

Geometrical 2D Face Rotation by Using Gabor-tensor-based Active Appearance Model

Sadi Vural

Osaka University, Division of Systems Science,
Department of Systems Innovation, Graduate School of Engineering Science
1-3, Machikaneyama-cho, Toyonaka, Osaka, Japan

Abstract: In this paper, we present a novel face rotation methodology, which generates multi-pose faces from one single frontal face image. The proposed method has two major parts: the first part automatically detects twelve initial landmark points, on a given face image for the initialization of the facemask, and the second part applies Gabor-tensors to represent the texture and shape modelling. The proposed method uses both texture and shape information. Texture information is obtained by using spatial frequency in different sizes and orientations. Shape information is obtained by applying the Gaussian model. Experimental results on publicly available face databases show that the proposed approach gives high recognition rates on non-frontal images and it impressively outperforms the performance of other state-of-the-art methods such as active appearance model (AAM), 3D morphable models and their expanded methodologies.

Keywords: Feature, Gabor-tensor, Gaussian analysis, morphological techniques, morphable models, spatial-frequency, shape, texture

I. INTRODUCTION

Although there has been much interest in face recognition, several problems have raised such as face pose [1] and illumination [2],[3], [4], [5]. Recovering the geometry and texture of a human face from outdoor images remains a very important, yet a challenging issue. Zhao et al. [6] conducted a survey on real-world problems. Among them, especially face pose and illumination have been reported as two major problems. Furthermore, Chutorian et al. [7] conducted a survey on face pose methodologies. He compared eight different methodologies. He says that model-based tracking approaches provide promising results under high-resolution conditions. Although there are many works done for pose invariant feature extraction, the methods fail in real world conditions [8], [9], [10]. To overcome these problems, many studies on invariant feature extraction under various face poses [11] as well as the estimation of face angle [12], [13] have been proposed. Breitenstein et al. [14] proposed a real-time estimation of a face pose by using range images. They used nose tip for angle determination and a precompiled face database for angle estimation.

Feature extraction methods use either holistic face features or local features. Holistic approaches are fast and stable if the testing and training environments are the same [15]. Local feature approaches are computationally expensive and weak to illumination conditions [16].

As an alternative feature extraction, Active Appearance Model (AAM) has been introduced by Edwards et al. [17], and AAM has been expanded by Cootes et al. [18], [19]. AAM is a generative parametric model that describes both

shape and appearance of a given object. AAM is originally inherited from Active Contour Model (ACM) [20] and Active Shape Model (ASM) [21], and it integrates the idea of Eigen-faces [22]. It was initially used for medical image processing to trace the activity of a brain [23], [24]. AAM is now widely used for various purposes such as face alignment [18], face feature tracking [25], [26], feature classification [27], feature extraction [28], [29] and pose correction [30], [31], [32]. Hyung offered a different method of AAM called Tensor-AAM [33]. He used a tensor structure to have variation specific vectors. He obtained tensor structure by a multiplication of training coefficient vectors with constituent basis vector mixture. His method improves the originally proposed AAM [34]. However, small fitting errors in his approach cause significant performance drop. To improve the fitting errors, Scott et al. [22] proposed gradient orientation instead of intensity. The approach is quite different from other proposed approaches [35], [36], [37]. In his method, although convergence speed is faster, the failure rate is high. Direct appearance models (DAMs) predict shape parameters directly from texture [38]. Li et al. [39] extended DAM for multi-view face alignment. Yan proposed texture-constrained ASMs (TC-ASMs), where the shape update predicted by an ASM is combined with the shape from a global texture model such like the one in AAM [40]. Jiao et al. [41] and Yu et al. [42] used Gabor texture representation in AAM. Furthermore, Jiao used a generic algorithm to model the Gabor feature distribution for improving the errors that occur during illumination and



expression changes. Su et al. [43] proposed a texture representation method by combining Gabor wavelets and local binary patterns (LBP). Blanz et al. [44] used an optical flow for building the Morphable Model. Morphable models have been studied in the field of face analysis and shape reconstruction [45]. John et al. [46] proposed 3D morphable models for recognizing faces. It is successful in certain environments but the performance degrades due to the fitting errors in unconstrained environments. Wang et al. [47] introduced an alternative method for face reconstruction by using morphable models. Asthan et al. [54] proposed a pose invariant face recognition by using 3D pose normalization. The approach fully depends on 3D technique and the processing speed is slow compared to AAM-based approaches.

In this paper, we propose a novel approach for rotating a frontal face into various angles. We detect 12 initial data points for facemask initialization. The detection of twelve points is done by Haar-like features and then refined by Scale-Invariant Feature Transform (SIFT) in order to have a better representation of face images. We compute the derivatives of the residuals for each precompiled model, input face image by extracting the Gabor-tensor features, and use the derivatives to compute the update steps via an iterative matching algorithm. We apply illumination normalization proposed by Tan and Triggs [48] before the feature extraction.

The paper is organized as follows: In section 2, we explain overall architecture and in section 3, the technical details of the texture, shape and frequency modelling are given in detail. Section 4 gives the experimental results. The results of face rotation based on the proposed method here are done by using major face databases and the results and rotation degree and success rates are given in this section. Finally, section 5 concludes the work.

