

# Spatial Feature Interdependence Matrix (SFIM): A Robust Descriptor for Face Recognition

Anbang Yao<sup>1</sup> and Shan Yu<sup>2</sup>

<sup>1</sup> National Laboratory of Pattern Recognition, Institute of Automation,  
Chinese Academy of Science, Beijing, 100090, China

<sup>2</sup> National Institute for Research in Computer Science and Control, France  
abyao@nlpr.ia.ac.cn, shanyu329@yahoo.com

**Abstract.** In this paper, a new face descriptor called spatial feature interdependence matrix (SFIM) is proposed for addressing representation of human faces under variations of illumination and facial expression. Unlike traditional face descriptors which usually use a hierarchically organized or a sequentially concatenated structure to describe the spatial arrangement of features in different facial regions, SFIM is focused on exploring inherent spatial feature interdependences among separated facial regions in a face image. We compute the feature interdependence strength between each pair of facial regions as the Chi square distance between two corresponding histogram based feature vectors. Once face images are represented as SFIMs, we then employ spectral regression discriminant analysis (SRDA) to achieve face recognition under a nearest neighbor search framework. Extensive experimental results on two well-known face databases demonstrate that the proposed method has superior performance in comparison with related approaches.

**Keywords:** Face recognition, spatial feature interdependence matrix, spectral regression discriminant analysis, object representation.

## 1 Introduction

Urged by the fact that human face is one of the most potential physiological biometrics [1] for many applications such as public security, surveillance, human computer interaction (HCI) and multimedia, automatic face recognition has been an active research area in computer vision community for over three decades [2], [3], [4]. From the perspective of practical application, a desirable 2-D image based face recognition system should be the one that is able to identify or verify a human face under variations of facial expression, illumination, pose and occlusion in an accurate and efficient manner. To this end, a myriad of approaches have been proposed so far. However, face recognition under above challenges is still far from being effectively resolved [3], [5], [6]. One critical reason that prevents this from happening is the lack of reliable and generic methods to represent face instead of image data itself. The fact of the matter is that the information contained in source image is usually highly redundant or non-discriminative for face recognition regardless of feature extraction.

A lot of seminal research works have already been done on the issue of how to build up an effective and compact face representation from 2-D image feature space. The current state of the art approaches differ vastly in terms of principles and techniques applied. Following a high-level categorization guideline suggested from the psychological studies on how human recognize objects, the available face representation approaches can be classified into two categories—on the basis of holistic information and on the basis of local features [3], [7], [8], [9].

The category of holistic face representation approaches is characterized by a family of subspace methods originated from the Eigenface approach [10], which employs principal component analysis (PCA) to project high dimensional face feature vectors onto a significantly low dimensional feature space. By probing the projection directions that maximize the total scatter of all labeled face data of the same person, the underlying Euclidean space structure is discovered. Different from PCA, linear discriminant analysis (LDA) [11] pays particular attention to the discrimination between face classes in a linear separable space without the prerequisite of orthogonal bases. Although PCA and LDA are the two most popular techniques for face recognition, they cannot recover the non-linear structure of data set. To address this problem, the Laplacianface method [12] employs locality preserving projections (LPP) to find an embedding that preserves local information, and obtains a face subspace that best detects the essential face manifold structure. Similar to the approaches of [10], [11] and [12], other popular holistic face representation approaches including kernel based methods [13], [14], independent component analysis (ICA) [15], 2-dimensional principal component analysis (2-D PCA) [16] and so on, are also designed to address the problem of linear/nonlinear dimensionality reduction. Apart from these traditional holistic representation approaches [10], [11], [12], [13], [14], [15], [16], a probe face image can also be expressed as various linear combinations of gallery set [17], [18], [19]. In summary, all above mentioned holistic face representation methods directly use typical intensity images of human face as the inputs. That is, they mainly focus on seeking a linear/non-linear subspace or a linear combination that can best represent the structure of face data in favor of the class with minimal reconstructive error, while pay less attention to the fact that gray-scale images are generally not discriminative enough owing to the effect of factors such as redundancy, rotation, noise, lighting and albedo.

