Int. J. Biometrics, Vol. 7, No. 3, 2015 191

Face recognition using a novel image representation scheme and multi-scale local features

Qing-Chuan Tao

School of Electronics and Information Engineering,
Sichuan University,
Chengdu, Sichuan, China
Email: taoqingchuan@scu.edu.cn

Zhi-Ming Liu*

Department of Electrical and Biomedical Engineering,
University of Nevada,
Reno, Nevada, USA
Email: zl9@njit.edu
*Corresponding author

George Bebis

Department of Computer Science and Engineering,
University of Nevada,
Reno, Nevada, USA
Email: bebis@cse.unr.edu

Muhammad Hussain

Department of Computer Science,
King Saud University,
Riyadh, Saudi Arabia
Email: mhussain@ccis.edu.sa

Abstract: This paper presents a new method for improving face recognition performance under difficult conditions. Specifically, a new image representation scheme is proposed which is derived from the YCrQ colour space using principal component analysis (PCA) followed by Fisher linear discriminant analysis (FLDA). A multi-scale local feature, LBP-DWT, is used for face representation which is computed by extracting different resolution local binary patterns (LBP) features from the new image representation and transforming the LBP features into the wavelet domain using discrete wavelet transform (DWT) and Haar wavelets. A variant of non-parametric discriminant analysis (NDA), called regularised non-parametric discriminant analysis (RNDA) is introduced to extract the most discriminating features from LBP-DWT. The proposed methodology has been evaluated using two challenging face databases (FERET and multi-PIE). The promising experimental results show that the proposed method outperforms two state-of-the-art methods, one based on Gabor features and the other based on sparse representation classification (SRC).

Keywords: face recognition; colour image; local binary patterns; LBP; discrete wavelet transform; DWT; Fisher linear discriminant analysis; FLDA; non-parametric discriminant analysis; NDA.

Reference to this paper should be made as follows: Tao, Q-C., Liu, Z-M., Bebis, G. and Hussain, M. (2015) 'Face recognition using a novel image representation scheme and multi-scale local features', *Int. J. Biometrics*, Vol. 7, No. 3, pp.191–212.

Biographical notes: Qing-Chuan Tao is an Associate Professor at the School of Electronics and Information Engineering, Sichuan University, China. He received his BS and MS in Electronics Engineering and PhD in Optics from Sichuan University, China. His research interests mainly include image processing, pattern recognition, information optics and biomedical image analysis, etc. He has published more than 30 articles in academic journals.

Zhi-Ming Liu is a Research Associate at the Department of Electrical and Biomedical Engineering, University of Nevada, Reno, USA. He received his BS and MS in Electronics Engineering from Sichuan University, China, and PhD in Computer Science from New Jersey Institute of Technology, USA. His research interests are in computer vision and pattern recognition, especially face recognition, object recognition, and gaze tracking.

George Bebis received his BS in Mathematics and MS in Computer Science from the University of Crete, Greece, in 1987 and 1991, respectively, and PhD in Electrical and Computer Engineering from the University of Central Florida, Orlando, in 1996. Currently, he is Professor and Chair of the Department of Computer Science and Engineering at the University of Nevada, Reno (UNR), Director of the UNR Computer Vision Laboratory (CVL). His research interests include computer vision, image processing, pattern recognition, machine learning, and evolutionary computing.

Muhammad Hussain is a Professor at the Department of Computer Science, King Saud University, Saudi Arabia. He received his MSc and MPhil, both from University of the Punjab, Lahore, Pakistan, in 1990 and 1993 respectively. In 2003, He received his PhD in Computer Science from Kyushu University, Fukuoka, Japan. His current research interests include image processing, pattern recognition and computer graphics. He is an Associate Editor of *Journal of Computer and Information Sciences*, King Saud University and has served on the program committees of various international conferences.

1 Introduction

Feature selection and feature extraction are crucial to many pattern classification problems (e.g., face recognition). The key objective is transforming the original data into a new domain where more effective features can be extracted, giving rise to significant differences between classes. Using a good set of features is vital for improving classifier performance. This paper presents a new method for face recognition which capitalises on fusing features across colour, spatial, and frequency domains for enhancing recognition performance. Specifically, the proposed method uses features extracted by employing three different modules: the image representation module, the image feature extraction module, and the statistical feature extraction module.

While popular methods in face recognition mainly focus on extracting features either from the image space [e.g., using techniques such as Gabor kernels (Daugman, 1985) and local binary patterns (LBP) (Ojala et al., 2002)], or from transformed spaces [e.g., using statistical methods such as PCA (Turk and Pentland, 1991) and FLDA (Belhumeur et al., 1997)], there is little work on studying the role of image representation in face recognition. Using an appropriate image representation scheme can be considered as an important feature extraction stage. Most face recognition methods start with an image representation that is a linear combination of the three primary colours, R, G, and B, for example, using the luminance image Y = 0.299R + 0.58G + 0.114B or the intensity image I = (R + G + B) / 3. However, more theoretical work and experimental results are in great demand to support that such image representation schemes are optimal for image analysis, especially for face recognition. Ideally, novel image representation schemes that improve class separability are highly desirable in order to simplify classifier design and improve face recognition performance. Motivated by recent work on fusing colour information to improve face recognition accuracy (Liu and Liu, 2008; Yang et al., 2010), we propose a new image representation scheme, which is derived from the YCrQ colour space using PCA followed by FLDA.

Local image features, such as Gabor and LBP, have demonstrated remarkable success in improving face recognition performance (Liu and Wechsler, 2002; Ahonen et al., 2006). Since Gabor filters are sensitive to different scales and orientations, Gabor features have been used to capture salient visual properties and cope with image variations such as illumination changes. Motivated by these studies, we propose a multiple-scale LBP feature that is computed by transforming the LBP features into the wavelet domain using Haar wavelets. This new type of feature, which will be referred to as LBP-DWT, extends the traditional single-resolution LBP feature to a multiple-scale space across both the spatial and frequency domains.

Image features, or low-level features, alone are not sufficient to deal with face recognition under complex conditions, due to the fact that image features are usually redundant and noisy. It is widely accepted that statistical learning techniques, such as PCA and FLDA, can find low-dimensional spaces that contain more effective features, offering enhanced discriminating ability and computational efficiency. Motivated by non-parametric discriminant analysis (NDA) (Fukunaga, 1990), large margin nearest neighbour (LMNN) classification (Weinberger and Saul, 2009), and eigenfeature regularisation (Jiang et al., 2008), we present a variant of NDA using regularisation (RNDA) to extract the most discriminating LBP-DWT features for enhancing face recognition performance.

The proposed face recognition approach has been extensively evaluated on two challenging face databases, FER-ET (Phillips et al., 2000) and CMU Multi-PIE (Gross et al., 2010). FERET images were acquired during different photo sessions where illumination conditions and face size may vary. CMU Multi-PIE images have large variations across pose, illumination, and expression (PIE). The experimental results show that the proposed method improves face recognition performance significantly, and outperforms two state-of-the-art methods, one based on Gabor features (Liu and Wechsler, 2002) and the other based on SRC (Wright et al., 2009).

The rest of the paper is organised as follows: Section 2 reviews methods related to our approach. Section 3 presents the novel image representation scheme while Section 4 introduces the LBP-DWT features and the regularised non-parametric discriminant analysis (RNDA) approach. Section 5 reports our experimental results and comparisons. Finally, Section 6 contains our conclusions and directions for future research.

2 Background

Face recognition has attracted a lot of attention recently due to the complexity of the problem and the enormous number of applications both in the commercial and government sectors (Bowyer et al., 2006; Jain et al., 2004; Zhao et al., 2003). Numerous methods have been proposed in the past two decades to address the problems of face recognition with large image variations due to illumination, expression, blurring, etc., which are often confronted in real-world face recognition applications.