II. OVERALL SYSTEM FLOW

Fig. 1 shows the overall structure of the proposed method. In our proposed method, we propose an automated face rotation methodology, which includes automatic face detection, landmark detection, and face synthesis (fig. 1)

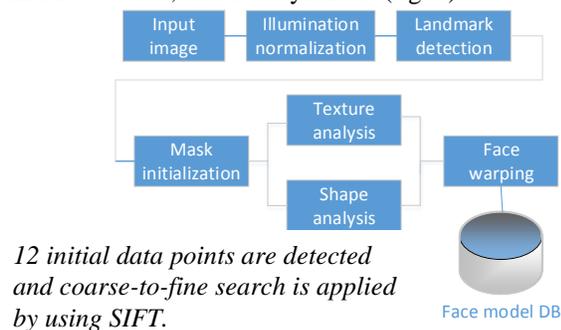


Fig. 1. Overall structure of the proposed method

Input image is the cropped 2D face image. Landmark detection is performed by pretrained haar-like cascades based on Lienhart's technique [56]. Because of haar-like cascades, a total of 12 points consisting of face coordinates, nose, mouth corners and the center of mouth is detected (fig. 2). Furthermore, we apply SIFT to do a coarse-to-fine-search around these points. Coarse-to-fine-search approach is necessary step to determine the exact location of the face components. The more accuracy we obtain, the less iteration we use for the model fitting. Hence, we save time for model fitting while increasing the accuracy of the face warping.

III. FACE MODELING BY TEXTURE AND SHAPE

A. Model training

In order to make a model, which is robust to environmental changes, we need a number of marked face images. We train the images and create a training model. We use Gabor-tensor filters to extract multi-scale and multi orientation information from face images. Each face part is trained and grouped. For example, eye patterns from all images are extracted and their features are transformed into spatial-frequency subdomain. After the points from all other face parts including nose, mouth, and chin areas are detected, the features are extracted and unified to create one feature domain from subdomains. The quality of the model training is directly related to the orientation and scale of the Gabor-tensor filters. The robustness of Gabor-tensor filters comes from their circular structure. In addition, the amount of shape and texture variation that is caused by changes in illumination is significantly less for the proposed method than other proposed AAMs.

The training presented in this paper covers poses with yaw angles from -45 to $+45$ degree and pitch angles from -30 to $+30$ degree. Face images in this range were trained using data from the USF Human ID DB [49], Extended Yale DB, CMU-PIE database and the MultiPIE DB [50]. From the Multi-PIE DB, we used 200 people in poses 05 1, 05 0, 04 1, 19 0, 14 0, 13 0, and 08 0 to capture the shape and texture information induced by pose, and 50 people in 18 illumination conditions to capture the texture variation induced by different illumination conditions. For the training, we manually marked 120 images and 57 points were marked in each face image. These 57 points are selected from lower face boundary, mouth, nose, eyes, and eyebrows. We used fully automated process during testing. No manual marking or selection has been done during the testing.

B. Model building



After model training, we need to construct a model for face images. Constructed model is fit to the face and it does not displace its position so that proper face features are extracted. Extraction of proper face features is specifically important if subject is moving. Proposed techniques prevents the effects of movement by using tracking algorithms. As we consider a face manipulation for many-to-many face surveillance, tracking of multiple faces is not always satisfactory. In our method, we establish a model that fits to the face. Fitting quality depends on model building and model training.



(a) Initial detection (b) Initial points (c) Face mask
Fig. 2. Face landmark locations

Fig.2 (a) shows the initial landmark points, which are automatically detected by using Haar-like filters and refined by SIFT. Fig.2 (b) is the points only. Outer four points are the face boundary points and internal points are the landmark coordinates. Fig. 2(c) is the face shape model. Fig. 3 shows an entire procedure of the proposed technique.

Determination of twelve landmark points is automatically done by using the same procedure during face initial point marking. Fitting an AAM to a new image is generally accomplished in an iterative way and it requires accurate model initialization to avoid converging to bad local minima. Let \bar{t} denote mean texture vector in the reference image, \bar{s} denote the mean shape vector and G_t, G_s be the feature matrices of Gabor wavelet for the shape and texture dispersions computed from the training set.

$$\begin{aligned} \mathbf{t} &= \bar{\mathbf{t}} + \mathbf{G}_t \mathbf{P}_t + \theta \\ \mathbf{s} &= \bar{\mathbf{s}} + \mathbf{G}_s \mathbf{P}_s + \theta \end{aligned} \quad (1)$$

where, \mathbf{t} is the texture data of the synthesized face, \mathbf{s} is the shape data of the synthesized face. \mathbf{P} is the adjusting parameter of the annotation mask and θ is the angle variation.

Let \mathbf{P}_t be face texture features, \mathbf{P}_s be share features. To obtain the combined shape and texture data, \mathbf{P}_t and \mathbf{P}_s are combined as

$$\mathbf{P} = \begin{bmatrix} \mathbf{W}_s \mathbf{P}_s \\ \mathbf{P}_t \end{bmatrix} = \begin{bmatrix} \mathbf{W}_s \mathbf{G}_s^T (\mathbf{s} - \bar{\mathbf{s}}) + \theta \\ \mathbf{G}_t^T (\mathbf{t} - \bar{\mathbf{t}}) + \theta \end{bmatrix} \quad (2)$$

where \mathbf{W}_s represents the Gabor-tensor features in matrix shape. A face can be described by G_s and G_t .

