An Efficient and Accurate Hierarchical ICIA Fitting Method for 3D Morphable Models

Bong-Nam Kang
Department of Computer Science, Yonsei University, Korea
bongnam@yonsei.ac.kr

Daijin Kim
Department of Computer Science & Engineering, POSTECH, Korea
dkim@postech.ac.kr

Hyeran Byun
Department of Computer Science, Yonsei University, Korea
hrbyun@yonsei.ac.kr

Abstract

We propose the efficient and accurate hierarchical ICIA fitting method for 3D Morphable Models (3DMMs). The conventional ICIA fitting method for 3DMMs requires a long computation time because the 3D face model contains a large number of vertices and it also requires to compute the Hessian matrix using the visible vertices every iteration. For the efficient fitting, we use the hierarchical fitting that use a set of multi-resolution 3D face model and the Gaussian image pyramid. For more accurate fitting, we use a two-stage parameter update that only update the rigid and the texture parameters and then update all parameters after the initial convergence. We present several experiment results to prove that our proposed method shows better performance than previous works.

1. Introduction

Recently, many researchers have been interested in the human face analysis such as face detection, face recognition and facial expression recognition. To conduct research in these topics, we need to perform the face modeling and fitting method of the model to the input image. There are two kinds of face modelings: the 2D face modeling and the 3D face modeling.

In the 2D-based face modeling, there are many approaches such as the active contour models (ACMs), the active shape models (ASMs), and the active appearance models (AAMs). ACMs are only used to extract the facial contour of frontal-view faces and the performance crucially depends on the weight for which manually parameters tuning may be inevitable in many applications [9, 10]. In ASMs, the linear characteristics of ASMs limit their application to them small range of shape variations [14, 4]. AAMs contain the shape and the texture models. AAMs have the efficient fitting methods such as the inverse compositional simultaneous update (ICSU) [6] and the inverse compositional project out (ICPO) [11]. They do not fulfill the two requirements such as the stability and the operating speed at the same time. The above 2D-based face modeling and their fitting method have still the limitations that the face modeling are not robust to the pose and illumination variations and the fitting methods are unstable and inaccurate for representing the input facial images.

To reduce these limitations, many researchers have proposed the 3D face models whose shape or texture can be controlled by a compact set of parameters [2, 5, 8]. For example, the 3DMMs uses the most detailed 3D shape consisting of thousands of vertices and the linear texture. In the 3DMMs, there are two famous fitting methods such as a stochastic Newton optimization (SNO) [3] and an inverse compositional image alignment (ICIA) [12]. While more accurate, the SNO requires a huge amount of computation time because it is necessary to recompute the Jacobian and the Hessian matrices of the warping function at the pixel coordinate every iteration. On the other hand, the ICIA is a time-efficient method because the Jacobian and the Hessian matrices are pre-computed once for all iterations. Despite of this merits, the ICIA took about 30 seconds [12].

To reduce the computation time of the conventional ICIA fitting method, we proposed the efficient and accurate hierarchical ICIA fitting method. It requires a set of multi-resolution 3D face models and the Gaussian image pyramid of the input face image and fitting has been performed hierarchically from the low resolution fitting to the high resolution fitting. We take the two-stage parameter update for obtaining the more accurate performance that the rigid and the texture parameters are updated at the first stage and all parameters are updated after the initial convergence.

This paper is organized as follows: Section 2 describes the existing ICIA fitting method for the 3DMMs. Section 3 describes the multi-resolution 3D face models, a Gaussian image pyramid, and the proposed hierarchical ICIA fitting

978-1-4244-2154-1/08/$25.00 2008 IE

method. Section 4 describes the experimental results that evaluate the performance of the proposed hierarchical ICIA fitting method in terms of the histogram of shape errors, the average fitting error, the convergence rate, and the convergence time. Finally, Section 5 draws a conclusion.

