Andrew Soon
Won-Sook Lee

Shape-based detail-preserving exaggeration of extremely accurate 3D faces*

A. Soon (✉)
The Department of Systems and Computer
Engineering, 1125 Colonel By Drive,
Carleton University, Ottawa, Ontario,
Canada K1S 5B6
asoon@sce.carleton.ca

A. Soon · W.-S. Lee
School of Information Technology and
Engineering (SITE), University of Ottawa,
800 King Edward Avenue, P.O. Box 450,
Stn A, Ottawa, Ontario, Canada K1N 6N5
wslee@uottawa.ca

Abstract We present an approach to automatically exaggerate the distinctive features of extremely detailed 3D faces. These representations comprise several million triangles and capture skin detail down to the pores. Despite their high level of realism, their size makes visualization difficult and real-time mesh manipulation infeasible. The premise of our methodology is to first remove the detail to obtain low resolution shape information, then perform shape-based exaggeration on a low resolution model and finally reapply the detail onto the exaggeration to recover the original resolution. We also present the results of applying this methodology to a small set of faces.

Keywords Exaggeration · High resolution · 3D faces · Morphable model · Detail reconstruction

1 Introduction

Exaggeration is commonly used in art and entertainment to varying degrees and for different reasons. For many people, exaggeration holds connotations of excessive usage for the purpose of humour because it is most commonly experienced in this manner. A caricaturist deliberately uses exaggeration copiously to elicit humour and an actor may deliver an overly animated performance for the same reason. Exaggeration also makes the subject more distinctive and recognizable. In the case of caricatures, the distinctive characteristics associated with an individual are made even more pronounced, strengthening the association. The mannerisms or the style of the overly animated actor distinguishes him from other actors. This increase in recognizability and individuality makes exaggeration a useful technique for other applications as well.

Two example applications are the production of small-scale replicas of real-world objects and the creation of computer-generated (CG) work for digital entertainment.

In manufacturing, digital 3D models could conceivably be used in the production of certain items (e.g., souvenirs resembling small objects such as seashells). It then becomes advantageous to exaggerate certain features of the models to increase their recognizability since they become less recognizable as small-scale replicas. In CG productions, exaggeration may be used in conjunction with other techniques to make characters distinct from each other.

In the two cases described above, one might see the benefit of using very highly detailed models. There is an inherent loss of detail associated with miniaturization, and this can in turn cause the product to lose some of its visual appeal. A high level of detail in the models can partially offset the detail loss and even enhance the miniature's appearance. Photorealistic CG sequences in motion pictures starring human actors are feasible with the aid of highly accurate models as demonstrated by movies such as "The Matrix Reloaded".

Many studies have been conducted into the use of exaggeration in digital works. However, many of these earlier studies like [3, 5] and [14] were conducted in the

context of 2D images while others like [1] and [9] which dealt with 3D faces did not cover highly detailed 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 Inc. (http://www.xyzrgb.com) utilize a combination of technologies developed by the National Research Council of Canada [18] capable of capturing surface detail in the order of about 100 microns. The resulting human face 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 [7] to achieve higher resolution renderings or models. Previous work such as [2, 11, 17] and [19] demonstrated that the low polygon models used in this practice can be constructed in many different ways. However, methodologies such as [2], 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 accurate 3D faces [18]. 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 nine sections. An overview of our proposed approach is presented in Sect. 2. Section 3 describes the derivation of the low resolution model from a generic head model using surface fitting. We summarize the results of our preliminary evaluation of detail capture techniques in Sect. 4. Section 5 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 Sect. 6 and the process of constructing the high resolution exaggeration is covered in Sect. 7. Sample results achieved with our methodology are presented in Sect. 8 and are followed by a conclusion in Sect. 9.

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 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. The simplified model is introduced to act as a bridge between the working model and the dense model. The simplified model is a linker to the working model where sparse point distribution in both models makes the correspondence easy to find (less than 10K). The simplified model is also a linker to the dense model in shape with different number of triangles (one with less than 10K and the other more than 1M). Only one simplified model is needed regardless of the dense model's resolution.
– *Detail Capture*: Captures the high resolution detail of the dense model using model parameterization. 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 that 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.
– *Detail Reconstruction*: A two-step procedure that 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.
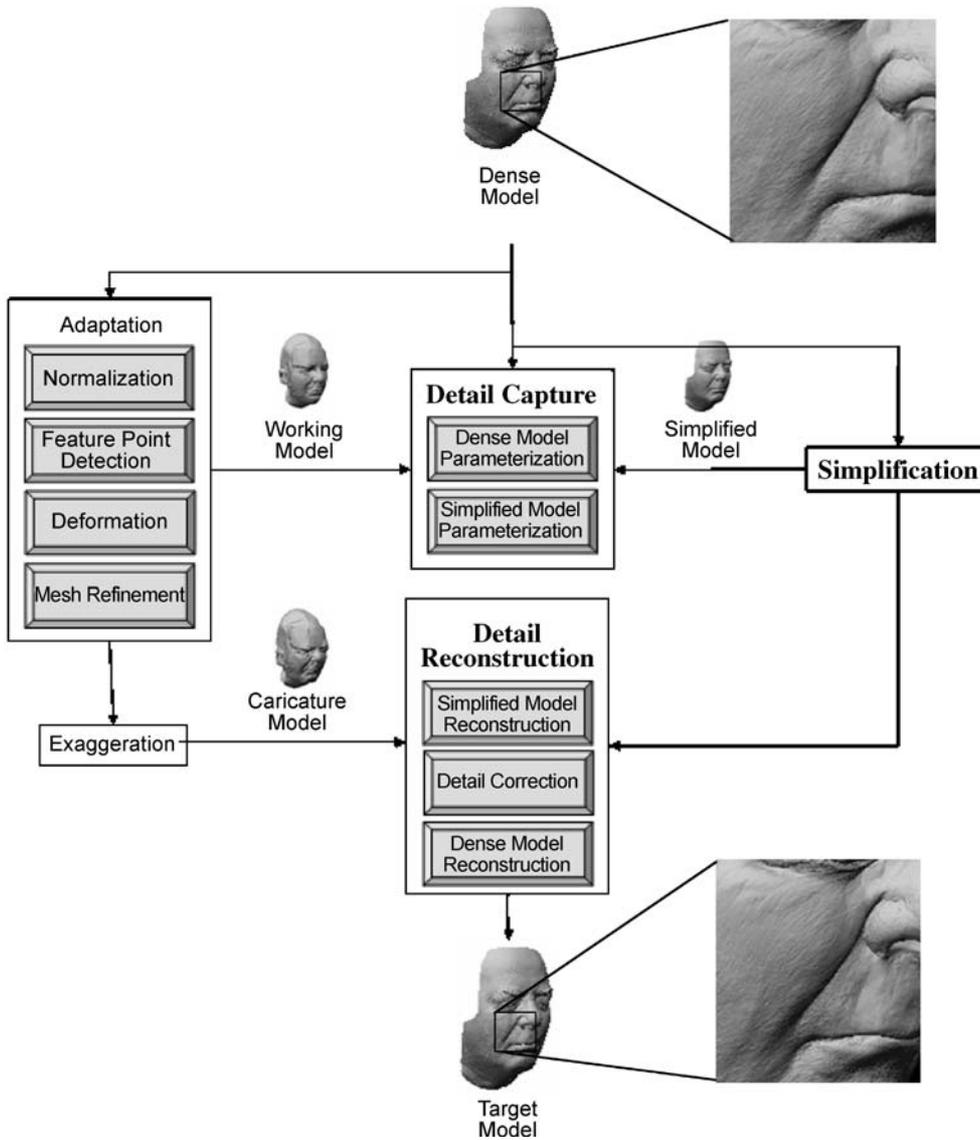
Fig. 1. Flow diagram for producing exaggerations of extremely accurate 3D faces

– *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 a modified version of the mesh adaptation procedure described in [13]. Instead of using two orthogonal (front and side view) photos of the subject, the high resolution dense model acts as the reference to which a generic head model is fitted. Radial basis function (RBF) networks [4, 16] are also employed instead of Dirichlet free form deformation

(DFFD) [13] to globally adapt the generic model to the high resolution surface.