G_s and G_t are Gabor-tensor features which are obtained by the convolution of impulse function and Gaussian envelope. G_s and G_t are used to get fourth-order tensor in $G \in R^{n_1 \times n_2 \times \mu \times \gamma}$. n_1, n_2 gives the pixel location. The entries of 4th tensor are complex numbers. There are 40 components in 4th tensor and each of its components contains the magnitude part of the output.

$$\begin{aligned} G_{\mu, \gamma}(\vec{z}) &= \frac{\|\mathbf{k}_{\mu, \gamma}\|^2}{\sigma^2} \exp\left(-\frac{\|\mathbf{k}_{\mu, \gamma}\|^2 \|\vec{z}\|^2}{2\sigma^2}\right) \\ &\times \left[\exp(i\mathbf{k}_{\mu, \gamma} \times \mathbf{z}) - \exp\left(-\frac{\sigma^2}{2}\right) \right] \end{aligned} \quad (3)$$

where μ is the scale, γ is the orientation and typically, $\mu = 5$, $\gamma = 8$. \vec{z} is variable in z domain. \mathbf{k} is the frequency vector which determines the scale and orientation. The term $\exp\left(-\frac{\sigma^2}{2}\right)$ is used to make DC components zero to minimize the illumination effects. σ determines the oscillation under Gaussian envelope and it is set to 2π to derive the Gabor representation by convolving face images during testing. In fourth-order tensor, first two values are the pixel values, third value is the scale and fourth value is the orientation. The entries of the fourth-order tensor are complex numbers and the magnitude part of the tensor is defined as Gabor feature. We compute the Gabor for both scale and orientation separately to remove the computational complexity. Let G_s be the summation of all Gabor filter convolution in vertical axis and G_t be the summation of all filters output in horizontal axis. Horizontal axis is the summation of the scale parameter and vertical axis is the summation of the lambda orientation parameter. Then,

$$G_{sum} = G_s + G_t \quad (4)$$

where $G_s \in R^{n_1 \times n_2 \times \mu}$, $G_t \in R^{n_1 \times n_2 \times \gamma}$ and $G_{sum} \in R^{n_1 \times n_2}$. G_s and G_t are 3rd order tensor and G_{sum} is the 2nd order tensor.

After obtaining the shape and texture model, all the testing images are warped to the mean shape and a vector set is obtained as seen in (2).

To use Gabor-tensor-based Active Appearance Model (GAAM), we need an automatically iterating fitting method, namely \mathbf{P} in (3). Cootes proposed an iterative scheme to find \mathbf{P} . It starts with an initial value for \mathbf{p} , and converges to the optimal value by minimizing the difference between the target image and the image synthesized by the model. \mathbf{p} is the estimation value of \mathbf{P} .

During the model iteration step, the current face residual between the training shape ($G_s(\mathbf{p})$) and the training texture ($G_t(\mathbf{p})$), the mean shape is recomputed by using



$$r(p) = G_s(p) + G_t(p) \quad (5)$$

By using the Gauss-newton method, the approximation by using first-order Taylor expansion can be done by

$$r(p + \delta p) \approx r(p) + \frac{\partial r}{\partial p} \delta p$$

(6)

Building the derivative with respect to p and setting it to zero give the following equation

$$\delta p = - \left(\frac{\partial r}{\partial p}^T * \frac{\partial r}{\partial p} \right)^{-1} \frac{\partial r}{\partial p}^T r(p)$$

(7)

Training is done once and each parameter is displaced from its value. During the iteration, the current situation is subtracted from the future predictions. Furthermore, we trained right faces, left faces, up and down faces from Extended Yale database, USF Human ID DB, MultiPIE DB and CMU-PIE database, which contain light-controlled face images. We selected approximately 30 images from



each set.

(a) left face samples



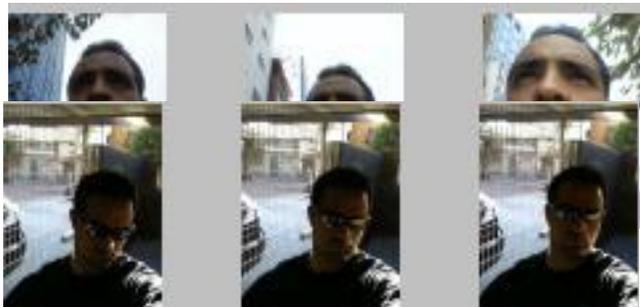
(b) right face samples

(c) Faces from down view

(d) Faces from top view

Fig. 1. Training set for face synthesis

In fig.3, there are four groups. We construct a multi-view



training set model. This leads an improved speed and robustness in accuracy of the model generation. It is because of the fact that train set is separated into a number of subsets. Each subset consists of small training set and less feature variation. We build a separate texture and frequency model by sampling the pixels of the face and its frequency components by using Gabor tensor features. For simplicity, we use the same set of landmarks in each group, namely twelve points as given in fig. 2(b). In our training set, we used four subsets and each subset contains 30 images which leads 120 images in four subsets.

