Blur-Robust Face Recognition via Transformation Learning

Jun Li, Chi Zhang, Jiani Hu, Weihong Deng

Beijing University of Posts and Telecommunications, Beijing, China

Abstract. This paper introduces a new method for recognizing faces degraded by blur using transformation learning on the image feature. The basic idea is to transform both the sharp images and blurred images to a same feature subspace by the method of multidimensional scaling. Different from the method of finding blur-invariant descriptors, our method learns the transformation which both preserves the manifold structure of the original shape images and, at the same time, enhances the class separability, resulting in a wide applications to various descriptors. Furthermore, we combine our method with subspace-based point spread function (PSF) estimation method to handle cases of unknown blur degree, by applying the feature transformation corresponding to the best matched PSF, where the transformation for each PSF is learned in the training stage. Experimental results on the FERET database show the proposed method achieve comparable performance against the state-of-the-art blur-invariant face recognition methods, such as LPQ and FADEIN.

1 Introduction

Face recognition plays an important role in the field of computer vision. Previous works on this topic [1][2][3][4] mainly focuses on recognizing faces under controlled imaging conditions. However, in practice, the performance of recognition algorithms tends to suffer from image degradation [5]. One example of degradation is blur as a consequence of out-of-focus lens, atmospheric turbulence, and relative motion between the camera and the objects. The process of such image blur could be modeled as:

$$g(n_1, n_2) = (I * H)(n_1, n_2) + n(n_1, n_2) \tag{1}$$

where (n_1, n_2) is the pixel location at which a convolution $*$ is performed between the original sharp image I and a point spread function(**PSF**) H, which has the same size with image I. n denotes the additive image noise coming from quantization, or other camera-induced errors.

The performances of face recognition could suffer from blur for the fact that blur leads to two main problems[6]:

(i) the facial appearance of an individual changes drastically due to blur. (ii) different individuals tend to appear more similar when blurred.

A few existing approaches have been proposed to handle these problems. All of them are based on the formulation of the process of blur,and could be roughly classified into three categories: (i)converting sharp images to match the blurred ones via blurring methods, (ii)converting blurred images to match sharp images via deblurring methods, (iii)finding blur-invariant descriptors.

One typical method of class (i) could be found in [7], where Stainvas & Intrator artificially blurred the sharp gallery images to match a blurred query image through a hybrid network architecture. Although problem(i) can be thoroughly handled, problem(ii) still remains. Besides, the gallery images may be blurred themselves.

The deblurring approaches are widely used to estimate I from the observed blurred image g, which means solving the inverse problem of formulation (1). But this is a difficult task considering unknown types of blur kernel. For instance, the PSF H of ideal motion and out-of-focus blur tends to be rectangular according to [8][9], while atmospheric and optical blurs are more likely to be Gaussian blur[10][9]. These approaches could also been treated as a way of image restoration. When the model for blur is known, known as non-blind deconvolution[11], Levin et.al [12] learn priors on clean image statistics, Fergus [13] use coded-computational photography to remove motion blur. Blind image deconvolution assumes nothing about the blur kernels is known[14]. Most of the existing works perform blind image deconvolution based on a single image[15]. A PSF is inferred through total variation regularization[16], the variation of Gaussian scale in the edges[17][18][19], or variation in the wavelet domain[20][21]. Recently, M. Nishiyama[6] has revealed that deblurring from a single image is an ill-posed problem and these deblurring methods are insufficient for accurate face recognition. Other methods deduce a PSF using multiple images[22][23], and M. Nishiyama et.al [6][24] propose to build multiple PSFs and use the best match as the final PSF. Even if the PSF could be correctly estimated, deblurring could not be very robust due to unknown noise n.

As regards blur-invariant descriptors, researchers have tried to find blur insensitive descriptors. Useful descriptors include local binary pattern 'LBP'[25][26], the subspace [27], manifold [28], sparse representation [29] and the local phase quantization 'LPQ'[30]. The 'LPQ' descriptor performs best when the blur is centrally symmetric[31], which is invariant with respect to blur effects. Recently, Gopalan et.al [32] proposed a new blur-robust descriptor using subspace techniques. These blur-invariant-descriptor relevant methods could also be viewed as direct methods.

