reconstruction algorithm to produce the target model.



a. b.

Figure 10: a. Simplified model reconstructed using an incomplete parameterization; b. Reconstructed simplified model after surface approximation correction

7. Results

The results presented in this section were produced using an implementation of the proposed methodology running on a desktop PC. The freeware tools VizUp and UVMapper were also used to produce the simplified models and assign UV maps to models, respectively. The high polygon models shown here were provided by XYZ RGB. All simplified models produced for our experiments contained roughly 2,000 triangles and all exaggerations were performed with respect to the average face shown in **Error! Reference source not found.**.

XYZ RGB supplied us with five scanned faces for our testing. The highest resolution model available contained 2,000,000 triangles (Figure 11a and Figure 12a). The corresponding working model (Figure 11b) was exaggerated by a factor of 75% ($c = 0.75$) to produce the caricature model (Figure 11c). The subsequent target model (Figure 11d and Figure 12b) obtained using detail reconstruction comprised of 1,990,616 triangles. Although there were some missing polygons in the central region of the face, the majority of the artifacts occurred around the perimeter of the model. The remaining four scanned faces were provided at a resolution of 30,000 triangles and the results achieved using $c = 0.75$ are given in Figure 13.

The suitability of the proposed methodology was also tested by using several values of the exaggeration factor c . The target models shown in Figure 14 were produced using $c = 0.3$, $c = 0.6$, $c = 0.9$ and $c = 1.2$.

8. Conclusion

An approach to automatically exaggerate the distinctive features in extremely detailed 3D faces is discussed in this paper. Two low polygon approximations of the detailed face are prepared: a working model created by fitting a generic head model to the high resolution data and a simplified model produced by any model simplification algorithm. The high resolution detail is then captured in a two-step procedure by performing model parameterization. Point-to-surface mapping is employed to parameterize the dense model with respect to the simplified model and to parameterize the simplified model with respect to the working model. The working model is exaggerated using a vector-based caricature algorithm which automatically enhances the prominent features by comparing the working model to an average face. Finally, the exaggerated working model and the simplified model parameterization are used to generate an exaggerated simplified model, which is in turn used with the dense model parameterization to produce the exaggerated version of the highly detailed face.

The proposed approach possesses several benefits. It is a suitable solution in applications which require high resolution geometry to be produced instead of merely high resolution renderings. The use of low polygon models keeps execution times and resource requirements modest. With full point correspondence established between working models, calculations to generate the average face and to exaggerate the working model can be performed quickly, making it possible to immediately display these results to a user. The animation structure found in each working model makes it possible to produce animated exaggerations.

The approach does, however, possess some limitations as well. The non-feature points in the eyes and mouth areas of the generic model are not adapted to the surface of the dense model because of large differences between the models' structures in these regions. As a result, these areas are only roughly approximated by the working model and can cause minor errors to appear when large exaggeration factors are used. Issues could also potentially arise because of the inability to achieve full dense model parameterizations, although we did not observe any glaring aberrations in the important areas of the faces we produced.

The results presented in this paper demonstrate that the methodology is capable of handling high resolution models which contain several million triangles. Equally as important is the fact that it is capable of producing models which retain the original level of detail and have minimal noticeable errors. Results can be improved by increasing the number of faces that contribute to the average face to give a truer mean, which in turn allows the system to better identify the characteristic features of each input face.

Acknowledgements

The authors wish to acknowledge the financial support of Materials and Manufacturing Ontario and the in-kind contribution of XYZ RGB Inc. to this research. The contributions of team members Lijia Zhu and Quan Zhou from the University of Ottawa are also recognized.

References

[1] S. Brennan, The caricature generator, *Leonardo, 18*, 1985, 170-178.

[2] P.Y. Chiang, W.H. Liao and T.Y. Li, Automatic caricature generation by analyzing facial features, *Proc. 2004 Asian Conference on Computer Vision*, Jeju Island, Korea, 2004.

[3] L. Liang, H. Chen, Y.Q. Xu and H.Y. Shum, Example-based caricature generation with exaggeration, *Proc. 10th Pacific Conference on Computer Graphics and Applications*, Beijing, China, 2002, 386-393.

[4] J. Taylor, J.A. Beraldin, G. Godin, L. Cournoyer, M. Rioux and J. Domey, NRC 3d imaging technology for museums & heritage, *Proc. First International Workshop on 3D Virtual Heritage*, Geneva, Switzerland, 2002, 70-75.

[5] G. Borshukov and J.P. Lewis. Realistic human face rendering for "The Matrix Reloaded", *Proc. SIGGRAPH 2003 Conference on Sketches & Applications: in conjunction with the 30th Annual Conference on Computer Graphics and Interactive Techniques*, San Diego, CA, 2003.

[6] A. Hilton, J. Starck and G. Collins, From 3d shape capture to animated models, *Proc. First International Symposium on 3D Data Processing Visualization and Transmission*, Padova, Italy, 2002, 246-257.

[7] W. Sun, A. Hilton, R. Smith and J. Illingworth, Layered animation of captured data, *Visual Computer*, *17*(8), 2001, 457-474.

[8] Y. Zhang, T. Sim and C.L. Tan, Adaptation-based individualized face modeling for animation using displacement map, *Proc. Computer Graphics International 2004*, Crete, Greece, 2004, 518-521.

[9] W.S. Lee and N. Magnenat-Thalmann, Fast head modeling for animation, *Journal Image and Vision Computing*, *18*(4), 2000, 355-364.

