

Facial Parts-Based Face Hallucination Method

Kaori Kataoka¹, Shingo Ando², Akira Suzuki¹, and Hideki Koike¹

¹ NTT Cyber Space Laboratories, NTT Corporation
1-1 Hikarinooka Yokosuka-Shi Kanagawa 239-0847, Japan

² Research and Development Center, Nippon Telegraph and Telephone
West Corporation
6-2-82 Shimaya Konohana-ku Osaka 554-0024, Japan
kataoka.kaori@lab.ntt.co.jp

Abstract. Face hallucination produces high-resolution facial images from low-resolution inputs. In this paper, we propose a facial-parts-based face hallucination method. Since our goal is face recognition rather than face reconstruction, the contour information of facial-parts (such as eyes) is important. This method reconstructs facial parts as entities instead of dividing them into small blocks. We obtain the contours of facial parts by using the Active Appearance Model (AAM), and transform training images based on contours. We confirm that the proposed method significantly enhances face recognition performance.

Keywords: face hallucination, contour of facial-parts, Active Appearance Model.

1 Introduction

Surveillance cameras are becoming more and more common in places such as banks, stores, and parking lots. In most cases, however, the face sizes are very small because the cameras are configured to capture very wide views. The low resolution of these face images is a primary obstacle to effective face identification and recognition. To gain the detailed facial features needed for recognition, it is necessary to process the low-resolution image captured so as to generate a high-resolution face image. This technique, called face hallucination, was first proposed by Baker and Kanade[1].

A number of learning-based image hallucination techniques that generate a high-resolution image from one low-resolution image have been proposed recently. They employ a training database consisting of pairs of high- and low-resolution image samples to output hallucinated high-resolution faces. These methods are based on either global models or the patch approach. Wang and Tang[2] developed an efficient scheme based on a global-model-based face hallucination method. They use PCA to express the input face image as a linear combination of low-resolution training face images, and the final high-resolution image is synthesized by replacing the low-resolution training face images with their high-resolution counterparts; the same combination weights are used. Unfortunately, their technique ignores local details to focus on global information, so the resulting images are unclear and lack detailed features.

Chang et al.[3] apply the manifold learning method by assuming that small equivalent patches in the low- and high-resolution images form manifolds with similar local geometry. This method requires fewer training examples than other learning-based hallucination methods because the generation of a high-resolution image patch depends on multiple nearest neighbors. Their approach is similar to LLE for manifold learning and so does not depend on just one of the nearest neighbors in the training set. They represent each low- or high-resolution image as a set of small overlapping image patches. Liu et al.[4] proposed a two-step approach to face super-resolution. First, a global linear model learns a transfer function from low-resolution faces to high-resolution faces, then a patch-based non-parametric Markov network locally reconstructs high-frequency content. Li and Lin[5] have improved this approach by applying PCA to both low- and high-resolution images in the first step and also using a MAP framework instead of the MRF model for finding the optimal local face. Since patch-based methods divide the image into small blocks that are uniformly overlapped, and hallucination is performed on each block, the reconstructed high-resolution image exhibits significant blocking artifacts.

In this paper, we propose a new facial-parts-based method that preserves local details. This method reconstructs each facial part(such as eyes) as an entity instead of dividing the facial parts into small blocks. To obtain the contours of facial parts, we construct a Active Appearance Model (AAM)[6] for each facial part in the training images, and locate the best match AAM parameters for each facial part. Next, we transform the training images according to the contours. We confirm that our approach yields significant benefits in face recognition.

2 Active Appearance Model

Our method uses AAM to obtain the contours of facial parts, so we must first construct AAM from the training images. An AAM is a statistical model that can describe the variation of shape and appearance with lower dimensionality.

A statistical model of shape variation can be generated from a training set of labeled images. The labeled points on a single object describe the shape of that object. Each shape is represented by vector x which consists of the aligned coordinates of the labeled points. By applying principal component analysis (PCA) to the data, any example can then be approximated using

$$x = \bar{x} + P_s b_s \quad (1)$$

where \bar{x} is the mean shape, P_s is a set of orthogonal modes of variation and b_s is a set of shape parameters.

To build a statistical model of object appearance, grey level information g is sampled from the shape-normalized image by warping, by using a triangulation algorithm, each training image so that its labeled points match the mean shape. By applying PCA to the normalized data, we can obtain the following linear model:

$$g = \bar{g} + P_g b_g \quad (2)$$

where \bar{g} is the mean normalized grey-level vector, P_g is a set of orthogonal modes of variation and b_g is a set of grey-level parameters.