2. 3D ICIA Fitting Method

The fitting method performs the matching of the 3DMM to a 2D face image and seeks the parameters of the 3DMM that express the model texture as close to the input image as possible. In the 3DMMs, the fitting methods aim to find the model parameters $\boldsymbol{\alpha}$, $\boldsymbol{\rho}$, and $\boldsymbol{\beta}$ that explain shape parameters, rigid parameters, and texture parameters of an input face image by a *Maximum a Posteriori* (MAP) estimator which maximizes $p(\boldsymbol{\alpha}, \boldsymbol{\rho}, \boldsymbol{\beta}|\mathbf{I}_{input})$. Applying the Bayes rule, the posterior probability can be represented as: $p(\boldsymbol{\alpha}, \boldsymbol{\rho}, \boldsymbol{\beta}|\mathbf{I}_{input}) \sim p(\mathbf{I}_{input}|\boldsymbol{\alpha}, \boldsymbol{\rho}, \boldsymbol{\beta}) \cdot p(\boldsymbol{\alpha}) \cdot p(\boldsymbol{\rho}) \cdot p(\boldsymbol{\beta})$, where $p(\mathbf{I}_{input} \mid \boldsymbol{\alpha}, \boldsymbol{\rho}, \boldsymbol{\beta})$ is the likelihood, and $p(\boldsymbol{\alpha})$, $p(\boldsymbol{\rho})$, and $p(\boldsymbol{\beta})$ are the prior probabilities of the shape, rigid, and texture parameter, respectively. The prior probabilities of $p(\boldsymbol{\alpha})$ and $p(\boldsymbol{\beta})$ are given by the process of the building 3D face model by PCA, and the prior probability of $p(\boldsymbol{\rho})$ is assumed to be a Gaussian distribution.

In the 3D ICIA fitting method, the cost function $E_{\mathbf{I}}$ is an iteratively minimized log-likelihood that is the sum of the squares of the difference between the model texture and the image texture defined in the reference frame $\mathbf{u} = (u, v)$ as

$$E_{\mathbf{I}}(\delta\boldsymbol{\alpha}, \delta\boldsymbol{\rho}, \delta\boldsymbol{\beta}, \boldsymbol{\alpha}^d, \boldsymbol{\rho}^d, \boldsymbol{\alpha}^t, \boldsymbol{\rho}^t, \boldsymbol{\beta}^t, \boldsymbol{I}) = = \sum_{\Omega(\boldsymbol{\alpha}^d, \boldsymbol{\rho}^d)} \Big[\mathbf{t}(\mathbf{p}^{-1}(\mathbf{p}(\mathbf{u}_i; \boldsymbol{\alpha}^d + \delta\boldsymbol{\alpha}, \boldsymbol{\rho}^d + \delta\boldsymbol{\rho}); \boldsymbol{\alpha}^d, \boldsymbol{\rho}^d); \delta\boldsymbol{\beta}) - \mathbf{t}^{-1}(\mathbf{I}(\mathbf{p}(\mathbf{u}; \boldsymbol{\alpha}^t, \boldsymbol{\rho}^t)); \boldsymbol{\beta}^t)\Big]^2, \quad (1)$$

where the superscripted parameters by d refer to the parameters at which the derivatives are computed, the parameters $\boldsymbol{\alpha}^t$, $\boldsymbol{\rho}^t$, and $\boldsymbol{\beta}^t$ are the current shape, rigid, and texture parameters, respectively. $\Omega(\boldsymbol{\alpha}^d, \boldsymbol{\rho}^d)$ is a set of the visible vertices. Here, the first and second terms in the summation are the model texture that is the result of a texture update with respect to the parameters $\delta\boldsymbol{\beta}$ composed with a shape projection update with respect to the parameters $\delta\boldsymbol{\alpha}$ and $\delta\boldsymbol{\rho}$, and the image texture that removes the modes of texture variations by the inverse texture using the current texture parameters $\boldsymbol{\beta}^t$, respectively.

To obtain the updates of the model parameters, we need to compute the derivatives of the ICIA cost function with respect to the shape parameters $\delta\boldsymbol{\alpha}$, the rigid parameters $\delta\boldsymbol{\rho}$, and the texture parameters $\delta\boldsymbol{\beta}$ at the condition $(\mathbf{0}, \mathbf{0}, \mathbf{0}, \boldsymbol{\alpha}^d, \boldsymbol{\rho}^d, \boldsymbol{\alpha}^t, \boldsymbol{\rho}^t, \boldsymbol{\beta}^t)$.