3.1 Generic model

The generic model originally used in [13] (Fig. 2a) comprises 1,473 triangles and also contains structures for head animation. In an attempt to achieve more precise low resolution approximations, the polygon count was increased to 5,892 by applying Loop subdivision [15] to the generic model (Fig. 2b).

3.2 Dense model normalization

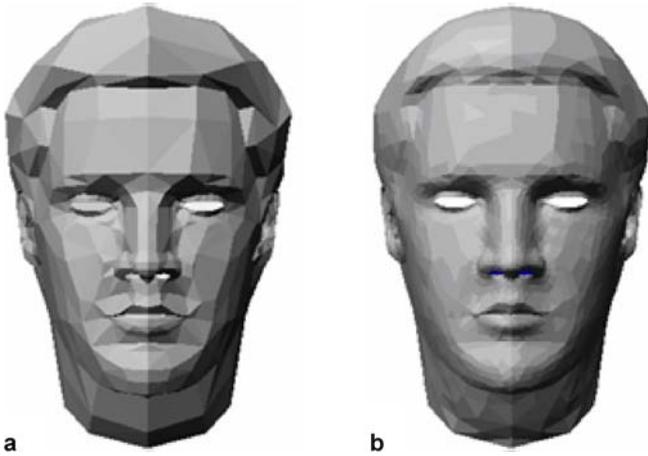Normalization is performed as a preprocessing step to transfer the dense model from object space to predefined

Fig. 2a,b. Generic models before and after subdivision. **a** Old (1473 triangles); **b** New (5892 triangles)
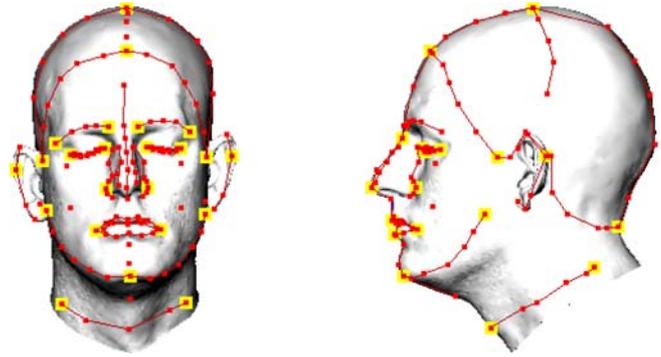

Fig. 3. Feature points marked on the front and side views of a head. Major feature points are marked in yellow-red while minor ones are marked in only red

feature point (characteristic point or landmark) space. This is achieved by applying the procedure reported in [10] to scale and translate the dense model. The normalized result is then used for all subsequent work. For this reason, we can refer to the normalized dense model as simply the dense model to facilitate discussion.

3.3 Feature point detection

The feature point detection step is essentially identical to the image-based technique originally presented in [13]. The generic model has a set of 163 vertices on the 3D head (front, sides and back) classified as feature points which represent the most characteristic points used for human recognition. Of these, 41 are considered major while the remaining points are minor. In order to produce the working model, these feature points must first be defined on the dense model. They are semi-automatically marked on 2D front and side view images of the dense model (Fig. 3) so that 3D feature points can be calculated. The markers for the minor feature points are automatically set based on the placement of major feature point markers.

3.4 Generic model deformation

The generic model is adapted to the dense surface by performing deformation with Radial Basis Functions as described in [4] and [16]. Generic model deformation is performed at a global level by treating the feature points as centers of the RBF. Three RBF networks – one for each coordinate of 3D space – are established to represent the generic model. The Hardy multiquadric basis function is used for each of the networks.

The networks are trained using the vector of points $\overline{X} = \begin{bmatrix} \bar{x}_1 & \cdots & \bar{x}_{163} \end{bmatrix}^{\mathrm{T}}$ containing the initial positions of the feature points in the generic model. The stiffness constants are calculated using the minimum separation approach suggested in [8]. The weight vector for each network is computed by ensuring that the networks evaluate point i in $\overline{X}$ to point i in $\overline{X}' = \begin{bmatrix} \bar{x}_1' & \cdots & \bar{x}_{163}' \end{bmatrix}^{\mathrm{T}}$, where $\overline{X}'$ contains the new positions of the feature points as determined during feature point detection.

After training and weight vector computation have been completed, the new positions of the non-feature points are calculated by evaluating the RBF networks with the initial position of each of these points. We term the resulting deformed generic model the working model. It should be obvious that the working model contains the same point and polygon structures as those of the generic model.

3.5 Working model refinement

In order to obtain a more accurate representation of the dense geometry, the approximated positions of the working model's non-feature points need to be refined so that they lie on the dense surface. Before doing so, the working model must first be scaled and translated to bring it from generic model space into feature space. This step aligns the dense model with the working model (Fig. 4).

After alignment is achieved, ray casting is applied on the working model to project its non-feature points onto the dense model's surface. The direction of the normal at each vertex is used to determine the direction of the ray. The ray casting concept is illustrated in Fig. 5. A user-definable threshold is imposed to ensure that the distance by which each vertex is displaced is not overly large. If the displacement is beyond the threshold or if an intersection point cannot be found, the working model vertex remains in its original location. If multiple intersections are found, the closest intersection point is chosen. Figure 6 demonstrates the benefit of performing this step by contrasting examples of working models present before and after mesh refinement.
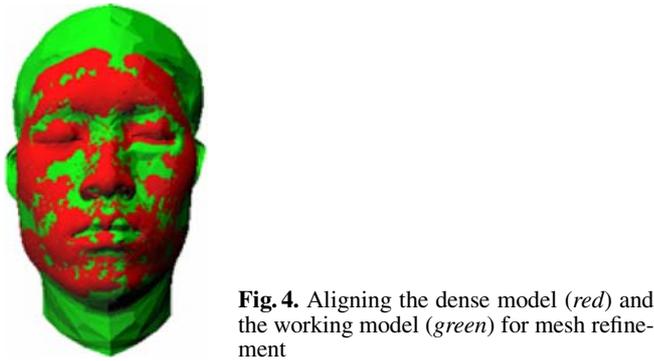
Fig. 4. Aligning the dense model (*red*) and the working model (*green*) for mesh refinement
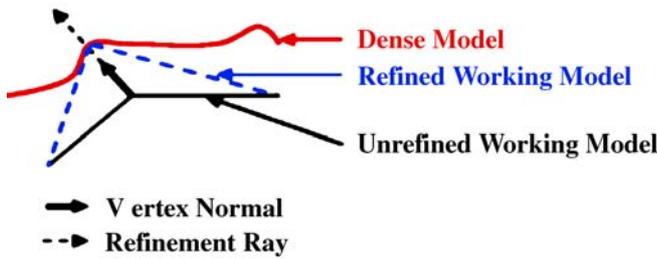

Fig. 5. 2D interpretation of mesh refinement using ray casting

Mesh refinement is not performed in the nose and mouth regions because the triangles in those areas tend to be small and ray casting typically leads to a poorer representation.