Like any pattern recognition problem, face recognition relies heavily on the features employed for classification. Most state-of-the-art face recognition methods usually start with extracting image-level features that are desired to be invariant to image variations in order to facilitate subsequent statistics-level feature extraction for classification. Two popular image-level features used in face recognition are Gabor (Liu and Wechsler, 2002) and LBP (Ahonen et al., 2006). Gabor features model quite well the receptive field profiles of cortical simple cells (Daugman, 1985). Thus, they capture salient visual properties such as spatial localisation, orientation selectivity, and spatial frequency characteristic. The success of LBP in face recognition has been attributed to its robustness in terms of grey-level monotonic transformation since faces can be seen as a composition of micropatterns that are well described by this operator (Ahonenet al., 2006). In general, LBP features describe small face details while Gabor features capture facial shape at larger scales. The complementary information provided has been exploited for enhancing face recognition performance (Zhang et al., 2005; Liu and Liu, 2010; Tan and Triggs, 2010).

Methods using grey-level images have difficulties dealing with complex image variations, such as severe illumination changes, as evidenced in a recent face recognition grand challenge (FRGC) competition (Phillips et al., 2005). Instead of developing more sophisticated feature extraction methods and classifiers, a few researchers investigated the potential of using colour information to address some of the challenges in face recognition (Liu, 2008; Liu and Liu, 2008, 2010; Yang and Liu, 2008; Yang et al., 2010; Choi et al., 2011; Wang et al., 2011; Sultana and Garilova, 2014). Using colour information for face recognition has been motivated by the fact that, for any given colour face image, a natural yet powerful complementary image representation can be provided using multiple colour components that can enhance the diversity of misclassifications in statistical pattern recognition (Kittler et al., 1998). As a result, feature-level or decision-level fusion can be used to improve the performance of face recognition.

Although most researchers have ignored the role of colour information in face recognition, its promising effect on significantly boosting face recognition performance has been validated using several colour face databases, such as FRGC (Liu, 2008; Liu and Liu, 2008, 2010), FERET (Choi et al., 2011), PIE (Choi et al., 2011), XM2VTSDB (Choi et al., 2011), AR (Yang et al., 2010; Wang et al., 2011), and GT (Wang et al., 2011). Several researchers (Liu, 2008; Yang and Liu, 2008; Yang et al., 2010; Wang et al., 2011) have proposed creating new colour spaces that are more discriminative than existing ones for face recognition, by applying statistical learning methods such as PCA, FLDA, independent component analysis (ICA), and tensor-based discriminant analysis. On the other hand, it has been also demonstrated that performance gains in face recognition can be achieved by extracting different features from different colour spaces (Liu and Liu, 2008; Liu and Liu, 2010; Choi et al., 2011).

Low-level features, such as Gabor and LBP, usually reside in a high-dimensional space. As it is more likely to enhance the generalisation accuracy and robustness of classification in a lower dimensional space (Jiang, 2011), there is a great demand to find meaningful and compact patterns to improve the robustness of pattern recognition. Therefore, applying dimensionality reduction is an essential step for designing robust face recognition paradigms. Commonly used methods for dimensionality reduction include PCA and FLDA. PCA reduces a large set of correlated variables to a smaller number of uncorrelated components, which can be used for recognition or face reconstruction. On the other hand, FLDA can achieve higher separability than PCA among different classes by deriving a projection basis that separates data in different classes as much as possible and compresses data in each class as much as possible. A fundamental limitation of FLDA arises from the parametric nature of scatter matrices, whose computation relies on the assumption that the samples in each class follow a Gaussian distribution. To address the case of a non-Gaussian distribution, NDA (Fukunaga, 1990) has been proposed which computes the between-class scatter matrix using samples near the boundary between classes. The margin, which is derived from the boundary between classes, plays an important role in classification since the larger the margin the lower the generalisation error. A large margin between classes can be artificially maintained locally using a linear transformation to pull same labelled samples closer together and push different labelled samples further apart (Weinberger and Saul, 2009; Gu et al., 2010). As the high dimensionality of face image features and the small number of training samples usually cause the small sample size (SSS) problem, dimensionality reduction is considered as an important method to overcome this obstacle. It should be mentioned that in dimensionality reduction, the eigenspace of the within-class scatter matrix usually contains an unstable subspace due to noise and finite number of training samples. Eigenfeature regularisation based on an eigenspectrum model can be used to alleviate the problems of instability, overfitting, and poor generalisation (Jiang et al., 2008).

3 A novel image representation

Over the past two decades, much effort has been made on developing different methods in face recognition using grey-scale image representations. Except in face detection, the effectiveness of colour information in face recognition has been largely ignored. Recently, a few works have considered using the complementary characteristics residing in colour spaces to improve face recognition performance (Liu, 2008; Liu and Liu, 2008, 2010; Yang and Liu, 2008; Yang et al., 2010; Choi et al., 2011; Wang et al., 2011). While these methods improve recognition performance significantly, they also have higher computational complexity due to using multiple colour components. To provide a trade-off between performance and computational efficiency, we present an optimum colour space conversion scheme based on PCA and FLDA, namely a conversion from a 3-D colour space to a 2-D image space.

Assuming a $\mathbf{C}^1\mathbf{C}^2\mathbf{C}^3$ colour space, an image of resolution $m \times n$ consists of three colour components $\mathbf{C}^1$, $\mathbf{C}^2$, and $\mathbf{C}^3$. Without loss of generality, let $\mathbf{C}^1$, $\mathbf{C}^2$, and $\mathbf{C}^3$ be column vectors: $\mathbf{C}^1, \mathbf{C}^2, \mathbf{C}^3 \in \mathbb{R}^N$, where $N = m \times n$. Each vector $\mathbf{C}$ is normalised to have zero mean and unit variance. A data matrix $\mathbf{X} \in \mathbb{R}^{3\times Nl}$ can be formed using the training images:

$$\mathbf{X} = \begin{bmatrix} \mathbf{C}_1^1, & \mathbf{C}_1^2, & \mathbf{C}_1^3 \\ \mathbf{C}_2^1, & \mathbf{C}_2^2, & \mathbf{C}_2^3 \\ \vdots, & \vdots, & \vdots \\ \mathbf{C}_l^1 & \mathbf{C}_l^2, & \mathbf{C}_l^3 \end{bmatrix}^t \tag{1}$$

where l is the number of training images. In $\mathbf{X}$, each column is an observation and each row is a variable. The covariance matrix Σ_X may be formulated as $\Sigma_X = \frac{1}{Nl-1}\tilde{\mathbf{X}}\tilde{\mathbf{X}}^t \in \mathbb{R}^{3\times 3}$, where $\tilde{\mathbf{X}}$ is the centred data matrix. PCA is used to factorise Σ_X into the following form: $\Sigma_X = \Phi\Lambda\Phi^t$, where $\Phi = [\Phi_1, \Phi_2, \Phi_3]\mathbb{R}^{3\times 3}$ is an orthonormal eigenvector matrix and $\Lambda = diag\{\lambda_1, \lambda_2, \lambda_3\}\mathbb{R}^{3\times 3}$ a diagonal eigenvalue matrix with diagonal elements in decreasing order ($\lambda_1 \geq \lambda_2 \geq \lambda_3$).

By projecting the three colour components $\mathbf{C}_1$, $\mathbf{C}_2$, $\mathbf{C}_3$ of an image onto the eigenvector 1, a new image representation $\mathbf{U} \in \mathbb{R}^N$ can be obtained:

$$\mathbf{U} = \left[\mathbf{C}^1, \mathbf{C}^2, \mathbf{C}^3\right]\Phi_1 \tag{2}$$

According to the properties of PCA, $\mathbf{U}$ is optimal for data representation but not for data classification. To obtain a representation that improves discrimination, one needs to handle within- and between-class variations separately. FLDA is a well-known subspace learning method that yields an optimal subspace, separating different classes as far as possible and compressing same classes as compactly as possible. Thus, a more powerful image representation can be obtained based on $\mathbf{U}$ via the FLDA framework.