The Jacobian matrices of the shape, rigid, and texture parameters are combined and denoted as $\mathbf{J} = [\mathbf{J}^s\ \mathbf{J}^\rho\ \mathbf{J}^t]$. Then, the increments of the model parameters are computed by the Gauss-Newton formula as

$$\begin{pmatrix} \delta\boldsymbol{\alpha} \\ \delta\boldsymbol{\rho} \\ \delta\boldsymbol{\beta} \end{pmatrix} = -\mathbf{H}^{-1} \cdot \mathbf{J}^T \cdot \mathbf{E}, \quad (2)$$

where $\mathbf{H}$ is the Hessian matrix that is defined by $\mathbf{H} = \mathbf{J}^T \cdot \mathbf{J}$. After obtaining the increments of the model parameters, they are updated by the following procedures. First, the shape projection is updated using the increments of the shape and projection parameters as

$$\mathbf{p}^{t+1}(\mathbf{u}_i) = \mathbf{p}(\mathbf{p}^{-1}(\mathbf{u}_i; \boldsymbol{\alpha}^d + \delta\boldsymbol{\alpha}, \boldsymbol{\rho}^d + \delta\boldsymbol{\rho}); \boldsymbol{\alpha}^d, \boldsymbol{\rho}^d); \boldsymbol{\alpha}^t, \boldsymbol{\rho}^t) \quad (3)$$

Second, the updated shape parameters $\boldsymbol{\alpha}^{t+1}$ and the updated projection parameters $\boldsymbol{\rho}^{t+1}$ are obtained from the Kernel-based selective method [13]. Third, the texture parameters are updated by the additive manner as $\boldsymbol{\beta}^{t+1} = \boldsymbol{\beta}^t + \delta\boldsymbol{\beta}$. Table 1 summarizes the overall procedure of the ICIA fitting method.

Table 1. The overall process of ICIA fitting algorithm.

Procedure ICIA_fitting($\mathbf{I}$, $\mathbf{S}$, $\mathbf{T}$, $\boldsymbol{\alpha}^0$, $\boldsymbol{\rho}^0$, $\boldsymbol{\beta}^0$)
$t = 0$.
Set the initial parameters:
$\boldsymbol{\alpha}^t = \boldsymbol{\alpha}^0, \boldsymbol{\rho}^t = \boldsymbol{\rho}^0, \boldsymbol{\beta}^t = \boldsymbol{\beta}^0, \mathbf{E}^{old} = \infty$.
Iterate:
Warp the current face image $\mathbf{I}(\mathbf{p}(\mathbf{u}; \boldsymbol{\alpha}^t, \boldsymbol{\rho}^t))$.
Compute the error $\mathbf{E}^t = \mathbf{t}(\mathbf{u}; \mathbf{0}) - \mathbf{t}^{-1}(\mathbf{I}(\mathbf{p}(\mathbf{u}; \boldsymbol{\alpha}^t, \boldsymbol{\rho}^t)); \boldsymbol{\beta}^t)$.
Check stop condition:
If($(t > max\_iter)$ or $(\lvert\mathbf{E}^{old} - \mathbf{E}^t\rvert < E_{th})$), Stop.
Compute the Jacobian matrix $\mathbf{J} = [\mathbf{J}^s\ \mathbf{J}^\rho\ \mathbf{J}^t]$.
Compute the Hessian matrix $\mathbf{H} = \mathbf{J}^T \cdot \mathbf{J}$.
Compute $\delta\boldsymbol{\alpha}$, $\delta\boldsymbol{\rho}$, $\delta\boldsymbol{\beta}$ using Eq. (2).
Update $\boldsymbol{\alpha}^{t+1}$, $\boldsymbol{\rho}^{t+1}$ using the kernel-based selective method [13].
Update $\boldsymbol{\beta}^{t+1} = \boldsymbol{\beta}^t + \delta\boldsymbol{\beta}$.
$\mathbf{E}^{old} = \mathbf{E}^t$.
Set $t = t + 1$ and go to Iterate.
End

3. The Proposed Hierarchial ICIA Fitting Method

Although the ICIA fitting method is time-efficient due to the pre-computation of the derivatives in the initial step, it requires a lot of computation time($\approx$ 30 seconds) for the 3DMM fitting. (1) we handle thousands of vertices in the 3DMMs, (2) we need to select the visible vertices at each iteration, and (3) we need to change the derivatives by adding the new vertices and discarding the the invisible vertices accordingly and re-compute the derivatives at each iteration.

The time complexity of the computation of the existing ICIA fitting method is $O((N_s + N_t) \cdot N_{vv} + N_{vv} \log N_{vv} +$

$(N_s + N_t)^2 \cdot N_{vv})$, where N_s, N_t, and N_{vv} are the dimension of the shape model, the dimension of the texture model and the number of the visible vertices, respectively, and the first terms are due to the update of the shape and texture parameters, the second term is due to the computation of the median of absolute deviations (MAD) which is necessary for sorting the residuals to designate the outliers, and the third term is due to the re-computation of the Hessian matrix. As shown by, the time complexity of the ICIA fitting algorithm, we need to reduce the number of vertices as many as possible to reduce the overall fitting time. To meet this requirement, we propose the hierarchical ICIA fitting method that uses a set of multi-resolution 3D face model and the Gaussian image pyramid of the input face image.