Figure 7 gives an example of the refined working model produced from a dense model. Note that since the dense model only captures the face of the subject, rough estimations had to be made for the remainder of the head during the construction of the working model.

Even though we use two steps to construct the working model (feature-based deformation and model refinement), its approximation of the dense model's shape can still be insufficient to acceptably capture high resolution detail (discussed in the proceeding section). Large differences between the dense and working models exist in areas such as the eyes and mouth. For this reason, the simplified model is introduced in our methodology to bridge the gap between the working and dense models.

4 Evaluating detail capture techniques

Many different detail capture techniques were tested in order to find one that could faithfully capture the fine detail of the dense geometry. We sought a method which would allow us to reconstruct the original model (i.e., with the original polygon structure), but methods for which this is not possible were also considered. Regardless of the actual geometrical structure, extremely high visual quality

Fig. 6a–c. Comparison of results before and after mesh refinement. **a** Dense model; **b** Working model before refinement; **c** Working model after refinement
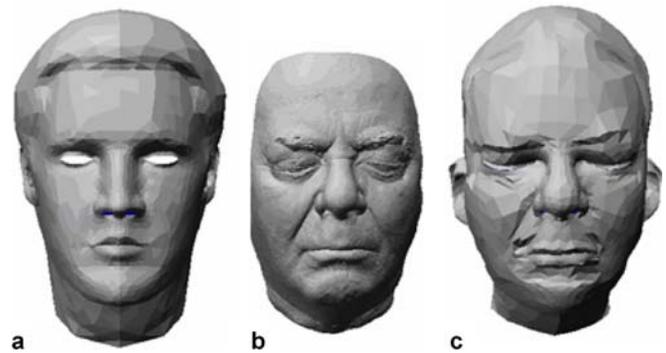

Fig. 7. a Generic model; **b** Dense model; **c** Working model

in the reconstructed models was our primary concern. The proceeding subsections summarize our experiences with traditional displacement mapping, which did not satisfy our needs, and model parameterization, which we adopted for our process.

4.1 Displacement mapping

Displacement mapping [7] is a technique commonly used to capture the detail of high resolution geometry relative to a low resolution surface. Unlike other schemes like bump mapping and normal mapping, the low resolution geometry is modified to increase the level of detail.

In our preliminary experiments, we attempted to use the technique described in [19] which combines hierarchal mesh refinement and displacement map sampling. We stored the detail of a 2 000 000 triangle face model relative to a cropped working model in a high resolution displacement map as illustrated in Fig. 8a-c. The working model was cropped so that it could be trivially UV mapped with as few overlapping textures as possible. To produce the high polygon model, the cropped model was iteratively subdivided five times using the quaternary scheme (Fig. 9) and every vertex in the final model was displaced according to the values in the map. The refined subdivided mesh is shown in Fig. 8d.

Disregarding the glaring aberrations in the reconstructed model, a close-up comparison with the original dense model (Fig. 10) also showed imperfections at a finer level. The triangles of the reconstructed model can be seen much more clearly than those of the dense model, which made the surface look unnatural. Furthermore, this deficiency was evidenced in many regions of the face in which
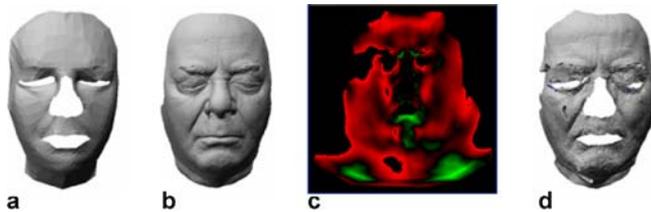

Fig. 10a,b. Close-up comparison of the right cheek of the original and reconstructed dense models shown in Fig. 8. **a** Original model; **b** Reconstructed model

point density was high. Overall, the reconstructed model's appearance was not comparable to the original dense model. This result led us to believe that displacement mapping was incapable of producing satisfactory geometry for our needs.

The chief reason for displacement mapping's failure to satisfactorily reconstruct the dense model geometry is its inability to properly recreate the vertices of the model, which ultimately determines the appearance of the reconstructed surface. This inability can be attributed to two factors. The first is the way in which a displacement map is generated. A displacement value is calculated for each pixel of the (cropped) working model's UV map. Each pixel corresponds to a point on the surface of the working model and the displacement is measured between the working model and the dense model in the direction of the surface normal at said point. Since it is impossible to record a displacement for every point on the dense model surface, some dense model vertices are not accounted for in the displacement map (Fig. 11a). As a result, reconstructing the model using the displacement map can produce a substandard result (Fig. 11b). The second factor attributed to the downfall is the way in which the reconstructed model is achieved. As mentioned previously, quaternary subdivision is used to increase the resolution of the working model whose surface is then perturbed by sampling the displacement map at each vertex. Even if the location of every vertex in the dense model was successfully encoded in the displacement map, there is no guarantee that the dense model vertices were captured using the working model vertices (Fig. 11c). As a result, the reconstructed model will differ from the dense model.

Essentially, a different technique that can capture (and then consequently reproduce) the original vertices of the dense model is needed. One might argue that the limitation shown in Fig. 11c may be overcome by recording the locations of the points used to capture the dense model vertices. However, one must keep in mind that in such a case, there is a great deal of computation time needed to produce a displacement map and much of this work is essentially wasted since only a relatively small set of dis-

Fig. 8a–d. Detail reconstruction result using displacement map sampling. **a** Cropped working model (3780 triangles); **b** Dense model (2 000 000 triangles); **c** High resolution RGB displacement map (2048 × 2048 pixel resolution). Positive displacements are encoded in the red channel while negative displacements are stored in the green channel; **d** High polygon reconstruction (2 002 944 triangles)
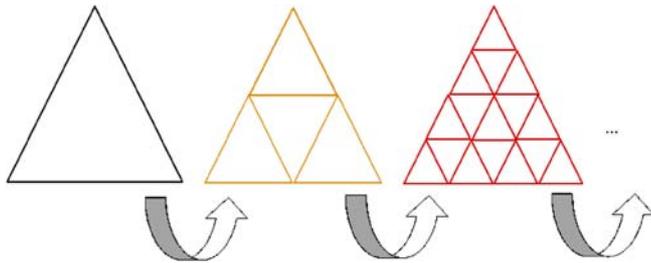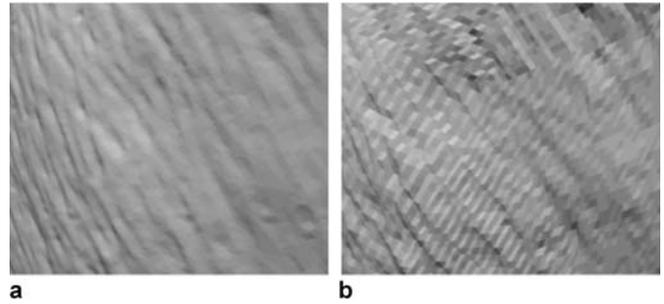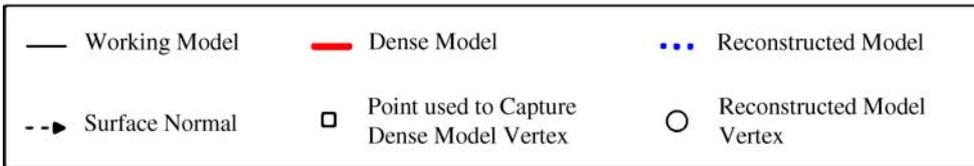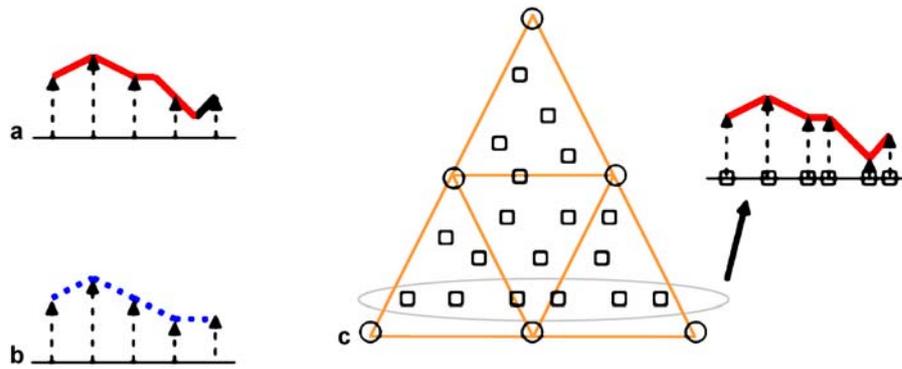

