Face Recognition Based on Improved Space Variant

Zhang Ting-man  
Xi’an Innovation College, Yanan University Shaan Xi, Xi’an, 710100, China  
Email: Zhangtingman1974@163.com

Wan Cho  
Chonbuk National University, South Korea  
Email: ghch@chonbuk.ac

Abstract—The fusion of imaging model and differential geometric is used to research the recognition problem of blurred-image in this paper. According to some assumptions, the established subspace results from the convolution of an image with some complete orthonormal basis functions with a predefined maximum size. Therefore, we demonstrate that the corresponding subspace created from a clear image and its blurred version is equal under the ideal case of zero noise and properties of blur kernels. Then, the paper studies the application of invariance in direct face recognition algorithm. We view the subspaces as points on the Grassmann manifold and adopt the subspace representation method to perform recognition of blurred image. In addition, we also provide the recognition rate of blurred image, where the blur variable is both homogeneity and space-variant. Finally, simulation experiment results show that the proposed algorithm can effectively improve the accuracy, and achieve a higher face recognition rate in comparison with the existing face recognition algorithms.

Index Terms—Blur Basis; Subspace; Face Recognition; Grassmann Manifold; Space Variant; Complete Orthonormal Basis

1. INTRODUCTION

The blur effect is usually caused by various reasons. For example, in various applications, images may contain blur, which can result from atmospheric turbulence, out-of-focus, or relative motion between the sensor and the target. The blur is inevitable and unavoidable during the formation of image. It is an important problem of image analysis applications to understand the influence of fuzziness, such as face recognition. Imaging process of blur image can be denoted as follows:

\[ y(n_1,n_2) = (y * k)(n_1,n_2) + \eta(n_1,n_2) \]  

(1)

where \((n_1,n_2)\) is a spatial coordinate position of two-dimensional convolution between the unknown point spread function (PSF) \(k_{(n_1,n_2)}\) and the clean image \(y_{(d_1,d_2)}\) with size of \(d_1 \times d_2\). The point spread function represents the blur, while other degradations are denoted by the noise term. \(\eta_{(d_1,d_2)}\) is additive noise, where the ubiquitous noise is an interference signal. The noises are caused mainly by the quantization error and the external interference. In the field of face recognition, the existing algorithms for solving the problem can be classified into two categories: inverse method based on deblurring and direct method based on affine statistical invariants. The aim of deblurring is to convert the blurred image \(y\) into the clean image \(\hat{y}\). However, even if the information of blur basis is completely mastered, the reverse problem of Equation (1) for attaining \(y\) is also an ill-posed problem, because the property of noise is unknown and has to be estimated. Nevertheless, under normal condition, it is not possible to completely understand the blur basis. During the past years, a lot of image restoration technologies have been appeared in the field of image processing, which contain some very important methods, such as blind deconvolution and non-blind deconvolution [1-2], regularization method based on total variation and the Tikhonov regularization method [3-4]. These methods have also been applied to the blurred face recognition. Effective methods that can tackle the ill-posed problems are the so-called regularization methods for the inverse process. Comparing with these methods, the direct method based on blur-invariability mainly searches some invariants in blurred image to realize the image deblurring. Nevertheless, the analysis of the blurred images can be performed without deblurring by using invariant features. This method could be applied to extract blur invariants so as to provide information for the following image processing task. Most of algorithms pursue to find a blur PSF of central symmetry related to the atmospheric effect. Its main idea can be described as follows: assume \(y_F, y_f, k_F\) denote the fourier transform of \(y, y, k\), respectively. In ideal conditions without noise, Equation (1) can be written as

\[ y_f(u,v) = y_f(u,v) k_F(u,v) \]

where \((u,v)\) is the frequency-domain coordinate. The phase relationship among these signals can be described as \(\angle y_f(u,v) = \angle y_f(u,v) + \angle k_F(u,v)\). Since the phase
among these symmetry signals and functions is either 0 or \( \pi \), so \( \tan(\angle y_f(u,v)) = \tan(\angle y_f(u,v)) \) is a blur invariant. According to the property, the invariant moment in space and frequency domain can be obtained. The linear movement invariant is proposed in [8]. And the algorithm has been extended by literature [9], which attains some invariants, such as rotation, translation, affine transformation etc.

