

Separability Oriented Preprocessing for Illumination-Insensitive Face Recognition

Hu Han^{1,2}, Shiguang Shan¹, Xilin Chen¹, Shihong Lao³, and Wen Gao^{4,1}

¹ Key Lab of Intelligent Information Processing
of Chinese Academy of Sciences (CAS),

Institute of Computing Technology, CAS, Beijing 100190, China

² Department of Computer Science and Engineering,

Michigan State University, East Lansing, MI 48824, U.S.A.

³ Omron Social Solutions Co., LTD., Kyoto, Japan

⁴ Institute of Digital Media, Peking University, Beijing 100871, China

{hhan,sgshan,xlchen}@jd1.ac.cn, lao@ari.ncl.omron.co.jp, wgao@pku.edu.cn

Abstract. In the last decade, some illumination preprocessing approaches were proposed to eliminate the lighting variation in face images for lighting-invariant face recognition. However, we find surprisingly that existing preprocessing methods were seldom modeled to directly enhance the separability of different faces, which should have been the essential goal. To address the issue, we propose to explicitly exploit maximizing separability of different subjects' faces as the preprocessing objective. With this in mind, a novel approach, named by us Separability Oriented Preprocessing (SOP), is proposed to enhance face images by maximizing the Fisher separability criterion in scale-space. Extensive experiments on both laboratory-controlled and real-world face databases using different recognition methods show the effectiveness of the proposed approach.

Keywords: Separability oriented, illumination preprocessing, lighting-invariant, face recognition.

1 Introduction

Face recognition has wide potential applications in commercial and law enforcement applications [1], mainly owing to its contact-free and non-intrusive characteristics. However, a practical face recognition system has to handle the grand challenges caused by the variations in face appearance due to lighting, head pose, facial expression, partial occlusion, and so on. Among them, harsh lighting conditions such as underexposure, overexposure and shadow, are regarded as one of the bottlenecks for robust face recognition. Attempts for handling lighting variation include lighting-invariant feature extraction, lighting variation modeling and image enhancement (or illumination normalization). Approaches of the first category try to extract facial features that are robust against lighting variations [7,2,3,5] while approaches of the second category tend to build a representation that models arbitrary lighting variations [9,10,8,13,11,12] and approaches of the

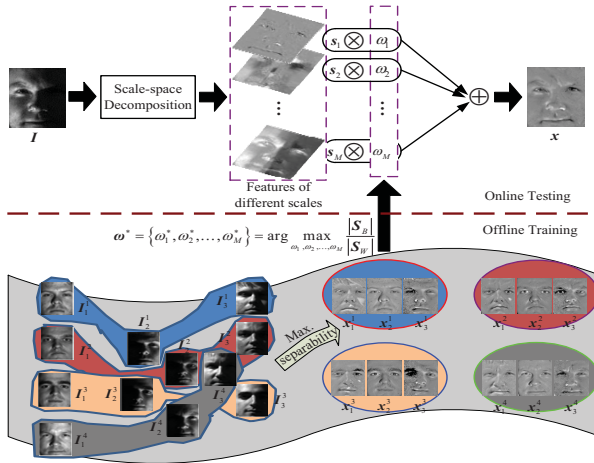


Fig. 1. An overview of Separability Oriented Preprocessing (SOP) for illumination-invariant face recognition. $s_l, l = 1, 2, \dots, M$ are the extracted facial features distributed from fine (such as the eyes and mouth in s_1) to coarse (such as the shape and shading in s_M) sales. The enhanced face image x is computed as the linear combination of the decomposed components of distinct scales, in which the weights are determined by an optimization framework via maximizing the separability of training samples.

third category preprocess face images by eliminating their lighting variations before the subsequent recognition algorithm [16,14,15,6,21].

Compared with those approaches of the former two categories, approaches of the third category have attracted much attention due to the properties of high efficiency and general purpose usage. However, we could find that there is no objective function in existing illumination preprocessing methods that guarantees separability improvement after preprocessing. This should be the reason why existing preprocessing methods are reported to degrade the recognition performance on face images with normal lighting [18]. This intuitive analysis inspires our basic idea that higher separability of different faces should be the explicit goal of preprocessing. With this basic idea in mind, a novel image enhancement framework, Separability Oriented Preprocessing (SOP), is proposed to normalize face images by maximizing the separability. Specifically, in this study, as shown in Fig. 1, an input face image with unknown lighting is first decomposed through scale-space decomposition. Then, the enhanced face image is computed as a linear combination of the decomposed coarse-to-fine bands with weights learned offline from training set by maximizing the Fisher separability criterion. Global optimization based on simulated annealing is exploited to solve the optimization problem in SOP.

The proposed approach is extensively evaluated on laboratory-controlled databases, i.e., Multi-PIE [23] and Extended YaleB [9], as well as a real-world dataset, i.e., FRGC Ver2.0 [26]. Impressive improvement in recognition accuracy is achieved compared with several well-known preprocessing methods, which is validated by using three distinct face recognition methods, i.e., Fisherfaces [8], LBP [4] and

Eigenfaces [27]. Furthermore, cross-database training on heterogeneous databases also shows that the proposed method has outstanding generalization ability.