Our Contributions: All of the approaches mentioned above could be viewed as preprocessing for feature extraction. We offer a new perspective that handling the effects of blur and extracting features at the same time via Transformation Learning. Based on the intuition that face recognition algorithms perform best when the gallery images and their corresponding blurred ones are in the same kind of feature spaces, we transform the features of sharp gallery images and blurred query images to a same space. We adopted Multidimensional Scaling (MDS)[33] to learn the desired transformation from training images us-

ing iterative majorization algorithm[34]. Extensive experiments on FERET and CMU-PIE datasets show that our MDS-based feature extraction could efficiently weaken the degradation caused by blur.

Outline of the Paper: The rest of the paper is organized as follows. In Section 2, we present our method for blur-robust face recognition. In Section 3, we describe the detailed experiments and show that our method not only outperforms others, but also demonstrates excellent capability of generalization. Finally, a brief summary and discussion is provided in Section 4.

2   Our Approaches

We try to find a subspace into which the sharp original images and their blurred counterparts are mapped in the same way. What's more, the performance of face recognition algorithms has been taken into consideration by combining classifiers.

We introduce Multidimensional Scaling (MDS) method to conduct transformation learning to solve the blur problem. It should be noticed that MDS assumes already knowing the blur degree, as the transformation is performed between the sharp images and those with certain blur degrees.

For an image with unknown blur degree, we first infer the PSF with the method in [6]. Then, we select the corresponding transformation matrix from the previously learned MDS matrices. Finally, classification methods are applied in the transformed space.
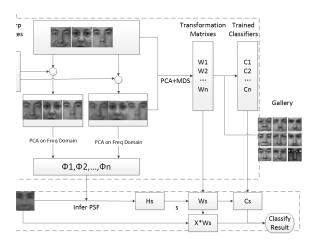
An overview of our method could be found in Fig 1.

Fig. 1. Overview of our method

Recently, Biswas et.al [35] also adopted MDS to solve various pose and resolution problem in face recognition, but we handle clearly different tasks in face recognition. Moreover, our work differs from [35] in the following two respects:

(i) In [35], an uniform W is trained for all kinds of variations in pose and resolution, while we differentiate different blur levels, and train a series of transformations $W_1 \ldots W_n$. This enables us to acquire the most accurate transformation for the blurred query image from the series of matrices (see Fig.1). For the blur problem, all kinds of blur have certain blur kernels [2], which means detailed PSF quantization is feasible and effective for solving the problem. (ii) Traditional multidimensional scaling is adopted in [35] to transform data from a high dimension to a low dimension. In contrast, we transform the data to a higher dimension, which is indeed a linear kernel trick.

2.1 PSF Inference

We adopt the methods proposed in [6] to infer the PSF. The basic idea is to find a subspace which is much more sensitive to the variance caused by blur than differences between individuals. The algorithm is shown as follows, which finally gives the subscript s of the inferred PSF:

Algorithm 1 PSF Inference

1: Extract feature image $x(\epsilon^{'}, \eta^{'})$ from a blurred image $g(u,v)$ as:

$$x(\epsilon^{'}, \eta^{'}) = [log(|g(\epsilon, \eta)|)] \downarrow \qquad (2)$$

where $g(\epsilon, \eta)$ is the Fourier transform of $g(u,v)$, $||$ takes the amplitude, and $[] \downarrow$ stands for down-sampling. It should be mentioned that x is normalized so that $||x||_2 = 1$.

2: Running PCA in the training set. Training images are blurred with each PSF H_i, We extract the feature images for the whole blurred training set, forming correlation matrix $A_i = \frac{1}{M} \Sigma_{k=1}^{M} x^{'}_{ik}(x^{'}_{ik})^T$, where $x^{'}_{ik}$ represents the feature image extracted from image g_{ik} which is blurred with H_i. A subspace $\theta_i = \{b_{ij}\}_{j=1}^{D}$ is got with the first D eigenvectors by decreasing eigenvalue.