Fig. 9. Iterative quaternary subdivision. Each existing edge is split at its midpoint

Fig. 11a–c. Generation and usage of displacement map. **a** Displacement map generation (sampling at regular intervals); **b** Dense model reconstruction using displacement map from **a**; **c** Illustration of inability to reconstruct original dense model vertices despite having the values in the displacement map

placement values would actually be used during model reconstruction.

4.2 Model parameterization using point-to-surface mapping

Point-to-surface mapping, as reported in [17] and [19], enables a point in 3D space to be mapped to a surface (triangle) using an interpolated surface normal. It is possible to capture the positions of a model's vertices by applying this mapping to every point. The global application of this scheme to the vertices of a high resolution model with respect to a low resolution model was discussed in [11] where the process was referred to as model parameterization. Point-to-surface mapping and model parameterization are explained to a greater extent in the next section.

The suitability of this scheme was evaluated once again using the 2 000 000 triangle mesh. However, since we were not concerned with UV mapping, the full working model served as the base surface for detail capture. The dense model was parameterized with respect to the working model and this parameterization was applied onto the base geometry to produce the result shown in Fig. 12c.

The obvious errors in the reconstruction were ignored once again and only well reconstructed areas were judged. The side-by-side comparison (Fig. 13) of the same cheek area originally shown in Fig. 10 proved that this capture technique was able to yield a smooth and faithful reconstruction of the original surface with only a relatively small number of localized errors. Unlike displacement mapping, model parameterization is capable of reproducing the original vertices of the dense model and this is the primary reason for its success in surface reconstruction.

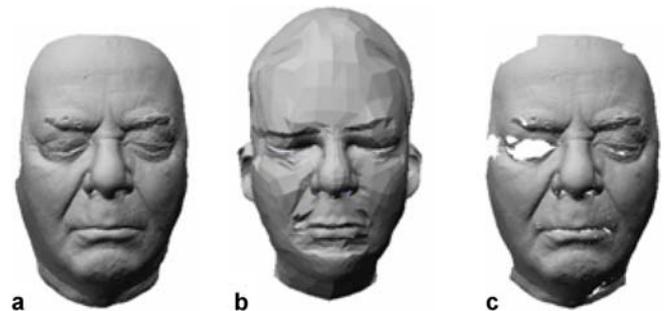To ensure that the captured detail would fare equally well when applied to altered geometry, we performed lo-

Fig. 12a–c. Detail reconstruction result using model parameterization. **a** Dense model (2 000 000 triangles); **b** Full working model (5892 triangles); **c** Reconstructed dense model (1 758 403 triangles)
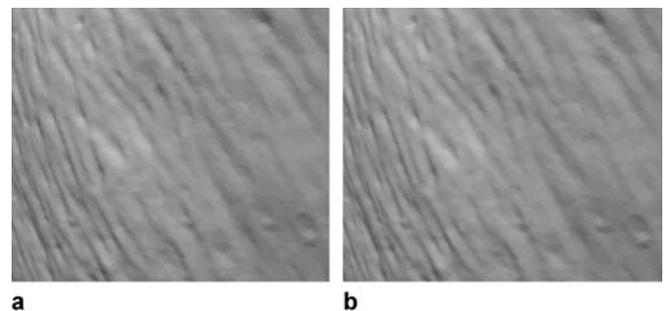

Fig. 13a,b. Close-up comparison of the right cheek of the original and reconstructed dense models shown in Fig. 12. **a** Original model; **b** Reconstructed model

calized deformations as described in [16] on three areas of the working model: the cheeks and the chin (Fig. 14a). The parameterization previously achieved was then ap-
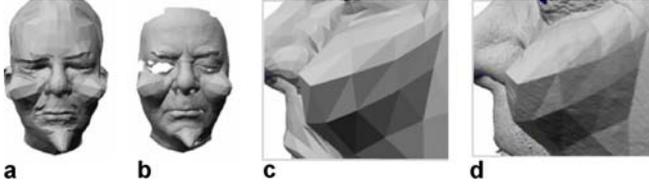
Fig. 14. a, **c** Full working model with localized deformations; **b**, **d** Reconstructed dense model

plied to the deformed model to obtain the new surface (Fig. 14b). It was immediately evident that the edges of the low polygons manifested themselves in the high resolution reconstruction (Fig. 14c–d). Despite the presence of these artifacts, the skin detail in the regions bordered by these edges was reproduced correctly. These results testified that model parameterization using point-to-surface mapping had the potential to give us visually impressive results and was at the very least worth some more in-depth investigation. In the work that followed this initial evaluation, we learned and developed techniques that dramatically reduced the presence of holes and edges in the dense model reconstruction. The discussion of these techniques is deferred to the proceeding section.

5 Capturing high resolution detail

5.1 Point-to-surface mapping

The point-to-surface mapping scheme outlined in [17] and [19] is based upon simplification envelopes [6]. This scheme maps a point V in 3D space to a triangle ABC using an interpolated vertex normal (Fig. 15). The ray origin P and the interpolated vertex normal $\overline{N}_P$ are given by

$$P = (1 - u - v)A + uB + vC$$
$$\overline{N}_P = (1 - u - v)\overline{N}_A + u\overline{N}_B + v\overline{N}_C, \qquad (1)$$

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

$$V = P + d\frac{\overline{N}_P}{|\overline{N}_P|}, \qquad (2)$$

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

To calculate the values of u and v in Equation (1), a new triangle $A^{par}B^{par}C^{par}$ that is parallel to $\triangle$ABC and whose vertices are coplanar with V is defined (Fig. 15). The vertices of this new triangle are obtained by finding
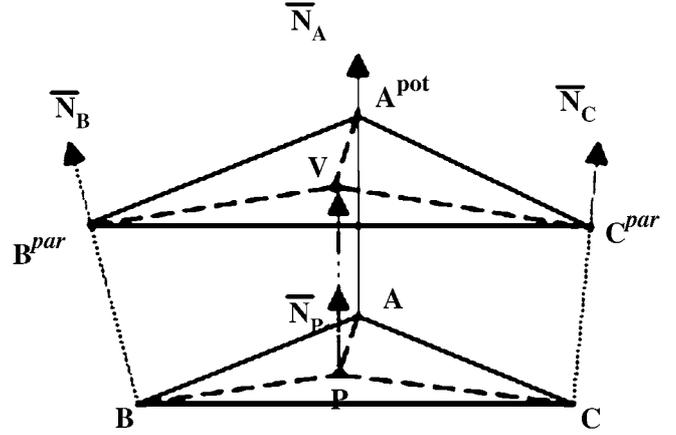