The category of local face representation approaches has recently also gained attention due to its robust capability to handle difficulties such as rotation and lighting. In general, this category approach benefiting from local information assumes that individual features extracted from prominent facial regions (e.g., eyes, nose, chin and mouth) are more vital to face recognition than the identification of holistic information. One of the pioneering works is elastic bunch graph matching (EBGM) [20], which describes faces using Gabor filter responses in 25 facial landmarks and uses a graph structure to represent the spatial locations of these landmarks. A modified EBGM algorithm is presented in [21] where the Gabor features in all facial landmarks are replaced by the histograms of oriented gradients (HOG) [22]. Instead of using a graph structure to represent face, the authors of [23] consider face images as a composition of micro patterns over small regions and take spatially weighted local binary pattern (LBP) histograms as the descriptors to represent face images. Later on, local Gabor binary pattern histogram sequence (LGBPHS) [24] and

histogram of Gabor phase patterns (HGPP) [25] are proposed for robust face recognition by the same research group. LGBPHS is intended to use the magnitude parts of Gabor filter and LBP operator simultaneously, while HGPP is mainly designed to jointly encode local and global Gabor phase patterns. In the approaches of [23], [24] and [25], a face image is divided into different regions, from which respective histograms are independently extracted and further concatenated into an extended histogram vector to represent the target face. It is worth noting that these approaches put more emphasis on the spatial arrangement of features in facial regions yet pay less attention to the spatial feature interdependences among different facial regions.

In this paper, we explore inherent spatial feature interdependence between any two different facial regions in a face image. By encoding all pair-wise spatial feature interdependence strengths over separated facial regions inside a face image, we propose a new face descriptor called spatial feature interdependence matrix (SFIM). In contrast to previous face descriptors [23], [24], [25], SFIM explicitly depicts an inherent spatial feature interdependence network among the facial parts of a face image. Furthermore, it provides a bridge to the association of sophisticated learning approaches [10], [11], [12], [13], [14], [15], [16], [17], [18], [19] and discriminative local features [22], [23], [24], [25]. In our approach, the feature interdependence between two different facial regions is estimated as the Chi square distance between a corresponding histogram vector pair. Once face images are represented as SFIMs, spectral regression discriminant analysis (SRDA) [26] is further employed to achieve face recognition under a nearest neighbor search framework. Comparative experiments are confined to frontal human face recognition, and the variations of illumination and facial expression are mainly addressed. Extensive experimental results on the extended Yale B [27] and the cropped AR [28] face databases validate the efficacy of the proposed approach.

The rest of this paper is organized as follows. Section 2 presents a detailed description of our SFIM based face descriptor. Section 3 describes the face recognition algorithm. Section 4 presents the comparative experiments on two publicly available face databases. Section 5 summarizes the paper and makes an outlook of possible future extensions.

## 2 Spatial Feature Interdependence Matrix

In this section, we will define the concept of SFIM, describe SFIM based descriptor for face object and show the properties of SFIM.

### 2.1 Definition of SFIM

The idea of SFIM is inspired by the following two facts. First, a standard human face is a whole unit consisting of different facial part structures. A successful application example of this fact is the well-known facial action coding system (FACS) [29], in which 46 different action units are defined to account for changes in facial expression. Second, as for human visual system, face recognition is not only dependent on visual information extracted from prominent facial parts (e.g., eyebrows, eyes, nose, mouth

and chin) but also assisted by latent information contained in the non-prominent facial parts [2], [3], [5], [18]. The aim of our proposed SFIM is not focused on precise segmentation of facial parts but on exploring spatial feature interdependence between any two different facial regions in a face image. We believe that the spatial feature interdependence between each pair of facial regions is a potential cue to identify face images of a certain person, and a careful handling of spatial feature interdependences may provide a completely new face representation approach. In accordance with above description, SFIM is defined as follows.

Let  $I$  represent a face image containing  $M$  separated regions  $\{R_1, R_2, \dots, R_M\}$ , the SFIM of  $I$  is defined as a square symmetric matrix of size  $M \times M$ .

$$A = [a_{ij}]_{M \times M} = \begin{bmatrix} 0 & a_{12} & a_{13} & \cdots & a_{1M} \\ a_{21} & 0 & a_{23} & \cdots & a_{2M} \\ a_{31} & a_{32} & 0 & \cdots & a_{3M} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{M1} & a_{M2} & a_{M3} & \cdots & 0 \end{bmatrix}, \quad (1)$$

where  $a_{ij}$  is the feature interdependence between facial regions  $R_i$  and  $R_j$ ,  $0 \leq a_{ij} \leq 1$ ,  $1 \leq i, j \leq M$ . The diagonal corresponds to the feature interdependence of a facial region itself, and it is composed of zeros.