Let $\bar{\mathbf{U}}_i$ be the mean vector of class ω_i and $\bar{\mathbf{U}}$ be the grand mean vector. Then, the between- and within-class scatter matrices $\mathbf{S}_b$ and $\mathbf{S}_w$ are defined as follows.

$$\mathbf{S}_b = \sum_{i=1}^{k} P(\omega_i)\left(\bar{\mathbf{U}}_i - \bar{\mathbf{U}}\right)\left(\bar{\mathbf{U}}_i - \bar{\mathbf{U}}\right)^t \tag{3}$$

$$\mathbf{S}_w = \sum_{i=1}^{k} P(\omega_i)\varepsilon\left\{\left(\mathbf{U} - \bar{\mathbf{U}}_i\right)\left(\mathbf{U} - \bar{\mathbf{U}}_i\right)^t |\omega_i\right\} \tag{4}$$

where $P(\omega_i)$ is the prior probability of class ω_i, and k is the number of classes, and $\mathbf{S}_b, \mathbf{S}_w \in \mathbb{R}^{N\times N}$. The general Fisher criterion in the $\mathbf{U}$ image space can be defined as follows.

$$J(\mathbf{P}) = \frac{\left|\mathbf{P}^t\mathbf{S}_b\mathbf{P}\right|}{\left|\mathbf{P}^t\mathbf{S}_w\mathbf{P}\right|} \tag{5}$$

Maximising this criterion can be solved by deriving the optimal transform matrix $\mathbf{P} = [\psi_1, \psi_2, \ldots, \psi_d]\mathbb{R}^{N\times d}$, where $\psi_1, \psi_2, \ldots, \psi_d$ are chosen from the generalised eigenvectors of $\mathbf{S}_b\Psi = \lambda\mathbf{S}_w\Psi$ corresponding to the d largest eigenvalues.

Let $\mathbf{C} = [\mathbf{C}_1, \mathbf{C}_2, \mathbf{C}_3]$ be the original colour configuration with the normalisation of zero mean and unit variance for each vector. By projecting $\mathbf{C}$ onto the matrix $\mathbf{P}$, one can define the general colour-space between-class scatter matrix $\mathbf{L}_b$ and within-class scatter matrix $\mathbf{L}_w$ as follows (Yang and Liu, 2008).

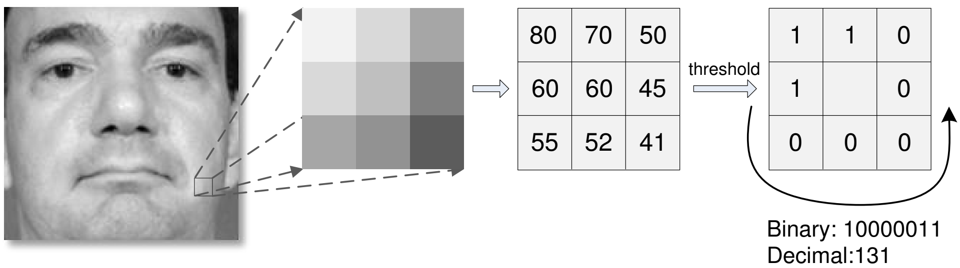
$$\mathbf{L}_b = \sum_{i=1}^{k} P(\omega_i)\left(\bar{\mathbf{C}}_i - \bar{\mathbf{C}}\right)^t \mathbf{PP}^t\left(\bar{\mathbf{C}}_i - \bar{\mathbf{C}}\right) \tag{6}$$

$$\mathbf{L}_w = \sum_{i=1}^{k} P(\omega_i)\varepsilon\left\{\left(\bar{\mathbf{C}}_i - \bar{\mathbf{C}}\right)^t \mathbf{PP}^t\left(\bar{\mathbf{C}}_i - \bar{\mathbf{C}}\right)\middle|\omega_i\right\} \tag{7}$$

where $\bar{\mathbf{C}}_i$ is the mean of class ω_i and $\bar{\mathbf{C}}$ is the grand mean, and $\mathbf{L}_b, \mathbf{L}_w \in \mathbb{R}^{3\times 3}$. By obtaining the generalised eigenvectors ξ_1, ξ_2, ξ_3 of $\mathbf{L}_b\Xi = \lambda\mathbf{L}_w\Xi$, ξ_1 corresponding to the largest eigenvalue is selected as the optimal colour transformation, which can generate the discriminating image representation suitable for face recognition. Finally, a novel image representation $\mathbf{D} \in \mathbb{R}^N$ can be derived by projecting the three colour component images $\mathbf{C}_1, \mathbf{C}_2, \mathbf{C}_3$ of an image onto ξ_1:

$$\mathbf{D} = \left[\mathbf{C}^1, \mathbf{C}^2, \mathbf{C}^3\right]\xi_1 \tag{8}$$

Figure 1 An example of the basic LBP approach



4 Face recognition using LBP-DWT and RNDA

In this section, first we present the LBP-DTW approach that yields multi-scale local features both in spatial and frequency domains. Then, we present the RNDA approach that is used to extract the most discriminating features from LBP-DWT. For completeness, we provide a brief review of LBP.

4.1 Local binary patterns

Local binary patterns (LBP) which was originally introduced for texture analysis (Ojala et al., 1996), has been successfully extended to describe faces as a composition of micro-patterns. LBP has demonstrated good success in face recognition. In a 3 × 3 neighbourhood of the image, the basic LBP operator assigns a binary label 0 or 1 to each surrounding pixel by thresholding its grey value using the central pixel value as threshold. Then, the value of the central pixel is replaced with a decimal number obtained by concatenating the binary labels of the neighbours. Formally, the LBP operator is defined as follows.

$$LBP = \sum_{p=0}^{7} 2^p s\left(i_p - i_c\right) \tag{9}$$

where $s(i_p - i_c)$ equals 1 if $i_p - i_c \geq 0$ and 0 otherwise. Figure 1 illustrates an example of the basic LBP operator. Two extensions of the basic LBP were developed in Ojala et al. (2002) to enhance the discrimination ability of the texture descriptor. The first extension allows LBP to deal with any neighbourhood size by using circular neighbourhoods and bilinearly interpolating the pixel values. The second extension defines the so-called uniform patterns. Refer to Ojala et al. (2002) for the details.

The aforementioned extensions allow LBP is to capture image information at different space resolutions. For example, $LBP_{8,1}^{u2}$ operator describes face microstructure while $LBP_{8,3}^{u2}$ encodes information in larger regions. Similarly to Gabor features that encode information from multiple scales to improve face recognition, multiple-resolution LBP features contain face information from multiple scales.

A straightforward method to generate multiple-resolution LBP features is to concatenate histogram features derived from different LBP operators. This, however, leads to high dimensional vectors that can degrade classifier performance. To address this concern, we proposed using DWT.

4.2 LBP-DWT

DWT takes a series of n observations and produces n new values (i.e., wavelet coefficients). An advantage of DWT is its ability to decompose the variance of the observations on a scale-by-scale basis. Specifically, DWT decomposes an input signal x simultaneously using a low pass filter $g(k)$ and a high-pass filter $h(k)$. The output consists of two parts, the approximation and detail coefficients. Mathematically, DWT is defined as follows.

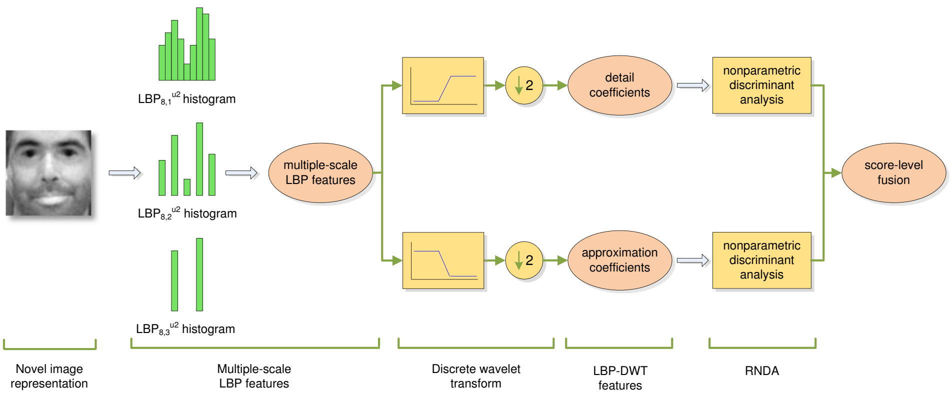
$$y_{\text{low}}(k) = \sum_n x(n) \cdot g(n-2k) \tag{10}$$

$$y_{\text{high}}(k) = \sum_n x(n) \cdot h(n-2k) \tag{11}$$

The decomposition is applied recursively to the low-pass output until a desired number of iterations are reached.