C. Face synthesis

Major point of the face synthesis is the accurate synthesizing of frontal faces from one single face photo image. The process consists of five major stages for face synthesizing. The first stage is that face is detected by face detection, the landmark detector determines twelve initial locations from eyes, nose and mouth region. The second step is that potential interest points are identified over location and scale by constructing a Gabor-tensor pyramid and searching for local peaks (termed keypoints). In the third stage, candidate points are localized to sub-pixel accuracy and eliminated if there is any unstable point. The fourth step identifies the dominant points from each candidate points based on its local image patch. Finally, local image descriptors are built for each point. Local gradient data is used to create point descriptors. The gradient information is rotated to line up with the orientation of the point and then weighted by a Gabor-tensor features with variable scale. This data is then used to create a set of histograms over a window centered on the points. In this paper, we use SIFT as local image descriptor. Furthermore, to make the model robust to illumination changes, we use an illumination normalization technique proposed by Tan and Triggs.

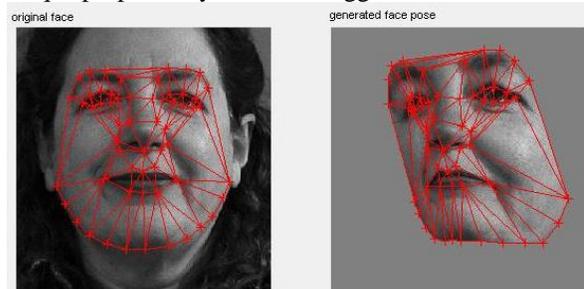


Fig. 2. Face rotation and feature points

GAAM is a Delaunay triangle sets connected to each other as shown in fig. 4. The feature points are projected into a different feature space by comparing it with the database set. All the points are automatically detected.

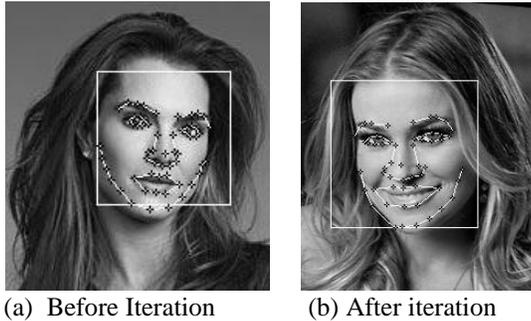


Fig. 3. GAAM masking error and iteration error correction

In GAAM, we fit a GAAM to a 2D frontal face and show the resulting model with warped texture. If the GAAM mask is directly applied to the face without using initial data point, we get some misaligned face data points as in fig. 5(a). However, consecutive iterations helps to correct these as observed in fig. 5(b). We apply the GAAM to the initial data points to reduce the misalignment of the GAAM in the initial steps.

IV. EXPERIMENTAL RESULTS

We perform extensive evaluations on the proposed method for face recognition. Publicly available face databases, CMU PIE [52] and Extended Yale-B [53], are used.

A. Data preparation for Face recognition evaluation

The CMU PIE database contains 41,368 images of 68 subjects with 500+ images for each. The face images were captured by 13 synchronized cameras and 21 flashes, under varying pose, illumination, and expression. For each subject, we manually select 168 images that cover large illumination variation, pose variation, and moderate variety in expression, constituting a challenging face database for recognition task. We randomly choose the images. Their selection is not critical for the testing.

The Yale face database B contains 5,760 single-light-source images of 10 subjects, each under 576 viewing conditions (9 poses 64 illumination conditions). The extended Yale Face Database B contains 16,128 images of 28 human subjects with the same condition and data format as in the previous database. We combine these two databases to include 98 subjects in total. The images are automatically cropped and resized to 50 x 50 pixels.

B. Comparison results

TABLE 1 Face recognition performance comparison by using FRGC v2.0

Faces	GAAM	AAM	AAM-CCA
Number of people in gallery	625	625	625
Number of people in probe	50	50	50
Recognition rate	95%	85%	90%
FRR in FAR=1%	4%	7%	5%
EER Rate	1.7%	4%	3%

For measuring the recognition performance, we used one image in gallery and around fifty images per person in probe. We chose 98 people under 10 different lighting conditions for gallery by combining CMU-PIE DB, Yale DB(extended Yale DB and YaleB-DB). We used extra images from rest of people by combining FRGC V2.0 DB set. Table 2 shows these results.

TABLE 2 Face recognition performance comparison by using CMU-PIE DB+Yale DB

Faces	GAAM	AAM	AAM-CCA
Number of people in gallery	98	98	98
Number of people in probe	10	10	10
Recognition rate	93%	76%	88%
FRR in FAR=1%	5%	10%	7%
EER Rate	3%	6%	4%

We used Gabor+PCA approach as the face recognition algorithm. We used one image per person for gallery and several images per person in probe. The input image (One frontal image) is taken as input and 12 rotated faces from this image are used for enrollment. Probe images are not rotated, not processed by GAAM. We generate 10 images which have pose angles with 5 degree (-30, -25, -20, -15, -10, -5, +5, +10, +15, +20, +25, +30). We rotate the input image from left to right. We do not include up/down, pitch angles in this test.