2 Previous Preprocessing Methods by Scale-space Decomposition and Their Problems

In many computer vision tasks, the face imaging procedure is formulated as a reflection-illumination model

$$\mathbf{I} = \mathbf{R}\mathbf{L} \quad (1)$$

where \mathbf{R} is the reflectance component which corresponds to intrinsic facial features and \mathbf{L} is the illumination component which corresponds to lighting variations. Based on this face imaging model, illumination normalization for a face image is usually performed by estimating the illumination component \mathbf{L} and subtracting it to get the reflectance component \mathbf{R} . It is convenient to perform the above normalization based on an additive model; however, the face imaging model is multiplicative. Therefore, logarithmic transformation is usually applied first to convert the multiplicative imaging model into an additive one

$$\log \mathbf{I} = \log \mathbf{R} + \log \mathbf{L} \quad (2)$$

For convenience, we rewrite the above equation as $\mathbf{f} = \mathbf{u} + \mathbf{v}$, where $\mathbf{f} = \log \mathbf{I}$, $\mathbf{u} = \log \mathbf{L}$ and $\mathbf{v} = \log \mathbf{R}$. Then, illumination preprocessing can be performed as

$$\hat{\mathbf{v}} = \mathbf{f} - \hat{\mathbf{u}} \quad (3)$$

where $\hat{\mathbf{u}}$ is the estimation of the illumination component in the input face image, and it can be generally represented as

$$\hat{\mathbf{u}} = \mathcal{F}(\mathbf{f}) \quad (4)$$

Based on the observation that illumination variations mainly lie in the low-frequency domain of a face image [10,19], in practice, a low-pass filtering is commonly used as $\mathcal{F}(\cdot)$ to estimate $\hat{\mathbf{u}}$. For instance, [5] uses a Gaussian kernel \mathcal{G} to get a smoothed version of \mathbf{f} , which is a convolution procedure

$$\mathcal{F}(\mathbf{f}) = \mathcal{G} * \mathbf{f} \quad (5)$$

In [16], discrete cosine transform (DCT) is utilized to implement $\mathcal{F}(\cdot)$ by removing the DC coefficient and several large AC coefficients scanned in zigzag pattern. Recently, total variation model (TV-L1) [20] is utilized to estimate $\hat{\mathbf{u}}$ in [21], in which the illumination component in image \mathbf{f} is estimated by solving the following variational problem

$$\mathcal{F}(\mathbf{f}) = \arg \min_{\mathbf{u}} \int |\nabla \mathbf{u}| + \lambda \|\mathbf{f} - \mathbf{u}\|_{L^1} \quad (6)$$

where $\int |\nabla \mathbf{u}|$ is the total variation of \mathbf{u} , λ is a scalar constant controlling the truncation scale between \mathbf{u} and \mathbf{v} , and $\|\cdot\|_{L^1}$ denotes $L1$ norm.

In all the above methods, only features in small-scale band are believed helpful for face recognition while those in the large-scale band are not and removed.

One problem of this strategy is that it is hard to decide where to truncate the scale space. In the literature, it is generally set empirically, which is quite tedious and not easy to guarantee optimum. Lighting aware preprocessing (LAP) [18] was proposed to adaptively determine the truncation scale. A more serious problem with the above methods is that there are also some features in the component of large-scale which are helpful to distinguish different faces. Another common issue in above image enhancement approaches is that different feature bands are considered to be of the same importance for face recognition. However, discriminant facial features are not necessarily distributed uniformly among each band in scale space. So, treating them equally is not desirable.

To summarize, the essential problem of the previous preprocessing methods with scale-space decomposition is: no clear and formal objective is designed to guide the enhancement to keep or even improve the discriminative capacity of the resulting images. With the above analysis, in this work, we propose a novel approach that establishes an explicit relationship between the goal of image enhancement and separability improvement by appropriately exploiting the features in all the scales.

3 Separability Oriented Preprocessing (SOP)

3.1 Scale-space Decomposition

As above mentioned, signals at different bands play different roles for face recognition. However, a rough division of features in a face image into one large-scale and small-scale band pair is insufficient for understanding the roles played by different feature bands. To address this issue, a group of band-pass filters can be carefully designed to extract the facial features in different bands. In fact, such a careful design is unnecessary as many well-studied existing approaches can be utilized for scale-space decomposition.