Due to concerns related to computational efficiency, we perform only one-level DWT decomposition using the Haar wavelet. The resulting LBP-DWT features contain multiple-scale information from both the spatial and frequency domains. Figure 2 illustrates the process for generating the LBP-DWT features from the novel image representation. The detail and approximation parts of LBP-DWT, which are complementary to each other, will be fed to RNDA and the classification outputs will be fused at the score level to further improve classification results.

Figure 2 Illustration of the proposed face recognition method (see online version for colours)



4.3 *Regularised non-parametric discriminant analysis*

FLDA is a commonly used classification method for face recognition. However, when the samples follow a non-Gaussian distribution, it often leads to performance degradation. Another drawback of FLDA is that the number of features assuming L classes has an upper bound of $L - 1$. To overcome these problems, Fukunaga (1990) proposed a non-parametric technique, known as NDA. The basis of the extension to NDA from FLDA is a non-parametric between-class scatter matrix that measures between-class scatter using k-nearest neighbours (kNN) on a local basis, instead of merely the class centres. In two-class NDA, the between-class scatter matrix is defined as

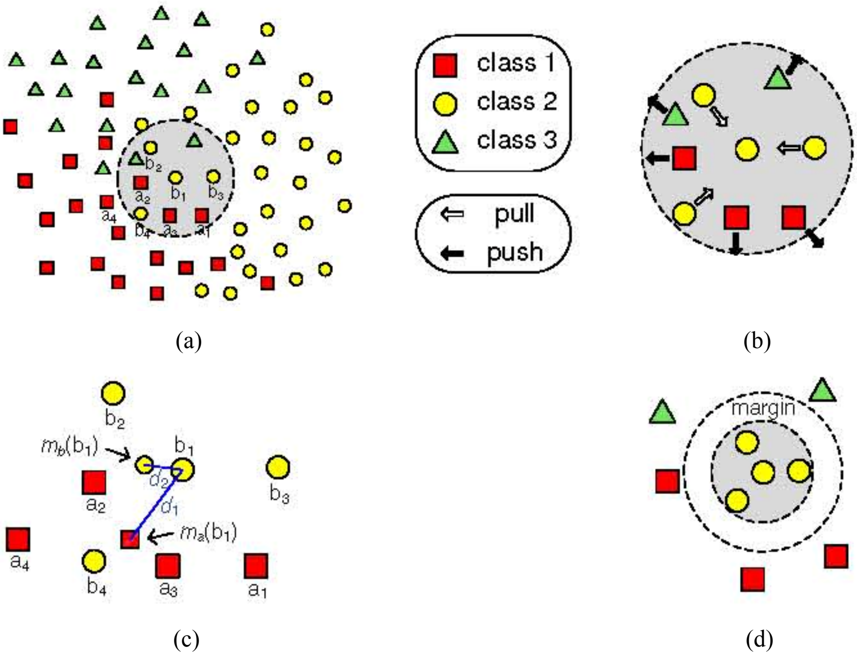
$$\begin{aligned}\mathbf{S}_b = &\sum_{p=1}^{n_1} w(1,p)\left(x_p^1 - m_2\left(x_p^1\right)\right)\left(x_p^1 - m_2\left(x_p^1\right)\right)^t \\ &+\sum_{p=1}^{n_2} w(2,p)\left(x_p^2 - m_1\left(x_p^2\right)\right)\left(x_p^2 - m_1\left(x_p^2\right)\right)^t\end{aligned} \tag{12}$$

where x_p^i denotes the p^{th} sample of class i and $m_j(x_p^i)$ is the local kNN mean that is defined by

$$m_j\left(x_p^i\right) = \frac{1}{k}\sum_{l=1}^{k} NN_l\left(x_p^i, j\right) \tag{13}$$

where $NN_l(x_p^i, j)$ is the l^{th} nearest neighbour from class j to the sample x_p^i. $w(i, p)$ is the value of the weighting function. Under the new formulation and preferably weighting the samples that are near the classification boundary, NDA inherently yields features preserving the structure that is important for classification (Fukunaga, 1990). The original NDA is restricted to two-class problems; recent work has extended NDA to the general L-class problem that is suitable for face recognition (Li et al., 2009; Gu et al., 2010).

Figure 3 Schematic illustration of the linear transformation to enlarge the margin between three classes in a two-dimensional input space. A local neighbourhood centred at sample b_1 contains its three nearest neighbours (3NNs) in class 2, as shown in, (a) a linear transformation can be learned (b) to pull same labelled samples closer together and push differently labelled samples further apart, such that after transformation there exists a finite margin between class 2 and other classes in such a local region, as shown in (d) (c) the effect of the linear transformation, which enlarges the margin between class 1 and class 2 (d) can be approximated by reducing the distance d_2 between sample b_1 and local mean $m_b(b_1)$ and meanwhile increasing the distance d_1 between sample b1 and local mean $m_a(b_1)$ (see online version for colours)



The margin, which is defined as a hyperplane representing the largest separation between classes, is widely used in several machine learning algorithms. A well-known example is support vector machines (SVMs), which choose a subset of input vectors (i.e., support vectors) to construct the decision hyperplane by maximising the margin for achieving good generalisation. In practice, samples from different classes are likely to mix with each other near the boundary. Such overlapping samples bring considerable difficulty for the learning process and are the primary cause of classification errors. One solution to this problem is to learn a linear transformation of the input space, which pulls same labelled samples closer together and pushes differently labelled samples further apart in a local region (Weinberger and Saul, 2009; Gu et al., 2010). This allows a large margin between different classes to be created in overlapping regions. For visualisation purposes, in Figure 3 we use a three-class problem in a two-dimensional input space to illustrate the main idea behind using a linear transformation to enlarge the margin. Figure 3(a) shows that the samples of three classes mix with each other in a local neighbourhood. An appropriate linear transformation is learned to pull the same-class samples close while push the different-class samples away. After the transformation has been applied on the data, there exists a finite margin between class 2 and the other classes in the local region, as shown in Figure 3(d). In particular, the effect of the linear transformation, which enlarges the margin between class 1 and class 2 in Figure 3(d), can be approximated by reducing the distance d_2 between sample b_1 and local mean $m_b(b_1)$ while increasing the distance d_1 between sample b_1 and local mean $m_a(b_1)$, as shown in Figure 3(c) (Gu et al., 2010).

Motivated by the aforementioned methods (Weinberger and Saul, 2009; Gu et al., 2010), we present a variant of NDA. Let x_p^i be the p^{th} sample of class i. Its kNNs in class j can be found and the local mean is denoted as $m_j(x_p^i)$. Based on $m_j(x_p^i)$, its kNNs in class i can be found and the local mean is denoted as $m_j(x_p^i)$.. An example is given in Figure 3(c), where $k = 4$ is used to define kNNs and $m_a(b_1) = m_j(x_p^i)$, $m_b(b_1) = m_i(x_p^i)$. During the learned linear transformation, pushing $m_j(x_p^i)$ away from x_p^i as far as possible and pulling $m_i(x_p^i)$ toward x_p^i as close as possible are approximated by maximising the ratio of between-class scatter to within-class scatter (Gu et al., 2010).

