Improved Face Recognition With Expressions
By Warping To The Best Neutral Face

Chayanut Petpairote, and Suthep Madarasmi

Abstract—In face recognition, the face gallery database usually consists of expressionless frontal face images. However, the probe image is often a non-frontal facial image with expressions such as surprise, happiness, sorrow, anger, fear, disgust, and so on. Such expressions often cause the face recognition algorithm to fail. In this paper, we present a novel method to improve face recognition by creating a neutral face from a given expression face based on the best face reference found. We first use a variation of the classical principal component analysis (PCA) and local binary pattern (LBP) algorithms for classification to get the best neutral face reference in the database. We then employ a modified version of the thin plate splines warping to convert the given expression face into its corresponding neutral face reference. This effectively removes the expression from our probe image and helps improve the correctness in face recognition. We evaluate our proposed method using the classical face recognition algorithms including appearance-based PCA and feature-based LBP with two well-known expression databases; namely, the AR-Face and MUG-FED. The experimental results show that our proposed method significantly improves the accuracy of face recognition under expression variations for both the probe images and their corresponding neutral face reference. This effectively removes the expression from our probe image and helps improve the correctness in face recognition.

Keywords—Face Recognition, Local Binary Pattern, Principal Component Analysis, Facial Expression, Neutral Face.

I. INTRODUCTION

Face recognition is still a challenge when faced with images containing expressions, head pose variations, and partial occlusion. Face recognition happens to be the most common, natural method for biometric authentication, police investigation, surveillance, and etc. Two commonly used approaches to face recognition include appearance-based and feature-based methods. The most popular approach for appearance-based face recognition is the principal component analysis (PCA) [4-6]. A commonly used feature-based approach is the local binary pattern (LBP) algorithm which partitions the image into blocks and uses histogram features around the center pixel of each image block. LBP is often also used for the identification of facial expressions [7]. A face recognition algorithm usually consists of the following components in this order: face detection, face alignment [14], feature extraction, and face identification.

An important challenge in face recognition is when the probe face has expressions such as surprise, happiness, sorrow, anger, fear, disgust, and so on. Expressions is the cause of failure for many face recognition algorithms, since the face database or gallery generally consists of neutral (expressionless, frontal) face images such as those obtained from ID cards, driver’s licenses, passports, or mug shots. Many researches have tried to improve face recognition for images with expression by removing the facial expression to obtain a neutral face, similar to the sample images in a particular database of faces. One such example is the work in converting a smiling face to a neutral face to improve face recognition by Ramachandran et al. [8]. They employed a generic 3D Candide model [17] to register onto the smiling image to then change the appearance of the smiling face using a piecewise affine warp to move each triangular mesh into a neutral state. The warped neutral face looks realistic without any wrinkles, although the lip position is often incorrect relative to the corresponding neutral face, since the triangles somewhat changed the configuration of the face. The computation needed to generate the mesh models with different configurations of the lip is rather expensive. The tested the images with expressions using PCA plus LDA face recognition algorithm on the FERET face database to report an accuracy improvement of about 6%; i.e., from 67.5% accuracy to the improved 73.8% accuracy. Later, Lee et al. [10] presented a method to transform an expression face image into its corresponding neutral facial. Here, the face image was transformed via direct and indirect processes after using the active appearance model (AAM) to extract facial features. Facial expression was also matched and then the closest found neutral face to the probe face was used. They then transformed the feature vector expression image to that corresponding neutral image. Their result improved face recognition for the PF07 database using the LDA recognition algorithm from 76.67% accuracy for no expression transformation to 74.17%, 80%, and 95% accuracy for the direct transformation, the direct transformation using a bilinear model, and the indirect transformation, respectively.

Another approach that uses optical flow to transform an expression face image into a neutral face image was proposed by Hsieh et al. [12, 9]. Their work [12] used matching flow field vectors between the expression face and the corresponding neutral face to warp the expression. They use the radial basic function (RBF) interpolation on the motion vector at the facial feature points to map to a new synthesized set of neutral coordinates. They then applied optical flow to
warp all neutral faces [9] in the gallery to the expression probe image by using a lip masked synthesized image to remove artifacts created by the warping. Their results improved expression-invariant face recognition but the optical flow computation have problems with illumination. Besides, the computation time was rather expensive in the warping process, especially in the later work [9] where all neutral faces in the gallery needed to be warped. They evaluated face recognition improvement with the BU-3DFE database using the commonly employed PCA and LDA recognition methods. Their result improved from 56.88% (no warping) to 58.19% (after warping) for PCA and from 83.39% to 87.69% for LDA. Their later work [11] proposed an AdaBoost.M1 boosting framework to convert different expression faces into their corresponding pseudo-neutral face for the preprocessing stage by Tasai et al. Their experimental results improved face recognition for PCA and LDA algorithms on the JAFFE face database but their synthesized pseudo neutral image did not look like a realistic face.