As shown in Fig. 2, we decompose a face image into pairs of large-scale and small-scale components: $\{\mathbf{u}_l, \mathbf{v}_l\}, l = 1, 2, \dots, M$ by varying the truncation scale between \mathbf{u}_l and \mathbf{v}_l . It is not difficult to find that \mathbf{v}_1 is flattened as all the face shape information is lost and there are mainly facial features of smallest scale in \mathbf{v}_1 . From \mathbf{v}_1 to \mathbf{v}_M more and more face shape information will be included. As mentioned above, the decomposition of \mathbf{f} into \mathbf{u} and \mathbf{v} is based on an additive model, therefore, the relationship between all \mathbf{v}_l will be

$$\mathbf{v}_1 \subset \mathbf{v}_2 \subset \dots \subset \mathbf{v}_M \quad (7)$$

Then, as illustrated in Fig. 2, a face image can be decomposed into features of various bands in scale-space as below

$$\mathbf{s}_l = \begin{cases} \mathbf{v}_1 & \text{if } l = 1 \\ \mathbf{v}_l - \mathbf{v}_{l-1} & \text{otherwise} \end{cases} \quad (8)$$

Theoretically, all the methods mentioned in Sect. 2 can be used for scale-space decomposition for a face image, and in our experiments below, TV-L1 model [20] is utilized.

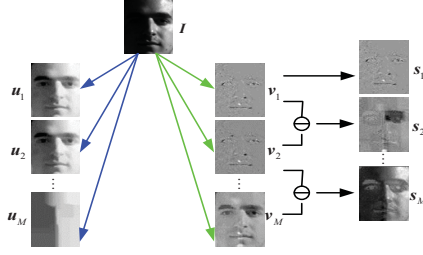


Fig. 2. Scale-space decomposition for a face image. A face image is first decomposed into several pairs of $\{u_l, v_l\}$, $l = 1, 2, \dots, M$ by varying the truncation scale between large-scale and small-scale bands and then features in different bands are calculated. s_1 band contains the smallest-scale facial features such as the eyes and mouth, while s_M band contains the largest-scale facial features such as the face shape and shading.

3.2 Formulation of SOP

Once the facial features of different bands are extracted, enhanced face images that are expected to have improved separability can be calculated by emphasizing the feature scales of greater importance. Therefore, the preprocessed face image can be formulated as

$$\begin{aligned}
 \mathbf{x} &= \mathbf{S}\boldsymbol{\omega} = \sum_{l=1}^M \omega_l \mathbf{s}_l \\
 \text{s.t.} & \\
 &\sum_{l=1}^M \omega_l = 1 \\
 &\omega_l \geq 0, \quad l = 1, 2, \dots, M
 \end{aligned} \tag{9}$$

where \mathbf{x} is the enhanced face image in vectorized form. $\boldsymbol{\omega} = \{\omega_1, \omega_2, \dots, \omega_M\}^T$ is a group of weights for different feature bands, which determines the importance of each band played for face recognition. Another reason why we formulate the preprocessing as (9) is to follow the principle of image enhancement: the enhanced results are still face images, not abstract features extracted from images.

From (9), it is not difficult to find that $\boldsymbol{\omega}$ determines the final image enhancement results, which thus will affect the effectiveness for robust face recognition across varying illumination. In the proposed approach, $\boldsymbol{\omega}$ is optimized by maximizing the separability of samples based on Fisher separability criterion

$$\begin{aligned}
 \boldsymbol{\omega}^* &= \{\omega_1^*, \omega_2^*, \dots, \omega_M^*\}^T \\
 &= \arg \max_{\omega_1, \omega_2, \dots, \omega_M} \frac{|\mathbf{S}_B|}{|\mathbf{S}_W|} \\
 \text{s.t.} & \\
 &\sum_{l=1}^M \omega_l = 1, \quad \omega_l \geq 0, \quad l = 1, 2, \dots, M
 \end{aligned} \tag{10}$$

where \mathbf{S}_B and \mathbf{S}_W are respectively the between- and within-class scatter matrices calculated from the training set. Given a training set with C classes, we denote

the vectorized face images after enhancement as $\mathbf{X} = \{\mathbf{X}^1, \mathbf{X}^2, \dots, \mathbf{X}^C\}$, in which $\mathbf{X}^i = \{\mathbf{x}_1^i, \mathbf{x}_2^i, \dots, \mathbf{x}_{N_i}^i\}$. Then, the between- and within-class scatter matrices are defined as

$$\mathbf{S}_B = \sum_{i=1}^C N_i (\boldsymbol{\mu}_i - \boldsymbol{\mu})(\boldsymbol{\mu}_i - \boldsymbol{\mu})^T \quad (11)$$

and

$$\mathbf{S}_W = \sum_{i=1}^C \sum_{\mathbf{x}_k \in \mathbf{X}^i} (\mathbf{x}_k - \boldsymbol{\mu}_i)(\mathbf{x}_k - \boldsymbol{\mu}_i)^T \quad (12)$$

where $\boldsymbol{\mu}_i$ is the mean image of class \mathbf{X}^i

$$\begin{aligned} \boldsymbol{\mu}_i &= \frac{1}{N_i} \sum_{j=1}^{N_i} \mathbf{x}_j^i \\ &= \omega_1 \boldsymbol{\mu}_i^{s_1} + \omega_2 \boldsymbol{\mu}_i^{s_2} + \dots + \omega_M \boldsymbol{\mu}_i^{s_M} \\ &= (\boldsymbol{\mu}_i^{s_1} \boldsymbol{\mu}_i^{s_2} \dots \boldsymbol{\mu}_i^{s_M}) (\omega_1 \omega_2 \dots \omega_M)^T \end{aligned} \quad (13)$$