Fig. 15. Example of point-to-surface mapping where point V is mapped to $\triangle$ABC. $\triangle A^{par}B^{par}C^{par}$ and $\triangle$ABC lie on parallel planes and A^{par}, B^{par}, C^{par} and V are all coplanar

the intersection of $\overline{N}_A$, $\overline{N}_B$ and $\overline{N}_C$ with the parallel plane. Computing the barycentric coordinates of V with respect to $\triangle A^{par}B^{par}C^{par}$ yields the values of u and v. In other words

$$V = (1 - u - v)A^{par} + uB^{par} + vC^{par}. \qquad (3)$$

5.2 Model parameterization

The parameterization of a high resolution model with respect to a low resolution model using point-to-surface mapping is discussed in [11]. 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 [17] and [19] 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 $u + v \leq 1$ were enforced, smaller values of d were favored and a user-defined threshold on the value of d was used.

5.3 Address of model parameterization deficiencies

As originally stated in Sect. 4.2, two primary issues were encountered during preliminary experiments with model

parameterization: the appearance of low polygon edges in the reconstructed geometry and an inability to capture some areas of the dense model. The techniques used in our approach to address these problems are discussed here.

The edge anomaly was remedied by heeding the suggestion made in [12] to use a smooth domain surface when deforming a control mesh. Prior to performing model parameterization, we iteratively subdivided the control model twice using Loop's algorithm [15] in order to achieve the desired smoothness. We applied this new strategy to the deformed geometry test described in Sect. 4.2 to verify its usefulness. The detail of the 2 000 000 triangle mesh was recaptured using a smooth working model, and we also subdivided the deformed model shown in Fig. 14 before applying the recaptured detail. The resulting reconstructed model (Fig. 16) has a much smoother appearance although the edges are still visible because of the high degree of deformation. In our exaggeration work, the amount of perturbation to the working model never reached the level used in the preliminary test. Our exaggeration results showed a much greater benefit from adopting the subdivision strategy as low polygon edges are virtually non-existent (Fig. 17).

The other deficiency uncovered during evaluation was that full model parameterization was difficult to achieve (i.e., it was not possible to obtain a set of mapping parameters for every vertex in the dense model). This was primarily a result of imposing the displacement threshold on detail capture in the presence of offsets which exceeded this threshold. As mentioned earlier, mesh refinement is also governed by a displacement threshold and is not performed in the eyes and mouth regions, which result in larger discrepancies between the models.

Our early work revealed that the direct parameterization approach (dense model with respect to the working model) led to high resolution exaggerations that contained many errors (Fig. 18b). We relaxed the constraints on the barycentric coordinates and used a liberal value for the
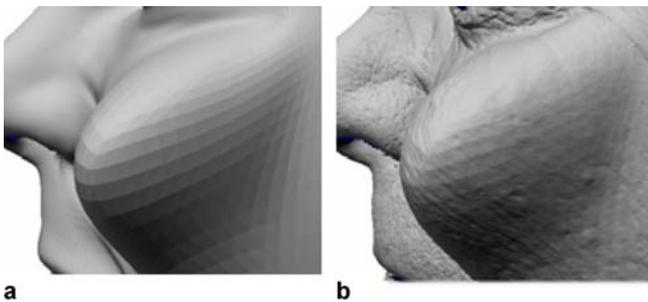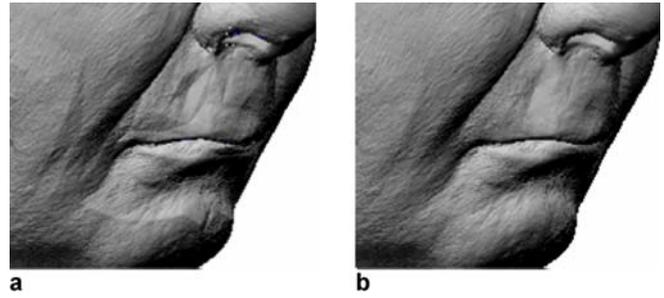

Fig. 17a,b. High resolution exaggerated models achieved using different model parameterization strategies. **a** Control models are not subdivided; **b** Control models are subdivided
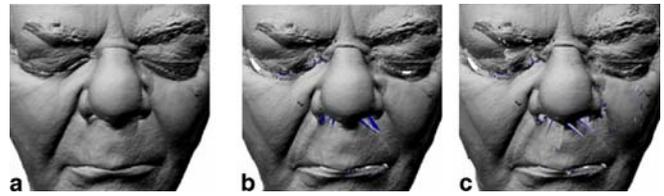

Fig. 18a–c. High resolution exaggerated models achieved using different parameterization techniques. **a** Two-stage parameterization using both the working model and a simplified model; **b** Direct parameterization on the working model; **c** Direct parameterization with relaxed constraints and a liberal threshold value for displacements on the working model

displacement threshold, but this only led to more errors with no significant improvements to the level of detail (Fig. 18c). The surface error issues were resolved by introducing a simplified model into the detail capture process. The simplified model may be constructed using any model simplification algorithm which produces a good approximation that preserves the general shape of the original (Fig. 19). The high resolution detail of the dense model is then 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 configuration, 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.

5.4 Dense model parameterization

The dense model is parameterized with respect to the simplified model to achieve the dense model parameterization. The most significant benefit of using the simplified model is that since it is typically a better approximation than the working model, the notoriously problematic areas (eyes, nose and mouth) are reconstructed much better (Fig. 18a) because the detail is captured more precisely. In general, it should be possible to achieve an almost full dense model parameterization.
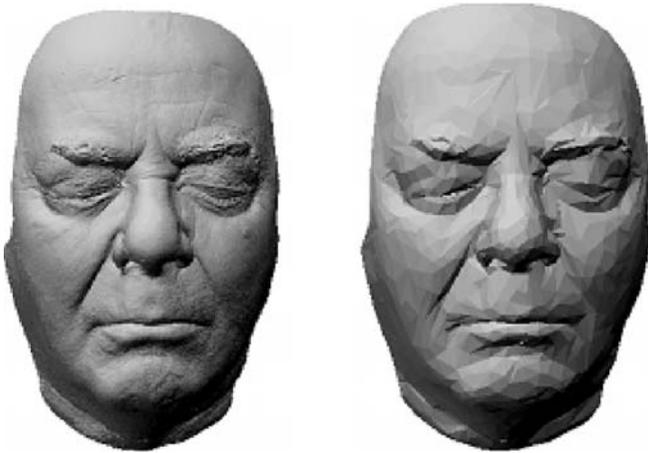

Fig. 16a,b. Revisiting the deformed geometry test. The base geometry is subdivided before detail capture and detail reconstruction. **a** Full working model with localized deformations (94 272 triangles); **b** Reconstructed dense model (1 767 986 triangles)

Fig. 19. Example of a dense model (*left*) and its shape-preserving simplified model (*right*, 2000 triangles)

5.5 Simplified model parameterization

The other parameterization needed for detail capture is the simplified model parameterization, which is obtained by parameterizing the simplified model with respect to the working model. This parameterization suffers from the same limitation as the original direct parameterization; that is, it is not always possible to achieve full model parameterization. This problem is much more significant in the two-stage scheme than the direct parameterization scheme because the dense model parameterization is based on a full simplified model. An incomplete parame-

terization means that the full simplified model cannot be reconstructed later. We resolve this problem by statically recording the 3D coordinates of points which could not be parameterized and then using a detail correction scheme to make any necessary adjustments to the positions of these points. The correction algorithm is discussed later in Sect. 7.