The paper proposes an algorithm based on blur invariants, which is a direct method. Different from traditional algorithm, the adopted blur basis in this paper has arbitrary properties. The blur basis is spanned by linear combination of a set of orthogonal basis in allowable space. A novel blur invariant is proposed in this paper, and can solve more types of blur phenomenon. In addition, the paper views the subspaces of invariant as points in the Grassmann manifold, and proposes a new explanation for the space spanned by invariant based on modern differential geometry. Then, a novel algorithm producing from the new understanding method can be adopted to recognize the blurred face image. Compared with the existing methods, experimental results have shown the proposed algorithm has some advantages.

II. BLURRED-FACE RECOGNITION

The purpose of the section is to obtain the invariant of arbitrary \( k \) on image \( y \). Assumptions used in our proposed algorithm are as follows: without noise in the system; known maximum size of blur basis; BTTB matrix corresponding to the unknown blur PSF is full-rank.

As to two dimensional, the size of square-integrable and shift-invariant is defined as \( b_1 \times b_2 \), and its kernel \( k \) is denoted as follows:

\[
k = \sum_{i=1}^{N} \alpha_i \phi_i
\]  

where \( \{\phi_i\}_{i=1}^{N} \) is a complete orthogonal basis on \( R^{b_1 \times b_2} \), and \( \{\alpha_i\}_{i=1}^{N} \) is the corresponding combination coefficient. Therefore, under various noise conditions, the Equation (1) can be rewritten as follows

\[
y = y \ast \sum_{i=1}^{N} \alpha_i \phi_i
\]  

The specific form of \( k \) is related to \( \{\alpha_i\}_{i=1}^{N} \). At the same time we can construct a dictionary:

\[
D(y) = \left[ (y \ast \phi_1) \ (y \ast \phi_2) \ldots (y \ast \phi_N) \right]
\]  

where, the size of the dictionary is \( d \times N \), with \( d = d_1 \times d_2 \), \( d > N \), and \( \ast \) is denoted as victimization operator. The space spanned by the column of \( D(y) \) contains all the subspace of the convolution of image \( y \) and the arbitrarily function with the maximum size \( b_1 \times b_2 \). Note that all the blurred images only span a part of the subspace.

Conclusion 1. \( \text{span}(D(y)) \) is a blur invariant of \( y \), namely,

\[
\text{span}(D(y)) = \text{span}(D(y))
\]

The adopted method has an important advantage, namely, it is not necessary in this paper to do any constrain on the blur function’s shape, because the basis function can span any space of fuzzy function with the maximum known size. As to the assumption of rank \( K_y \), this paper needs to show that although the PSF of some blurs are irreversible, the BTTB matrices are usually full rank. These BTTB matrices may have very large condition number, but this paper doesn’t need to calculate the inverse of a matrix. Therefore, it will not face the question that the condition number is too large in the fuzzy method. This paper notices that noises are everywhere in the reality practice, which may make the assumption in this paper untenable. However, the following analysis has shown that these invariants have robustness on the noises.

This paper considers an image library containing \( M \) individuals, where \( \{y_{i1}\}_{i=1}^{8} \) denotes face database from all individuals. That is to say, it may be clean or fuzzy, \( y \) denotes the blurred training image of someone. The problem that should be solved in this paper is to find which individual the image of \( y \) belongs to under the situation that \( y_i \) have been given, and \( i \in \{1, 2, \ldots, M\} \). The paper first establishes its own dictionary \( D(y), D(y) \) according to the face database and the training image and then compares its spanning space \( Y_i, Y \) to realize the recognition.

A. Grassmann Manifold

Because the paper here adopts the linear subspaces whose dimension is \( N \) in \( R^d \), the recognition problem can come down to the recognition problem based on Grassmann manifold. Grassmann manifold \( G_{N,d} \) is a manifold corresponding to any subspaces whose dimension is \( N \) in \( R^d \). The blur invariant \( Y \) is a node in \( G_{N,d} \), which is shown in Figure 1. In order to understand the geometric characteristics of Grassmann manifold, literature [10] has been for intensive study and these geometric characteristics have been applied to the recognition problems [11] with the constraints from subspaces. Now this paper uses these conclusions to calculate the distances among blur invariants.