where $\boldsymbol{\mu}_i^{s_l}$, $l = 1, 2, \dots, M$ is the mean image of class i calculated in band s_l . $\boldsymbol{\mu}$ is the mean image of all the training samples

$$\boldsymbol{\mu} = (\boldsymbol{\mu}^{s_1} \boldsymbol{\mu}^{s_2} \dots \boldsymbol{\mu}^{s_M}) (\omega_1 \omega_2 \dots \omega_M)^T \quad (14)$$

where $\boldsymbol{\mu}^{s_l}$, $l = 1, 2, \dots, M$ is the mean image of all samples calculated in band s_l . Combining (11), (13) and (14), we have

$$\begin{aligned} \mathbf{S}_B &= \sum_{i=1}^C N_i \\ &\quad \left((\boldsymbol{\mu}_i^{s_1} \boldsymbol{\mu}_i^{s_2} \dots \boldsymbol{\mu}_i^{s_M}) - (\boldsymbol{\mu}^{s_1} \boldsymbol{\mu}^{s_2} \dots \boldsymbol{\mu}^{s_M}) \right) \\ &\quad (\omega_1 \omega_2 \dots \omega_M)^T (\omega_1 \omega_2 \dots \omega_M) \\ &\quad \left((\boldsymbol{\mu}_i^{s_1} \boldsymbol{\mu}_i^{s_2} \dots \boldsymbol{\mu}_i^{s_M}) - (\boldsymbol{\mu}^{s_1} \boldsymbol{\mu}^{s_2} \dots \boldsymbol{\mu}^{s_M}) \right)^T \\ &= \sum_{i=1}^C N_i \mathbf{A}_i \boldsymbol{\omega}^T \boldsymbol{\omega} \mathbf{A}_i^T \end{aligned} \quad (15)$$

where $\mathbf{A}_i = (\boldsymbol{\mu}_i^{s_1} \boldsymbol{\mu}_i^{s_2} \dots \boldsymbol{\mu}_i^{s_M}) - (\boldsymbol{\mu}^{s_1} \boldsymbol{\mu}^{s_2} \dots \boldsymbol{\mu}^{s_M})$. Similarly, combining (9), (12) and (13), we will get

$$\mathbf{S}_W = \sum_{i=1}^C \sum_{\mathbf{x}_k \in \mathbf{X}^i} \mathbf{B}_{\mathbf{x}_k} \boldsymbol{\omega}^T \boldsymbol{\omega} \mathbf{B}_{\mathbf{x}_k}^T \quad (16)$$

where $\mathbf{B}_{\mathbf{x}_k} = (\mathbf{s}_1^{\mathbf{x}_k} \mathbf{s}_2^{\mathbf{x}_k} \dots \mathbf{s}_M^{\mathbf{x}_k}) - (\boldsymbol{\mu}_i^{s_1} \boldsymbol{\mu}_i^{s_2} \dots \boldsymbol{\mu}_i^{s_M})$. Finally, by combining (10), (15) and (16), the optimization problem defined in (10) can be rewritten as

$$\begin{aligned} \boldsymbol{\omega}^* &= \{\omega_1^*, \omega_2^*, \dots, \omega_M^*\}^T \\ &= \arg \max_{\omega_1, \omega_2, \dots, \omega_M} \frac{|\sum_{i=1}^C N_i \mathbf{A}_i \boldsymbol{\omega}^T \boldsymbol{\omega} \mathbf{A}_i^T|}{|\sum_{i=1}^C \sum_{\mathbf{x}_k \in \mathbf{X}^i} \mathbf{B}_{\mathbf{x}_k} \boldsymbol{\omega}^T \boldsymbol{\omega} \mathbf{B}_{\mathbf{x}_k}^T|} \end{aligned} \quad (17)$$

s.t.

$$\sum_{l=1}^M \omega_l = 1, \quad \omega_l \geq 0, \quad l = 1, 2, \dots, M$$

By solving the optimization problem in (17), we get the optimum estimation ω^* . Then image enhancement for a face image can be performed based on (9)

$$\mathbf{x} = \mathbf{S}\omega^* = \sum_{l=1}^M \omega_l^* \mathbf{s}_l \quad (18)$$

3.3 Solving the Optimization Problem

Theoretically, the best image enhancement results can be achieved using the weight ω learnt by solving the optimization problem in (17); however, it is difficult to get an analytical solution to the optimization problem defined in (17). In order to get an approximation to the global optimum solution to the above optimization problem, we propose to solve it based on simulated annealing [25] algorithm. Simulated annealing, which is based on the concepts and techniques from physics, specifically statistical mechanics, has a highly developed and rigorous theory and achieves many successes in pattern recognition [24]. Although simulated annealing is an iterative algorithm, it is a heuristic method; thus it makes a good balance between approximating the global optimum for a given function with a large search space and the computational cost.