6 Constructing the caricature model

6.1 Average face model

The purpose of calculating an average face model is to obtain some measure of an average human face which is also void of individual characteristics. Individual faces are exaggerated by making comparisons with the average face to determine which features are distinctive (i.e., those that deviate from the average). The true average human face can only be calculated by considering the world's population in its entirety. This is clearly not feasible and so we use a limited number of faces to obtain the average.

An average face model is constructed by taking the average of a set of working models (Fig. 20). 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.
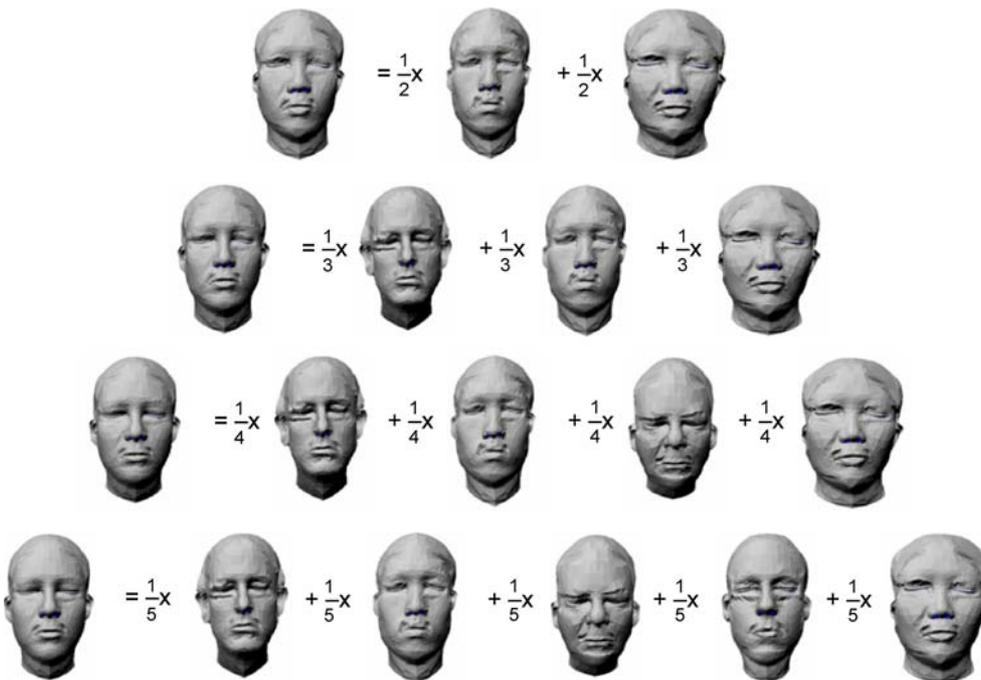

Fig. 20. Examples of average face models constructed from working models

The average face also inherits the same point and polygon structure as the generic model.

Figure 20 shows examples of average faces constructed using two faces, three faces, four faces and five faces. The two-face and three-face averages contain some features inherited from their respective sets. The four-face and five-face averages are quite similar and do not bear significant resemblance to any of their respective constituents. We concede that an average derived from five faces is not an accurate representation, but it is still acceptable for emphasizing the characteristics of individual faces as (i) it is very near to the true average face and (ii) it does not contain individual characteristics.

6.2 Vector-based exaggeration

The caricature procedure originally proposed by Brennan [3] is used to exaggerate the working model to produce the *caricature model*. Each feature vector $\bar{v}_i$ defined as

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

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

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

as shown in Fig. 21. 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).

It is worthwhile to point out that this exaggeration algorithm operates on a point-by-point basis and, as is, does not take into consideration the changes applied to the other vertices. The model being exaggerated can experience self-penetration problems when c becomes very large. In such cases, collision detection algorithms can be employed to detect their occurrence and
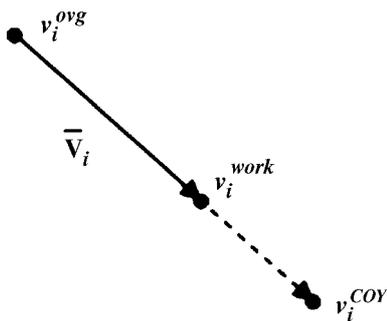
context-dependent solutions can be used to resolve the issues.

7 Constructing the target model

7.1 Model reconstruction

The premise of the model reconstruction algorithm is to use the model parameterization (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 (Fig. 22). Equations (1) and (2) can then be restated using the new labels in Fig. 22 as

$$P' = (1 - u - v)A' + uB' + vC'$$
$$\overline{N}'_{P'} = (1 - u - v)\overline{N}'_A + u\overline{N}'_B + v\overline{N}'_C \qquad (6)$$

and

$$V' = P' + d\frac{\overline{N}'_{P'}}{|\overline{N}'_{P'}|}, \qquad (7)$$

7.2 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
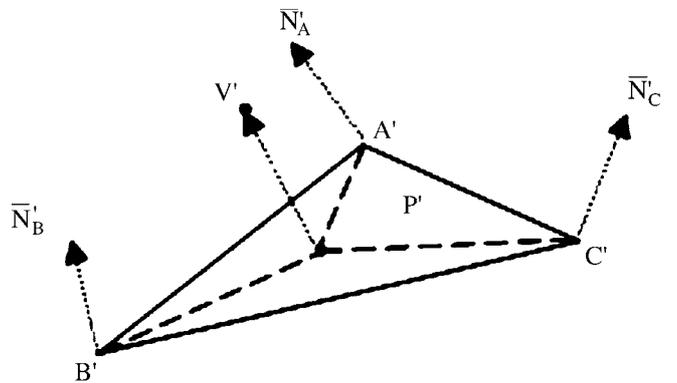

Fig. 21. Vector-based exaggeration

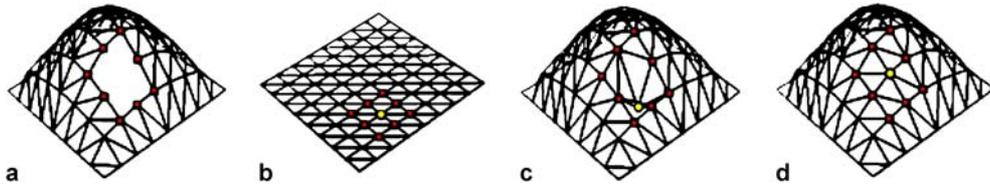**Fig. 22.** Reconstruction of a point V. Its new position is denoted by V'

Fig. 23. a Curved surface with hole; **b** Flat surface from which the curved surface in a is derived; **c** Curved surface with statically captured point added; **d** Curved surface after detail correction

model. As mentioned previously, however, this model cannot be used directly to reconstruct the dense model if the simplified model parameterization is incomplete. In such an instance, the holes in the model's surface must be repaired appropriately. The statically captured points (described in Sect. 5.5) must be used to reconstruct the missing areas. This is the responsibility of the detail correction algorithm, which is described in the proceeding subsection. After undergoing correction, the reconstructed simplified model and the dense model parameterization are used by the reconstruction algorithm to produce the target model.