3.1. Multi-Resolution 3D Face Models

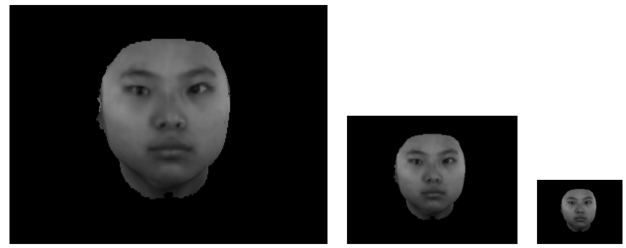
The multi-resolution 3D face models are constructed by sub-sampling the model shape $\mathbf{S}$ that is formed by the mean shape and the shape basis vectors, formed by the mean texture and the texture basis vectors. The model shape $\mathbf{S}_{levle-3}$ and the model texture $\mathbf{T}_{levle-3}$ at the level-3 are just the model shape $\mathbf{S}$ and the model texture $\mathbf{T}$, respectively. Next, the shape and the texture at the level-3 are sum-sampled at the 2:1 sampling rate to construct the model shape $\mathbf{S}_{level-2}$ and the model texture $\mathbf{T}_{level-2}$. Finally, the model shape and the model texture at the layer-2 are sub-sampled at the 2:1 sampling rate to construct the model shape $\mathbf{S}_{level-1}$ and the model texture $\mathbf{T}_{level-1}$. Table 2 summarized the overall procedure of constructing the proposed multi-resolution 3D face models.

Table 2. The overall procedure of constructing the multi-resolution 3D face models.

(1) Construct the level-3 3D face model using PCA.
(2) Construct the level-2 3D face model:
  (a) Subsample the mean face of the level-3 face model by 2.
  (b) Subsample the basis vectors of face model
    using the indices that are obtained in the above mean face subsampling.
(3) Construct the level-1 3D face model:
  (a) Subsample the mean face of the level-2 face model by 2.
  (b) Subsample the basis vectors of face model
    using the indices that are obtained in the above mean face subsampling.

3.2. Gaussian Image Pyramid

We also generate the Gaussian image pyramid, which is a hierarchy of the low-pass filtered versions of the input face image such that the successive level corresponds to the lower frequency image. Figure 1 shows the example of the Gaussian image pyramid of an input face image, where Figure 1-(a) is the original face image (level-3), Figure 1-(b) is the sub-sampled image with a half of the number of pixels at the level-3, and Figure 1-(c) is the sub-sampled image with a half of the number of pixels at the level-2.



(a) Level-3 (b) Level-2 (c) Level-1

Figure 1. A Gaussian image pyramid of the input face image.

3.3. Hierarchical ICIA Fitting Method

After generating the multi-resolution 3D face models and the Gaussian image pyramid of the input face image, we apply the proposed hierarchical ICIA fitting method. At the initialization, the ICIA fitting method requires the correspondences between some of the model vertices and the input face image manually [12]. But the proposed fitting method tries an automatic initialization by aligning the pre-designated eye's positions in the face model and the detected eye's positions by the revised modified census transform (RMCT)-based face detector [7].