From (17), it can be noticed that in our problem, each ω_l , $l = 1, 2, \dots, M$ is analog; therefore, deterministic simulated annealing can be performed, which will be more efficient than a stochastic one [24]. Before conducting simulated annealing in solving the optimization problem, a *cost* or *energy* function must be specified. In our method, the minimization of the energy function should correspond to the improvement of the separability of faces. Therefore, the following energy function is calculated in the t -th iteration of simulated annealing

$$E(\omega_{(t)}) = \left(\frac{1}{F(\omega_{(t)})} \right)^2 \quad (19)$$

where

$$F(\omega_{(t)}) = \frac{|\sum_{i=1}^C N_i \mathbf{A}_i \omega_{(t)}^T \mathbf{A}_i^T|}{|\sum_{i=1}^C \sum_{\mathbf{x}_k \in X_i} \mathbf{B}_{\mathbf{x}_k} \omega_{(t)}^T \mathbf{B}_{\mathbf{x}_k}^T|} \quad (20)$$

With the energy function defined, the learning stage in SOP can be performed offline. Simulated annealing is initialized with equal band weights and stops when the change of separability is less than $1e-3$ for 500 times. Moreover, to improve the generalization ability, we apply simulated annealing in a cross-validation pattern by dividing the training set into multiple folds. The offline learning of SOP is detailed in the supplemental material.

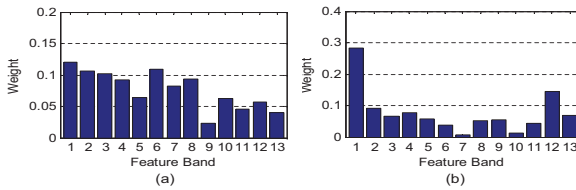


Fig. 3. The optimum weights for different feature scales that are learnt in SOP on (a) Multi-PIE and (b) Extended YaleB face databases

4 Evaluation

4.1 Roles of Features in Different Bands

In this section, the roles that features in different bands play for face recognition are analyzed based on the weights learnt by the proposed SOP from different face databases. In scale-space decomposition, we use 13 ($M = 13$) feature bands by linearly selecting lambda for TV-L1 model. Fig. 3 shows the weights for different bands that are respectively learnt from Multi-PIE [23] and Extended YaleB [9] databases. From Fig. 3, it can be noticed that the weights learnt from Multi-PIE is dramatically different from those learnt from Extended YaleB. For example, more importance is assigned to the smallest and largest bands for Extended YaleB. That is because the shadow in Extended YaleB is much more severe than that in Multi-PIE; thus, for many face images in Extended YaleB, features in many bands are corrupted. In this case, features of the smallest- and largest-scale bands, which mainly correspond to sharp edges and face shading respectively, are usually least affected by shadow.

Through the comparison of weights learnt by SOP on Multi-PIE and Extended YaleB, possible conclusions can be summarized as below:

- Different from the viewpoint of most previous illumination preprocessing methods which mainly preserve features of small-scale band, the proposed SOP shows that features of large-scale band are also important for face recognition. In fact, as shown in Fig. 2, the features of the largest-scale band, i.e. S_M , mainly correspond to face shading.
- The importance of one specific feature band is not fixed across different face databases, as the distributions of lighting in different face databases usually vary from one another. Therefore, for different face databases, only appropriately determined weights for each feature scales can lead to optimum illumination preprocessing and satisfying face recognition performance.

4.2 Experiments on Laboratory-controlled Database

The proposed SOP is first evaluated on two representative laboratory-controlled face databases: Multi-PIE [23] and Extended YaleB [9]. Multi-PIE database contains as many as 755,370 images from 337 subjects, imaged under 15 view

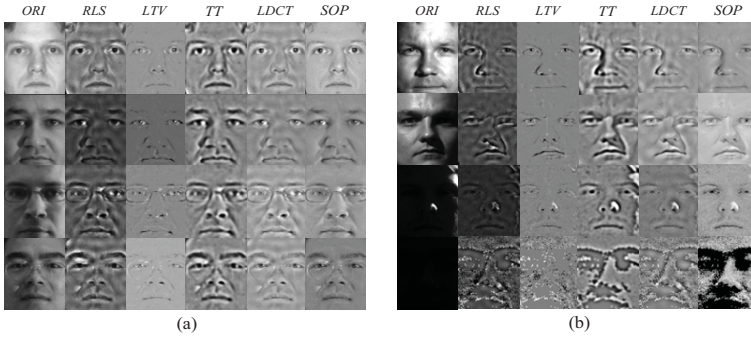


Fig. 4. Illumination preprocessing using different approaches is performed on (a) Multi-PIE and (b) Extended YaleB face databases

points and 19 illumination conditions in up to four recording sessions. According to the official testing protocol, face images of 14 randomly selected subjects are used for training and images of all the other 323 subjects are used for testing. Among all the testing images, only one image of each individual recorded without flashes is used as gallery. Extended YaleB database contains 21,888 face images from 28 human subjects under 9 poses and 64 illumination conditions. Seven face images of each individual are used for training and all the other images are used for testing. Following most of the image enhancement approaches [22,21,16,6], face images in frontal pose are used for experiments.