3: Infer the PSF. For a new query image with unknown blur, we calculate the closest subspace with cosine distance to determine the PSF H_s:

$$s = arg \max_{i} \Sigma_{j=1}^{D}(b_{ij}^{T}x)^2 \qquad (3)$$

2.2 Feature Selection

Feature selection is an issue worth considering for blurred face recognition. Previous researchers attempted to find a blur invariant descriptor, however, this is

an ill-posed problem and we could only find approximately invariant ones. For instance, Local Phase Quantization(LPQ) is such a descriptor, and only when the facial image is boundless could it be invariant to blur. More details about how LPQ is applied in face recognition can be found in [30]. In fact, blur is closely related to scaling, making global descriptors inappropriate for the problem. Instead, local descriptors like LPQ, LBP and EF describe a pixel in an image with its neighborhood content, and are therefore capable of handling the invariants of blur to some extent. In our experiments, we first choose uniformed LBP as our feature. For a new coming face image, we first resize it to 64×64. Then, we compute the uniformed LBP(ranging from 0 to 58) for the pixels of the whole image. Later, each image is divided into 8×8 cells, and we would obtain the histograms for each cell. Combining these histograms together will result in a feature with dimension 59×8×8. Besides, we also extract LPQ features similarly, while the cell size is set to be 6×6 and cell is slid with a gap of 4 pixels.

2.3 Transformation Learning

High-dimensional transformation learning is always terribly high in time and space complexity. For instance, the dimension of using uniformed LBP is 3776, and the raw pixel dimension is 4096. Therefore, before we perform transformation learning, PCA (Principal Component Analysis), which is a standard technique to create low-dimensional representation of high-dimensional data, is always required. It helps to reduce the complexity significantly while preserving the performance of Transformation Learning. Besides, PCA is also effective for reducing noises. When the training set contains enough images for each individual, LDA (Linear Discriminant Analysis) is performed to make the distances between facial images of different classes as large as possible while preserving the distances between same-class images. Based on this, LDA further reduces the dimensions of the data. However, later we would demonstrate that PCA and LDA only help promote the recognition performance – it is the multidimensional scaling that really counts.

Objective Function We denote f as the transformation from R^l to R^t, here l represents the dimension of the input space, while t is the dimension of the transformed space. f could be written as:

$$f(x) = W^T x \tag{4}$$

where $x \in R^l$ is the feature vector we got from LBP, PCA and LDA. W is a $l \times t$ matrix, denoting the weights to be learned. Traditional metric multidimensional scaling tends to set $t = 2$ or 3 for visualization; while on the contrary, we find that mapping data to a higher dimension may lead to better separability, just as Support Vector Machine reveals. Suppose we have N original training images, all of which are processed with feature extraction (always followed with PCA for high dimensional feature), and we denote them as $\{x_1^s, x_2^s, \ldots x_N^s\}$, and

their corresponding blurred images are $\{x_1^b, x_2^b, \ldots x_N^b\}$. We apply the objective function of [34] to solve the problem:

$$\min J(W) = \lambda J_{SP}(W) + (1 - \lambda)J_{CS}(W) \tag{5}$$

where $J_{SP}(W)$ is a structure-preserving term. Our objective is to make sure the distance of transformed feature space between the original sharp image $f(x_i^s)$ and the blurred image $f(x_j^b)$ approximates the distance of input space between original sharp images d_{ij}^s as well as possible. We define $J_{SP}(W)$ as:

$$J_{SP}(W) = \sum_{i=1}^{N} \sum_{j=1}^{N} (q_{ij}(W) - d_{ij}^s)^2 \tag{6}$$

where

$$q_{ij}(W) = |f(x_i^s) - f(x_j^b)| = |W^T(x_i^s - x_j^b)| \tag{7}$$

stands for the distance between the sharp and the blurred images in the transformed space. d_{ij}^s is the distance between two sharp images in the input space, usually defined as $|x_i^s - x_j^s|$. This formulation could effectively handle the two main problems we have presented.