Following the definition of the non-parametric scatter matrix (Fukunaga, 1990; Li et al., 2009), the between-class and within-class scatter matrices are defined as

$$\mathbf{S}_b = \sum_{i=1}^{c}\sum_{\substack{j=1\\ j\neq 1}}^{c}\sum_{p=1}^{n_i} w(i,j,p)\left(x_p^i - m_j\left(x_p^i\right)\right)\left(x_p^i - m_j\left(x_p^i\right)\right)^t, \tag{14}$$

$$\mathbf{S}_w = \sum_{i=1}^{c}\sum_{p=1}^{n_i}\left(x_p^i - m_i\left(x_p^i\right)\right)\left(x_p^i - m_i\left(x_p^i\right)\right)^t \tag{15}$$

where c is the number of classes and n_i is the number of samples in class i. The value of the weighting function, denoted as $w(i, j, p)$, is defined as

$$w(i,j,p) = \frac{\min\left(\left\|x_p^i - m_i\left(x_p^i\right)\right\|_1^{\alpha}, \left\|x_p^i - m_j\left(x_p^i\right)\right\|_1^{\alpha}\right)}{\left\|x_p^i - m_i\left(x_p^i\right)\right\|_1^{\alpha} + \left\|x_p^i - m_j\left(x_p^i\right)\right\|_1^{\alpha}} \tag{16}$$

where $\|\cdot\|_1$ is l_1-norm and α is a parameter between zero and infinity, which controls the changing speed of the weight with respect to the distance ratio. By using the Fisher criterion, the optimum linear transformation thus is composed of the leading eigenvectors of $\mathbf{S}_w^{-1}\mathbf{S}_b$.

In face recognition, the dimensionality of the face image is usually very high compared to the number of available training samples, which may make the within-class scatter matrix singular. The inverse of the within-class scatter matrix often becomes unstable and problematic. A popular solution to address this problem is using PCA first for dimensionality reduction such that the within-class scatter matrix becomes non-singular; then, FLDA is applied in the lower dimensional space. However, FLDA discards the null space of the within-class scatter matrix, which is also deemed to contain important discriminatory information (Chen et al., 2000). To address this issue, an eigenfeature regularisation subspace approach was proposed in Jiang et al. (2008) to extract the discriminatory information residing in both the principal and null spaces of the within-class scatter matrix. Integrating eigenfeature regularisation with NDA leads to the proposed RNDA approach that can be summarised as follows.

Training stage:

1 Apply PCA to a set of training samples and select the leading eigenvectors to construct the principal subspace $\mathbf{W}_{\text{pca}} = \{\psi_1, \psi_2, \ldots, \psi_n\}$ so as to avoid the singularity problem and reduce noise.

2 Project the training sample x into $\mathbf{W}_{\text{pca}}$ and whiten the features, $u = \mathbf{W}_{\text{wpca}}^T x$, where $\mathbf{W}_{\text{wpca}} = \{\psi_1 / \sqrt{\lambda_1}, \psi_2 / \sqrt{\lambda_2}, \ldots, \psi_n / \sqrt{\lambda_n}\}$.

3 Use u to compute the non-parametric $\mathbf{S}_b$ and $\mathbf{S}_w$ by (14) and (15).

4 Subspace decomposition of $\mathbf{S}_w$.

   a Apply PCA to $\mathbf{S}_w$ for deriving the eigenspace $\Phi = \{\phi\}_{i=1}^{l}$ and the eigenvalues that are sorted in descending order $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_l$.

   b Split Φ into two subspaces: a principal subspace $\{\phi\}_{i=1}^{m}$ and its complement, a trailing subspace $\{\phi\}_{i=m+1}^{m}$. The m can be estimated by using a median-based operation (Jiang et al., 2008).

5 Eigenspectrum regularisation. In the trailing subspace $\{\phi\}_{i=m+1}^{l}$, the largest eigenvalue λ_{m+1} is used to replace the remaining eigenvalues $\lambda_i(m + 2 \leq i \leq l)$. As a result, a new eigenvalue matrix Γm is formed, whose diagonal elements are the union of the eigenvalues in $\{\phi\}_{i=1}^{m}$ and the replaced ones in $\{\phi\}_{i=m+1}^{l}$.

6 Whiten $\mathbf{S}_w$ and the new between-class scatter matrix is $\mathbf{S}_b' = \Gamma_m^{-1/2} \Phi^t \mathbf{S}_b \Phi \Gamma_m^{-1/2}$.

7 Compute the eigenspace Θ of $\mathbf{S}_b'$ and finally generate the transformation matrix $\mathbf{U} = \mathbf{W}_{\text{wpca}} \Phi \Gamma_m^{1/2} \Theta$.

Recognition stage:

1 project a probe sample x_p into $\mathbf{U}$ to extract the probe feature vector $y_p = \mathbf{U}^T x_p$

2 apply a classifier to recognise the probe feature vectors.

The classifier employs the nearest neighbour rule for classification using cosine distance measure between a probe feature vector y_p and a gallery feature vector y_g:

$$\delta_{\cos}\left(y_p, y_g\right) = -\frac{y_p^t y_g}{\|y_p\| \|y_g\|} \tag{17}$$

5 Experiments

Extensive experiments have been carried out to evaluate the proposed face recognition method using two challenging data sets: FERET (Phillips et al., 2000) and CMU Multi-PIE (Gross et al., 2010). Specifically, we demonstrate the effectiveness of the proposed method using three sets of experiments. The first set evaluates the effectiveness of the novel image representation scheme for face recognition. The second set evaluates the discrimination power of LBP-DWT features and compares them with traditional LBP features. The third set compares the proposed face recognition method with two state-of-the-art methods: Gabor-based recognition (Liu and Wechsler, 2002) and RSC (Yang et al., 2011).

5.1 Datasets

FERET. The colour FERET database (Phillips et al., 2000) consists of more than 13,000 facial images corresponding to more than 1,500 subjects. Since images were acquired during different sessions, the illumination conditions and the size of the face may vary. The diversity of the FERET database is across gender, ethnicity, and age. The images were acquired without any restrictions imposed on facial expression and with at least two frontal images shot at different times during the same session. The FERET database has become the de facto standard for evaluating face recognition technologies.

The datasets used in the experiments contain four frontal face image sets, fa, fb, dup1, and dup2. Specifically, 1,660 images comprising of 830 persons are first selected from the fa and fb sets to form a set, each person having one fa image and one fb image. From this set, a subset of 1,000 images corresponding to 500 persons then is randomly selected to construct the training set. The gallery set consists of 967 fa images. The two most challenging FERET probe sets, dup1 and dup2, which are used for analysing the effect of ageing on the face recognition performance, are evaluated as the testing set in the experiments. The image numbers of dup1 and dup2 are 722 and 228. Since FERET has provided ground truth eye positions, all images are processed by the normalisation consisting of the following procedures: first, the centres of the eyes are annotated by using ground truth; second, rotation and scaling transformations align the centres of the eyes to predefined locations and fixed interocular distance; finally, a subimage procedure crops the face image to the size of 64×64 to extract the facial region, where the locations of left eye and right eye are (17, 17) and (47, 17).

Multi-PIE. The PIE database was collected at CMU in the fall of 2000 (Sim et al., 2003), in order to support research for face recognition across pose and illumination. Despite its popularity in the research community, the PIE database has some shortcomings. Particularly, there are only 68 subjects that were recorded in a single session, displaying a small range of expressions (neutral, smile, blink, and talk) (Gross et al., 2010). To address these issues, the Multi-PIE database was collected between October 2004 and March 2005. This new database improves upon the PIE database in a number of categories. Most notably, a substantially larger number of subjects were imaged (337 vs. only 68 in PIE) in up to four recording sessions (Gross et al., 2010). In addition, the recording environment of the Multi-PIE database has been improved in comparison to the PIE collection through usage of a uniform, static background and live monitors showing subjects during the recording, allowing for constant control of the head position (Gross et al., 2010).

The experiments use only frontal images, i.e., pose 05 1. In particular, all of the 249 subjects presented in session 1 serve as the training set. Each subject has six images that cover illuminations 12, 14, and 16. The number of training images thus is 1,494. The gallery set consists of 249 images that were recorded in a neutral expression without flash illumination, i.e., illumination 0. Of all 337 subjects in Multi-PIE 129 were recorded in all four sessions. The probe set comprises the images of such 129 subjects in sessions 2, 3, and 4, covering illuminations 4, 8, and 10. The image number of the probe set is 3,483.