Several state-of-the-art illumination preprocessing methods, e.g. LTV [21], LDCT [16], RLS [17] and TT [6], are used to compare with the proposed SOP. For fair comparison, we exploited the best parameters settings recommended in the original literature proposing the corresponding methods. For convenience, we denote “ORI” as the original input face images without illumination preprocessing. Before illumination preprocessing, all the face images are geometrically normalized into the size of 64×80 with 35 pixels distance between two eyes.

Preprocessed face images from Multi-PIE and Extended YaleB are illustrated in Fig. 4. Compared with previous methods which usually performed for the goal of better visualization, the preprocessed face images with SOP also show impressive visual appearance. For face images with severe shadow in the last two rows of Fig. 4 (b), the proposed SOP preserves more discriminative facial features in the preprocessed face images.

Face recognition is then performed on the preprocessed face images to evaluate the effectiveness of different illumination preprocessing methods in improving the robustness of a face recognition method. As what we concern is the comparison among different illumination preprocessing methods, Fisherfaces [8] is fixed as the recognition method in this experiment. $C - 1$ dimensions are used in Fisherfaces, where C is the class number in the training set. Recognition performance of Fisherfaces with different methods for illumination preprocessing is detailed in Table 1. The recognition rate of 52.8% on original face images reveals the challenge of the testing on Multi-PIE. Face recognition with SOP

Table 1. Face recognition performance of Fisherfaces on the enhanced images preprocessed by different approaches on Multi-PIE and Extended YaleB face databases. The training set for SOP is from the same database with the testing set.

Approach	Recognition Rate (%)	
	Multi-PIE	Extended YaleB
ORI	52.8	54.2
RLS	56.9	75.0
LTV	60.8	78.0
TT	61.2	71.6
LDCT	61.8	74.1
SOP	75.0	84.1

Table 2. Generalization ability of our proposed SOP is evaluated on Multi-PIE and Extended YaleB face databases in a cross-database training pattern. With cross-database training, the training and testing sets for SOP are NOT from the same database.

Training Set	Testing Set	Recognition Rate (%)
Extend YaleB	Multi-PIE	73.8
Multi-PIE	Extend YaleB	83.1

for illumination preprocessing gets a recognition rate as high as 75.0%, which is 13% higher than the best of the state-of-the-art (61.8%). The effectiveness of the proposed SOP can also be observed from the recognition performance on Extended YaleB. Face recognition with SOP for illumination preprocessing gets a recognition rate as high as 84.1%, achieving 6% improvement compared with the best of the state-of-the-art (78.0%).

As there is no separate gallery set in Extended YaleB, more experiments are performed in a cross-database training pattern to verify the generalization ability of the proposed SOP. For example, SOP is trained on the training set from Extended YaleB but tested on the testing set from Multi-PIE, and vice versa. Face recognition is then performed on the preprocessed face images and comparisons with that without cross training/testing are listed in Table 2. As clearly shown in Table 2, even when SOP is performed in a cross training/testing pattern, it still shows to be effective for illumination-insensitive face recognition, e.g. 73.8% and 83.1% recognition rates are achieved in Multi-PIE and Extended YaleB, which are impressively higher than the best of previous methods. Both experiments show satisfying generalization ability of the proposed SOP.

4.3 Experiments on Real-world Database

In Multi-PIE and Extended YaleB, the lighting is strictly controlled and other factors, e.g. pose and expression, are mainly excluded in face imaging. From a practical standpoint, it would also be valuable to verify the effectiveness of the proposed SOP on a real-world database. Therefore, the dataset for Exp. 4 in FGRC Ver2.0 [26] is employed in this evaluation, as the query set in FGRC

Table 3. Performance comparison of different preprocessing methods in terms of VRs at FAR = 0.1%, 1% and 10% on real-world database FRGC Ver2.0 Exp. 4.

Preprocessing methods	VRs (%) at different FARs		
	0.1%	1.0%	10%
ORI	40.1	67.9	93.5
RLS	32.9	54.2	79.6
LTV	28.5	56.2	86.6
TT	44.5	69.1	91.1
LDCT	32.3	58.6	84.9
SOP	46.2	72.2	93.7

Ver2.0 Exp. 4 is quite challenging due to uncontrolled lighting conditions as well as sensor or photon noise, defocus, motion blur and even small pose and expression variations. According to the protocol for FRGC Ver2.0 Exp. 4, there are respectively 12,776, 16,028 and 8,014 face images in the training, target and query sets. All the face images in FRGC Ver2.0 Exp. 4 are also normalized in the same way as that in Multi-PIE and Extended YaleB face databases.

Illumination preprocessing is first performed on FRGC Ver2.0 Exp. 4 dataset using different preprocessing methods, and then face verification with Fisherfaces is conducted on the preprocessed face images. As shown in Table 3, Verification Rates (VR) at different False Acceptance Rates (FAR) are reported for comparison. It can be noticed that existing preprocessing methods may even degrade the face verification performance. By contrast, the proposed SOP shows to be more effective in improving face verification performance on FRGC Ver2.0 Exp. 4. Our preprocessing method learns to enhance images by increasing the separability of different subjects' faces, but other variations caused by non-illumination factors such as noise, blur, etc., can also be suppressed to some degree. This should be the reason why our method also proves to be useful for uncontrolled conditions.