The shape and appearance of any example can be summarized by the vectors b_s and b_g . For each example we generate the concatenated vector

$$b = \begin{pmatrix} W_s b_s \\ b_g \end{pmatrix} = \begin{pmatrix} W_s P_s (x - \bar{x}) \\ P_g (g - \bar{g}) \end{pmatrix} \quad (3)$$

where W_s is a diagonal matrix of weights for each shape parameter, which handles the difference in units between the shape and grey models.

Since there may be correlations between the shape and grey-level variations, we again apply PCA to the data to obtain

$$b = Qc \quad (4)$$

where Q are the eigenvectors and c is a vector of appearance parameters that determines both the shape and grey-level of the model.

Next, we synthesize the object image by matching a statistical model of object shape and appearance to the input image. We minimize the difference between the input image and the one synthesized by AAM. Difference vector δI can be defined as:

$$\delta I = I - I_m \quad (5)$$

where I is the vector of grey-level values of the input image, and I_m is the vector of grey-level values for the current model parameters. To locate the best match between the model and the input image, we minimize the magnitude of the difference vector, $\Delta = |\delta I|^2$, by varying model parameters c . Then, we can obtain coordinates of contour by obtaining most appropriate parameter c .

3 Facial Parts-Based Face Hallucination Method

Our facial parts-based method is based on the assumption that AAM can extract the contours of facial parts. We describe our method below.

First, we construct an AAM for each facial part. The procedure is illustrated in Fig.1.

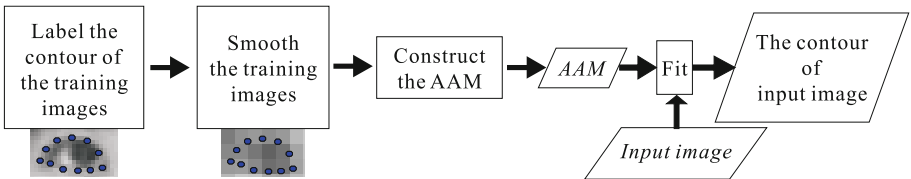


Fig. 1. The diagram of the facial parts contour extraction

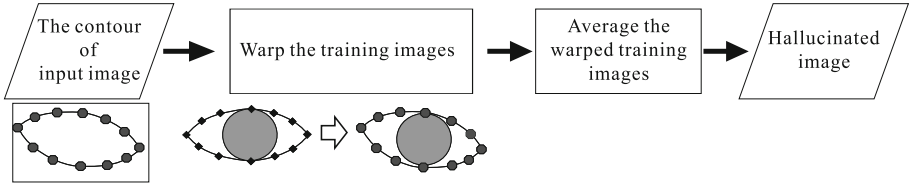


Fig. 2. The diagram of our synthesis method

1. Prepare training sets in which the contours of facial-parts are labeled.
2. Downsize the training images to the input image size and then scale them back up to the original size using the linear interpolation algorithm to yield blurred images similar to the input image.
3. Build an AAM of each facial-part such as eyes from the blurred training images and the contour coordinates.

Next, we locate the best match between the model and the input image to obtain the contour coordinates.

4. Scale up the input image(low resolution image) with linear interpolation algorithm.
5. Locate parameter c that minimizes the difference vector between the synthesized image and the scaled-up input image.
6. Obtain the contour of facial-parts from the best match parameter c .

Finally, we synthesize the high-resolution image using the obtained contour coordinates. The synthesis procedure is illustrated in Fig.2.

7. Warp the original high resolution training images of facial parts to fit the obtained contour and average the wrapped training images. This average image is output as the high-resolution facial parts image.

4 Experiments and Results

We decided to construct one AAM for each facial part to better preserve the diversity in their appearances. Since eyes are especially important in facial recognition[6], our experiments focused on them.

4.1 Building an AAM

4.1.1 Data Set

Our experiments utilized the Japanese Face Image Database HOIP[8], which contains 300 people. From the HOIP frontal facial database, we chose 103 people (three of whom were used as training data) whose eyes were not occluded by glasses, accessories, or hair. The high-resolution face images were scaled to 80×80 which yields the eye size of about 25×10 . The input images were created by blurring the face images (averaging neighboring pixels) and downsizing them low-resolution images with size of 20×20 (eye size is about 6×3).