Since Multi-PIE does not provide the ground truth of eye locations, face images in the experiments are geometrically rectified by first automatically detecting face and eye locations through the Viola-Jones face and eye detectors (Viola and Jones, 2004) and then aligning the centres of the eyes to predefined locations for extracting 64 × 64 face region. However, due to the occlusion of sunglasses and the presence of the dark glasses frames, the Viola-Jones eye detector fails in detecting the locations of the eyes for a small portion of face images. In such cases, manual annotation is used to locate the eyes.

Figure 4 Example images from the FERET and Multi-PIE databases (see online version for colours)

Notes: The top row shows examples of FERET images. The bottom row shows examples of Multi-PIE images, in which the illumination labels are 0, 4, 8, 10, 12, 14, and 16 from left to right.

Figure 4 shows some example FERET and Multi-PIE images used in the experiments that are already cropped to the size of 64 × 64 in face region. The FERET database features various ethnicity and acquisition environments, while the Multi-PIE database has different lighting directions and extreme expressions. All images used in the experiments are original colour images without any photometric normalisation.

5.2 Evaluation of the novel image representation scheme

The essence of the novel image representation scheme is learning a transformation from a space of uncorrelated colour representations to a space of most discriminative image. Previous research (Liu et al., 2009; Liu and Tao, 2009) has shown that the YCrQ colour space provides more discriminatory information than other colour spaces in the case of the FRGC database (Phillips et al., 2005). Therefore, we have used YCrQ in our experiments. To derive the novel image representation from YCrQ, the transformation coefficients given by (8) are first obtained using PCA followed by FLDA using a set of training images corresponding to the union of the FERET and Multi-PIE training sets. In this case, the transformation coefficients are [0.5768, 0.8156, 0.0446]. It should be noted that the Y, Cr, and Q colour images should be normalised to zero mean and unit variance before applying the linear combination. Figure 5 shows some examples using the novel image representation scheme. Interestingly, the resulting representation looks more like the Cr colour image than the grey-scale image Y. It is the colour-related information that helps the novel image representation scheme to outperform grey-level-based face recognition.

Figure 5 Examples of the novel image representation

Note: The images from left to right are: Y, Cr, Q, and novel image representation.

Three classification methods, Eigenfaces (PCA) (Turk and Pentland, 1991), Fisherfaces (FLDA) (Belhumeur et al., 1997), and SRC (Wright et al., 2009), were used in our experiments to evaluate the effectiveness of the novel image representation scheme for face recognition. The Eigenfaces and Fisherfaces methods are classical face recognition methods, which apply the global scatter information of the training samples whereas largely ignore the sample structures nonlinearly embedded in the high-dimensional face space. Although face images are naturally very high dimensional, they lie on or near low-dimensional sub-spaces or submanifolds. If a collection of representative samples is found for the distribution, one should expect that a typical sample has a very sparse representation with respect to such a basis (Wright et al., 2010). Inspired by this finding, Wright et al. proposed a SRC approach for face recognition (Wright et al., 2009). In SRC, a given test image is first coded as a sparse linear combination of the training samples via l_1-norm minimisation. Then, recognition can be accomplished by evaluating which class from the training samples results in the smallest reconstruction error for the test image. In our experiments, we used all 1,660 training images from the FERET database in SRC to improve the likelihood that for any test image, there exist similar training images. In addition, PCA was applied before SRC to reduce the dimensionality of input images.

Table 1 shows the recognition accuracy using the novel image representation, the grey-level image Y, the colour image Cr, and the colour image Q on the two databases using different recognition methods. Generally speaking, the Cr colour image yields better accuracy than the grey image Y on both databases. This finding is also confirmed on the FRGC database (Liu and Liu, 2010). The new image representation, defined as a linear combination of Y, Cr, and Q, thus inherits the discrimination power of Y, Cr, and Q to further improve face recognition performance.

Table 1 The performance of the novel image representation, grey image Y, colour image Cr, and colour image Q on the FERET and Multi-PIE databases

Image	*Method*	*FERET dup1*	*FERET dup2*	*Multi-PIE*
Y image	PCA	34.76%	10.96%	19.37%
	FLDA	50.69%	28.00%	55.29%
	SRC (Wright et al., 2009)	55.95%	48.24%	56.04%
Cr image	PCA	48.75%	24.12%	38.38%
	FLDA	60.66%	47.80%	59.57%
	SRC (Wright et al., 2009)	56.64%	47.36%	64.05%
Q image	PCA	43.76%	27.19%	17.97%
	FLDA	63.40%	49.60%	41.65%
	SRC (Wright et al., 2009)	59.00%	52.63%	52.56%
Novel representation	PCA	54.98%	41.66%	33.90%
	FLDA	72.71%	72.80%	62.13%
	SRC (Wright et al., 2009)	67.45%	66.66%	63.04%

5.3 Evaluation of LBP-DWT

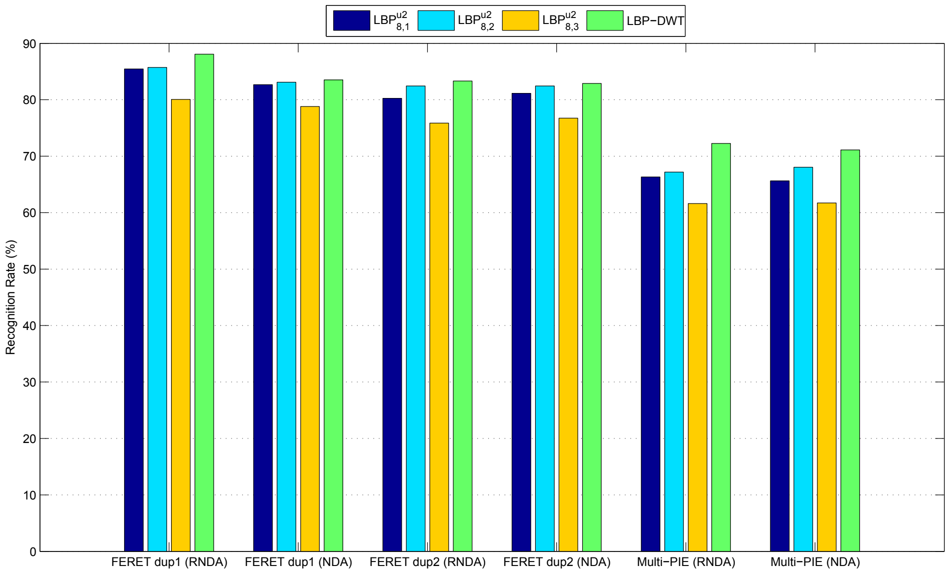
To extract the LBP features, the face image, which is of size 64 × 64, is divided into overlapping patches. The size of each patch is 9 × 9 and there are three overlapping pixels between adjacent patches. This leads to 144 patches with each patch yielding LBP features of length 59. As a result, each of the three LBP operators, $LBP_{8,1}^{u2}$, $LBP_{8,2}^{u2}$, and $LBP_{8,3}^{u2}$, yields LBP histograms of length 8,496.

In the case of the new image representation, first we extract the LBP features using the three LBP operators, and then RNDA is applied to extract a set of discriminating features from the detail and approximation parts of LBP-DWT, and finally the classification outputs are fused by the sum rule at the score level to derive the final classification results, as shown in Figure 2. For comparison, multiclass NDA (Li et al., 2009) is also implemented in the experiments. There are two important parameters in RNDA and NDA. First, the number of nearest neighbours k and second, the weight α. As α does not influence recognition performance significantly, we set its value to 2 for both databases. However, k is training-set specific and affects performance. If k is chosen to the training number of each class and $w(i, j, p)$ in (12) is set to one, NDA is essentially a generalisation of parametric discriminant analysis (PDA) (Li et al., 2009). To avoid this, in our experiments we set k to one half of the training number of each class, i.e., 1 and 3 for the FERET and Multi-PIE databases, respectively.