## 2.2 SFIM Based Face Descriptor

From equation (1), it is clear the exact patterns of SFIM directly rely on the features and measures used to compute  $a_{ij}$ . For easy implementation, in our approach, face images are divided into spatially non-overlapped rectangular facial regions of same size. As the facial regions  $\{R_1, R_2, \dots, R_M\}$  have been determined, a feature vector is computed in each facial region independently. Various features are available in the literature for describing an image region of interest. We employ widely used histogram features to describe each rectangular facial region mainly due to two reasons. First, as the quantized and compact distributions of particular contents (e.g., intensity, gradient, phase, texture and high order filter response) in an image region of interest, histogram features have already been shown to be robust to noise, local image transformation, partial occlusion, etc. Second, histogram features can be calculated in a highly efficient way.

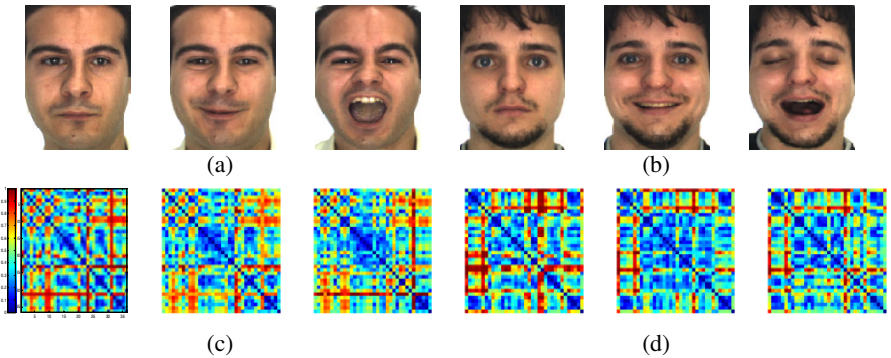
Let  $h_i$  and  $h_j$  be the normalized N-bin histograms of facial regions  $R_i$  and  $R_j$ , respectively, where  $\sum_{n=1}^N h_i(n) = 1$  and  $\sum_{n=1}^N h_j(n) = 1$ . The corresponding feature interdependence between facial regions  $R_i$  and  $R_j$  is computed as

$$a_{ij} = \sum_{n=1}^N \frac{(h_i(n) - \bar{h}(n))^2}{\bar{h}(n)}, \quad (2)$$

where

$$\bar{h}(n) = \frac{h_i(n) + h_j(n)}{2}. \quad (3)$$

In the literature, above defined interdependence is known as Chi square distance between a histogram feature vector pair. It measures how unlikely histogram distribution  $h_j$  is drawn from the population represented by histogram distribution  $h_i$ .



**Fig. 1.** The discriminative capability of SFIM on different face objects

Fig. 1 shows some examples to demonstrate the power of SFIM for representing faces under different facial expressions. The source images of two persons in three different facial expressions are displayed in Fig. 1(a) and Fig. 1(b), respectively. The corresponding SFIMs are shown in Fig. 1(c) and Fig. 1(d), respectively. The source images are taken from the cropped AR face database [28]. To compute corresponding SFIMs, the converted gray-scale images are divided into  $6 \times 6$  spatially non-overlapped rectangular facial regions, and 32-bin intensity histogram is used to describe each facial region. Note that the patterns of the resulting SFIMs of each person are relatively similar across different facial expressions, but are different across two persons. These results preliminarily show that the proposed SFIM can effectively handle the difficulties resulted from facial expression changes. The detailed experiments described in section 4 will further show the effectiveness of our SFIM based face descriptor, and the choice of dimensionality of SFIM is also clarified in section 4.

### 2.3 Properties of SFIM

The above section presents the definition of SFIM and describes the approach to represent a face image as an SFIM. In this section, we will shed more light on the properties of our SFIM based face descriptor.