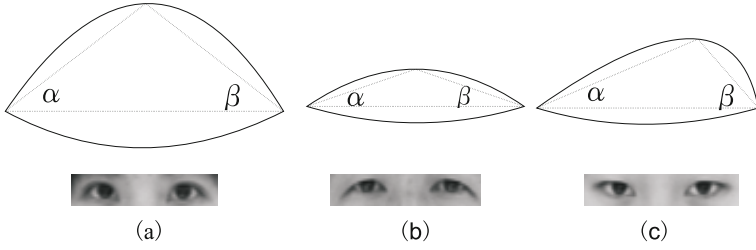


Fig. 3. Training eye images

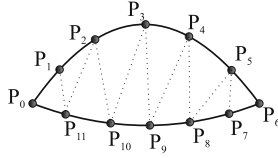


Fig. 4. Labelled eye image

4.1.2 A Set of Training Images

As the training data, we chose three eye images that exhibited different types of eye contour; this was done to challenge AAM with a wide variety of eye curve. In this experiment, the eye types reflect the angle of eye corners as shown as Fig.3. Fig.3(a) shows the curve type wherein angle α and β are almost the same and large. The angles in Fig.3(b) are small and similar. The angles in Fig.3(c) are quite different. We expect that these training data well represent various eye types such as drooping eyes, narrowed eyes, wide-open eyes, and slanting eyes.

4.1.3 A Set of Labeled Points on the Training Images

We manually labeled 12 landmark points on each eye and set triangular patches as shown as Fig.4.

4.2 Modified Method Using Residue Images

We tried to modify the obtained facial parts images by using the residue images as shown as Fig.5.

Suppose I_l is the low-resolution(LR) input image and I_h is the corresponding obtained high-resolution(HR) image, the low-resolution residue image R_l can be denoted as $R_l = D(I_h) - I_l$, where $D(\cdot)$ is the down-sampling function. Given bias is 128 when we display R_l , $D(I_h)$ is darker than I_l when pixel value of R_l is smaller than 128 and $D(I_h)$ is lighter than I_l when pixel value of R_l is larger than 128. Since contour line is darker than skin color, we expect to be able to correct the position of contour line based on R_l .

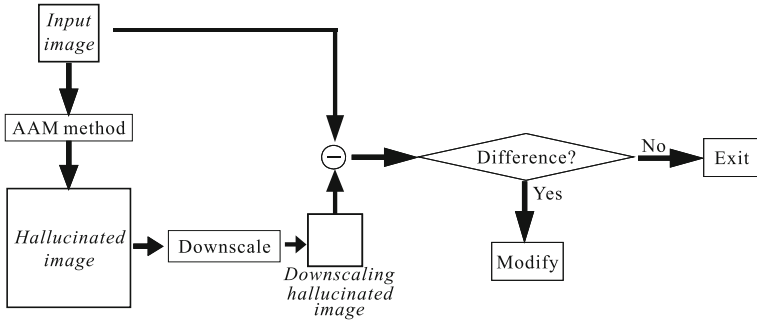


Fig. 5. The procedure of modification using the residue image

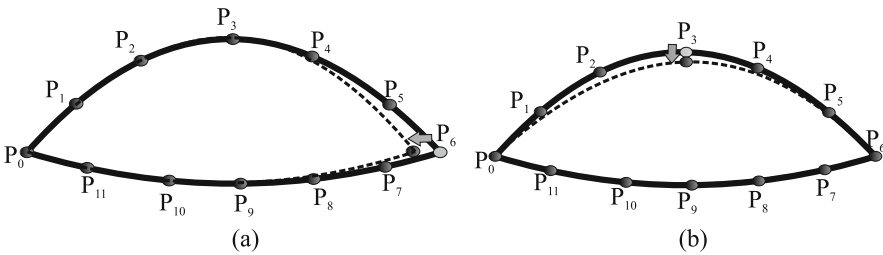


Fig. 6. Examples of the modified method based on residue image

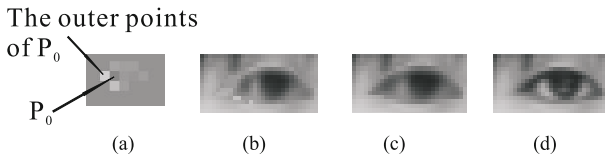


Fig. 7. An example of modified eye image. (a)Residue eye image, (b)Eye image before modification, (c)Eye image after modification, (d)Original HR eye image.

