SVM-Based Face Recognition through Difference of Gaussians and Local Phase Quantization

Chi-Kien Tran*, Thanh-Hoa Ngo, Cam-Ngoan Nguyen, and Lan-Anh Nguyen

***Abstract*—This paper addresses the problem of improving face recognition accuracy for local phase quantization (LPQ) descriptor, introduced by Ojansivu et al. in 2008, when recognizing face images under varying conditions. To do this, we propose to apply difference of Gaussians (DoG) for normalizing face images before encoding the obtained images by LPQ and classifying by support vector machines. Experimental results on three databases (the FEI, FERET, and ORL database of faces databases) demonstrated the improvement of the proposed approach from 0.89% to 17.50% compared to LPQ and other descriptors (CS-LBP, LBP, LDP, LTP, and RLBP) and a combination of them with illumination preprocessing methods (DoG, histogram equalization, Gradient faces, self-quotient image, Tan and Triggs, and Weber-face) using the same classification technique. These results indicated that the introduced approach was robust against variations in illumination, pose, expression, occlusion, scale, and age.**

***Index Terms*—Face recognition, difference of Gaussians, local phase quantization, support vector machines.**

I. Introduction

The problems of face recognition under varying environmental conditions remain great challenges (variation of lighting, blur, age, and pose) for computer vision [1], [2]. To overcome these challenges, face recognition systems need robust feature extraction methods, which can reduce the effect from mentioned variations.

Recently, local phase quantization (LPQ) [3], [4], first introduced by Ojansivu et al. in 2008, has been shown to be robust to blur and to be better than local binary patterns (LBP) [5] in texture classification and face recognition [6]–[8]. However, it still exists limits when recognizing face images under variations in lighting, age, and pose. In [7], authors proposed to use LBP and LPQ for a histogram-based feature extraction and concatenated into a feature vector to be used as a face descriptor. The results conducted on two databases (Yale and AR) indicated that this approach was better than single methods respect to illumination, expression, and occlusions changes. In [9], authors combined Gabor filter, LBP, and LPQ to represent the face images. The results on the CMU-PIE and Yale B databases shown that this approach is robust to illumination and expression changes. In [1], Tran et al. indicated that face images preprocessed by the Tan and Triggs (TT) [10] method and extracted characteristic by LPQ was more robust to varying illumination conditions. As can be seen, most of the previous researches focus on combining LPQ with other methods to improve the performance of face recognition under several varying conditions.

In this study, a difference of Gaussians (DoG) method [11]–[13] is applied to preprocess face images before extracting the obtained images by LPQ and classifying by support vector machines (SVM) [14]–[17]. The efficiency of the proposed approach was demonstrated by comparing to center-symmetric local binary patterns (CS-LBP) [18], LBP, local directional pattern (LDP) [19], local ternary patterns (LTP) [10], and rotated LBP (RLBP) [20] descriptors and a combination of them with DoG, histogram equalization (HE) [21], gradientfaces (GF) [22], self-quotient image (SQI) [23], TT, and Weber-face (WF) [24] methods. Experimental results on the FEI [25], [26], FERET [27], and ORL database of faces (ORL) [28] databases indicated that the proposed approach yielded higher results compared to mentioned methods.

The rest of paper is organized as follows: In the next section, a brief review of related works is presented. Materials and methods are described in Sections III. Section IV presents the experimental results and analysis. Finally, the conclusion is drawn in Section V.

Manuscript received June 27, 2020; revised September 10, 2020.

Chi-Kien Tran*, Thanh-Hoa Ngo, Cam-Ngoan Nguyen, and Lan-Anh Nguyen are with Faculty of Information Technology, Hanoi University of Industry, Bac Tu Liem district, Hanoi, Vietnam (e-mail: chikien.tran@haui.edu.vn).

II. Related Work

A. Difference of Gaussians

Difference of Gaussians (DoG) [11], [12] is an edge-feature enhancement algorithm from an input image based on the subtraction of one blurred version of an original image from another, less blurred, version of the original. The blurred images are obtained by convolving the original images with Gaussian kernels having two different standard deviations.