First, considering each facial region as a node, the most characteristic property of SFIM is that it depicts a feature interdependence network among the separated facial regions in a face image. That means all pair-wise spatial feature inconsistencies are explicitly encoded via SFIM. Consequently, different from traditional face descriptors which usually employ a hierarchically organized [20], [21] or a sequentially concatenated structure [23], [24], [25] to describe the spatial arrangement of features in different facial regions, SFIM is designed to make a careful handling of inherent spatial feature interdependence between any two different facial regions in a face image. Second, SFIM has a form of square symmetric matrix and its entries are computed from histogram features in different facial region pairs, which make the resulting SFIMs of face images can be easily used as the inputs of any sophisticated learning approaches [10], [11], [12], [13], [14], [15], [16], [17], [18], [19] to achieve face recognition goal. In this paper, although we employ histogram features and Chi square distance to calculate the entries of SFIM, other features (e.g., covariance matrices [30] and Gabor based region covariance matrices [31]) and measures (e.g., Kullback-Leibler divergence and earth mover's distance) can also be used to compute SFIM. Therefore, the proposed SFIM provides a bridge to the association of sophisticated learning approaches and discriminative local features. Third, according to the description provided in section 2.2, it is clear that the computation of square symmetric SFIM is intrinsically efficient. Finally, the notations of video self-similarity matrices [32], [33], [34] used for describing periodic motion in temporal space are most closely related to our proposed SFIM in form. However, the main focus of SFIM is on the use of a structured layout of spatial feature interdependences over all pairs of facial regions for face representation. To our knowledge, this problem is not clearly addressed in available face representation methods [35], [36], [37], [38].

## 3 SFIM Based Face Recognition

Once every face image is represented as an SFIM with its entries computed as Chi square distances between histogram vectors extracted from any pair of non-overlapped rectangular facial regions, face recognition can be achieved in various ways. One possible way is to use a matrix measure (e.g., matrix cosine similarity and Frobenius norm) to directly compare the dissimilarity between a probe SFIM and every gallery SFIM. Another possible way is to use the SFIMs as the inputs of a traditional subspace method [10], [11], [12], [13], [15], [26] or a linear optimization approach [17], [18], [19]. Considering that SRDA [26] casts discriminant analysis into a regression framework that facilitates efficient computation, we use it as a test case to evaluate the performance of SFIM based face descriptor under a nearest neighbour search framework. Recall that SFIM is a square symmetric matrix whose diagonal entries are zeros, the lower/upper triangular entries of SFIM are discarded.

Let us consider a gallery set of  $N$  normalized face images belonging to  $C$  classes, and assume that  $n_k$  represents the number of training images of the  $k$ th face class

( $\sum_{k=1}^C n_k = N$ ). The face recognition algorithm is described as follows.

1. Compute the SFIM of every gallery face image according to section 2.
2. Rearrange the upper triangular entries (excluding diagonal) of each SFIM into a vector. Let  $x_1, x_2, \dots, x_N \in \mathfrak{R}^{(M^2-M)/2}$  be the resulting vector set, where  $M \times M$  is the dimensionality of SFIM.

3. Let  $y_k = \left[ \underbrace{0, \dots, 0}_{\sum_{i=1}^{k-1} n_i}, \underbrace{1, \dots, 1}_{n_k}, \underbrace{0, \dots, 0}_{\sum_{i=k+1}^C n_i} \right]^T, k=1, \dots, C$  and  $y_0 = [1, 1, \dots, 1]_{1 \times N}^T$ , find  $C-1$  basis vectors  $\{v_k\}$  by solving

$$v_k = \arg \min_v \left( \sum_{i=1}^N (v^T x_i - y_i^k)^2 + \alpha \|v\|^2 \right), \quad (4)$$

where  $v_k \in \mathfrak{R}^{(M^2-M)/2}$ ,  $y_i^k$  is the  $i$ th element of Gram-Schmidt orthogonalized  $y_k$ .

4. Embed each training  $x_i$  into the  $C-1$  dimensional subspace by

$$z_i = V^T \begin{bmatrix} x_i \\ 1 \end{bmatrix}, \quad V = \begin{bmatrix} v_1, v_2, \dots, v_{C-1} \\ 1, 1, \dots, 1 \end{bmatrix}. \quad (5)$$

5. Given a probe face image, compute corresponding SFIM according to step 1 and rearrange it into a vector  $x$  according to step 2.
6. Get the embedding  $z$  of  $x$  by equation (5).
7. Compute the Euclidean distance between  $z$  and each  $z_i$ , and take the class of the gallery face image with the minimum distance to label probe face image.