7.3 Detail correction

Detail correction displaces the statically captured points so that they conform to the rest of the simplified model's surface. RBF networks are deployed in the same manner as in generic model deformation (Sect. 3.4) to calculate new positions for these points. The premise of this simple correction technique is illustrated in Fig. 23. Suppose the curved surface in Fig. 23a was derived from the flat surface in Fig. 23b. The point marked with the yellow circle on the flat surface needs to be recreated in the curved surface to patch the hole. The missing point is introduced into the curved surface in Fig. 23c and is marked with the same yellow circle. However, the position of this point was obtained from the flat surface and hence the point does not lie on the curved surface. To remedy this, the correction algorithm first seeks out a set of points on the original flat surface that are in the vicinity of the point marked with the yellow circle. These neighbours (marked with red squares) serve as centers for the RBF networks

to represent the desired region in the flat surface. The new positions of these centers, taken from the curved surface, are needed during network training to calculate the weight vectors. Once the networks have been trained to represent the region surrounding the missing point on the curved surface, they are able to calculate the point's approximate position, which results in the surface shown in Fig. 23d.

An actual example of this algorithm in practice is given in Fig. 24. It is quite clear that the presented correction algorithm is capable of fixing the incomplete simplified model.

8 Results

The results presented in this section were produced using an implementation of the proposed methodology running on a desktop PC. Our testing focused on producing 3D geometry with extremely high visual fidelity, so no head animations were performed. 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 Inc. All simplified models produced for our experiments contained roughly 2000 triangles and all exaggerations were performed with respect to the five-face average shown in Fig. 20.

We used five faces scanned by XYZ RGB for our testing. The highest resolution model available contained $2\,000\,000$ triangles (Fig. 25a and Fig. 26a). The corresponding working model (Fig. 25b) was exaggerated by a factor of 75% ($c = 0.75$) to produce the caricature model
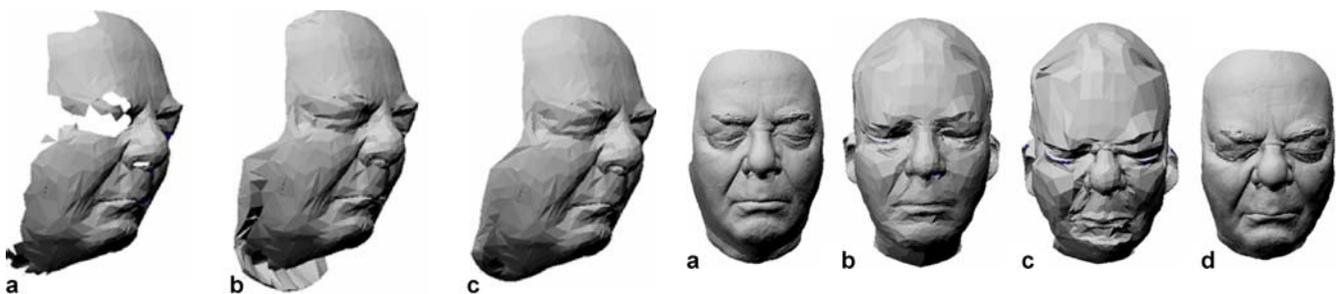

Fig. 24. a Simplified model constructed using an incomplete parameterization; **b** Simplified model with statically captured points; **c** Simplified model after detail correction

Fig. 25. a Dense model ($2\,000\,000$ triangles); **b** Working model (5892 triangles); **c** Caricature (75%) model (5892 triangles); **d** Target (detail recovered, 75% caricatured) model ($1\,990\,616$ triangles)
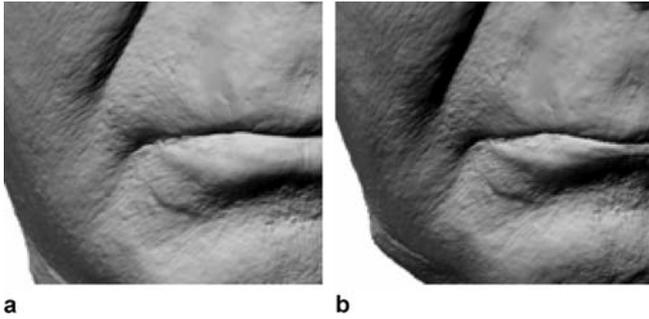
Fig. 26a,b. Comparison of detail in **a** dense model (2 000 000 triangles); **b** target (detail recovered, 75% caricatured) model (1 990 616 triangles)

(Fig. 25c). The subsequent target model (Fig. 25d and Fig. 26b) 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 Fig. 27.

The suitability of the proposed methodology was also tested by using several values of the exaggeration factor c.

The target models shown in Fig. 28 were produced using $c = 0.3$, $c = 0.6$, $c = 0.9$ and $c = 1.2$.

9 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's characteristics are exaggerated using a vector-based caricature algorithm that 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.
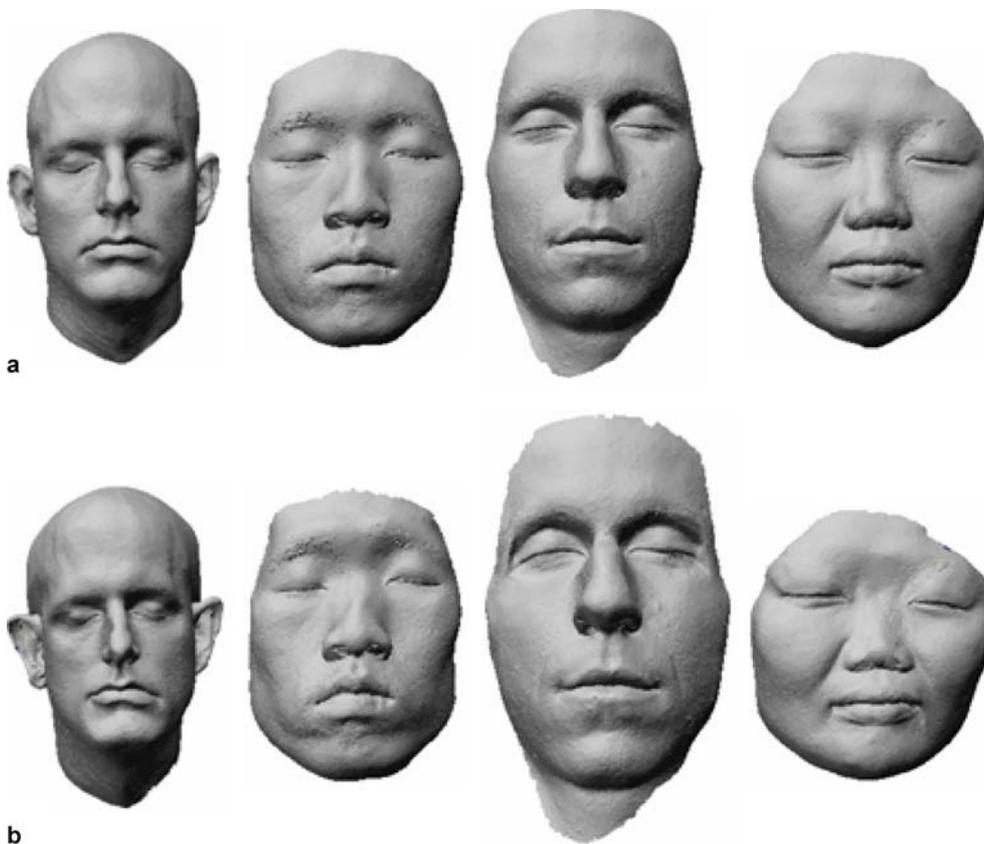