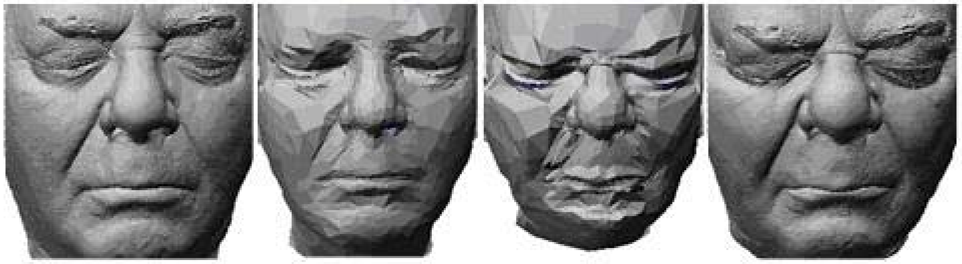
[10] T.D. Bui, M. Poel, D. Heylen and A. Nijholt, Automatic face morphing for transferring facial animation, *Proc. 6th IASTED International Conference on Computer Graphics and Imaging*, Honolulu, HI, 2003, 19-24.

[11] J.Y. Noh, D. Fidaeo and U. Neumann, Animated deformations with radial basis functions, *Proc. ACM Symposium on Virtual Reality Software and Technology*, Seoul, Korea, 2000, 166-174.

[12] J. Cohen, A. Varshney, D. Manocha, G. Turk, H. Weber, P. Agarwal, F. Brooks and W. Wright, Simplification envelopes, *Proc. 23rd Annual Conference on Computer Graphics and Interactive Techniques*, New Orleans, LA, 1996, 119-128.

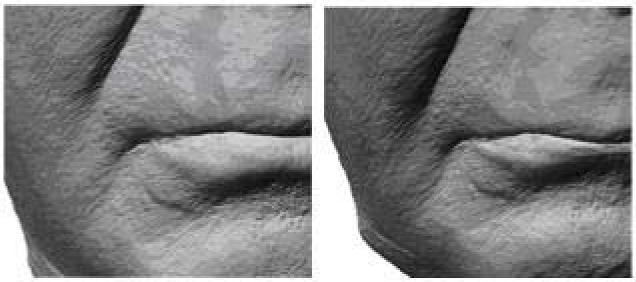
[13] A. Lee, H. Moreton and H. Hoppe, Displaced subdivision surfaces, *Proc. 27th Annual Conference on Computer Graphics and Interactive Techniques*, New Orleans, LA, 2000, 85-94.

[14] C. Loop, *Smooth subdivision surfaces based on triangles* (Master's thesis, University of Utah, Department of Mathematics, 1987).



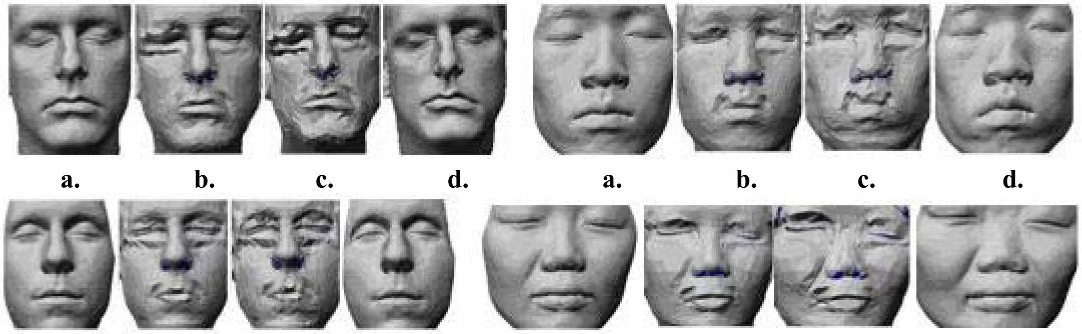
a. b. c. d.

Figure 11: a. Dense model (2,000,000 triangles); b. Working model (5,900 triangles); c. Caricature (75%) model (5,900 triangles); d. Target (detail recovered, 75% caricatured) model (1,990,616 triangles)



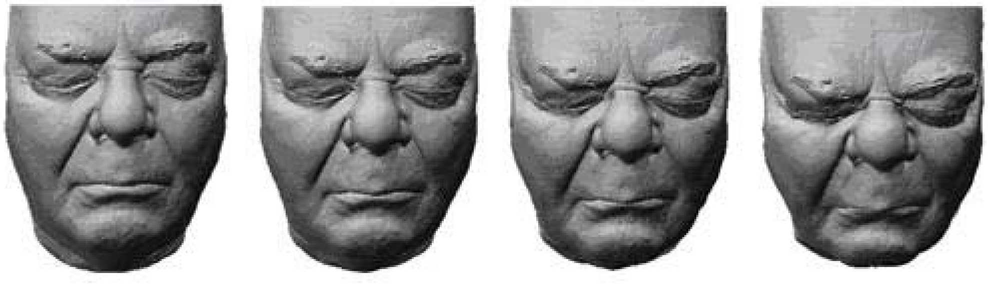
a. b.

Figure 12: Comparison of detail in a. dense model (2,000,000 triangles); b. target (detail recovered, 75% caricatured) model (1,990,616 triangles)



a. b. c. d. a. b. c. d.

Figure 13: a. Dense models (30,000 triangles); b. Working model; c. Caricature (75%) model; d. Target (detail-recovered, 75% caricatured) model



a. b. c. d.

Figure 14: Target (detail-recovered caricatured) models generated using exaggeration factors of. a. 30% ($c = 0.3$); b. 60% ($c = 0.6$); c. 90% ($c = 0.9$); d. 120% ($c = 1.2$)