Smoothed images g_1 and g_2 obtain by convolving the original image F with Gaussian kernels of two different certain widths σ_1 and σ_2 ($\sigma_2 > \sigma_1$).

$$g_1(x,y) = \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left(-\frac{x^2+y^2}{2\sigma_1^2}\right) * F(x,y). \quad (1)$$

$$g_2(x,y) = \frac{1}{\sqrt{2\pi\sigma_2^2}} \exp\left(-\frac{x^2+y^2}{2\sigma_2^2}\right) * F(x,y). \quad (2)$$

The resulting image D is defined as

$$\begin{aligned} D_{\sigma_1,\sigma_2}(x,y) &= g_1(x,y) - g_2(x,y) \\ &= \left(\frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left(-\frac{x^2+y^2}{2\sigma_1^2}\right) - \frac{1}{\sqrt{2\pi\sigma_2^2}} \exp\left(-\frac{x^2+y^2}{2\sigma_2^2}\right)\right) * F(x,y) \\ &= \mathrm{DoG} * F(x,y). \end{aligned} \quad (3)$$

DOI: 10.7763/IJCTE.2021.V13.1282

B. Local Phase Quantization Descriptor

The local phase quantization (LPQ) descriptor was first proposed by Ojansivu and Heikkila for texture description [4]. It is designed based on quantizing the Fourier transform phase in local neighborhoods. Recently, it has widely used in face recognition and proven to be a very efficient descriptor to blurred face image compared to other descriptors such as LBP, LDP [1], [7], [9].

1) Fourier transform phase

In digital image, a discrete model for spastically shift-invariant blurring of an ideal image $f(z)$ can be expressed as

$$g(z) = f(z) * h(z) + n(z), \tag{4}$$

where z is a vector of coordinates $[x,y]^T$, $g(z)$ is an observed image, $h(z)$ is the point spread function (PSF) of the system, $n(z)$ is an additive noise function, and * denotes 2-D convolution.

In the LPQ feature model, the PSF is the centrally symmetric and the additive noise is ignored, so Formula 4 can be written as

$$g(z) = f(z) * h(z), \tag{5}$$

In the Fourier domain, Formula 5 corresponds to

$$G(u) = F(u) * H(u), \tag{6}$$

where $G(u)$, $F(u)$, and $H(u)$ are the discrete Fourier transforms (DFT) of the observed image $g(z)$, the ideal image $f(z)$, and the point spread function $h(z)$. Formula 6 can be separated into the magnitude and phase parts, which can be expressed as follows:

$$\begin{aligned} |G(u)| &= |F(u)| \cdot |H(u)|, \\ \angle G(u) &= \angle F(u) + \angle H(u). \end{aligned} \tag{7}$$

In LPQ, the PSF $h(z)$ is considered as centrally symmetric, namely, $h(z) = h(-z)$, its Fourier transform is real-valued and its phase is only a two-valued function, given by

$$\angle H(u) = \begin{cases} 0 & if \quad H(u) \geq 0 \\ \pi & if \quad H(u) < 0 \end{cases}. \tag{8}$$

As above mentioned, the LPQ method is designed based on the blur invariance property of the Fourier phase spectrum. It uses a short-term Fourier transform (STFT) to extract the local phase information, which is given by the following equation:

$$F(u,z) = \sum_{y \in N_z} f(z-y) e^{-j2\pi u^T y}, \tag{9}$$

where u is a vector of frequency coordinates $[u,v]^T$ and N_z is a local neighborhood at each pixel position z.

In the LPQ method, only four complex coefficients are considered, such as $u_1 = [k, 0]^T$, $u_2 = [0, k]^T$, $u_3 = [k, k]^T$, and $u_4 = [k, -k]^T$, where k is a scalar frequency that satisfies $H(u) > 0$. For each pixel position, this results in a vector as follows:

$$\begin{aligned} F_z &= [F(u_1,z), F(u_2,z), F(u_3,z), F(u_4,z)], \\ G_z &= [\mathrm{Re}\{F_z\}, \mathrm{Im}\{F_z\}]^T. \end{aligned} \tag{10}$$

where Re{.} and Im{.} return real and imaginary parts of a complex number, respectively.