We describe the modified method below.

case1 outer corner of eyes:

If the pixel corresponding to point P_0 on R_l is darker than the threshold as is the pixel adjacent to point P_0 that lies outside and normal to the contour, the hallucinated eye is larger than the original, and thus P_0 is transformed inward. In association with this transformation, P_1 and P_{11} are transformed inward. We move P_0 4 pixels inward and P_1 and P_{11} 2 pixels inward. If the pixel adjacent to point P_0 that lies outside and normal to the contour is lighter than the threshold and point P_0 is darker, the hallucinated eye is smaller than original one, therefore P_0 is transformed outward.

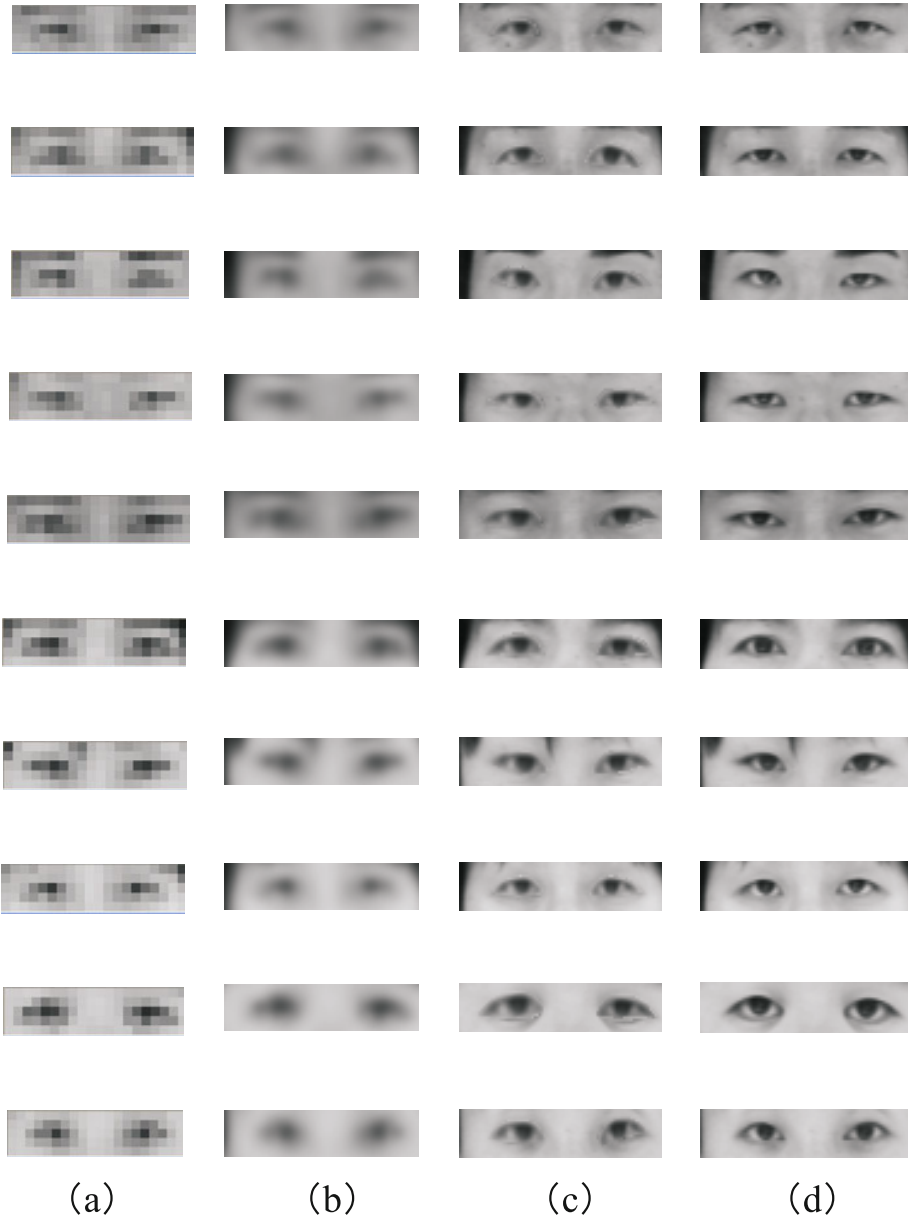


Fig. 8. Exsamples of hallucinated images. (a) Input LR images, (b) Cubic B-spline interpolation images, (c) Hallucinated HR images by our method, (d) Original HR images.