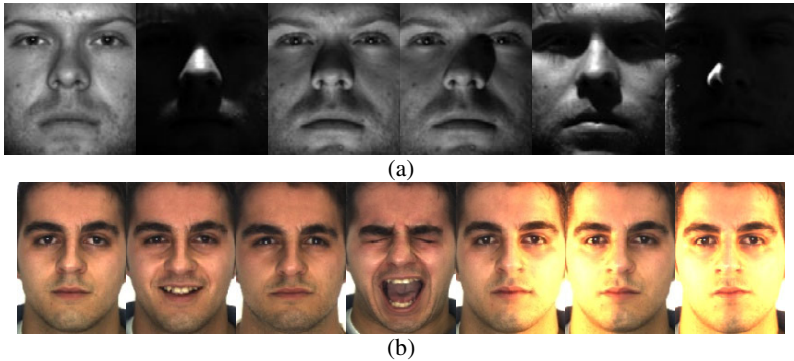
## 4 Experiments and Results

In this section, extensive experiments on two challenging face databases (i.e., the extended Yale B [27] and the cropped AR [28] face databases) are carried out to demonstrate the efficacy of the proposed SFIM. In the experiments, two kinds of histogram feature (i.e., 32-bin intensity histogram and histogram of LBP) are employed to calculate SFIM. A comprehensive comparison of the performance of our approach and the published results on each test database is presented. We want to point out here that the reported results grouped in [18] (including Eigenface [10], Fisherface [11], Laplacianface [12], SVM+Laplacianface [14] and sparse representation [18]) are shown for a large range of feature dimensions. For the sake of fair comparison, we just pick the best results. The results of spatially weighted LBP

[23], LGBPHS [24] and the histogram of monogenic binary pattern (HMBP) [39] on the cropped AR face database are collected from [39]. Considering that we embed SFIM based descriptor in SRDA algorithm to achieved face recognition, the results of standard SRDA [26] are also presented. As for standard SRDA, the results are computed for an exact same range of feature dimensions to that of [18], and only the best results are shown for comparison (see the supplementary file for details).

#### 4.1 Experiments on the Extended Yale B Face Database

The extended Yale B face database contains 2414 images of 38 subjects showing 1 frontal pose under 64 laboratory-controlled illumination conditions. Some examples are shown in Fig. 2(a). The cropped and normalized gray-scale image has a resolution of  $192 \times 168$  pixels. For each subject, half of the images are selected for training while the other half are adopted as testing dataset. All these parameters are same to those of the published reference results grouped in [18].



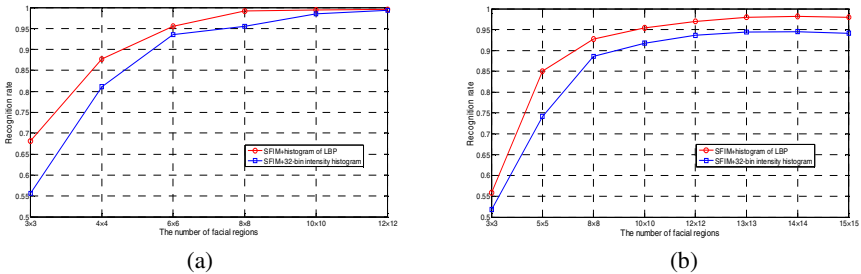
**Fig. 2.** Examples of representative subjects from: (a) the extended Yale B face database; (b) the cropped AR face database

**Table 1.** Comparison of the top recognition accuracy of different approaches

Methods	Recognition rate on the extended Yale B face database (%)	Recognition rate on the cropped AR face database (%)
Eigenface [10]	88.40	80.50
Fisherface [11]	87.60	86.80
Laplacianface [12]	90.70	89.70
SVM+Laplacianface [14]	97.70	95.70
Sparse representation [18]	98.26	94.99
Spatially weighted LBP [23]	N/A	97.71
LGBPHS [24]	N/A	97.29
HMBP [39]	N/A	98.57
Standard SRDA [26]	94.38	78.71
SFIM+32-bin intensity histogram	99.34	94.52
SFIM+histogram of LBP	99.59	98.71