2) Decorrelation process

To remove the correlations but leaves variances intact, authors [3], [4] proposed to decorrelate G_z by using a whitening transform:

$$G_z = V^T G_z, \tag{11}$$

where V is an orthonormal matrix derived from the singular value decomposition (SVD) of the matrix D that is formulared as follows:

$$\begin{aligned} D = WCW^T \,|\, W = [&\mathrm{Re}\{u_1\}; \mathrm{Im}\{u_1\}; \mathrm{Re}\{u_2\}; \\ &\mathrm{Im}\{u_2\}; \mathrm{Re}\{u_3\}; \mathrm{Im}\{u_3\}; \mathrm{Re}\{u_4\}; \mathrm{Im}\{u_4\}]^T \end{aligned}, \tag{12}$$

where C is the covariance matrix of all $M \times M$ samples in N_z, which can be expressed as bellows:

$$C = \begin{bmatrix} 1 & \sigma_{12} & \cdots & \sigma_{1M} \\ \sigma_{21} & 1 & \cdots & \sigma_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{M1} & \sigma_{M2} & \cdots & 1 \end{bmatrix}, \tag{13}$$

where σ_{ij} is the covariance between positions z_i and z_j. It is computed by

$$\sigma_{ij} = \rho^{\|z_i - z_j\|}, \tag{14}$$

where ρ is the correlation coefficient between adjacent pixel values and $\|\cdot\|$ denotes L_2 norm. In this study, ρ is set to 0.9.

3) Local phase quantization

The local phase quantization is processed with two steps:

Step 1: The phase information in the Fourier coefficients at pixel z is represented as binary values based on the signs of the real and imaginary parts of each component in G_z (with or without the decorrelation process).

$$q_j(z) = \begin{cases} 1, & g_j(z) \geq 0 \\ 0, & otherwise \end{cases}, \tag{15}$$

where $g_j(z)$ is the j-th component of the vector G_z.

Step 2: the LPQ pattern of pixel z is encoded by converting into a decimal number. It is done as follows:

$$f_{LPQ}(z) = \sum_{j=1}^{8} q_j(z) 2^{j-1}. \tag{16}$$

C. Support Vector Machines

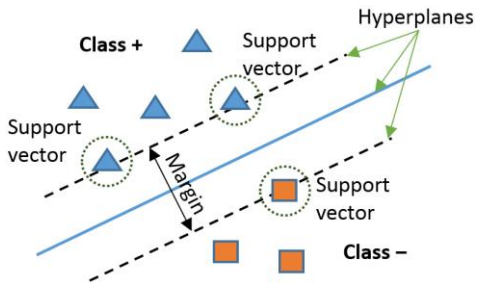

Fig. 1. Maximum-margin hyper-plane and margins for an SVM trained with samples from two classes.

This section will briefly introduce about the key points of support vector machines (SVM) for binary and the extension of it to resolve general multi-class classification problems. For details of the SVM method, the reader can be found e.g. in [14]–[16], [29], [30]. Fig. 1 is a visual illustration about maximum-margin hyper-plane and margins for an SVM trained with samples from two classes.

1) Binary classification

SVM performs a classification between two classes by finding a hyper-plane that has maximum distance to the closest points in the training set which are termed support vectors. Given a labeled training set of points $\mathcal{D} = \{(x_i, y_i) | x_i \in R^n, y_i \in Y\}_{i=1}^m$, where x_i is a data point, $Y = \{-1, +1\}$, and m is the total number of data points for two class case. In the case of classification of linearly separable data, the learning stage will finds the parameters w and b (w is a vector and b is a bias) of a hyper-plane (a discriminant or decision function), which separating the dataset and satisfying:

$$w \cdot x + b = 0, \tag{17}$$

Subject to $y_i(w \cdot x_i + b) \geq 0, \forall i = 1, \ldots, \mathrm{m}$.

The aim of the learning stage of SVM is to find an optimal separating hyper-plane with a maximum margin between our two hyper-planes equals $2/\|w\|$. This means is we will have found the couple (w,b) for which $\|w\|$ is the smallest possible and the constraints are met. In order to find the maximum margin, we need to solve the following optimization problem:

$$\underset{w,b}{\text{minimize}} \quad \frac{1}{2}\|w\|^2, \tag{18}$$

Subject to $y_i(x_i \cdot w + b) \geq 1, \forall i = 1, \ldots, \mathrm{m}$.

Instead of solving (18) one deals with the Lagrange function:

$$L(w, b, \alpha) = \frac{1}{2}\|w\|^2 - \sum_{i=1}^{m} \alpha_i \left[y_i \left(w \cdot x_i + b \right) - 1 \right], \tag{19}$$

Subject to $\alpha_i \geq 0, \forall i = 1, \ldots, m$,

where α_i denotes Lagrange variables.

Formula 19 can be solved by the Wolfe dual problem [31]:

$$\underset{\alpha}{\text{maximize}} \quad \sum_{i=1}^{m} \alpha_i - \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} \alpha_i \alpha_j y_i y_j (x_i \cdot x_j), \tag{20}$$

Subject to $\alpha_i \geq 0, \forall i = 1, \ldots, m$ and $\sum_{i=1}^{m} \alpha_i y_i = 0$.