The first method adopts the distances among nodes and there is only one image for each person in the image library. Usually, the Riemann distance between two subspaces is the minimum distance between two nodes connecting Grassmann manifold. A method to obtain the distance is to calculate the direction matrix \( A \), so as to
make the measurement line from \( Y_1 \) to \( Y_2 \) can arrive in unit time. \( A \) can be mapped through the inverse index. However, the analytical expression mapped by Grassmann manifold’s inverse index doesn’t exist and it doesn’t adopt the number method shown in algorithm 1 of this paper. The length of \( A \) denotes the distance \( d_{\alpha} \) between \( Y_1 \) and \( Y_2 \). \( \text{trace}(A_A') \) is adopted as a metric to calculate the length. More commonly, if \( \phi_{n,x} \) denotes the directional matrix of \( C \) and \( Y \), then

\[
d_{\alpha}(Y_1, Y_2) = \text{trace}(A_{n,x}A_{n,x}' )
\]

(5)

Finally, this paper can adopt the nearest neighbor clustering method to realize recognition.

B. The Training Based on the Data \( G_{n,t} \)

If the image library contains more useful information about the image change, the statistical method based on the data \( G_{n,t} \) can be adopted. Because the dimension of the blur invariant is \((d-N)\times N \) and \( d \) is much larger than \( N \), then the training needs lots of samples to be conducted and various distribution of each kind of images have been obtained. Therefore, this paper adopts the method in literature [12] to conduct the linear discriminant analysis of the blur invariant’s kernel function based on projection kernel:

\[
k_p(\bar{D}(y_1), \bar{D}(y_2)) = \left\| \begin{bmatrix} 12,yy & P_{11} \\ 12,pp & P_{22} \end{bmatrix} \bar{D}(y_1) \bar{D}(y_2) \right\|_F = \text{trace} \left[ \begin{bmatrix} 12,yy & P_{11} \\ 12,pp & P_{22} \end{bmatrix} \bar{D}(y_1) \bar{D}(y_2) \right]
\]

(6)

The concrete implementation of the algorithm is shown in literature [12].

C. Blurred Face Recognition

When \( k \) doesn’t change in all pixels \( (n_i,n_k) \) for \( d_i \times d_j \) image \( y \), it can perform the nearest neighbor clustering method based on the previous Riemann distance \( d_{\alpha} \) and the Euclidean distance of low dimension space calculated through algorithm 2. So we can obtain the category of probe image \( y \) as follows:

\[
i^* = \arg \min_i SD(D(y_i), D(y_j))
\]

(7)

Now we will research the more complicated problem in this paper: the blur kernel function \( k \) has the feature of space variance. It primarily happens in the situation where various parts of the scene have been influenced by different fuzziness, for example, the focused fuzzy variable and the motion fuzzy variable. Under this situation, the image acquiring equation can be denoted as:

\[
y_i = y_n \ast k_n
\]

(8)

where \( n \) denotes the location of the pixel. Since the blur kernel happens in a relative small local space, it makes each pixel in the space changes and further makes the problem hardly be constrained. A common assumption is that the fuzzy variable is locally balanced. This is also quite effective in most situations. Based on this kind of assumption, the fuzzy variable is balanced in the range of \( d_i \times d_j (d_i > b_i, d_j > b_j) \), so this paper can divide the image into \( T \) non-overlapping graphics with the size \( d_i \times d_j \) to be recognized:

\[
i^* = \arg \min_i \sum_{i=1}^N SD(D(y_i), D(y_j))
\]

(9)

where \( t \) denotes the recognition of image block. The potential assumption of this method is that the images have been in alignment with each other. However, as to the excessive blocks among the blur kernels, the column vector space of \( D(\ast) \) is not the blur invariant. The ratio of this situation relates to the feature of fuzzy variable’s space-variance, which will be explained in the following experiment.

III. EXPERIMENTAL VERIFICATION

Two kinds of experiments in this paper are performed to verify the robustness of blur invariant on noise. The shown noise \( \eta \) in Equation (1) is divided into \( \eta = \eta_q + \eta_f \), where \( \eta_q \) denotes the quantization noise and other noises related to sensor and \( \eta_f \) denotes the facial variation caused by some factors, such as, illumination and expression, etc. The research purpose of this paper is to find out the \( \eta_q \)’s influence on the recognition when blur is the only source leading to the variation between image library and probe images, and also analyzes the effect based on the data training when there are other face variation \( \eta_f \) between image library and the probe images. In all experiments, the only parameter controlled by the users is the maximum size of the blur kernel function, which decides the number of the dictionary’s column. If \( (b_i',b_j') \) is \( b_i,b_j \) for all possible maximum in sum function, then the value of \( N \) should be \( b_i' \times b_j' \). At the same time, if \( N < d \) is true, the dictionary is often incomplete in the real world. Otherwise, the base will span the whole image space. Therefore, in the experiment of this paper, it is set to \( N = \lfloor (d_j / 2 \times d_i / 2) \rfloor \). The column of unit matrix \( I_N \) can be used to denote \( \phi_i^{(N)} \), though it can adopt the random complete orthogonal base.