TABLE 3 Performance comparison in terms of fitting speed with 400 training set

Faces	GAAM	AAM	AAM-CCA
Training samples	400	400	400
Number of annotation	57	57	57
# of average iteration	10	10	10
Fitting speed	55ms	88ms	41ms



Fitting error in pixel	3.2%	8.6%	6.3%
------------------------	-------------	------	------

TABLE 4 Performance comparison in terms of fitting speed with 1000 training set

Faces	GAAM	AAM	AAM-CCA
Training samples	1000	1000	1000
Number of annotation	57	57	57
# of average iteration	10	10	10
Fitting speed	105ms	242ms	176ms
Fitting error in pixel	0.7%	4.9%	2.8%

We compared the GAAM with the default AAM and AAM-CCA under different criteria. AAM is the default one, which was proposed by Cootes. AAM-CCA is the reimplementation of the AAM by using canonical correlation analysis. We implemented both AAM-CCA and AAM in matlab and compared them with our methods.

We compared GAAM with these two methods and its results are given in table 1 to table 5. As seen in table 3, 4, two training sets are used. One is small set, which includes 400 images, and other one is big training set with 1000 images. In table 3, we used 400 images in training set. GAAM fitting speed was 55ms. AAM gave 88ms with worse error rate. AAM-CCA performed the fitting faster than GAAM and AAM but the error rate was worse than GAAM and better than AAM.

The results in table 4 show that the fitting speed becomes slower but the error rate significantly improves for all methods. GAAM gives 105ms with better convergence time while AAM gives 242ms and AAM-CCA gives 176ms. In table 3, it was 55ms in GAAM, 88ms in AAM and 41ms in AAM-CCA. Although fitting speed by using big training set as in table 4 is the slowest one compared with table 3, the error rate is the lowest.

TABLE 5 Comparison of GAAM with the existing methods

Faces	GAAM	3D-SFS [55]	3D Morphable model [47]
Training samples	1000	200	1200
Number of annotation	57	57	57
Max. Angle(possibility)	67.5	50	90
Speed	105ms	250ms	270ms
Error rate	0.7%	7%	2.3%

We also compared the GAAM with the 3D face reconstruction methodology by analyzing the shape-from-shading (SFS). The basic idea of SFS is to infer the 3D surface of object from the shading information. In addition, we compared our method by using 3D morphable model,

which is one of the state-of-the-art methodologies. Unlike 3D-SFS, there is no restriction with respect to illumination models or reflectivity functions in 3D morphable models. In 3D-SFS, the training images are just two. However, for each angle, one training sample must be prepared. The 3D-SFS methods extract the reflection model from the face surface. This is very time consuming process. In general, 3D reconstruction takes 250ms. GAAM spends only 55ms to generate the rotated faces from one single face. Furthermore, GAAM gives 3.2% error rate but 3D-SFS gives 12%. It is because 3D-SFS perfectly relies on features extracted under reflection. Reflection model computation is an ill-posed problem due to the complexity of the light. To prevent it, there are strict restrictions for the input face. However, GAAM does not have strict restrictions. On the other hand, 3D morphable model outperformed GAAM and 3D-SFS by providing the lowest fitting error rate. Most of errors happen due to the illumination issue. Error is computed by the deviation from the pixel values of frontal face.

C. Synthesized face results

The warped faces after the synthesizing are given below. After the face warping, the face geometry is kept proportional and only face angle changes. Head pose is the same for the all people as seen in below sample images. The warped part is drawn on the original face image. The boundary of the synthesized face seems like a distortion. However, it is especially kept and it is not a distortion. Here, perfect warping is done by keeping the face feature proportion for all face features. Keeping the face portion is very important for face recognition. Therefore, the quality of image warping is critical for keeping the recognition error as low as possible.

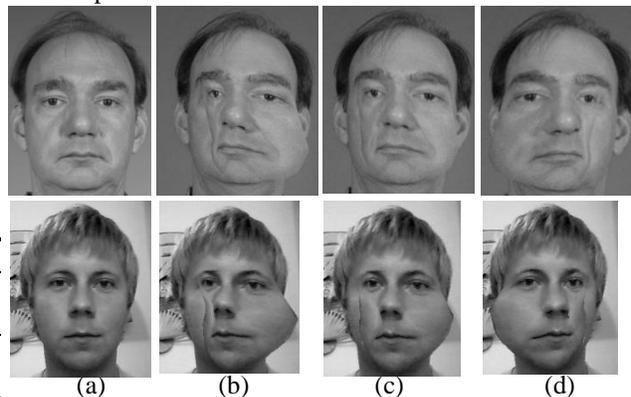


Fig. 4. Face reconstruction in actual images.

(a) original image, (b) 30 degree left-rotated, (c) 15 degree left-rotated, (d) 5 degree right rotated

As seen in fig. 6, GAAM is applied to the original images and the synthesized face is applied to the original face. Faces are natural and the faces after iteration are not distorted. More angles in different view aspects can be added and



warped to the original image. We only showed left and small right rotation faces. Especially Fig. 6(b) is rotated as much as 45degree.