The Karush Kuhn – Tucker theorem guarantees that the solution of (20) is exactly the same as of (19) and has the form as follows:

$$w = \sum_{i=1}^{m} \alpha_i y_i x_i, \tag{21}$$

Which is leading to a decision function of the form

$$f(x) = \operatorname{sgn}\left(\sum_{i=1}^{m} \alpha_i y_i (x \cdot x_i) + b \right). \tag{22}$$

2) The optimal hyper-plane for non-separable data

Real-life data is often noisy, so in order to this issue (non-separable cases), a soft margin is used:

$$y_i(w \cdot x_i + b) \geq 1 - \xi_i, 0 \leq \xi_i \leq 1, \forall i = 1, \ldots, \mathrm{m}, \tag{23}$$

where ξ_i is a slack variable. The objective function is also modified by adding a term that penalizes nonzero ξ_i and optimization becomes a trade-off between the margin and the penalty

$$\underset{w,b,\xi}{\text{minimize}} \quad \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{m} \xi_i, \tag{24}$$

Subject to $y_i(x_i \cdot w + b) \geq 1 - \xi_i, \xi_i \geq 0, \forall i = 1, \ldots, \mathrm{m}$,

where C is a parameter for balancing errors and generalization capability in the SVM.

Similarly as for the separable case, we can solve Formula 24 by maximizing the same Wolfe dual as before, under a different constraint [32]:

$$\underset{\alpha}{\text{maximize}} \quad \sum_{i=1}^{m} \alpha_i - \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} \alpha_i \alpha_j y_i y_j (x_i \cdot x_j), \tag{25}$$

Subject to $0 \leq \alpha_i \leq C, \forall i = 1, \ldots, m$ and $\sum_{i=1}^{m} \alpha_i y_i = 0$.

3) Non-linear SVM

Real-life data can be not separable and we cannot use SVM to classify. Here are non-separable and non-linearly separable cases. This problem can be solved by transforming the data into a higher-dimensional feature space based on a mapping function $\mathbf{\Phi}$ and replacing the dot products in (25) by the kernel function (26). This is hoped to can help the data become more easily separated or better structured.

$$K(x_i, x_j) = (\Phi(x_i), \Phi(x_j)). \tag{26}$$

$$\underset{\alpha}{\text{maximize}} \quad \sum_{i=1}^{m} \alpha_i - \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} \alpha_i \alpha_j y_i y_j K(x_i, x_j), \tag{27}$$

Subject to $0 \leq \alpha_i \leq C, \forall i = 1, \ldots, m$ and $\sum_{i=1}^{m} \alpha_i y_i = 0$.

Here are three widely used kernel functions:

- Polynomial kernel:

$$K(x_i, x_j) = (x_i \cdot x_j + 1)^d. \tag{28}$$

- Radial basis function kernel:

$$K(x_i, x_j) = \exp(-\gamma \left\| x_i - x_j \right\|^2). \tag{29}$$

- Sigmoid kernel:

$$K(x_i, x_j) = \tanh(\mathrm{k}\, x_i \cdot x_j - \delta). \tag{30}$$

4) Multi-class classification

There are two popular strategies for solving multi-class (k-class) problems with binary SVM classifiers: one-versus-one and one-versus-all approaches.

(i) In the one-versus-one (pairwise) approach, in order to classify k classes, we need to implement $k(k\text{-}1)/2$ different binary classifiers. Each of binary classifier (SVM) will separate a pair of classes. The pairwise classifiers are arranged in trees, where each tree node represents an SVM.

(ii) In the one-versus-all approach, in order to classify k classes, we need to construct k different binary classifiers (k SVM are trained). Each of SVM separates a single class from all remaining classes.

III. Materials and Methods

A. Local Descriptor Based Histogram Features

Once local descriptors are applied for every pixel, a histogram-feature vector of the encoded image can be constructed. However, in order to increase the classification accuracy of a face recognition system, the encoded image is first divided into $M \times N$ regions. A histogram is then constructed for every region. This means that each region is represented by a histogram. The regional histograms are final concatenated into one big histogram representing the whole image. Fig. 2 displays an example of this approach. Fig. 2(a) depicts for a facial image encoded by LPQ and divided into 10×10 non-overlapping regions. Fig. 2(b) is an example of the regional histograms. Fig. 2(c) illustrates for a single histogram representing the whole image.