A. The Role of Quantization Noise

The variation of recognition rate is researched at first when existing \( \eta_q \) and AWGN caused by sensors of various levels. The subspace distance of one person’s image is calculated in the image library under the condition of the uniform fuzziness, space variance fuzziness and the fuzzy image. The result in this paper is verified in two databases, CMU-PIE and YaleB [13-14].
It adopts the illumination face image library of PIE database which includes 21 different illumination conditions from each of 68 individual’s image and YaleB database which includes 64 different illumination conditions from each of 38 individual’s image. In this experiment, the image in library and probe images have the same illumination.

**B. Spatial Uniform Fuzziness**

This paper synthesizes the blurred image based on the following 4 principles: the motion blur, focused fuzziness, Gaussian fuzziness and the random fuzziness. The experiments are carried out in different blur kernel function sizes and illumination conditions. The experimental condition is: no noises, quantization noise, face variation noise and AWGN, respectively where the corresponding SNR is 50, 20, 10,5dB, respectively. The experiment is firstly conducted in the clear image and then performed in the blurred image, so as to obtain 12 different noise conditions. The recognition is realized by the comparison between the probe image and all the images in the image library. Therefore, apart from 12 different noise settings in the experiment, the image library and probe images separately include 68 images and 38 images, corresponding to PIE database and YaleB image library, respectively. The average value of several times experiments is shown in figure 1.

![Figure 1. Noise effect on recognition](image)

As we can see from Figure 1, the error between inside image and the outside image decreases with the increasing of noise. At the same time, the average error of inside image increases with the increasing of noise, which explains the reason why the recognition rate will decrease with the increasing of noise. Even in the condition without noise, whether for clear image or blurred image, the average error of correct matching will never be 0. Although the space is the same with the clear image in theory, the digitalization can lead to a certain noise. As for the noise conditions, the matching error of blurred image has the similar statistic characteristic with that of clear image. The primary reason is that the invariant contain the subspace of the image’s all blur type, so it will not be influenced whether the image is clear or fuzzy.

**C. Space-Variant Blur**

Next, this paper carries out the same verification with the above chapters, and the knowledge fuzzy variable has the feature of space-variant. The image blocks with arbitrary sizes are selected from images and one of the above four kernel functions is adopted to establish the synthetic blurred image. In order to recognize the face image, the paper selects the overlapped sub-images whose sizes are 75%, 50%, 40%, respectively. As to the above image blocks with different size, the size of the established dictionary is \( N = d_1 / 4 \times d_2 / 4 \). At the same time, we should notice that most of sizes which are applied to establishing the blurred image and function are much larger than \( d_1 / 4 \times d_2 / 4 \). This is mainly for retaining the generality of the innate character of the space variance fuzzy variable. Based on this kind of setting, the used distance parameter in recognition is \( d_G (5, 9) \) and the obtained result is shown in Figure 3.

According to the results in Figure 2, we can see: 1) Similar with the uniform blur variable, the recognition rate decreases with the increase of noise. As for the same noise setting, compared with the uniform blur variable, the recognition rate of the blurred images with feature of...
space-variant is lower and this is mainly caused by two factors: the size of blur kernel is usually larger than $d_i/4\times d_i/4$, which makes the space spanned by dictionary can’t contain blur kernel; there is transitional region among the blur kernels and $\text{span}(D(\bullet))$ is an invariant. 2) Suppose the fuzzy function’s equation is piecewise uniform, which its result is better than the result of fuzzy function with the original image, but there is not a fixed standard in the selection of block size.

C. Recognition Performance of Facial Change

In this chapter, we mainly research the recognition probability in the case of facial changes. The major objective of this paper is not to contain accurately the whole facial change, but to research the robust of invariant in the case of some other facial changes.

![Figure 3](image_url)

Figure 3. The comparison of different algorithms’ recognition performance adopting the Gaussian kernel to create blurred image.