$J_{CS}(W)$ is the class-separation term. Similar to Koontz and Fukunaga, we propose to define the separability term as:

$$J_{CS}(W) = \sum_i \sum_j \delta_{ij} q_{ij}^2(W) \tag{8}$$

However, our term differs from Koontz and Fukunaga in that both i and j are taken from 1 to N, and we define δ_{ij} as:

$$\delta_{ij} = \begin{cases} 0 & \text{if } w_i \neq w_j \\ 1 & \text{if } w_i = w_j (i \neq j) \\ 2 & \text{if } i = j \end{cases} \tag{9}$$

where w_i stands for the class label of image i. λ is the balance parameter between the structure-preserving term and class-separation term, whose value is taken in the range $[0, 1]$. Combining the equations (5)(6)(7)(8)(9), we could transform the objective function as:

$$J(W) = \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_{ij}(q_{ij}(W) - \beta_{ij} d_{ij}^s)^2 \tag{10}$$

where

$$\alpha_{ij} = (1 - \lambda)\delta_{ij} + \lambda \tag{11}$$

and

$$\beta_{ij} = \lambda/\alpha_{ij} \tag{12}$$

. We then adopt the iterative majorization algorithm [34][33] to solve the minimization problem (10).

2.4   Iterative Majorization

Here, we quickly review the iterative majorization algorithm, which is also called stress majorization. If we define:

$$J_m^2(W,V) = Tr\{W^T A W\} + \sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_{ij}d_{ij}^2$$
$$-2Tr\{V^T D(V)W\} \tag{13}$$

where

$$A = \sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_{ij}(x_i - x_j)(x_i - x_j)^T \tag{14}$$

and

$$D(V) = \sum_{i=1}^{N}\sum_{j=1}^{N}C_{ij}(V)(x_i - x_j)(x_i - x_j)^T \tag{15}$$

,
with

$$C_i j(V) = \begin{cases} \lambda d_{ij}^s/q_{ij}(V) & \text{if } q_{ij}(V) > 0 \\ 0 & \text{if } q_{ij}(V) = 0 \end{cases} \tag{16}$$

The solution for W that minimizes $J_m^2(W,V)$ satisfies

$$AW = D(V)V \tag{17}$$

based on the theory that $J(W) \leq J_m^2(W,V) \leq J(V)$. The detailed minimization procedure is:

Algorithm 2 Iterative Majorization

1: set $t = 0$, initial convergence precision ξ with a predefined value, initial W with random value in the range $[-1, 1]$.
2: set $V = W^t$
3: update W^t to W^{t+1} as
$$W^{t+1} = A^{-1}D(V)V \tag{18}$$
   where A^{-1} stands for the Moore-Penrose pseudoinverse of A.
4: check for convergence. If $|W^{t+1} - W^t| < \xi$, stop the iteration and output W; If otherwise, set $t = t + 1$ and go to step 2.

2.5   Recognition across Blur

The transformation above gives a common feature space for sharp and blurred images, in which distance metrics such as Euclidean distance and cosine distance can be utilized for recognition. Here we choose cosine distance for verification,

considering its best performance in verification, or choose Euclidean distance for face recognition combined with an NN classifier.

The gallery set and the query set consist of sharp and blurred images respectively. A query blurred image could be compared to all of the sharp gallery images using cosine distance metrics for verification, or be put into the trained classifier C_s with Euclidean distance for recognition.

3   Experiments and Results

The following experiments are conducted to demonstrate the effectiveness of our method. We assume that the blur degrees of the test images are the same and already known, thus we can construct training images with the same blur effect. We shall discuss how to deal with unknown degree of blur in the Conclusion part.

3.1   Databases

We perform different experiments based on the following two databases. It should be noted that blur does not matter much in face detection, as the current face detector could detect faces as small as 15×15 pixels.
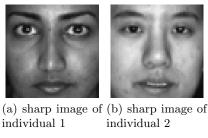
CMU-PIE Database [36] contains 68 subjects with 41368 images on the whole. For each subject, the illumination subset (C27), which contains 21 distinct sources of lights, was used in our experiment. For both datasets, all the images were first normalized by a similarity transformation that sets the inter-eye line horizontal and the two eyes 70 pixels apart, and then cropped to the size of 128×128 with the centers of the eyes located at (29, 34) and (99, 34) to extract the pure face region. No additional preprocessing procedure is required in our experiments.