The results on this database are shown in Table 1. From Table 1, it can be seen that the Fisherface method is worse than the others. This is partially due to the fact that the maximal number of valid Fisherfaces is one less than the number of face classes [11], [18]. Although the approaches of [10], [12] outperform the Fisherface method, their recognition rates are relatively low. The algorithm presented in [14] achieves a high recognition accuracy of 97.7%. More high result is reported for the sparse representation approach [18]. Taking the SFIMs computed from 32-bin intensity histogram and histogram of LBP as face descriptors, our approach yields recognition rates of 99.34% and 99.59%, respectively. These results are 1.08 and 1.33 percent better than the best result of comparison approaches, respectively. Furthermore, by utilizing intensity histogram based feature interdependence between each facial region pair in a face image, our approach outperforms standard SRDA (whose inputs are normalized intensity images) by 4.96 percent in accuracy. These results fairly verify our assumption that the spatial feature interdependence between each pair of facial regions is a potential cue to identify face images in difficult scenarios. Therefore, by properly encoding all pair-wise spatial histogram based feature interdependences over different facial regions inside a class-specific face image, SFIM can handle the difficulty resulted from illumination changes in a more effective way. The choice of dimensionality of SFIM on the extended Yale B face database is shown in Fig. 3(a). With respect to the SFIM based face descriptors computed from histogram of LBP, it can be seen that the recognition rate of our approach becomes constant ( $>99\%$ ) in the face partition range of  $8\times 8\sim 12\times 12$  rectangular regions. As for the other case, stable results ( $>98.5\%$ ) are achieved in the face partition range of  $10\times 10\sim 12\times 12$  rectangular regions.



**Fig. 3.** The choice of the dimensionality of SFIM on: (a) the extended Yale B face database; (b) the cropped AR face database

## 4.2 Experiments on the Cropped AR Face Database

The cropped AR face database consists of 2600 color images corresponding to 100 subjects (50 men and 50 women). Different facial expressions (neutral, smile, anger and surprise), illumination conditions (left light on, right light on and all lights on) and occlusions (glass, scarf, etc.) are included. Each subject participates in two sessions separated by two weeks. The same images are taken in both sessions. For each subject, 14 images with illumination and expression variations are considered in the experiments. The 7 images from session 1 are selected for training, and the other

7 images from session 2 are chosen for testing. The cropped  $165 \times 120$  color images (as shown for 7 example images in Fig. 2(b)) are converted to gray-scale images. Same to section 4.1, all above parameters are identical to those of the published results reported in [18], [39].

The results on this database are shown in Table 1. Similar to the results on the extended Yale B face database, the approaches of Eigenface, Fisherface and Laplacianface exhibit relatively low recognition rates, while the approaches of [14] and [18] demonstrate better performance. More high results are reported for the approaches of [23], [24], [39]. As for the SFIM based face descriptors computed from histogram of LBP, our approach yields the best recognition rate of 98.71%. With regard to the SFIM based face descriptors computed from 32-bin intensity histogram, the recognition rate of our approach is 94.52%. Compared with the results of standard SRDA, the results of our approach are 20 and 15.81 percent better, respectively. These results fairly show that typical intensity images of human face are generally not discriminative enough for face recognition in complex scenarios. Since the inherent spatial feature inconsistencies between any two different facial regions of a class-specific face image are encoded via SFIM, our approach demonstrates better capability to deal with the variations of facial expression and illumination. The choice of dimensionality of SFIM on the cropped AR face database is shown in Fig. 3(b). Note that the recognition rate of our approach becomes stable in the face partition range of  $13 \times 13 \sim 15 \times 15$  rectangular regions ( $>98\%$  and  $>94\%$  corresponding to two types of SFIM, respectively).

## 5 Conclusions

This paper presents a new face descriptor called SFIM. SFIM takes advantage of the form of a square symmetric matrix to explicitly encode inherent spatial feature interdependence between any two separated facial regions in a face image. We first use SFIMs computed from histogram features and Chi square distance as the descriptors to represent face images. Subsequently, SRDA is used as a test case to achieve efficient face recognition under a nearest neighbor search framework. Extensive experiment results on two well-known benchmark face databases verify the efficacy of the proposed SFIM.