Recently our work [16] presented a method to improve face recognition for images with expressions. In that approach, faces expressions are converted into neutral face prior to the face recognition process where the face gallery primarily consisted of neutral faces. We modified the thin plate splines warping [13, 14] approach to warp expressions into a neutral face. In that work we first normalized image positions by using the line between the center of the eyes as a preprocessing step. We then warped the probe expression image to a synthetic, neutral face landmark using our modified thin plate spline method. The result showed significant improvement in face recognition for both the AR-Face and the MUG-FED face databases, measured by PCA, LDA and LBP recognition algorithms. However, in that work the synthetic neutral face did not have the same shape as the probe’s real neutral face since we warped all expression faces to a single, standard, synthetic neutral face reference. In addition, we needed to warp the entire neutral face databases to our standard synthetic neutral face reference. This warped neutral face was used as the face gallery for recognition. Although done only once prior to the recognition process, this gallery preprocessing step was somewhat computationally intensive.

In this paper, we extend our previous work [16] by presenting a novel method to create a neutral face from a given expression face using the best neutral face reference found. The objective is to achieve face recognition improvement for images with expressions. We modified the PCA and LBP algorithms to find the best neutral face reference of each probe face as explained in section II below. The appearance-based PCA method is used to find Rank-1 to Rank-10 (i.e., best 10 matches) faces recognized in the neutral face gallery using the probe image containing expressions. Using these 10 best matched neutral faces, the feature-based LBP method is used to find the best neutral face reference. We then used the same approach as in our previous work called the double Thin Plate Splines Warping [16] (explained below in section III) to convert a probe expression face to that best neutral face reference obtained from the LBP step. This work does not need to transform the entire neutral face databases to the single neutral face reference as in our previous approach, reducing the total computation time. We then show that our proposed method leads to improved recognition results, while the warped expression face also looks more realistic than in our own previous work. Our proposed method is summarized in Fig. 1 that show an overview in training and matching step to create a synthesized neutral face for face recognition improvement.

![Fig. 1 An overview of our proposed method in obtaining a synthesized neutral face from a facial image with expression. This synthesized neutral face is then used in the recognition algorithm such as PCA or LBP to get improved results when compared to using an image with no expression removal.](image)

**II. FINDING THE BEST NEUTRAL FACE**

Here we propose a method to find the best neutral face reference by modifying the PCA and LBP algorithms. The PCA method is an appearance-based method that finds the best match for the probe image among the face database. In this work, we focus on improving face recognition under expression variations so the feature-based LBP approach is helpful in matching details for each feature.

**A. Rank-1 to Rank-10 Face Recognition Using PCA**

We first compute PCA [4] for the entire gallery to find the significant components in the face image subspaces and then reduce the data dimensionality of PCA to keep only 98% of the information content (via Eigenvalues). We then simply use the nearest neighbor (NN) classifier to obtain the top 10 matches. The chosen Rank-1 to Rank-10 face images for PCA face recognition is then used to match features in more detail for the next step.

**B. The Best Neutral Face Reference Matching Using LBP**

The Local Binary Pattern (LBP) method is an efficient feature-based classification method [7] based on texture analysis. The LBP method provides histogram features based on the center pixel of each image block where the classification is performed by computing simple histogram similarity measures. The LBP method is summarized in Fig. 2a-d. The LBP method based on facial feature-based histogram description is outlined as follows:
1) Face image for both the gallery and probe databases
2) Divide the face image into blocks
3) To compute the LBP histogram of each image block, first each pixel’s value is compared to its 8 nearest neighbors where each of the 8 neighbor’s values are set to 1 if it is greater than center pixel and 0 otherwise as shown in Fig. 2e. Each pixel’s 8 neighbor binary value is converted in to a decimal representation as shown in Fig 2e. The histogram based on this LBP decimal number is then computed for each block as shown in Fig. 2c.
4) Concatenate the histogram feature descriptors to form an appearance description of the face image.