FERET Database [37] is a standard database used for algorithm development and testing, which is divided into development portion and sequestered portion. The development portion provides a common database for designing algorithms, and the sequestered portion is for testing and evaluating face recognition algorithms. In our experiments, we select three subsets, 'bk','bj', and 'ba', from the development portion, and two subsets, 'fa' and 'fb', from the sequestered portion. The faces of all subsets are detected with a V-J face detector and the facial area is cropped to the size of 128×128.

3.2   Transformation Learning: Intuition

To give an intuition as to how T́ransformation Learningẃorks, we select two individuals from database CMU-PIE, and each individual consists of 21 images. First of all, we resize the images to 64×64 bilinearly. Then all the images are artificially blurred with Gaussian blur PSF: $H(u,v) = \frac{1}{Z}exp(\frac{-(u^2+v^2)}{2\sigma^2})$, where σ denotes the standard deviation, and Z is a normalization term. In our experiment, we set $\sigma = 3$ and the filter size is fixed to be $HSIZE = [5\,5]$. Typical

images of the two individuals and their corresponding blurred ones are shown in Fig.2.

(a) sharp image of individual 1   (b) sharp image of individual 2

(c) blur image of individual 1   (d) blur image of individual 2

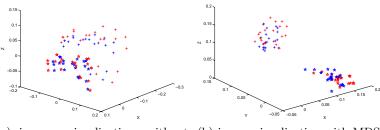**Fig. 2.** Example images from 2 different individuals. Each one is artificially blurred with a Gaussian blur filter whose size is 5×5 and the standard deviation is 3.

To exclude the effects of other factors, we simply choose raw pixel data as the feature vector. We use data of another 34 individuals to learn a PCA matrix and a transformation matrix. Then dimension reductions are conducted on all the original sharp images and their corresponding blurred ones, to 100 dimension after PCA and then to 900 dimension after MDS. The reason why PCA is used and why the dimension is 100 are that we want to perform MDS in a rapid way and to give a convenient visualization at the same time. The first 3 dimensions (the biggest 3 principals) are chosen for visualization. The result is shown in Fig.3.

It's obvious that the non-separable data in the original data space becomes linear-separable in the transformed space through MDS, which means a simple linear classifier would be sufficient to work well.

3.3 Face identification with and without MDS

Next, we conduct a face identification experiment on the CMU-PIE database. This time, all the cropped images are used. We randomly select 34 individuals for training, the remainings are used as gallery set. The training set images are blurred using Gaussian PSF, where σ is set to be 3, and $HSIZE$ is set to be [3 3]. After resizing the images to 64×64, we extract the LBP features of the images and then conduct PCA, LDA to reduce the dimension to 100. We use

(a) images visualization without MDS

(b) image visualization with MDS

Fig. 3. Visualization of the effectiveness of our method. Different shapes denote different individuals with 'cross' represents individual 1, and 'star' represents individual 2. And we mark the sharp images in blue color and the blurred ones in red.

the training set (34×21 images) to train PCA and MDS transformation matrix. Note that we set the initial MDS transformation matrixes W as equal values to give a fair comparison. Finally, given the test images and gallery ones, we adopt Cosine distance rather than Euclidean distance as the distance metric to get the identification results. See Equation 19.

$$\tilde{d}(i,j) = \frac{\tilde{x}_i \tilde{x}_j}{|x_i||x_j|} \tag{19}$$

The nature of the results would be affected by the threshold value. Take false detection rate as x axis, and true positive rate as y axis, and if we change the threshold value continuously, an ROC curve would be acquired. In our experiment, we change threshold value of identification for 505521 times, meaning that 505521 points (fp_i, tp_i) would be acquired. Here fp_i stands for false positive rate, while tp_i denotes true positive rate. The result is shown in Fig .4.

The result indicates that PCA and LDA help to improve the performance of identification, yet only marginally. The next experiment may reveal that PCA could even degrade the recognition rate sometimes. We calculate the Area Under Curve (AUC) of the ROC curve with the Trapezoidal Rule:

$$AUC = \sum_{i=1}^{n-1} (fp_{i+1} - fp_i) \frac{tp_{i+1} + tp_i}{2} \tag{20}$$

Here $n = 505521$, which denotes the number of sample points.