Our further work will be extended in two directions. First, how to incorporate more class-specific information into SFIM is an open issue. Second, SFIM can be generalized for representing other objects needed in applications such as object detection, multi-class object classification and content based image retrieval (CBIR).

## References

1. Jain, A.K., Ross, A., Parbhakar, S.: An Introduction to Biometric Recognition. IEEE Trans. Circ. Syst. Video Tech. 14(1), 4–20 (2004)
2. Samal, A., Iyengar, P.: Automatic Recognition and Analysis of Human Faces and Facial Expressions: A Survey. Pattern Recognition 25(1), 65–77 (1992)
3. Zhao, W., Chellappa, R., Phillips, P.J., Rosenfeld, A.: Face Recognition: A Literature Survey. ACM Computing Surveys 35(4), 399–458 (2003)

4. Bowyer, K.W., Chang, K., Flynn, P.: A Survey of Approaches and Challenges in 3D and Multi-Modal 3D + 2D Face Recognition. *Computer Vision and Image Understanding* 101(1), 1–15 (2006)
5. Sinha, P., Balas, B., Ostrovsky, Y., Russell, R.: Face Recognition by Humans: Nineteen Results All Computer Vision Researchers Should Know About. *Proceedings of the IEEE* 94(11), 1948–1962 (2006)
6. Zeng, Z., Pantic, M., Rosiman, G.I., Huang, T.S.: A Survey of Affect Recognition Methods: Audio, Visual, and Spontaneous Expressions. *IEEE Trans. Pattern Anal. Mach. Intell.* 31(1), 39–58 (2009)
7. Craw, I., Costen, N., Kato, T., Akamatsu, S.: How Should We Represent Faces for Automatic Recognition? *IEEE Trans. Pattern Anal. Mach. Intell.* 21(8), 725–736 (1999)
8. Tanaka, J.W., Farah, M.J.: Parts and Wholes in Face Recognition. *Quarterly Journal of Experiment Psychology* 46A(2), 225–245 (1993)
9. Tao, D., Tang, X., Li, X.: Which Components Are Important for Interactive Image Search? *IEEE Trans. Circ. Syst. Video Tech.* 18(1), 3–11 (2008)
10. Turk, M., Pentland, A.: Eigenfaces for Recognition. *Journal of Cognitive Neuroscience* 3(1), 71–86 (1991)
11. Belhumeur, P.N., Hespanha, J.P., Kriegman, D.J.: Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection. *IEEE Trans. Pattern Anal. Mach. Intell.* 19(7), 711–720 (1997)
12. He, X., Yan, S., Hu, Y., Niyogi, P., Zhang, H.: Face Recognition Using Laplacianfaces. *IEEE Trans. Pattern Anal. Mach. Intell.* 27(3), 328–340 (2005)
13. Scholkopf, B., Smola, A., Muller, K.R.: Nonlinear Component Analysis as a Kernel Eigenvalue Problem. *Neural Computation* 10(5), 1299–1319 (1998)
14. Liu, Q., Huang, R., Lu, H., Ma, S.: Face Recognition Using Kernel Based Fisher Discriminant Analysis. In: *5th IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 197–201 (2002)
15. Bartlett, M.S., Movellan, J.R., Sejnowski, T.J.: Face Recognition by Independent Component Analysis. *IEEE Trans. Neu. Net.* 13(6), 1450–1464 (2002)
16. Yang, J., Zhang, D., Frangi, A.F., Yang, J.: Two-dimensional PCA: A New Approach to Appearance-Based Face Representation and Recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* 26(1), 131–137 (2004)
17. Chai, X., Shan, S., Chen, X., Gao, W.: Locally Linear Regression for Pose-invariant Face Recognition. *IEEE Trans. Img. Proc.* 16(7), 1716–1725 (2007)
18. Wright, J., Yang, A.Y., Ganesh, A., Sastry, S.S., Ma, Y.: Robust Face Recognition via Sparse Representation. *IEEE Trans. Pattern Anal. Mach. Intell.* 31(2), 210–227 (2009)
19. Naseem, I., Togneri, R., Bennamoun, M.: Linear Regression for Face Recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* 32(11), 2106–2112 (2010)
20. Wiskott, L., Fellous, J., Kruger, N., von der Malsburg, C.: Face Recognition by Elastic Bunch Graph Matching. *IEEE Trans. Pattern Anal. Mach. Intell.* 19(7), 775–779 (1997)
21. Albiol, A., Monzo, D., Martin, A., Sastre, J., Albiol, A.: Face Recognition Using HOG-EBGM. *Pattern Recognition Letters* 29(10), 1537–1543 (2008)
22. Dalal, N., Triggs, B.: Histograms of Oriented Gradients for Human Detection. In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 886–893 (2005)
23. Ahonen, T., Hadid, A., Pietikäinen, M.: Face Description with Local Binary Patterns: Application to Face Recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* 28(12), 2037–2041 (2006)