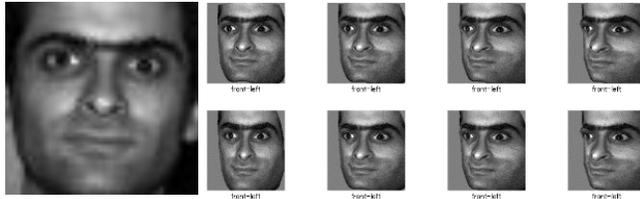


Fig. 5. Synthesized faces

In fig. 7, different face synthesized images are shown. Fig. 7(a) shows the left rotated faces, (b) shows the right and left-down rotated faces and (c) shows the left and left-down-up faces. We synthesized different illumination faces to see the rotation performance under reflection and shadow faces. In fig. 7(a), the training set trained by using left angle is used. In a similar way, the training set trained by using right angle is used.

In fig. 7(b), The shadows in fig. 7(b) are also kept unchanged after the synthesizing, Rotation to left-up, left-down, right-up, right down gave some deformations as it can be observed in fig. 7. These are due to the training database. We did not use such pose faces in training database. We forced the software to rotate the face to these orientations to see the deformations. We confirmed that the faces are generated in different poses successfully without much deformation. Gabor filters are round shaped and they are tolerable for different face poses. They can extract the local features without being affected by small face poses. This good feature of Gabor filters is used during training as well as testing steps. During the training step, there is no need to include all the different angles of faces. For example, face poses until 30degree left poses allow us to generate face posed images up to +15 degree. This degree varies with the Gabor scale and orientation. Our five scales and eight orientation gives up to +15 degree. We have not experienced the different Gabor sizes during our tests and it remains as a task for future direction of the research. Deformation rate of Face rotation is measured by comparing the deviation from the original face images on the same angle value. For example, we would like to rotate an image by 45degree. We rotate the face image and compare the resulting image with a rendered face images with 45 degree angle.

V. CONCLUSION

We introduced a face synthesis algorithm, which generates rotated faces from one single face image. To achieve this, we proposed Gabor based active appearance model (GAAM). GAAM extracts the features of input image and correlates them with the training efficiently and successfully. By using

its multiple training structures as well as the Gabor decomposition applied to the face image yields significant improvements of the active appearance model performance. After the computation of GAAM, both face mask fitting and face generation is done in very short time.

Empirical results show that GAAM is faster than original AAM but slower than AAM-CCA method. In addition to this, recognition performance using the images pre-processed by GAAM gives better performance. In our experiments, we used small dataset (400 faces) and big training set (1000 faces) to make sure of the performance. 1000 face training gives better performance with slow fitting speed. Furthermore, GAAM uses multi-cascade training to support pose variation. AAM and AAM-CCA use one single train set and large number of frontal images are needed to train. Extreme illumination as well as posed faces cause feature overlap. Instead of this, we use four different training sets, namely frontal, right, left, up-down faces. Therefore, the features between the training sets are very well separated which prevents GAAM from giving false fitting errors on different face poses. The structure is more compact and more descriptive. In addition, we used Gabor-tensor approach, which groups Gabor-tensor features separately in scale and orientation. Hence, it minimizes the feature overlap during model fitting. However, some of the rotated images are dramatically warped and distorted due to the illumination conditions of the face image before model is applied. These are the remaining works and improving the preprocessing before applying the GAAM mask will certainly prevent the distortion.

Evaluation results show that the GAAM outperforms state-of-the-art AAM techniques as well as 3D morphable models. Although we use the technique for face rotation, it can be extended to the rotation of other objects such as animals, general objects, face tracking, extraction of fine details of object features to be used for object recognition and medical imaging.

REFERENCES

- [1] A. Kumar and B.Yegnanaarayana, "Template Matching Approach for Pose Problem in Face Verification," Lecture notes in Computer Science of Springerlink, vol. 4105, pp. 191-198, 2006.
- [2] H.Satoshi, S.Hitoshi, and M.Naoki, "The Improvement for Illumination Problem on Face Recognition System," Joho Shori Gakkai Kenkyu Hokoku, vol. 2000, no. 9, pp. 25-30, 2000.
- [3] X. Zou, J. Kittler, and K. Messer, "Illumination Invariant Face Recognition: A Survey," IEEE 1st Conference on Biometrics: Theory, Applications and Systems, Washington, USA, 2007.
- [4] F. Xiea, L. Taoa, and G. Xu, "Estimating illumination parameters using spherical harmonics coefficients in frequency space," Tsinghua Science & Technology, vol. 12, no. 1, pp. 44-50, Feb. 2007.
- [5] J. Ruiz-del-Solar and J. Quinteros, "Illumination compensation and normalization in eigenspace-based face recognition: A comparative study of different pre-processing approaches," Pattern recognition letters, vol. 29, no. 14, pp. 1966-1979, Oct. 2008.
- [6] W.Y. Zhao, R. Chellappa, P.J. Phillips, and A. Rosenfeld, "A Literature Survey," ACM Computing Survey, pp. 399-458, Dec. 2003.