![Fig. 2 Each step of the LBP approach. (a) Face image (b) Face image divided into blocks (c) Each block has a LBP histogram (d) Concatenated feature histogram. Part (e) shows how each pixel’s LBP is obtained and converted to a decimal representation used for histogram computation.](image)

![Fig. 3 The process to find the best neutral face reference among the Rank-1 to Rank-10 neutral faces obtained by PCA by finding the most similar LBP feature histogram features.](image)

III. USING MODIFIED EXPRESSION WARPING METHOD

After we have the best neutral face reference for a given expression face image, we use the same approach as in our previous work called Double Thin Plate Splines warping [16] to create a synthesized neutral face by converting a given expression face to its best neutral face landmark. We use the AR-Face [1] and MUG-FED [3] databases which contain 80 and 130 [2] facial point landmarks, respectively. These feature points are point constraints to control the conversion step. A preprocessing step is first performed by comparing the line segment between the center of the 2 eyes for both the probe expression face and the neutral face to make sure both images have the same vertical alignment, horizontal alignment, scale, and rotation. We then use linear interpolation to fill more refined set of landmark points. In this case, we fill 4 additional interpolated points, resulting in 604 landmarks from the 130 landmarks in the AR-Face database and 365 landmarks from 80 landmarks for the MUG-FED database. Fig. 4 shows an example of the original facial landmark points and the increased landmark points through interpolation.

![Fig. 4 Original facial landmark points and (b) Added landmarks through interpolation [16].](image)

A. Double Thin Plate Splines Warping (Double -TPSW)

We then use double thin spline warping (double-TPSW) to transform a facial expression image to the synthesized neutral image by mapping all expression landmarks to its corresponding best neutral face landmarks. For the double-TPSW, this method used the original formulation of the Thin Plate Spline model that proposed by Bookstein [18] in warping computation for one part of upper lip portion and other part of lower lip portion. Double-TPSW method estimates the new coordinates that minimizes energy between expression points and the neutral reference points given by the energy function in (4) that includes the smoothing term shown in (1) below.

\[
I_f = \iint_x \left( \frac{\partial^2 f}{\partial x^2} \right)^2 + 2 \left( \frac{\partial^2 f}{\partial x \partial y} \right)^2 + \left( \frac{\partial^2 f}{\partial y^2} \right)^2 \, dx \, dy
\]

Given a set of coordinate probe landmarks \( P \), we define \( P = \{(x_1,y_1), (x_2,y_2), ..., (x_n,y_n)\} \) as a \( n \times n \) matrix where \( n \) is the number of landmarks then let the \( n \times n \) matrix \( K \) be defined by the kernel function \( U = r^2 \log(r^2) \) where \( r_p = |P_i - P_j| \) is the landmark distance.
\[
K = \begin{bmatrix}
0 & U(r_x) & \cdots & U(r_x) \\
U(r_u) & 0 & \cdots & U(r_u) \\
\vdots & \vdots & \ddots & \vdots \\
U(r_u) & U(r_u) & \cdots & 0
\end{bmatrix}
\] (2)

The matrices \(K\), \(P\), \(P^T\), and \(Z\) are combined to be matrix \(L\) where \(P^T\) is the matrix transpose operator and \(Z\) is a \(3 \times 3\) matrix of zeros. \(L\) is a \(n+3 \times n+3\) matrix. \(V = (v_1, \ldots, v_n)\) is defined as any \(n\)-vector and \(Y = (V000)^T\) is a column vector of length \(n+3\). Then \(W = (w_1, \ldots, w_n)\) is a vector of \(n\) weights and the coefficients \(a_i, a_x, a_y\) of affine transformation as shown in (3) is then be used to define \(f(x,y)\) by (4). The affine transformation is a composition of translation, rotation, dilation, and shear transformations.

\[
L^{-1}Y = (W | a_i a_x a_y)^T
\] (3)

\[
f(x,y) = a_i + a_x x + a_y y + \sum_{i=1}^n (w_i U \left[ P_i - (x,y) \right])
\] (4)

We estimate a smoothing TPSW to obtain the minimizing energy by function in (5), where the parameter \(\lambda\) is the smoothness constraint weight and \(I_i\) is the smoothing term given in (1).

\[
E_{\text{TPSW}} = \sum_{i=1}^n (v_i - f(x_i, y_i))^2 + \lambda I_i
\] (5)

After we obtain the new coordinates of the synthesized neutral points, we finally use bilinear interpolation \([15, 16]\) to fill the color for a new estimated coordinate set based on the color expression probe. The bilinear interpolation algorithm performs linear interpolation along the \(y\) direction and then again along the \(x\) direction to compute the color values.