As shown in Table 1, MDS significantly improves the performance of face identification, as the AUC value is much higher when MDS is employed.

3.4   Comparison with state-of-art methods

The state-of-art method is FADEIN+LPQ[6]. To make a fair comparison, we follow the same approaches. FERET database is adopted where all images could be
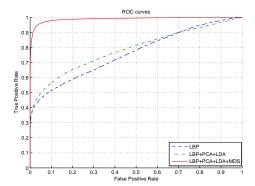
Fig. 4. Comparision of ROC curves.

Method	Accuracy (false positive rate=0.01)	AUC
LBP	37.66%	0.762
LBP+PCA+LDA	37.64%	0.787
LBP+PCA+LDA+MDS	**87.36%**	**0.989**

Table 1. Comparison of methods with and without MDS. We calculate the accuracy with false positive rate of 0.01. And the AUC values are based on the ROC curve shown in Fig 4

considered sharp. We choose 10001 images of 1001 individuals from subset 'fa' as the gallery set. Meanwhile, their corresponding images in 'fb' subset are selected to build up the target set, which have been filtered with Gaussian blur and added with $30dB$ Gaussian white noise. It should be mentioned that Gopalan et.al [32] improperly set the filter sizes to be the same ($hsize = 5$) for different standard deviations, while in fact, the filter size should increase with the standard deviation. Specifically, we set $hsize = 1.5 * \delta$, then around 87% energy would be preserved for all filters. Our objectives imply that only one image for each individual is not enough for learning the transformation matrix, so we combine 'bk','bj','ba' subsets together for training, with 3 images for each individual. We artificially blur these images as we did to those in the 'fb' subset. For the feature, we choose uniformed LPQ, where the cell size is set to be 6×6 and the sliding step to be 4. PCA is performed on the extracted features to reduce the dimension to 400, and a transformation to a dimension of 900 is learned through our MDS method, where we set $\lambda = 0.2$. Here we simply denote 'LPQ+PCA+MDS' as 'LPQ+MDS'. Finally, NN classifier is used for recognition, and the result is shown in Fig 5.

It can be seen clearly that our method could compete with the FADEIN+LPQ method. Both of them present considerable improvement over the pure LPQ
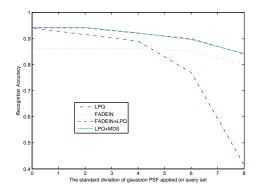
Fig. 5. Comparision with state-of-art methods.

method, which means that both of our MDS method (without deblurring) and FADEIN (with deblurring) are effective for improving the recognition accuracy.

3.5   Performance on mixed blur settings

The experiments above assume that all the training images and query images are blurred by PSFs with a certain standard deviation. The question is, what if there is no information available about the PSF? We solve this problem by adopting the PSF inference method by M.Nishiyama and A.Hadid[6]. While our experiment considers only Gaussian blur, the method could handle all kinds of blur indeed. First of all, we test the PSF inference accuracy on FERET database. The chosen datasets are the same as in section 2.4, and we still set $hsize = 1.5*\sigma$. We blur the whole query set with different Gaussian PSFs each time. The result is shown in Table 2.

standard deviationσ	Inference Accuracy
0	99.80%
2	98.49%
4	97.68%
6	98.39%
8	98.19%

Table 2. PSF inference accuracy for different standard deviation.

The average inference accuracy is 98.43%, proving the effectiveness of the PSF inference method.

Then, we blur the images with random standard deviation chosen from $\{0, 2, 4, 6, 8\}$, for each query image, we infer the PSF first, and then choose

the corresponding transformation matrix, using transformed feature to recognize with the corresponding classifier. The result is shown in Table 3.

Method	Recognition Rate
LPQ	83.56%
FADEIN	82.45%
FADEIN+LPQ	88.12%
LPQ+PCA+MDS	**88.56%**

Table 3. PSF inference accuracy for different standard deviation.

It can be noticed that without deblurring, our method performs slightly better than FADEIN combined with LPQ, thus demonstrating the effectiveness of our method.