In order to compare and assess the recognition performance between the proposed algorithm and other algorithms, we perform extensive experiments on two publically available face databases, namely, synthetic blurred images from FERET database and real blurred images from FRGC database. The experimental is set in accordance with literature [16]. We first use the images from FERET database. The database includes two subsets: ‘fa-images database’, and ‘fb-training database’. The database has 1001 images and there is a single image per person. The expression mode and alignment mode of the same person’s face image from fa and fb have a slight change. The size of original image is $128\times128$, but the scale of the image in practical application is resized as $64\times64$. Sharp faces are blurred by Gaussian kernel ($\sigma = 0\text{--}8$) with the size of $5\times5$, which get 9 different images from fb database. In addition, white Gaussian noise of 30dB is also added to the synthesized images. The recognition rate is shown in Figure 4. The nearest neighbor clustering method is also adopted to realize the image recognition, namely, no training, recognition with $d_c$ and recognition with training based on the kernel discriminated analysis. The result of experiment is shown in figure 5. As we can see, it is shown even if $\eta_f$ is not completely considered, the recognition rate will not increase with enhancing of the training images.

TABLE I. THE COMPARISON OF RECOGNITION PERFORMANCE

<table>
<thead>
<tr>
<th>Methods</th>
<th>Recognition rate (%)</th>
<th>blur type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Literature [16]</td>
<td>88.3</td>
<td>Gaussian</td>
</tr>
<tr>
<td>Hu and Haan [16]</td>
<td>82.5</td>
<td>Linear</td>
</tr>
<tr>
<td>Literature [21]</td>
<td>97.15</td>
<td>Both</td>
</tr>
<tr>
<td>Our algorithm</td>
<td>97.12</td>
<td>Both</td>
</tr>
</tbody>
</table>

Then, the recognition performance is assessed on FRGC database. The testing image library contains 608 images, and these images are acquired under non-uniform illumination conditions, where there are 306 blurred images. The recognition results are demonstrated in Table 2, where we compare the recognition results with and without illumination compensation. As can be seen, considering the illumination change, the recognition performance of the proposed algorithm is obviously improved.

TABLE II. RECOGNITION RATES OF DIFFERENT ALGORITHMS FOR REAL BLURRED IMAGES

<table>
<thead>
<tr>
<th>Methods</th>
<th>Recognition rate base on subset</th>
<th>Recognition rate based on the whole database</th>
</tr>
</thead>
<tbody>
<tr>
<td>Literature [12]</td>
<td>67.1</td>
<td>--</td>
</tr>
<tr>
<td>Literature [13]</td>
<td>73.5</td>
<td>--</td>
</tr>
<tr>
<td>Literature [3]</td>
<td>--</td>
<td>45.9</td>
</tr>
<tr>
<td>Literature [12]</td>
<td>--</td>
<td>74.5</td>
</tr>
<tr>
<td>(illumination</td>
<td></td>
<td></td>
</tr>
<tr>
<td>compensation)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Our algorithm</td>
<td>87.1</td>
<td>69.6</td>
</tr>
<tr>
<td>(illumination</td>
<td></td>
<td></td>
</tr>
<tr>
<td>compensation)</td>
<td>--</td>
<td>84.2</td>
</tr>
</tbody>
</table>

The paper considers the influences of $\eta_f$ on the recognition performance. It adopts UMD data here which includes the images of 17 individuals and is influenced by not only the blur with different degree but also the various illumination conditions, expression and alignment mode, etc. The adopted sample images in the experiment are shown in figure 4. The nearest neighbor clustering method is also adopted to realize the image recognition, namely, no training, recognition with $d_c$ and recognition with training based on the kernel discriminated analysis. The result of experiment is shown in figure 5. As we can see, it is shown even if $\eta_f$ is not completely considered, the recognition rate will not increase with enhancing of the training images.

Figure 4. Sample images
on recognition performance

The influences of $\eta_i$ on recognition performance

IV. CONCLUSIONS

Since the subspace results from the convolution of an image with some complete orthonormal basis functions, so it can be denoted as the blur kernel. Our research demonstrates the blur kernel is unchanged under certain assumptions and can also express many types of blur variables. Then, the paper studies the application of invariance in direct face recognition algorithm. We view the subspaces as points on the Grassmann manifold and adopt the subspace representation method to perform recognition of blurred image. In addition, we also provide the recognition case of blurred image, where the blur variable is both homogeneity and space-variant. Finally, we also analyze the role of noise and compare the recognition rate of different types of facial images with that of standard facial images.

REFERENCES


[20] Shotton, Jamie, et al. "Textonboost for image understanding: Multi-class object recognition and