Table 3 shows the overall procedure of the proposed hierarchical ICIA fitting method: (1) we generate the Gaussian image pyramid $\mathbf{I}_{level-1}$, $\mathbf{I}_{level-2}$, and $\mathbf{I}_{level-3}$ from the input face image, (2) we detect the face and eyes using the RMCT-based face detector. (3) we set the initial shape and texture parameters as $\boldsymbol{\alpha}^0 = \mathbf{0}$ and $\boldsymbol{\beta}^0 = \mathbf{0}$, and rigid parameter $\boldsymbol{\rho}$ are initialized by using the pre-designated eye positions of the face model and the detected eye positions, (4) we perform the first layer ICIA fitting process using the face image $\mathbf{I}_{level-1}$, the shape model $\mathbf{S}_{levle-1}$, the texture model $\mathbf{T}_{level-1}$, and the model parameters $\boldsymbol{\alpha}^0$, $\boldsymbol{\rho}^0$, and $\boldsymbol{\beta}^0$, (5) we set the model parameters obtained from the first layer ICIA fitting results as the initial parameters for the second layer ICIA fitting, where the superscript $*$ implies the parameter value after the ICIA fitting, (6) we perform the second layer ICIA fitting process using the face image $\mathbf{I}_{level-2}$, the shape model $\mathbf{S}_{level-2}$, the texture model $\mathbf{T}_{level-2}$, (7) we set the model parameters obtained from the second layer ICIA fitting results as the initial model parameters for the third layer ICIA fitting, (8) we perform the third layer ICIA fitting process using the face image $\mathbf{I}_{level-3}$, the shape model $\mathbf{S}_{level-3}$, the texture model $\mathbf{T}_{level-3}$, and (9) we obtain the synthesized face image using the obtained model parameters $\boldsymbol{\alpha}^*_{level-3}$, $\boldsymbol{\rho}^*_{level-3}$, and $\boldsymbol{\beta}^*_{level-3}$ as $\mathbf{S} = \mathbf{S}_0 + \sum_{i=1}^{i=N_s} \alpha^*_{level-3,i} \cdot \mathbf{S}_i$ and $\mathbf{T} = \mathbf{T}_0 + \sum_{i=1}^{i=N_t} \beta^*_{level-3,i} \cdot \mathbf{T}_i$.

Table 3. The overall procedure of the hierarchical ICIA fitting method.

Procedure Hierarchical_ICIA_fitting(I)
(1) Generate the GIP from the input image $\mathbf{I}$: $\mathbf{I}_{level-1}, \mathbf{I}_{level-2}, \mathbf{I}_{level-3}$.
(2) Detect face and eyes using the face detector.
(3) Set the initial parameters for the first layer ICIA fitting: $\boldsymbol{\alpha}^0 = \mathbf{0}, \boldsymbol{\beta}^0 = \mathbf{0}, \boldsymbol{\rho} = Face - detector$.
(4) Perform the first layer ICIA fitting: $\textbf{ICIA\_fitting}(\mathbf{I}_{level-1}, \mathbf{S}_{level-1}, \mathbf{T}_{level-1}, \boldsymbol{\alpha}^0, \boldsymbol{\rho}^0, \boldsymbol{\beta}^0)$.
(5) Set the initial parameters for the second layer ICIA fitting: $\boldsymbol{\alpha}^0 = \boldsymbol{\alpha}^*_{level-1}, \boldsymbol{\rho}^0 = \boldsymbol{\rho}^*_{level-1}, \boldsymbol{\beta}^0 = \boldsymbol{\beta}^*_{level-1}$
(6) Perform the second layer ICIA fitting: $\textbf{ICIA\_fitting}(\mathbf{I}_{level-2}, \mathbf{S}_{level-2}, \mathbf{T}_{level-2}, \boldsymbol{\alpha}^0, \boldsymbol{\rho}^0, \boldsymbol{\beta}^0)$.
(7) Set the initial parameters for the third layer ICIA fitting: $\boldsymbol{\alpha}^0 = \boldsymbol{\alpha}^*_{level-2}, \boldsymbol{\rho}^0 = \boldsymbol{\rho}^*_{level-2}, \boldsymbol{\beta}^0 = \boldsymbol{\beta}^*_{level-2}$
(8) Perform the third layer ICIA fitting: $\textbf{ICIA\_fitting}(\mathbf{I}_{level-3}, \mathbf{S}_{level-3}, \mathbf{T}_{level-3}, \boldsymbol{\alpha}^0, \boldsymbol{\rho}^0, \boldsymbol{\beta}^0)$.
(9) Obtain the 3D synthesized face using the fitted parameters. $\mathbf{S} = \mathbf{S}_0 + \sum_{i=1}^{i=N_s} \alpha^*_{level-3,i} \cdot \mathbf{S}_i$. $\mathbf{T} = \mathbf{T}_0 + \sum_{i=1}^{i=N_t} \beta^*_{level-3,i} \cdot \mathbf{T}_i$.

4. Experimental Results

We have performed several experiments that showed the validity of the proposed hierarchical fitting method. We define some performance measures to evaluate the fitting performance such as the normalized correlation and the root mean squared error (RMSE). First, we define the normalized correlation as

$$C = \frac{\boldsymbol{\alpha}^T \cdot \tilde{\boldsymbol{\alpha}}}{\|\boldsymbol{\alpha}\| \cdot \|\tilde{\boldsymbol{\alpha}}\|}, \tag{4}$$

where $\boldsymbol{\alpha}$ and $\tilde{\boldsymbol{\alpha}}$ are the ground truth model parameters and the recovered model parameters after the 3DMM fitting, respectively. Second, we define the root mean squared error (RMSE) as the average distance between the ground truth shape (or texture) and the fitted shape (or texture).