![Fig. 5 Example of double-TPSW method including TPSW for upper lip landmark portion and other for lower lip landmark portion to estimate the new neutral landmarks.](image)

The double–TPSW will estimate a new coordinate set for the expression probe that will minimize distance with result in a closed mouth solution for a face with open mouth. A single TPSW computation cannot provide this solution. The double–TPSW involves 2-steps of computing the TPSW, once for the upper lip and once for the lower lips. We have 2 sets of \(P_1\) and \(P_2\) that are matrix of \((3 \times u)\) and matrix of \((3 \times l)\), respectively, where \(n = u+l\). The \(u\) landmarks are all coordinate landmarks for the upper lip area and \(l\) landmarks correspond to the lower lip area. We then compute all TPSW steps for the two landmark portions and combine the new coordinates of both the upper lip (6) and lower lip (7) areas to obtain a synthesized neutral face (8). Fig. 5 illustrates the double-TPSW method including TPSW for upper lip landmark portion and other for lower lip landmark portion to estimate the new neutral landmarks. We then compute the final smoothing of the double TPSW by (5) and interpolate a synthesized neutral face by the bilinear interpolation method.

\[
f(x_i, y_i) = a_i + a_x x + a_y y + \sum_{i=1}^n (w_i U \left[ P_i - (x,y) \right])
\] (6)

\[
f(x_i, y_i) = a_i + a_x x + a_y y + \sum_{i=1}^n (w_i U \left[ P_i - (x,y) \right])
\] (7)

\[
f_{\text{Double-TPSW}}(x,y) = f(x_i, y_i) + f(x_j, y_j)
\] (8)

IV. EXPERIMENTAL RESULTS

We show the experimental results of our proposed method to find the best neutral face reference using the modified PCA and the LBP explained earlier. We then employ the double-TPSW method to warp the expression image into the best neutral face reference to obtain a synthesized neutral face.

We use the AR-Face Database \([1]\). This database contains frontal views of 112 people (58 men and 54 women), each with different facial expressions including neutral, smile, anger, and scream. There are 4 images per person, totaling 448 images in the database.

We also test on the MUG-FED Database \([3]\). This database contains frontal views of 25 people (18 men and 7 women), each with different facial expressions including neutral, anger, disgust, fear, happiness, sadness, and surprise. There are 1-3 images for each expression types per person, totaling 401 images in the database.

Experimental setup to find the best neutral face: We find the best neutral face reference by choosing Rank-1 to Rank-10 PCA face images then use LBP method to compute feature histogram descriptors. In this case, we collect 98% PCA data then use the nearest neighbor (NN) to classify Rank-1 to Rank-10 face images. For the AR-Face and MUG-FED database, we use the best neutral face reference using the modified PCA to find the best neutral face reference using the modified PCA and LBP classification. Retaining only 98% of the PCA data for AR-Face and MUG-FED database yields 82.1% and 95.8% accuracy, respectively. We choose Rank-1 to Rank-10 PCA face images and then compare LBP histogram descriptors between each top 10 PCA face image and the probe expression image by dividing the face image into 20x20 block sizes based on the center pixel's value and then combine all feature histograms to measure similarities. Table I shows the best neutral face reference accuracy by PCA and LBP classification. Retaining only 98% of the PCA data for AR-Face and MUG-FED database yields 82.1% and 95.8% accuracy, respectively. We choose Rank-1 to Rank-10 PCA face images and then compare LBP histogram descriptors between each top 10 PCA face image and the probe expression image. Our proposed method to find the best neutral face reference significantly improves Rank-1 by 91.5% and 98% accuracy for AR-Face and MUG-FED database, respectively. We then use the best neutral face reference to be the reference landmarks in the later warping step to transform an expression face into its synthesized neutral face.