Figure 6 shows the effectiveness of different LBP features. As it can be observed, the LBPu2 8;2 can achieve better performance than $LBP_{8,1}^{u2}$ and $LBP_{8,3}^{u2}$. The proposed LBP-DWT features consistently outperform LBP features; its accuracy is 88.08%, 83.33%, and 72.26%, respectively, when using RNDA. The improvement over $LBP_{8,2}^{u2}$ is

2.3%, 1.0%, and 5.0%. In particular, the gain is more significant for the Multi-PIE database that contains images having larger illumination variation. Generally, RNDA is marginally better than NDA for the LBP-DWT features.

Figure 6 Face recognition performance of traditional LBP features and proposed LBP-DWT features, both being extracted from the novel image representation



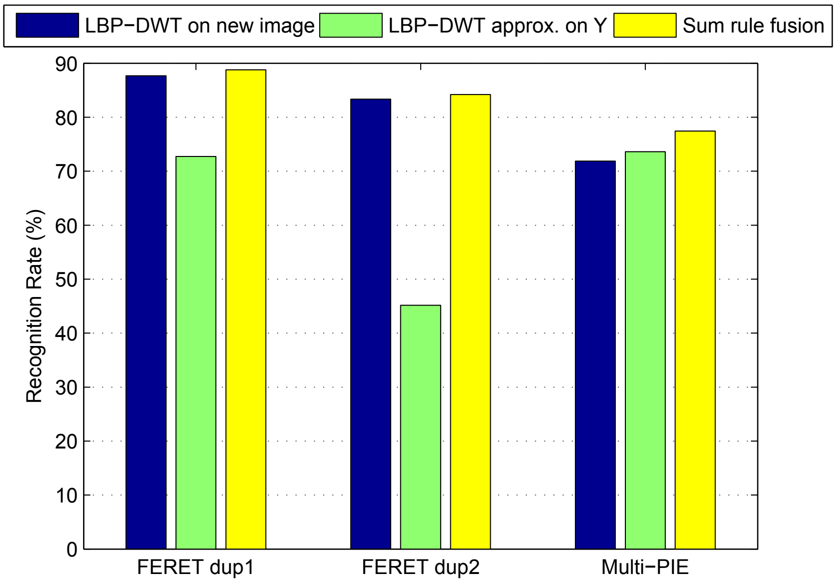
Notes: Our results indicate that $LBP^{u2}_{8,2}$ is better than $LBP^{u2}_{8,1}$ and $LBP^{u2}_{8,3}$ while LBP-DWT can further improve face recognition performance.

5.4 Further improvements and comparisons

The face recognition performance achieved in the previous subsection can be further improved by applying score-level fusion using the novel image representation, which encodes colour information, with the Y image, which encoded grey-scale information. Specifically, the sum rule was used to fuse the similarity score of the LBP-DWT approach using the novel image representation with the similarity score of the approximation part of the LBP-DWT approach using the grey-scale image Y. Figure 7 shows the obtained results, in which RNDA is used.

Table 2 compares the performance of the proposed approach with two state-of-the-art methods, Gabor-based classification (Liu and Wechsler, 2002) and robust sparse coding (RSC) classification (Yang et al., 2011). RSC models sparse coding as a sparsity-constrained robust regression problem and seeks a maximum likelihood estimation (MLE) solution of the sparse coding problem. The experiments on Multi-PIE show that RSC is more robust to outliers (e.g., occlusions and corruptions) than SRC (Yang et al., 2011). Our experimental results on FERET and Multi-PIE indicate that the proposed face recognition framework significantly outperforms the Gabor-based and RSC approaches.

Figure 7 Improving face recognition performance by fusing the similarity scores of the LBP-DWT approach using the new image representation and that of the approximation part of the LBP-DWT approach using the grey-scale image Y (see online version for colours)



Note: Note that RNDA is used in the experiments.

Table 2 Performance improvements and comparisons with state-of-the-art methods

Method	*FERET dup1*	*FERET dup2*	*Multi-PIE*
The method in Liu and Wechsler (2002) on Y[a]	51.38%	57.01%	66.43%
Gabor with NDA on Y	55.95%	60.08%	68.87%
Gabor with RNDA on Y	56.64%	61.84%	69.36%
RSC (Yang et al., 2011) on Y	45.15%	27.63%	64.85%
RSC (Yang et al., 2011) on the novel image	58.03%	48.24%	71.26%
The proposed method	88.78%	84.21%	77.43%

Notes: [a]Because Gabor features do not work well on colour images for face recognition (Liu and Liu, 2010), we extracted Gabor features from the grey-scale image Y only.

As local face details extracted by LBP are different from those extracted by Gabor, these two types of local features may be deemed to be complementary to each other. Therefore, their combination has the potential to boost face recognition performance (Liu and Liu, 2010; Tan and Triggs, 2010). To test this hypothesis, we used the sum rule to fuse the classification outputs of the proposed method with the one obtained using Gabor features using RNDA for classification. Table 3 shows that additional performance improvements can be obtained.

Table 3 Recognition results obtained by fusing the proposed method with Gabor-based classification

Method	*FERET dup1*	*FERET dup2*	*Multi-PIE*
Fusing the proposed method with Gabor	89.19%	85.52%	82.11%
RSC (Yang et al., 2011) with concatenation of Y-CrQ	73.13%	75.00%	81.73%
RSC (Yang et al., 2011)] on the grey image Y	45.15%	27.63%	64.85%

Note: Comparisons with RSC (Yang et al., 2011) using YCrQ and the grey-scale image Y.

We have also evaluated the effectiveness of RSC (Yang et al., 2011) using YCrQ. In particular, given face image, each colour component image of YCrQ is first resized to 32×32 and then normalised to zero mean and unit variance. The three normalised image vectors are then concatenated and used for RSC-based classification. Our experimental results shown in Table 3 indicate that YCrQ helps RSC to improve its performance remarkably, which implies that colour information is indeed a simple yet effective feature for face recognition. Nevertheless, the proposed face recognition framework fused with Gabor features is better than RSC using YCrQ.

6 Conclusions

Classical face recognition methods, such as Eigenfaces and Fisherfaces, have become inadequate in tackling complex face recognition tasks, as evidenced by our experiments using the FERET and Multi-PIE databases in this paper. Recently, efforts have focused on extracting multiple complementary feature sets and new research has witnessed the development of several successful face recognition systems using this idea (Su et al., 2009; Liu and Liu, 2010; Tan and Triggs, 2010). In this paper, we proposed a novel face recognition approach using multiple feature fusion across colour, spatial and frequency domains. We summarise below the main contributions of this work.

- First and foremost, a novel image representation scheme was derived from the uncorrelated YCrQ colour space by a transformation that maps the 3-D colour space to a 2-D feature space. Because uncorrelated colour components naturally decompose face colour images into multiple complementary image representations, such a good property is important for effective image representation used in face recognition. Therefore, the new face image inherently possesses enhanced discrimination for face recognition over grey image.
- Second, LBP-DWT features were designed to describe multiple-scale local face details. The multiple-resolution LBP features are extracted at three different resolutions and DWT decomposes the high-dimensional LBP features into multiple-scales across spatial and frequency domains, hence enhancing the complementary characteristic of the features. RNDA, a variant of NDA, was then introduced to extract a set of discriminative features from the LBP-DWT features.

Extensive experiments using the FERET and Multi-PIE databases demonstrated the feasibility of the proposed framework. In particular, our experimental results show that:

1 the novel image representation scheme performs better than the grey-scale image Y

2 the LBP-DWT features are more powerful than traditional LBP features

3 RNDA can extract more powerful features from LBP-DWT to improve face recognition performance

4 the proposed face recognition approach significantly outperforms state-of-the-art methods.

Recent research has shown that colour space normalisation (CSN) (Yang et al., 2010) techniques can significantly reduce the correlation between colour components, thus enhancing their discrimination power for face recognition. Some hybrid colour spaces, such as XGB, YRB, and ZRG, can achieve excellent performance in the context of face recognition, when CSN is applied (Yang et al., 2010). For future research, we plan to employ CSN techniques in order to create more discriminating image representations for face recognition. In addition, we plan to investigate the effectiveness of the new image representation scheme upon facial expression recognition and gender classification.