24. Zhang, W., Shan, S., Gao, W., Chen, X., Zhang, H.: Local Gabor Binary Pattern Histogram Sequence (LGBPHS): A Novel Non-statistical Model for Face Representation and Recognition. In: 10th IEEE International Conference on Computer Vision, vol. 1, pp. 786–791 (2005)
25. Zhang, B., Shan, S., Chen, X., Gao, W.: Histogram of Gabor Phase Patterns (HGPP): A Novel Object Representation Approach for Face Recognition. *IEEE Trans. Img. Proc.* 16(1), 57–68 (2007)
26. Deng, C., He, X., Han, J.: SRDA: An Efficient Algorithm for Large-Scale Discriminant Analysis. *IEEE Trans. Knowledge Data Eng.* 20(1), 1–12 (2008)
27. Georghiades, A., Belhumeur, P., Kriegman, D.: From Few to Many: Illumination Cone Models for Face Recognition under Variable Lighting and Pose. *IEEE Trans. Pattern Anal. Mach. Intell.* 23(6), 643–660 (2001)
28. Martinez, A.M., Kak, A.C.: PCA versus LDA. *IEEE Trans. Pattern Anal. Mach. Intell.* 23(2), 228–233 (2001)
29. Ekman, P., Friesen, W.V.: Facial Action Coding System (FACS). Consulting Psychologists Press, Palo Alto (1978)
30. Tuzel, O., Porikli, F., Meer, P.: Region Covariance: A Fast Descriptor for Detection and Classification. In: Leonardis, A., Bischof, H., Pinz, A. (eds.) *ECCV 2006*. LNCS, vol. 3952, pp. 589–600. Springer, Heidelberg (2006)
31. Pang, Y., Yuan, Y., Li, X.: Gabor-Based Region Covariance Matrices for Face Recognition. *IEEE Trans. Circ. Syst. Video Tech.* 18(7), 989–993 (2008)
32. Benabdelkader, C., Cutler, R.G., Davis, L.S.: Gait Recognition Using Image Self-similarity. *EURASIP Journal on Applied Signal Processing* 2004(1), 572–585 (2004)
33. Shechtman, E., Irani, M.: Matching Local Self-Similarities across Images and Videos. In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 1–8 (2007)
34. Junejo, I.N., Dexter, E., Laptev, I., Pérez, P.: Cross-View Action Recognition from Temporal Self-Similarities. In: Forsyth, D., Torr, P., Zisserman, A. (eds.) *ECCV 2008*, Part II. LNCS, vol. 5303, pp. 293–306. Springer, Heidelberg (2008)
35. Stanffer, C., Grimson, E.: Similarity Templates for Detection and Recognition. In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 221–228 (2001)
36. Savarese, S., Winn, J., Criminisi, A.: Discriminative Object Class Models of Appearance and Shape by Correlations. In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2, pp. 2033–2040 (2006)
37. Eleyan, A., Demirel, H.: Co-Occurrence Based Statistical Approach for Face Recognition. In: 24th International Symposium on Computer and Information Sciences, pp. 611–615 (2009)
38. Watanabe, T., Ito, S., Yokoi, K.: Co-Occurrence Histograms of Oriented Gradients for Pedestrian Detection. In: Wada, T., Huang, F., Lin, S. (eds.) *PSIVT 2009*. LNCS, vol. 5414, pp. 37–47. Springer, Heidelberg (2009)
39. Yang, M., Zhang, L., Zhang, L., Zhang, D.: Monogenic Binary Pattern (MBP): A Novel Feature Extraction and Representation Model for Face Recognition. In: 20th International Conference on Pattern Recognition, pp. 2680–2683 (2010)