We used the BJUT-3D Face Database [1]. And all experiments are performed with the initial parameters $\boldsymbol{\alpha}^0 = \mathbf{0}$, $\boldsymbol{\rho}^0 = \boldsymbol{\rho}_{Face-Detector}$, and $\boldsymbol{\beta}^0 = \mathbf{0}$, where $Face - Detector$ means that the parameter values are obtained from the results of detecting the face and the eyes.

All experiments have been conducted on the desktop PC that consists of a Pentium IV CPU with a clock speed of 3GHZ, a 4GB RAM, Window XP professional x64 Edition, and C++ and OpenCV development tools.

4.1. Comparison of the Fitting Performance Using Different Types of Image Pyramids

First, we compared the fitting performance with respect to two different types of image pyramid constructions: the sub-sampling image pyramid (SIP) and the Gaussian image pyramid (GIP) when the proposed hierarchical ICIA fitting method was used with the two-stage parameter updates (TSPU).

Table 4 compares the fitting performance of two different types of image pyramids: SIP and GIP. From this table, we know that the fitting performance of using GIP outperforms that of using SIP in all performance measures such as C_{shp}, C_{tex}, $RMSE_{shp}$, and $RMSE_{tex}$ where the normalized correlation of shape, normalized correlation of texture, RMSE of the shape, and RMSE of the texture.

Table 4. Comparison of the fitting performance between SIP and GIP.

	C_{shp}	C_{tex}	$RMSE_{shp}$	$RMSE_{tex}$
SIP	0.5623	0.9682	4.0623	3.0681
GIP	0.8200	0.9934	2.1677	1.6054

Figure 2 shows a histogram of the shape errors between the ground truths and the fitted shapes using SIP and GIP. It shows that GIP has the smaller mean and the smaller standard deviation of the shape errors than SIP.

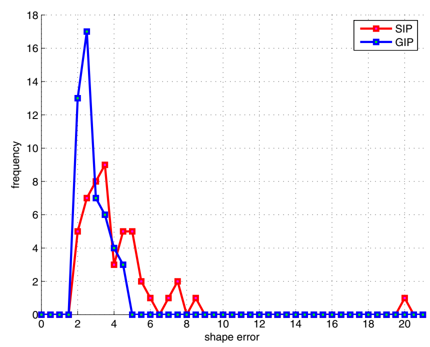


Figure 2. A histogram of the shape errors using SIP and GIP.

4.2. Comparison of the Fitting Performance Using Different Types of Parameter Updates

Second, we compared the fitting performance with respect to two different types of parameter updates: the single-stage parameter update (SSPU) and the two-stage parameter update (TSPU), where the former updates all model parameters in all layer ICIA fittings and the latter updates the rigid and texture parameters in the first layer ICIA fitting and then updates all model parameters in the second and third layer ICIA fitting when the proposed hierarchical ICIA fitting method was used with the GIP construction.

Table 5 compares the fitting performance of two different types of parameter updates: SSPU and TSPU. From this table, we know that the fitting performance of using the TSPU outperforms that of using the TSPU in all performance measures such as C_{shp}, C_{tex}, $RMSE_{shp}$, and $RMSE_{tex}$.

Figure 3 shows a histogram of the shape errors between the ground truths and the fitted shapes using SSPU and

Table 5. Comparison of the fitting performance between the SSPU and the TSPU.

	C_{shp}	C_{tex}	$RMSE_{shp}$	$RMSE_{tex}$
SSPU	0.5623	0.9682	4.0623	3.0681
TSPU	0.8200	0.9934	2.1677	1.6054

TSPU. It shows that TSPU has the smaller mean and the smaller standard deviation of the shape errors than SSPU.

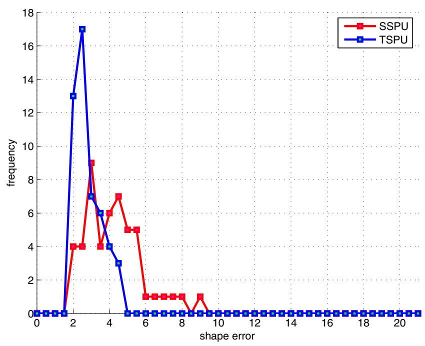


Figure 3. A histogram of the shape errors using SSPU and TSPU.