3.6   Time Costs

It is the training procedure that takes much time. When we set the convergence precision to be 0.1, it would take around 30 iteration steps to complete the transformation from 400 to 900 dimensions. In our Matlab implemention on 3.2GHZ core i5 CPU, about 3 hours was consumed to compute the matrix W. After obtaining W, however, it takes very short time for testing, thus making it suitable for real-time face recognition applications.

4   Conclusion

In this paper, we solve the blur problem of face recognition based on a new point of view, which is transforming both the sharp images and blurred ones to a common subspace. Our approach could be regarded as a procedure coming after feature extraction, thus compatible with various existing methods. One problem with our method is that we assume the blur degree to be already known, which could be solved by training a set of W with different blurs. Concretely, when a new test image is coming, we could first infer the blur degree with PSF inference methods, and then select the best matched W. Further research may focus on improving the transformation learning matrix through finding a subspace for various degrees of blur.

References

1. Deng, W., Hu, J., Guo, J.: Extended src: Undersampled face recognition via intraclass variant dictionary. IEEE Trans. Pattern Anal. Mach. Intell. **34** (2012) 1864–1870
2. Deng, W., Hu, J., Guo, J., Cai, W., Feng, D.: Robust, accurate and efficient face recognition from a single training image: A uniform pursuit approach. Pattern Recognition **43** (2010) 1748–1762
3. Deng, W., Hu, J., Zhou, X., Guo, J.: Equidistant prototypes embedding for single sample based face recognition with generic learning and incremental learning. Pattern Recognition **47** (2014) 3738–3749
4. Deng, W., Hu, J., Lu, J., Guo, J.: Transform-invariant pca: A unified approach to fully automatic face alignment, representation, and recognition. IEEE Trans. Pattern Anal. Mach. Intell. **36** (2014) 1275–1284
5. Deng, W., Hu, J., Guo, J., Cai, W., Feng, D.: Emulating biological strategies for uncontrolled face recognition. Pattern Recognition **43** (2010) 2210–2223
6. Nishiyama, M., Hadid, A., Takeshima, H., Shotton, J., Kozakaya, T., Yamaguchi, O.: Facial deblur inference using subspace analysis for recognition of blurred faces. Pattern Analysis and Machine Intelligence, IEEE Transactions on **33** (2011) 838–845
7. Stainvas, I., Intrator, N., Moshaiov, A.: Blurre face recognition via a hybrid network architecture. Pattern Recognition **2** (2000) 805–808
8. Cannon, M.: Blind deconvolution of spatially invariant image blurs with phase. Acoustics, Speech and Signal Processing, IEEE Transactions on **24** (1976) 58–63
9. Katsaggelos, A.K.: Digital image restoration. Springer Publishing Company, Incorporated (2012)
10. Kimia, B.B., Zucker, S.W.: Analytic inverse of discrete gaussian blur. Optical engineering **32** (1993) 166–176
11. Yuan, L., Sun, J., Quan, L., Shum, H.Y.: Progressive inter-scale and intra-scale non-blind image deconvolution. In: ACM Transactions on Graphics (TOG). Volume 27., ACM (2008) 74
12. Levin, A., Lischinski, D., Weiss, Y.: A closed-form solution to natural image matting. Pattern Analysis and Machine Intelligence, IEEE Transactions on **30** (2008) 228–242
13. Fergus, R., Singh, B., Hertzmann, A., Roweis, S.T., Freeman, W.T.: Removing camera shake from a single photograph. In: ACM Transactions on Graphics (TOG). Volume 25., ACM (2006) 787–794
14. Levin, A., Weiss, Y., Durand, F., Freeman, W.T.: Efficient marginal likelihood optimization in blind deconvolution. In: Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on, IEEE (2011) 2657–2664
15. Kundur, D., Hatzinakos, D.: Blind image deconvolution. Signal Processing Magazine, IEEE **13** (1996) 43–64
16. Chan, T.F., Wong, C.K.: Total variation blind deconvolution. Image Processing, IEEE Transactions on **7** (1998) 370–375
17. Hu, H., de Hann, G.: Low cost robust blur estimator. In: Image Processing, 2006 IEEE International Conference on, IEEE (2006) 617–620
18. Elder, J.H., Zucker, S.W.: Local scale control for edge detection and blur estimation. Pattern Analysis and Machine Intelligence, IEEE Transactions on **20** (1998) 699–716