Experimental setup to warp the probe face to create a neutral Face: All face images are first normalized using the line segment between the centers of the 2 eyes position of a standard face. For the AR-Face database the size of 240 x 270 is used, whereas the MUG-FED database uses an image size
of 440x480 pixels. In the warping process our double-TPSW method is used to warp each probe image to the best neutral face reference landmark provided from the previous step. After warping a probe expression image to a neutral image, we evaluate the improvement in face recognition for the original expression image vs. our synthesized neutral image. We check for face recognition by searching a neutral face gallery database using 2 well-known algorithms: PCA and LBP with nearest neighbor (NN) classifier. For the PCA method, only 95% of the information content is maintained in the comparison. For the LBP method, we divide the face image into 30x30 for the AR-Face database and 40x40 block sizes for the MUG-FED database. The LBP method computes a histogram based on the center pixel’s value and then combines all feature histograms for use in the classification step.

### Table I
THE BEST NEUTRAL FACE REFERENCE ACCURACY [%].

<table>
<thead>
<tr>
<th>Database</th>
<th>Step 1: Rank-1 PCA (98% data)</th>
<th>Step 2: Rank 1 - Rank 10 PCA + LBP</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR-Face</td>
<td>82.1</td>
<td>91.5</td>
</tr>
<tr>
<td>MUG-FED</td>
<td>95.8</td>
<td>98.0</td>
</tr>
</tbody>
</table>

A. Face Recognition Performance
The PCA and LBP face recognition methods are used to evaluate the face recognition improvement. There are 3 sets of results for each method:
1. Raw – no image normalization
2. Normalized – image normalization using the line between center of the eyes.
3. Warped – image normalization plus warping the probe face to the best neutral face reference.

Fig. 7a-b shows the recognition rate by rank number for the AR-Face database and Fig. 7c-d for the MUG-FED database for PCA and LBP methods with rank 1 to 25, respectively. Note that Rank of N means that the correct probe image was found in the Nth best match. The performance of face recognition with a synthesized neutral face by use our proposed method (Type Warped) is consistently better than the raw and normalized cases. A comparison between the face recognition rate for both the AR-Face and MUG-FED databases for all the face recognition methods: PCA and LBP as shown in Fig. 6. The synthesized neutral face using our proposed method for all cases has the highest accuracy for both the databases. The face recognition rate using feature-based LBP method has a higher than using appearance-based PCA algorithm. Table II summarizes the improvement accuracy for our proposed method and also compare it to our previous work. To compare with our previous work [16], the result of our proposed method provided the synthesized neutral face better and look like a real neutral face in databases. Besides, this work does not need to warp the neutral face databases to only average neutral face reference for recognition process as our previous work does so the computational time of this work is cheaper than our previous work.

### Table II
SUMMARIZING THE IMPROVEMENT ACCURACY [%] TO COMPARE WITH OUR PREVIOUS WORK BY PCA AND LBP METHODS.

<table>
<thead>
<tr>
<th>Database</th>
<th>Input Type</th>
<th>Accuracy by Method (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>PCA + NN</td>
</tr>
<tr>
<td>AR-Face</td>
<td>Warped in [16]</td>
<td>83.0</td>
</tr>
<tr>
<td></td>
<td>Our Proposed Warped</td>
<td>88.2</td>
</tr>
<tr>
<td>MUG-FED</td>
<td>Warped in [16]</td>
<td>99.3</td>
</tr>
<tr>
<td></td>
<td>Our Proposed Warped</td>
<td>98.4</td>
</tr>
</tbody>
</table>

B. Example of Our Synthesized Neutral Face
Examples of our warped images to convert expression faces with different expressions to neutral faces are shown in Fig. 8a-c for the AR-Face database and Fig. 8d-f for the MUG-FED database. Our synthesized neutral image provides a great texture similarity to the neutral face images; thus, a significant improvement in accuracy results. To compare with our previous work [16], the result of our proposed method provided the synthesized neutral face better and look like a real neutral face in databases. Besides, this work does not need to warp the neutral face databases to only average neutral face reference for recognition process as our previous work does so the computational time of this work is cheaper than our previous work.

V. Conclusions
In this paper, we have shown our proposed method to create a neutral face based on the best neutral face reference. We first obtain the best neutral face via PCA and LBP. We then warp
the probe face image to the neutral image to get a synthetic neutral image used for face recognition from the gallery. Our proposed method can significantly improve face recognition for images with facial expressions. The precision of the synthesized neutral face from the warped image depends on the number of facial landmarks used. The computation time also depends on the number of landmarks used. However, this 2D approach to warping and recognition is computationally less expensive than the 3D approaches. The probe expression face is then warped to the best neutral face reference using the double-TPSW warping method. The experimental results for face recognition of the AR-Face and MUG-FED databases present a significant accuracy improvement by using our proposed method.

REFERENCES