4.3. Comparison of the Fitting Performance with Different Types of Fitting Methods

Finally, we compared the fitting performance with respective to two different types of fitting methods: the conventional ICIA fitting method (CICIA)and the proposed hierarchical ICIA fitting method (HICIA) when the GIP is used to generate the image pyramid and the TSPU is used for the parameter updates.

Table 6 summarizes the fitting performances of two fitting methods in terms of the average number of iterations (N_{iter}), the average computation time for the ICIA fitting (T_{fit}), the normalized correlation of the fitted shape parameters (C_{shp}), the normalized correlation of the fitted texture parameters (C_{tex}), the RMSE of the shape errors (R_{shp}), and the RMSE of the texture errors (R_{tex}). This table indicates that (1) the proposed hierarchical ICIA fitting method is faster than the conventional ICIA fitting method by a speed up to 3, (2) the correlations of shape and texture parameters of using HICIA are much higher than those of using CICIA, and (3) the RMSEs of the shape and texture errors of using HICIA are reduced by 3-4 times than those of using CICIA.

Table 6. Comparison of fitting performance between the CICIA and the HICIA.

	N_{iter}	T_{fit}	C_{shp}	C_{tex}	R_{shp}	R_{tex}
CICIA	71.02	16.3	0.3754	0.8550	6.4540	6.0410
HICIA	57.52	5.7936	0.8200	0.9934	2.1667	1.6054

Figure 4 shows a histogram of the shape errors between the ground truths and the fitted shapes using CICIA and HICIA. It shows that HICIA has the smaller mean and the smaller standard deviation of the shape errors than CICIA.

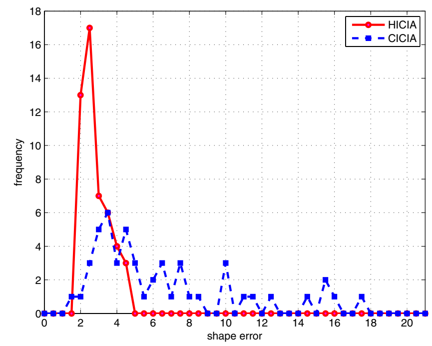


Figure 4. A histogram of the shape errors using the CICIA and the HICIA.

Figure 5 compares the convergence rates of five different fitting methods: TYPE1 (CICIA), TYPE2 (HICIA+SIP+SSPU), TYPE3 (HICIA+SIP+TSPU), TYPE4 (HICIA+GIP+SSPU), and TYPE5 (HICIA+GIP+TSPU). In this experiment, the input face image is successfully converged when the shape error of the fitting face image is smaller than a given threshold value and the convergence rate is defined by the ratio of the number of successfully converged face images over the total number of face images. This figure illustrates that the convergence rate of the proposed hierarchical ICIA fitting method with GIP construction and TSPU update outperforms those of other fitting methods.

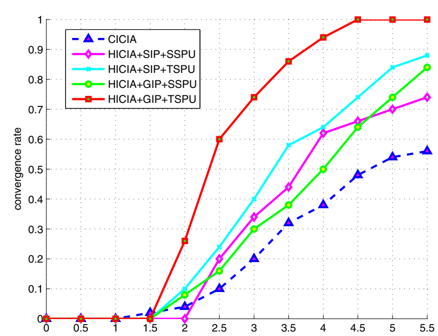


Figure 5. Convergence rates of the different fitting methods.

Figure 6 shows some fitting results of the proposed hierarchical ICIA fitting method, where the first, second, third, and fourth columns represent the input face images, the fitted shapes, the synthesized face images using the fitted model parameters, and the different views of the synthesized images of four different people, respectively. From this figure, we know that the proposed hierarchical ICIA fitting method provides good fitting performances, and thus it can be used for fitting a new subject that is not included in the training set.

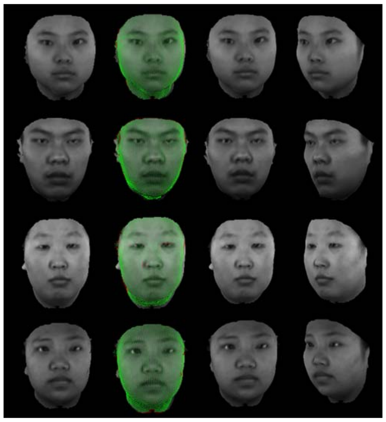


Figure 6. Some fitting results of the proposed hierarchial ICIA fitting method.