- [7] E. Chutorian and M. Trivedi, "Head Pose Estimation in Computer Vision: A Survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 4, pp. 607-626, Apr. 2009.
- [8] R. Chellappa and G. Aggarwal, "Pose and Illumination Issues in Face and Gait-Based Identification," *Advances in Biometrics: Sensors, Systems and Algorithms*, vol. 2, pp. 302-322, 2008.
- [9] R. Chellappa, C.L. Wilson, and S. Sirohey, "Human and Machine Recognition of Faces: A Survey," *Proceedings of the IEEE*, vol. 83, no. 5, pp. 705-740, May 1995.
- [10] L. Intra and H. Nissenbaum, "Facial Recognition Technology: A Survey of Policy and Implementation Issues," *Report of the Center for Catastrophe Preparedness and Response*, Apr. 2009.
- [11] T. Whangbo, J. Choi, M. Viswanathan, N. Kim, and Y. Yang, "Pose-Invariant Face Recognition Using Deformation Analysis," *Brain, Vision, and Artificial Intelligence*, vol. 3704, pp. 545-554, 2005.
- [12] P. Sankaran, S. Gundimada, R. C. Tompkins, and V. K. Asari, "Pose Angle Determination by Face, Eyes and Nose Localization," *FRGC, CVPR Workshop*, 2005.
- [13] Y. Hu, L. Chen, Y. Zhou, and H. Zhang, "Estimating Face Pose by Facial Asymmetry and Geometry," *Proceedings of the 6th IEEE International Conference on Automatic Face and Gesture Recognition*, Seoul, Korea, pp. 651-656, May 2004.
- [14] M. Breitenstein, D. Kuettel, T. Weise, and L. Gool, "Real-Time Face Pose Estimation from Single Range Images," *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, US, 2008.
- [15] J. Ruiz-del-Solar and P. Navarrete, "Eigenspace-Based Face Recognition: A Comparative Study of Different Approaches," *IEEE Transactions on Systems, Man and Cybernetics, Part C*, vol. 35, no. 3, pp. 315-325, Aug. 2005.
- [16] L. Wiskott, J.-M. Fellous, N. Kruger, and C.D. Von Malsburg, "Face Recognition by Elastic Bunch Graph Matching," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 775-779, Jul. 1997.
- [17] G. Edwards, C. Taylor, and T. Cootes, "Interpreting Face Images Using Active Appearance Models," *Proc. IEEE Int'l Conf. Automatic Face and Gesture Recognition*, pp. 300-305, 1998.
- [18] T. Cootes, C. Taylor, D. Cooper, and J. Graham, "Active Shape Models - Their Training and Application," *Computer Vision and Image Understanding*, vol. 61, no. 1, pp. 38-59, 1995.
- [19] T. Cootes, G. Edwards, and C. Taylor, "Active Appearance Models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 6, pp. 681-685, 2001.
- [20] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: active contour models," *International Conference on Computer Vision*, pp. 259-268, 1987.
- [21] K. Messer, J. Matas, J. Kittler, J. Luettin, and G. Maitre, "Xm2vtsdb: the extended m2vts database," *Audio- and Video-based Biometric Person Authentication*, pp. 72-77, 1999.
- [22] I.M. Scott, T.F. Cootes, and C.J. Taylor, "Improving Appearance Model Matching Using Local Image Structure," *Proc. Information Processing in Medical Imaging*, pp. 258-269, 2003.
- [23] M. B. Stegmann, H. Olafsdottir, and H. Larsson, "Unsupervised Motion-Compensation of Multi-Slice Cardiac Perfusion MRI," *Medical Image Analysis*, vol. 9, no. 4, pp. 394-410, 2005.
- [24] H. Wu, X. Liu, and G. Doretto, "Face Alignment via Boosted Ranking Model," *IEEE conference in Computer Vision and pattern recognition*, pp. 1-8, 2008.
- [25] M. Zhou, L. Liang, J. Sun, and Y. Wang, "AAM Based Face Tracking with Temporal Matching and Face Segmentation," *Computer Vision and Pattern Recognition*, pp.701-708, Jun. 2010.
- [26] J. Sung, T. Kanade, and D. Kim, "Pose Robust Face Tracking by Combining Active Appearance Models and Cylinder Head Models," *International Journal of Computer Vision*, vol. 80, no. 2, pp. 260-274, Jan. 2008.
- [27] R. Segulier, S.Gallou, G. Breton, and C.Garcia, "Adaptive Active Appearance Models," *EURASIP Journal on Image and Video Processing*, vol. 2009, Oct. 2009.
- [28] R. Gross, I. Matthews, and S. Baker, "Constructing and Fitting Active Appearance Models with Occlusion," *Computer Vision and Pattern Recognition Workshop*, 2004.
- [29] R. Larsen, M. B. Stegmann, S. Darkner, S. Forchhammer, T. F. Cootes, and B. K. Ersboll, "Texture Enhanced Appearance Models," *Computer Vision and Image Understanding*, vol. 106, pp. 20-30, 2007.
- [30] R. Ptucha and A. Savakis, "Facial Pose Estimation Using A Symmetrical Feature Model," *Proceedings of the 2009 IEEE international conference on Multimedia and Expo*, pp. 1664-1667, 2009.
- [31] C. Jang and K. Jung, "Human Pose Estimation Using Active Shape Models," *World Academy of Science, Engineering and Technology*, vol. 44, 2008.
- [32] A. Lanitis, C. J. Taylor, and T. F. Cootes, "Automatic interpretation and coding of face images using flexible models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 743-756, 1997.
- [33] H. Lee and D. Kim, "Tensor-Based AAM with Continuous Variation Estimation: Application to Variation-Robust Face Recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 6, Jun. 2009.
- [34] M. B. Stegmann, B. K. Ersboll, and R. Larsen, "FAME - A Flexible Appearance Modeling Environment," *IEEE Transactions on Medical Imaging*, vol. 22, no. 10, pp. 1319-1331, 2003.
- [35] M. Gleicher, "Projective Registration with Difference Decomposition," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 1997.
- [36] G. Hager and P. Belhumeur, "Efficient Region Tracking with Parametric Models of Geometry and Illumination," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 10, pp. 1025-1039, Oct. 1998.
- [37] T. Vetter, "Learning Novel Views to A Single Face Image," *Proc. Second Int'l Conf. Automatic Face and Gesture Recognition*, pp. 22-27, 1996.
- [38] X. Hou, S. Li, H. Zhang, and Q. Cheng, "Direct Appearance Models," *Proc. computer Vision and Pattern Recognition Conf.*, vol. 1, pp. 828-833, 2001.
- [39] S. Z. Li, Y. S. Cheng, H. J. Zhang, and Q. S. Cheng, "Multi-view face alignment using direct appearance models," in *Proc. IEEE Conf. Automatic Face and Gesture Recognition*, pp. 309-314, 2002.
- [40] S. Yan, C. Liu, S. Z. Li, H. Zhang, H. Y. Shum, and Q. Cheng, "Face alignment using texture-constrained active shape models," *Image Vision Computation*, vol. 21, no. 1, pp. 69-75, Jan. 2003.
- [41] F. Jiao, S. Li, H-Y. Shum, and D. Schuurmans, "Face Alignment Using Statistical Models and Wavelet Features," vol. 1, pp.321, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2003.
- [42] M. Yu and S. Li, "Face alignment based on statistical models and Gabor wavelets," *International Journal of Robotics and Automation*, vol. 23, no. 1, Jan. 2008.
- [43] Y. Su and D. Tao, "Texture representation in AAM using Gabor wavelet and local binary patterns," *IEEE international conference on Systems, Man and Cybernetics*, 2009.
- [44] V. Blanz and T. Vetter, "Face recognition based on fitting a 3d morphable model," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 25, no. 9, pp. 1063-1074, 2003.
- [45] V. Blanz, K. Scherbaum, and H. -P. Seidel, "Fitting a Morphable Model to 3D Scans of faces," *Proc. of Int. Conf. on Computer Vision*, 2007.
- [46] J. D. Bustard and M.S. Nixon, "3D morphable model construction for robust ear and face recognition," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 2582-2589, Jun. 2010.
- [47] X.Wang, W. Liang, and L. Zhang, "Morphable Face Reconstruction with Multiple Views," *International Conference on Intelligent Human-Machine Systems and Cybernetics*, pp. 250-253, Aug. 2010