19. Marziliano, P., Dufaux, F., Winkler, S., Ebrahimi, T.: A no-reference perceptual blur metric. In: Image Processing. 2002. Proceedings. 2002 International Conference on. Volume 3., IEEE (2002) III–57
20. Rooms, F., Pizurica, A., Philips, W.: Estimating image blur in the wavelet domain. In: IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS SPEECH AND SIGNAL PROCESSING. Volume 4., IEEE; 1999 (2002) 4190–4190
21. Tong, H., Li, M., Zhang, H., Zhang, C.: Blur detection for digital images using wavelet transform. In: Multimedia and Expo, 2004. ICME'04. 2004 IEEE International Conference on. Volume 1., IEEE (2004) 17–20
22. Yuan, L., Sun, J., Quan, L., Shum, H.Y.: Image deblurring with blurred/noisy image pairs. In: ACM Transactions on Graphics (TOG). Volume 26., ACM (2007) 1
23. Ancuti, C., Ancuti, C.O., Bekaert, P.: Deblurring by matching. In: Computer Graphics Forum. Volume 28., Wiley Online Library (2009) 619–628
24. Nishiyama, M., Takeshima, H., Shotton, J., Kozakaya, T., Yamaguchi, O.: Facial deblur inference to improve recognition of blurred faces. In: Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on, IEEE (2009) 1115–1122
25. Ojala, T., Pietikainen, M., Maenpaa, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. Pattern Analysis and Machine Intelligence, IEEE Transactions on **24** (2002) 971–987
26. Ahonen, T., Hadid, A., Pietikainen, M.: Face description with local binary patterns: Application to face recognition. Pattern Analysis and Machine Intelligence, IEEE Transactions on **28** (2006) 2037–2041
27. Deng, W., Liu, Y., Hu, J., Guo, J.: The small sample size problem of ica: A comparative study and analysis. Pattern Recognition **45** (2012) 4438–4450
28. Deng, W., Hu, J., Guo, J., Zhang, H., Zhang, C.: Comments on "globally maximizing, locally minimizing: Unsupervised discriminant projection with applications to face and palm biometrics". IEEE Trans. Pattern Anal. Mach. Intell. **30** (2008) 1503–1504
29. Deng, W., Hu, J., Guo, J.: In defense of sparsity based face recognition. In: Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on, IEEE (2013) 399–406
30. Ahonen, T., Rahtu, E., Ojansivu, V., Heikkila, J.: Recognition of blurred faces using local phase quantization. In: Pattern Recognition, 2008. ICPR 2008. 19th International Conference on, IEEE (2008) 1–4
31. Ojansivu, V., Heikkilä, J.: Blur insensitive texture classification using local phase quantization. In: Image and Signal Processing. Springer (2008) 236–243
32. Gopalan, R., Taheri, S., Turaga, P., Chellappa, R.: A blur-robust descriptor with applications to face recognition. Pattern Analysis and Machine Intelligence, IEEE Transactions on **34** (2012) 1220–1226
33. Cox, T.F., Cox, M.A.: Multidimensional scaling. CRC Press (2000)
34. Webb, A.R.: Multidimensional scaling by iterative majorization using radial basis functions. Pattern Recognition **28** (1995) 753–759
35. Biswas, S., Aggarwal, G., Flynn, P.J., Bowyer, K.W.: Pose-robust recognition of low-resolution face images. Pattern Analysis and Machine Intelligence, IEEE Transactions on **35** (2013) 3037–3049
36. Sim, T., Baker, S., Bsat, M.: The cmu pose, illumination, and expression (pie) database. In: Automatic Face and Gesture Recognition, 2002. Proceedings. Fifth IEEE International Conference on, IEEE (2002) 46–51

37. Phillips, P.J., Moon, H., Rizvi, S.A., Rauss, P.J.: The feret evaluation methodology for face-recognition algorithms. Pattern Analysis and Machine Intelligence, IEEE Transactions on **22** (2000) 1090–1104