5. Conclusion

We proposed an efficient and accurate hierarchical ICIA fitting method. The proposed fitting method is efficient because it generates the multi-resolution 3D face models and constructs the Gaussian image pyramid. Further the fitting is conducted hierarchically from the lower layer to the higher layer. As a side effect, this also improves the fitting performance because the fitted model parameters at the lower layer are used as the initial model parameters for the upper layer fitting.

We also proposed the two-stage parameter update method that updates the rigid parameter and the texture parameters only at the first layer fitting, and updated all parameters at the succeeding upper layer fitting. This update method reduces the computation time for fitting and improves the stability of fitting because it is very difficult to discriminate the movement of vertex positions due to the shape parameter changes and the rigid parameter changes in the very beginning of fitting.

We performed several experiments that validate the efficiency and accuracy of the proposed hierarchical ICIA fitting algorithm. From the experiment results, (1) it completed the fitting within about 57 iterations ($\approx$ 5 seconds), (2) its speed-up ratio is about 3, (3) the performance of the proposed fitting method outperformed that of the existing fitting method, (4) the two-stage parameter update showed better fitting performance than single-stage parameter update in terms of the fitting time and fitting error.

Acknowledgements

This research was supported by MIC, Korea under ITRC IITA-2008-(C1090-0801-0046). Also it was supported by grant No. R01-2007-000-11683-0 from the Basic Research Program of the Korea Science & Engineering Foundation.

References

[1] The BJUT-3D Large-Scale Chinese Face Database. Technical report, The Multimedia and Intelligent Software Technology Beijing Municipal Key Laboratory in Beijing University of Technology, 2005. 4

[2] J. Ahlberg. Using the active appearance algorithm for face and facial feature tracking. In *ICCV'01 Workshop on Recognition, Analysis and Tracking of Faces and Gestures in Real-Time Systems*, pages 68–72, 2001. 1

[3] V. Blanz and T. Vetter. Morphable Model for the Synthesis of 3D Faces. In *Proceedings of the 26th annual conference on Computer graphics and interactive techniques*, pages 187–194, 1999. 1

[4] T. Cootes, C. Taylor, D. Cooper, and J. Graham. Active Shape Models - their training and application. *Computer Vision and Image Understanding*, 61(1):38–59, 1995. 1

[5] S. Gokturk, J. Bouguet, C. Tomasi, and B. Girod. Model-based face tracking for view-independent facial expression recognition. In *FGR '02: Proceedings of the Fifth IEEE International Conference on Automatic Face and Gesture Recognition*, pages 272–278, 2002. 1

[6] R. Gross, I. Matthews, and S. Baker. Lucas-Kanade 20 Years on : A Unifying Framework: Part 3. Technical report, Carnegie Mellon University Robotics Institute, 2003. 1

[7] B. Jun and D. Kim. Robust Real-time Face Detection uing Face Certainty Map. In *In Proceeding of The 2nd International Conference on Biometrics*, volume 4642, pages 29–38, 2007. 3

[8] Y. Li, T. Gong, and H. Liddell. Modelling faces dynamically across views and over time. In *Proceedings of IEEE International Conference on Computer Vision*, 2001. 1

[9] A. W. M. Kass and D. Terzopoulos. Snakes: Active Contour Models. In *In IEEE International Conference on Computer Vision*, pages 259–268, 1987. 1

[10] A. W. M. Kass and D. Terzopoulos. Snakes: Active Contour Models. *International Journal of Computer Vision*, 1(4):321–331, 1987. 1

[11] I. Matthews and S. Baker. Active Apperance Models Revisited. *International Journal of Computer Vision*, 60(2):135–164, 2004. 1

[12] S. Romdhani and T. Vetter. Efficient, Robust and Accurate Fitting of a 3D Morphable Model. In *ICCV '03: Proceedings of the Ninth IEEE International Conference on Computer Vision*, pages 59–66, 2003. 1, 3

[13] N. C. S. Romdhani and T. Vetter. Selective vs. Global Recovery of Rigid and Non-rigid Motion. Technical report, CS Dept University of Basel, 2002. 2

[14] C. T. T. Cootes and A. Lanitis. Active Shape Models: Evalustion of a Multi-resolution Method for Improving Image Search. In *In British Machine Vision Conference*, volume 1, pages 327–336, 1994. 1