4.4 Combination with Other Recognition Methods

As mentioned in Sect. 1, illumination preprocessing method has the superiority of general purpose usage. Therefore, compared with feature extraction or model-based methods, preprocessing is not limited to work for only one specific feature. For example, the output of preprocessing is still a face image; thus, different features can further be extracted. Therefore, it would be helpful to verify that when the separability in original image space is improved, the separability in one specific feature space also increases. In this experiment, two more recognition methods besides Fisherfaces, e.g. LBP [4] and Eigenfaces [27], are utilized to verify the proposed SOP also works with different recognition methods. The choice of LBP and Eigenfaces covers both local-and global-based methods.

In our experiments, “uniform” pattern is utilized for LBP and 1,200 dimensions except the first ten are used for Eigenfaces. Face verification using LBP and Eigenface is performed on the preprocessed images with SOP and the performance is compared with that without illumination preprocessing. The VRs

Table 4. Performance comparison of different preprocessing methods for face verification using LBP and Eigenfaces.

Preprocessing methods	VRs (%) at different FARs (LBP/Eigenfaces)		
	0.1%	1%	10%
ORI	7.4/6.7	20.9/18.4	52.6/49.9
RLS	5.4/8.9	16.1/24.6	45.3/51.4
LTV	2.3/10.9	8.4/26.9	33.0/52.8
TT	8.9/11.6	22.0/27.1	54.0/55.9
LDCT	4.4/12.0	15.3/29.6	50.3/55.3
SOP	11.7/14.5	28.6/31.6	62.6/58.2

at different FARs are presented in Table 4. As shown in Table 4, the proposed SOP also shows to be more effective for LBP and Eigenfaces than state-of-the-art preprocessing methods. This experiment verifies that various face matching algorithms tend to achieve higher performance when the separability of different faces is improved with the proposed SOP. Therefore, performance resembling state-of-the-art performance on FRGC Ver2.0 Exp. 4 can be reasonably expected by combining our method with the best face recognition methods.

5 Conclusions

Instead of roughly decomposing facial features into a large-scale and small-scale band pair, the proposed SOP makes a fine decomposition of facial features in scale-space. Moreover, the proposed SOP establishes an explicit relationship between illumination preprocessing and separability improvement of faces. The weights learnt for different feature bands in SOP reveal that features of different bands play different roles in face recognition. However the importance of a specific feature band varies across different lighting conditions. In the proposed approach, weights of different feature bands are estimated for varying lighting conditions by maximizing the Fisher separability criterion.

Experiments on Multi-PIE and Extended YaleB reveal that the proposed SOP is superior to existing methods in preprocessing laboratory-controlled lighting. Evaluation with cross-database training shows that SOP has satisfying generalization ability. Experiments on FRGC Ver2.0 Exp. 4 further show the effectiveness of the proposed SOP in handling real-world illumination mixed with defocus, blur and slight pose or expression variations. The proposed method learns to enhance images by increasing the separability of different subjects' faces. Thus, SOP is also able to suppress other variations (such as pose, noise, and blur) than illumination to some degree. Face recognition using Fisherfaces, LBP, and Eigenfaces shows that when the separability in original image space is improved, the separability in a specific feature space also tends to improve. The effectiveness of proposed approach in conjunction with more face matchers, e.g. sparse representation-based classification (SRC) [28], will be investigated in the future work.

Acknowledgments. This work is partially supported by National Basic Research Program of China (973 Program) under contract 2009CB320902; Natural Science Foundation of China under contracts Nos. 61025010, U0835005, 60832004, and 60902068; and Beijing Natural Science Foundation (New Technologies and Methods in Intelligent Video Surveillance for Public Security) under contract No. 4111003. Shihong Lao is partially supported by “R&D Program for Implementation of Anti-Crime and Anti-Terrorism Technologies for a Safe and Secure Society”, Special Coordination Fund for Promoting Science and Technology of MEXT, the Japanese Government.