- [48] X. Tan and B. Triggs, "Enhanced local texture sets for face recognition under difficult lighting conditions," *IEEE Transactions on Image Processing*, vol. 19, no. 6, pp. 1635–1650, 2010.
- [49] V. Blanz and T. Vetter, "A morphable model for the synthesis of 3D faces," *SIGGRAPH*, Mar. 1999.
- [50] R.Gross, I. Matthews, J.F.Cohn, T. Kanade, S.Baker, "Multi-PIE," *Proceedings of the Eighth IEEE International Conference on Automatic Face and Gesture Recognition*, 2008.
- [51] R. Donner, M. Reiter, G. Langs, P. Peloschek, and H. Bischof, "Fast Active Appearance Model Search Using Canonical Correlation Analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 10, pp. 1690-1694, Oct. 2006.
- [52] T. Sim, S. Baker, and M. Bsat, "The CMU pose, illumination, and expression (PIE) database," in *Proc. IEEE Conf. Automatic Face and Gesture Recognition*, pp. 53-58, 2002.
- [53] A. S. Georghiadis, P. N. Belhumeur, and D. J. Kriegman, "From few to many: Illumination cone models for face recognition under variable lighting and pose," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 23, no. 6, pp. 643–660, Jun. 2001.
- [54] A. Asthana, T. K. Marks, M. J. Jones, K. H. Tieu and R. MV, "Fully Automatic Pose-Invariant Face Recognition via 3D Pose Normalization," *IEEE International Conference on Computer Vision*, Nov. 2011.
- [55] A.S.Georghiadis, P.N.Belhumeur, and D.J.Kriegman, "Illumination-based image synthesis: creating novel images of human faces under differing pose and lighting," *Multi-View Modeling and Differing of Visual Scenes, MVVIEW '99 Proceedings on IEEE Workshop*, pp. 47-54, 1999.
- [56] R. Lienhart and J. Maydt, "An extended set of haar-like features for rapid object detection," *Proc. of ICIP*, 1:900–903, 2002.