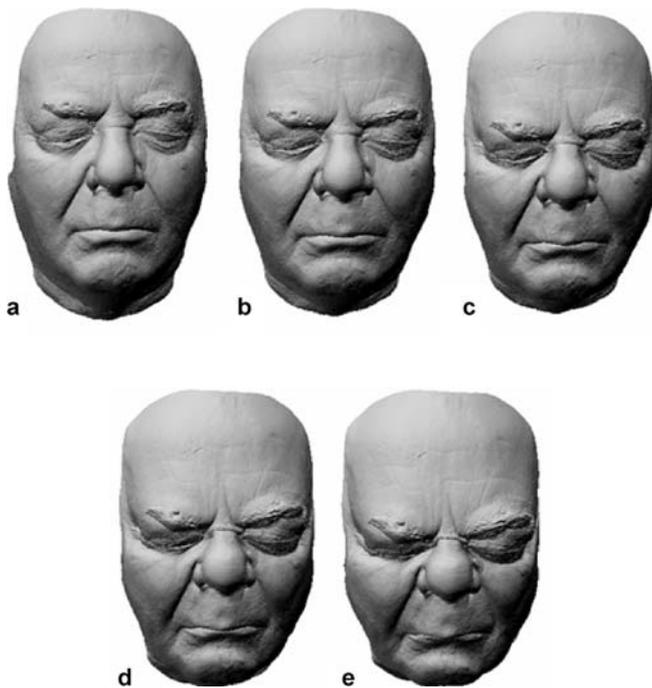**Fig. 27. a** Dense models (30 000 triangles); **b** Target (detail-recovered, 75% caricatured) model

Fig. 28. a Original dense model; Target (detail-recovered caricatured) models generated using exaggeration factors of. **b** 30% ($c = 0.3$); **c** 60% ($c = 0.6$); **d** 90% ($c = 0.9$); **e** 120% ($c = 1.2$)

The proposed approach possesses several benefits. It is a suitable solution in applications that 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. The rough approximation also makes it necessary to use a simplified model in order to obtain acceptable parameterizations of dense models. 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.

As a side note, the laser-scanned models always have closed eyes and mouths. However, it is still possible to open the models' eyes and mouths by using user-friendly modeling tools or programs dedicated to perform such functionality. This work only needs to be performed locally (eyes and mouth) while the remaining areas remain unchanged.

The results presented in this paper demonstrate that the methodology is sufficiently robust and flexible to handle high resolution models containing 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.

References

1. Blanz, V., Vetter, T.: A morphable model for the synthesis of 3D faces. In: Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques, pp. 187–194 (1999)
2. Borshukov, G., Lewis, J.P.: Realistic Human Face Rendering for "The Matrix Reloaded". In: Proceedings of the SIGGRAPH 2003 Conference on Sketches & Applications: in Conjunction with the 30th Annual Conference on Computer Graphics and Interactive Techniques, pp. 1–1 (2003)
3. Brennan, S.E.: The caricature generator. Leonardo **18**, 170–178 (1985)
4. Bui, T.D., Poel, M., Heylen, D., Nijholt, A.: Automatic face morphing for transferring facial animation. In:
   Proceedings of the 6th IASTED International Conference on Computers, Graphics and Imaging, pp. 19–23 (2003)
5. Chiang, P.-Y., Liao, W.-H., Li, T.-Y.: Automatic caricature generation by analyzing facial features. In: Proceedings of Asian Conference on Computer Vision (ACCV2005) (2005)
6. Cohen, J., Varshney, A., Manocha, D., Turk, G., Weber, H., Agarwal, P., Brooks, F., Wright, W.: Simplification envelopes. In: Proceedings of the 23rd Annual Conference on Computer Graphics and interactive Techniques, pp. 119–128 (1996)
7. Cook, R.L.: Shade trees. In: Proceedings of the 11th Annual Conference on Computer Graphics and Interactive Techniques, pp. 223–231 (1984)
8. Eck, M.: Interpolation methods for reconstruction of 3d surfaces from sequences of planar slices. CAD und Computergraphik **13**(5), 109–120 (1991)
9. Fujiwara, T., Koshimizu, H., Fujimura, K., Fujita, G., Noguchi, Y., Ishikawa, N.: 3D modeling system of human face and full 3d facial caricaturing. In: Proceedings of the Seventh International Conference on Virtual Systems and Multimedia, pp. 625–633 (2001)
10. Goto, T., Lee W.-S., Magnenat-Thalmann, N.: Facial feature extraction for quick 3d face modeling. Signal Process. Image Commun. **17**(3), 243–259 (2002)
11. Hilton, A., Starck, J., Collins G.: From 3d shape capture to animated models. In: IEEE Conference on 3D Data Processing,

Visualization Transmission, pp. 246–255 (2002)

12. Lee, A., Moreton, H., Hoppe, H.: Displaced subdivision surfaces. In: Proceedings of the 27th Annual Conference on Computer Graphics and interactive Techniques, pp. 85–94 (2000)

13. Lee, W.-S., Magnenat-Thalmann, N.: Fast head modeling for animation. Image Vision Comput. **18**(4), 355–364 (2000)

14. Liang, L., Chen, H., Xu, Y.-Q., Shum ,H.-Y.: Example-based caricature generation with exaggeration. In: Proceedings of the 10th Pacific Conference on Computer Graphics and Applications, pp. 386–393 (2002)

15. Loop, C.: Smooth subdivision surfaces based on triangles. Master's thesis, University of Utah, Department of Mathematics (1987)

16. Noh, J.-Y., Fidaeo, D., Neumann, U.: Animated deformations with radial basis functions. In: Proceedings of the ACM Symposium on Virtual Reality Software and Technology, pp. 166–174 (2000)

17. Sun, W., Hilton, A., Smith, R., Illingworth, J.: Layered animation of captured data. Visual Comput. **17**(8), 457–474 (2001)

18. Taylor, J., Beraldin, J.-A., Godin, G., Cournoyer, L., Rioux, M., Domey, J.: NRC 3D imaging technology for museums & heritage. In: Proceedings of The First International Workshop on 3D Virtual Heritage, pp. 70–75 (2002)

19. Zhang, Y., Sim, T., Tan, C.L.: Adaptation-based individualized face modeling for animation using displacement map. In: Proceedings of Computer Graphics International 2004, pp. 518–521 (2004)
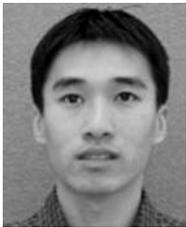
ANDREW SOON holds a B. Eng. in Computer Systems Engineering from Carleton University (Ottawa, Canada) and is currently an M.A.Sc. Electrical Engineering student in the Department of Systems and Computer Engineering at Carleton University. He is also a member of the DISCOVER Lab in the School of Information Technology and Engineering at the University of Ottawa. His research is in the area of 3D face modeling, but is also interested in other topics closely related to video games and computer-generated work for films. These include computer graphics, computer animation and 3D rendering.

WON-SOOK LEE is an Assistant Professor at SITE, Faculty of Engineering, University of Ottawa, Canada. She has studied and worked in companies and universities in South Korea, U.K., Singapore, Switzerland and U.S.A. She holds a Ph.D. in Computer Science from University of Geneva, Master's degrees from University of Geneva, National University of Singapore and Pohang University of Science and Technology (POSTECH) in Information Systems, Systems Science and Mathematics, respectively, and a B.Sc. in Mathematics from POSTECH. Her main research area is Computer Animation covering topics ranging from Computer Graphics to Computer Vision, but the main theme is VISUALIZATION and HUMAN.