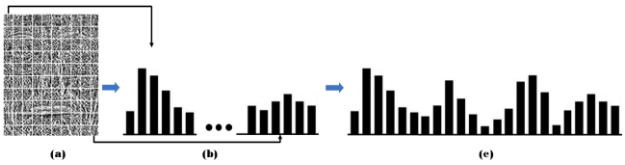

Fig. 2. An example of feature extraction approach.

B. Data Input

For evaluation of performance improvement, the proposed method was tested on three databases: the FEI, FERET, and ORL database of faces (ORL) databases. The brief information of them is presented as follows.

The FEI face database is a Brazilian face database that consists of 14 color images for 200 individuals. Original images are 640×480 pixels and contain variations in pose, expressions, lighting, and scale. In this study, a cropped frontal image set, which has 400 grayscale images of size 193×162 pixels for 200 individuals (2 images of each), was used. Sample images from this set are displayed in Fig. 3.

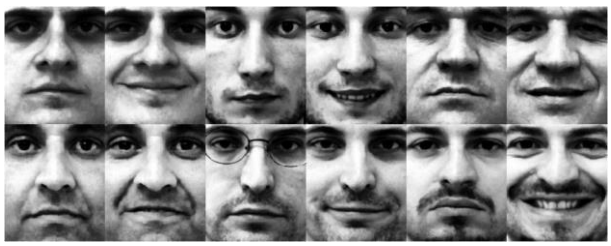

Fig. 3. The sample images of the cropped frontal image set.

The FERET face database is a complex and large-scale database. It has widely used to examine the facial recognition performance of methods. This database is partitioned into five sub-sets (*fa*, *fb*, *fc*, *dup I*, *dup II*) for standard evaluations. The *fa* set includes 1196 images and was used for training phase. The *fb* set consists of 1195 expression variation images. The *fc* set contains 194 illumination variation images. The d*up I* and *dup II* sets have 722 and 234 aging variation images, respectively. Sample images from five sub-sets are shown in Fig. 4.

Fig. 4. The images of the FERET database. First row is the images of fa set. Second row is the images of fb set. Third row is the images of fc set. Fourth row is the images of dup I set. Last row is the images of dup II set.

The ORL database consists of 400 grayscale face images 40 individuals. Images are of size 92×112 pixels and contain variations in pose, expressions, occlusion, and scale. Fig. 5 displays the sample images of this database.

Fig. 5. The sample images of the ORL database.

C. Proposed Approach

Difference of Gaussians (DoG) is an effective illumination preprocessing method to extract illumination components from the face images. Most of the previous researches focus on using this method to improve the preformation of face recognition under varying lighting conditions. In this paper, the DoG method is first applied to enhance edge–based features from face images. Next, the LPQ descriptor is applied after to extract the features of DoG-images. Final, the SVM classifier is performed for classification. The implementation of proposed approach is displayed in Fig. 6 as follows:

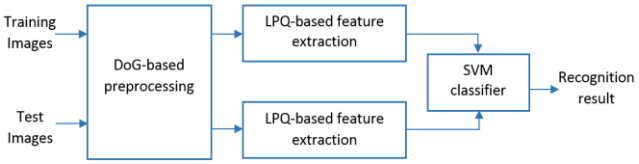

Fig. 6. Block diagram of the proposed face recognition system.

IV. Results and Discussion

A. Experimental Settings

In this study, the available functions of Matlab version R2014b were used. In order to classify the face images, the *fitcecoc* and *predict* functions with SVM models using the "one versus all" encoding scheme was used. The experimental results were implemented using the same classification technique. To extract the histogram features, the encoded images of the ORL, FERET and FEI databases were divided into 10 ×10 and 8 ×8 blocks.

For the DoG method, the parameters σ_1, σ_2 were set to 1 and 2, respectively. For the GF method, the number of neighbors was eight, and the parameter σ was set to 0.75. For the HE method, the number of discrete gray levels was set to 64. For the WF method, the number of neighbors was eight, the parameter σ was set to 1, the parameter α was set to 2, and the output result was normalized to the eight-bit interval. For the CS-LBP, LDP, and LTP methods, the value of parameters that correspond to the highest results was chosen.