Acknowledgements

This work was partially supported by the National Plan for Science and Technology, King Saud University, Riyadh, Saudi Arabia under project Number 10-INF1044-02.

References

Ahonen, T., Hadid, A. and Pietikinen, M. (2006) 'Face description with local binary patterns: application to face recognition', *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 28, No. 12, pp.2037–2041.

Belhumeur, P.N., Hespanha, J.P. and Kriegman, D.J. (1997) 'Eigenfaces vs. Fisherfaces: recognition using class specific linear projection', *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 19, No. 7, pp.711–720.

Bowyer, K.W., Chang, K. and Flynn, P.J. (2006) 'A survey of approaches and challenges in 3D and multi-modal 3D+2D face recognition', *Computer Vision and Image Understanding*, Vol. 101, No. 1, pp.1–15.

Chen, L., Liao, H., Lin, J., Ko, M. and Yu, G. (2000) 'A new LDA-based face recognition system which can solve the small sample size problem', *Pattern Recognition*, Vol. 33, No. 10, pp.1713–1726.

Choi, J.Y., Ro, Y.M. and Plataniotis, K.N. (2011) 'Boosting color feature selection for face recognition', *IEEE Trans. on Image Processing*, Vol. 20, No. 5, pp.1425–1434.

Daugman, J.G. (1985) 'Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional cortical filters', *J. Optical Soc. Am.*, Vol. 2, No. 7, pp.1160–1169.

Fukunaga, K. (1990) *Introduction to Statistical Pattern Recognition*, 2nd ed., Academic Press, San Diego.

Gross, R., Matthews, I., Cohn, J., Kanade, T. and Baker, S. (2010) 'Multi-PIE', *Image and Vision Computing*, Vol. 28, No. 5, pp.807–813.

Gu, Z., Yang, J. and Zhang, L. (2010) 'Push-Pull marginal discriminant analysis for feature extraction', *Pattern Recognition Letters*, Vol. 31, No. 15, pp.2345–2352.

Jain, A.K., Pankanti, S., Prabhakar, S., Hong, L. and Ross, A. (2004) 'Biometrics: a grand challenge', *Proc. IAPR Interna- tional Conference on Pattern Recognition (ICPR'04)*, pp.935–942.

Jiang, X. (2011) 'Linear subspace learning-based dimensionality reduction', *IEEE Signal Processing Magazine*, Vol. 28, No. 2, pp.16–26.

Jiang, X., Mandal, B. and Kot, A. (2008) 'Eigenfeature regularization and extraction in face recognition', *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 30, No. 3, pp.383–394.

Kittler, J., Hatef, M., Robert, P.W. and Matas, J. (1998) 'On combining classifiers', *IEEE Trans. Pattern Analysis and Ma- chine Intelligence*, Vol. 20, No. 3, pp.226–239.

Li, Z., Lin, D. and Tang, X. (2009) 'Nonparametric discriminant analysis for face recognition', *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 31, No. 4, pp.755–761.

Liu, C. (2008) 'Learning the uncorrelated, independent, and discriminating color spaces for face recognition', *IEEE Trans. on Information Forensics and Security*, Vol. 3, No. 2, pp.213–222.

Liu, C. and Wechsler, H. (2002) 'Gabor feature based classification using the enhanced Fisher linear discriminant model for face recognition', *IEEE Trans. on Image Processing*, Vol. 11, No. 4, pp.467–476.

Liu, Z. and Liu, C. (2008) 'Fusion of the complementary discrete cosine features in the YIQ color space for face recognition', *Computer Vision and Image Understanding*, Vol. 111, No. 3, pp.249–262.

Liu, Z. and Liu, C. (2010) 'Fusion of color, local spatial and global frequency information for face recognition', *Pattern Recognition*, Vol. 43, No. 8, pp.2882–2890.

Liu, Z. and Tao, Q. (2009) 'Face recognition using new image representations', *Proc. 2009 International Joint Conference on Neural Networks (IJCNN'09)*.

Liu, Z., Liu, C. and Tao, Q. (2009) 'Learning-based image representation and method for face recognition', *Proc. IEEE 3rd International Conference on Biometrics: Theory, Applications and Systems (BTAS'09)*.

Ojala, T., Pietikinen, M. and Harwood, D. (1996) 'A comparative study of texture measures with classification based on feature distributions', *Pattern Recognition*, Vol. 29, No. 1, pp.51–59.

Ojala, T., Pietikinen, M. and Menp, T. (2002) 'Multiresolution gray-scale and rotation invariant texture classification with local binary patterns', *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 24, No. 7, pp.971–987.

Phillips, P.J., Flynn, P.J., Scruggs, T., Bowyer, K.W., Chang, J., Hoffman, K., Marques, J., Min, J. and Worek, W. (2005) 'Overview of the face recognition grand challenge', *Proc. IEEE International Conference on Computer Vision and Pattern Recognition Workshop*.

Phillips, P.J., Moon, H., Rizvi, S. and Rauss, P. (2000) 'The FERET evaluation methodology for face recognition algorithms', *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 22, No. 10, pp.1090–1104.

Sim, T., Baker, S. and Bsat, M. (2003) 'The CMU pose, illumination, and expression database', *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 25, No. 12, pp.1615-1618.

Su, Y., Shan, S., Chen, X. and Gao, W. (2009) 'Hierarchical ensemble of global and local classifiers for face recognition', *IEEE Trans. on Image Processing*, Vol. 18, No. 8, pp.1885–1896.

Sultana, M. and Garilova, M.L. (2014) 'Face recognition using multiple content-based image features for biometric security applications', *Int. J. of Biometrics*, Vol. 6, No. 4, pp.414–434.

Tan, X. and Triggs, B. (2010) 'Enhanced local texture feature sets for face recognition under difficult lighting conditions', *IEEE Trans. on Image Processing*, Vol. 19, No. 6, pp.1635–1650.

Turk, M. and Pentland, A. (1991) 'Eigen faces for recognition', *Journal of Cognitive Neuroscience*, Vol. 13, No. 1, pp.71–86.

Viola, P. and Jones, M. (2004) 'Robust real-time face detection', *International Journal of Computer Vision*, Vol. 57, No. 2, pp.137–154.

Wang, S-J., Yang, J., Zhang, N. and Zhou, C-G. (2011) 'Tensor discriminant color space for face recognition', *IEEE Trans. on Image Processing*, Vol. 20, No. 9, pp.2490–2501.

Weinberger, Q. and Saul, L. (2009) 'Distance metric learning for large margin nearest neighbor classification', *Journal of Machine Learning Research*, Vol. 10, pp.207–244.

Wright, J., Ma, Y., Mairal, J., Sapiro, G., Huang, T. and Yan, S. (2010) 'Sparse representation for computer vision and pattern recognition', *Proceedings of the IEEE*, Vol. 98, No. 6, pp.1031–1044.

Wright, J., Yang, A., Ganesh, A., Sastry, S. and Ma, Y. (2009) 'Robust face recognition via sparse representation', *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 31, No. 2, pp.210–227.

Yang, J. and Liu, C. (2008) 'Color image discriminant models and algorithms for face recognition', *IEEE Trans. Neural Networks*, Vol. 19, No. 12, pp.2088–2098.

Yang, J., Liu, C. and Zhang, L. (2010) 'Color space normalization: enhancing the discriminating power of color spaces for face recognition', *Pattern Recognition*, Vol. 43, No. 4, pp.1454–1466.

Yang, M., Zhang, L., Yang, J. and Zhang, D. (2011) 'Robust sparse coding for face recognition', *Proc. IEEE International Conference on Computer Vision and Pattern Recognition*, pp.625–632.

Zhang, W., Shan, S., Gao, W., Chen, X. and Zhang, H. (2005) 'Local Gabor binary pattern histogram sequence (LGBPHS): a novel non-statistical model for face representation and recognition', *Proc. IEEE International Conference on Computer Vision (ICCV'05)*, pp.786–791.

Zhao, W., Chellapa, R., Philips, J. and Rosenfeld, A. (2003) 'Face recognition: a literature survey', *ACM Computing Surveys*, Vol. 35, No. 4, pp.399–458.