References

1. Li, S.Z., Jain, A.K. (eds.): Handbook of Face Recognition, 2nd edn. Springer, London (2011)
2. Chen, H.F., Belhumeur, P.N., Jacobs, D.W.: In Search of Illumination Invariants. In: IEEE CVPR, pp. 1254–1261. IEEE Press, South Carolina (2000)
3. Shashua, A., Riklin Raviv, T.: The Quotient Image: Class-Based Re-Rendering and Recognition with Varying Illuminations. IEEE Trans. PAMI 23, 129–139 (2001)
4. Ojala, T., Pietikäinen, M., Mäenpää, T.: Multiresolution Gray-scale and Rotation Invariant Texture Classification with Local Binary Patterns. IEEE Trans. PAMI 24, 971–987 (2002)
5. Wang, H., Li, S., Wang, Y.: Face Recognition under Varying Lighting Conditions Using Self Quotient Image. In: IEEE FG, pp. 819–824. IEEE Press, Seoul (2004)
6. Tan, X., Triggs, B.: Enhanced Local Texture Feature Sets for Face Recognition under Difficult Lighting Conditions. IEEE Trans. Image Process 19, 1635–1650 (2010)
7. Biswas, S., Aggarwal, G., Chellappa, R.: Robust Estimation of Albedo for Illumination-Invariant Matching and Shape Recovery. IEEE Trans. PAMI 31, 884–899 (2009)
8. Belhumeur, P.N., Hespanha, J.P., Kriegman, D.J.: Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection. IEEE Trans. PAMI 19, 711–720 (1997)
9. Georgiades, A.S., Belhumeur, P.N., Kriegman, D.J.: From Few to Many: Illumination Cone Models for Face Recognition under Variable Lighting and Pose. IEEE Trans. PAMI 23, 643–660 (2001)
10. Basri, R., Jacobs, D.W., Kriegman, D.J.: Lambertian Reflectance and Linear Subspaces. IEEE Trans. PAMI 25, 218–233 (2003)
11. Zhang, L., Samaras, D.: Face Recognition from a Single Training Image under Arbitrary Unknown Lighting Using Spherical Harmonics. IEEE Trans. PAMI 28, 351–363 (2006)
12. Zhou, S.K., Aggarwal, G., Chellappa, R., Jacobs, D.W.: Appearance Characterization of Linear Lambertian Objects, Generalized Photometric Stereo, and Illumination-Invariant Face Recognition. IEEE Trans. PAMI 29, 230–245 (2007)
13. Wang, Y., Liu, Z., Hua, G., Wen, Z., Zhang, Z., Samaras, D.: Face Re-Lighting from a Single Image under Harsh Lighting Conditions. In: IEEE CVPR, pp. 1–8. IEEE Press, Minnesota (2007)
14. Kim, J.Y., Kim, L.S., Hwang, S.H.: An Advanced Contrast Enhancement Using Partially Overlapped Sub-Block Histogram Equalization. IEEE Trans. CSVT 11, 475–484 (2001)

15. Solar, J., Navarrete, P.: Eigenspace-Based Face Recognition: A Comparative Study of Different Approaches. *IEEE Trans. SMC:C* 35, 315–325 (2005)
16. Chen, W., Er, M.J., Wu, S.: Illumination Compensation and Normalization for Robust Face Recognition Using Discrete Cosine Transform in Logarithm Domain. *IEEE Trans. SMC:B* 36, 458–466 (2006)
17. Xie, X., Zheng, W., Lai, J., Yuen, P.C.: Face Illumination Normalization on Large and Small Scale Features. In: *IEEE CVPR*, pp. 1–8. IEEE Press, Alaska (2008)
18. Han, H., Shan, S., Qing, L., Chen, X., Gao, W.: Lighting Aware Preprocessing for Face Recognition across Varying Illumination. In: Daniilidis, K., Maragos, P., Paragios, N. (eds.) *ECCV 2010, Part II. LNCS*, vol. 6312, pp. 308–321. Springer, Heidelberg (2010)
19. Ramamoorthi, R.: Analytic PCA Construction for Theoretical Analysis of Lighting Variability in Images of a Lambertian Object. *IEEE Trans. PAMI* 24, 1–12 (2002)
20. Chan, T.F., Esedoglu, S.: Aspects of Total Variation Regularized L1 Function Approximation. *SIAM J. Appl. Math.* 65, 1817–1837 (2005)
21. Chen, T., Yin, W., Zhou, X.S., Comaniciu, D., Huang, T.S.: Total Variation Models for Variable Lighting Face Recognition. *IEEE Trans. PAMI* 28, 1519–1524 (2006)
22. Beveridge, J.R., Draper, B.A., Chang, J., Kirby, M., Kley, H., Peterson, C.: Principal Angles Separate Subject Illumination Spaces in YDB and CMU-PIE. *IEEE Trans. PAMI* 31, 351–356 (2009)
23. Gross, R., Matthews, I., Cohn, J., Kanade, T., Baker, S.: Multi-PIE. *Image and Vision Computing* 28, 807–813 (2010)
24. Duda, R.O., Hart, P.E., Stork, D.G.: *Pattern Classification*, 2nd edn. Wiley Interscience, New York (2001)
25. Kirkpatrick, S., Gelatt, C.D., Vecchi, M.P.: Optimization by Simulated Annealing. *Science* 220, 671–680 (1983)
26. Phillips, P.J., Flynn, P.J., Scruggs, T., Bowyer, K.W., Chang, J., Hoffman, K., Marques, J., Min, J., Worek, W.: Overview of the Face Recognition Grand Challenge. In: *IEEE CVPR*, pp. 947–954. IEEE Press, San Diego (2005)
27. Turk, M., Pentland, A.: Eigenfaces for Recognition. *J. Cognit. Neurosci.* 3, 71–86 (1991)
28. Wright, J., Yang, A.Y., Ganesh, A., Sastry, S.S., Ma, Y.: Robust Face Recognition via Sparse Representation. *IEEE Trans. PAMI* 31, 210–227 (2009)