Table 1. Comparison of correlation value against original image and recognition performance

	<i>Cubic B – spline</i>	<i>Our method</i>
Correlation value	0.37	0.72
Recognition rate	6/100	49/100

case2 inner corner of eyes:

If point P_6 is darker than the threshold as is the pixel adjacent to of point P_6 that lies outside and normal to the contour, the hallucinated eye is larger than the original, and thus P_6 is transformed inward as shown in Fig.6(a). If the pixel adjacent to point P_6 that lies outside and normal to the contour is lighter than the threshold and P_6 is darker, the hallucinated eye is smaller than the original, and thus P_6 is transformed outward.

case3 upper curve:

If point P_3 is darker than the threshold as is the pixel adjacent to point P_3 that lies outside and normal to the contour, the hallucinated eye is larger than the original, and thus P_3 is transformed inward as shown in Fig.6(b). If the pixel adjacent to point P_3 that lies outside and normal to the contour is lighter and point P_3 is darker than the threshold, the hallucinated eye is smaller than the original, and thus P_3 is transformed outward.

case4 lower curve:

If point P_9 is darker than the threshold as is the pixel adjacent to point P_9 that lies outside and normal to the contour, the hallucinated eye is larger than the original, and thus P_9 is transformed inward. If the pixel adjacent to point P_9 that lies outside and normal to the contour is lighter than the threshold and point P_9 is darker, the hallucinated eye is smaller than the original one, and thus P_9 is transformed outward.

These rules were decided empirically, but the result of this modification shown in Fig.7((a) depicts a residue eye image, (b) depicts an eye image before modification, (c) depicts an eye image after modification, (d) depicts the original HR eye image), confirms that they are entirely valid for modifying hallucination images.

4.3 Face Recognition Experiments

We conducted a hallucination experiment on 100 eye regions. Some hallucination results are shown in Fig.8 in a comparison with the input LR faces images and the Cubic B-Spline interpolation results.

We compared the correlation of original HR eye region and hallucinated image, and interpolation results. Table 1 shows the average of 100 faces correlations. A perfect fit yields a coefficient value of 1.0. Thus the higher the correlation coefficient the better. We also compared the recognition rate of hallucinated images and interpolation results. Table 1 shows the recognition performances

for 100 individuals. The eye region edge images were created by a Sobel filter and estimation was based on normalized correlation values of the eye region.

5 Conclusion

In this paper, we proposed a new facial-parts-based hallucination method. Our method prevents facial-parts from being subdivided by obtaining the contour location of each facial-part by Active Appearance Model (AAM). Facial recognition experiments showed that the proposed method yields hallucination images that well approximate the original high resolution images, making our method effective for raising face recognition performance. In future work, we will conduct experiments on all facial-parts other than the eyes and synthesize them. We also will apply it to real-world images captured by surveillance systems because the current results were obtained by blurring and down-sampling original high resolution images.

References

1. Baker, S., Kanade, T.: Hallucinating faces. In: Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition, Grenoble, France, pp. 83–88 (2000)
2. Wang, X.G., Tang, X.O.: Hallucinating face by eigentransformation. *IEEE Trans. Syst. Man. Cybern.* 35(3), 425–434 (2005)
3. Chang, H., Yeung, D.Y., Xiong, Y.M.: Super-resolution through neighbor embedding. In: Proceedings of CVPR'04, Washington, DC, vol. 1, pp. 275–282 (2004)
4. Liu, C., Shum, H., Freeman, W.T.: Face hallucination: Theory and practice. *Int. J. Comput. Vis.* 75(1), 115–134 (2007)
5. Li, Y., Lin, X.: An improved two-step approach to hallucinating faces. In: Proc. IEEE Conf. Image and Graphics, December 2004, pp. 298–301 (2004)
6. Cootes, T.F., Edwards, G.J., Taylor, C.J.: Active appearance models. In: Proc. of the 5th European Conference on Computer Vision, vol. 2, pp. 484–498. Springer, Heidelberg (1998)
7. Moriyama, T., Kanade, T., Xiao, J., Cohn, J.F.: Meticulously detailed eye region model and its application to analysis of facial images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 28(5), 738–752 (2006)
8. <http://www.softopia.or.jp/rd/facedb.html>