The accuracy of each method was calculated as the number of correct classifications divided for the total of testing images, which is computed as follows:

$$\text{Accuracy}(\%) = \frac{\#\text{of correct classifications}}{\#\text{of total testing images}} \times 100. \quad (31)$$

B. Experimental Results on the FEI Database

Table I lists the results of methods experimented on the FEI database. First column (None) displays the results of implementation of methods without using illumination preprocessing methods. The other are the results of methods using illumination pretreatment methods. The results of them, which have percentage accuracy lower than that of methods without using illumination preprocessing methods, are put in round bracket.

TABLE I: Recognition Rates (%) of Methods Experimented on the FEI Database

Method	Illumination preprocessing method						
	None	**DoG**	**GF**	**HE**	**SQI**	**TT**	**WF**
CS-LBP	63.00	**71.50**	**71.50**	(62.50)	65.00	(58.50)	(59.50)
LBP	65.00	66.00	**72.50**	66.50	(61.00)	(60.00)	(58.00)
LDP	62.00	(58.00)	(60.00)	**63.00**	(56.00)	(54.00)	(55.00)
LTP	65.00	65.50	**72.50**	(63.50)	(60.00)	(60.50)	(58.00)
RLBP	47.00	54.50	(46.50)	48.00	**56.00**	(43.50)	47.00
LPQ	65.00	**73.50**	‖	‖	‖	‖	‖

Abbreviations: None: no use illumination preprocessing methods. DoG: difference of Gaussians; GF: Gradientfaces method; HE: histogram equalization method; SQI: self-quotient image method; TT: Tan and Triggs method; WF: Weber-face method; ‖: Not available result.

As shown in the first row of Table I, the CS-LBP combined with DoG and GF methods reached the highest recognition rate of 71.50%, while this method was combined with TT method got the lowest recognition rate of 58.50%. For the LBP method, similarly, it obtained the highest recognition rate of 72.50% when combined with GF method and the lowest recognition rate of 58.00% when was combined with WF method. For the LDP method, it achieved the best performance, 63.00%, when was combined with HE method and the lowest performance, 54.00%, when was combined with TT method. For the LTP method, it delivered the highest recognition rate of 72.50% when was combined with GF method and the lowest recognition rate of 58.00% when was combined with WF method. For the RLBP method, it acquired the highest recognition rate of 56.00% when was combined with SQI method and the lowest recognition rate of 43.50% when was combined with TT method. For the LPQ method, a combination of DoG and LPQ provided the highest recognition rate of 73.50%, which was 8.50% higher than that of LPQ without using illumination treatment methods. Of the highest accuracies obtained, our method (DoG + LPQ) achieved a higher recognition accuracy than the related methods of 1.00% to 17.50%.

C. Experimental Results on the FERET Database

Tables II and III show the recognition rates of methods conducted on the FERET database. In this study, the *fa* set were used for training, while the other sets was chosen for testing. The result was an average value of the results obtained from each testing set. From the presented results in Table II, it could be seen that the CS-LBP method achieved the highest recognition rate of 69.80% when combined with TT method and the lowest recognition rate of 58.00% when combined with GF method. For the LBP method, it reached the highest recognition rate of 72.38% when combined with TT method and the lowest recognition rate of 58.83% when combined with GF method. For the LDP method, it had the highest recognition rate of 69.42% when combined with TT method and the lowest recognition rate of 54.85% when combined with GF method. For the LTP method, it got the highest recognition rate of 73.48% when combined with DoG method and the lowest recognition rate of 58.31% when combined with GF method. For the RLBP method, it got the highest recognition rate of 67.87% when combined with DoG method and the lowest recognition rate of 34.91% when combined with SQI method. Comparing the highest results, it could be seen that the LTP method (TT followed by LTP) achieved the best performance, with recognition rates of 73.48% compared to 69.80%, 72.38%, 69.42%, and 67.87% for CS-LBP (TT followed by CS-LBP), LBP (TT followed by LBP), LDP (TT followed by LDP), and RLBP (DoG followed by RLBP), respectively.

Table III shows the accuracy of the LPQ and LPQ combined with DoG methods. In order to demonstrate the performance of proposed approach compared with LPQ, the recognition rate of sub-sets are presented in this table. The results revealed that the proposed approach got recognition rates significantly better than LPQ.

A comparison of the results from Tables II and III shown that the average accuracy of our approach achieved 76.39% while the highest average accuracy of CS-LBP (TT), LBP (TT), LDP (TT), LTP (DoG), and RLBP (DoG) reached 69.80%, 72.38%, 69.42%, 73.48%, and 67.87%, respectively. These results indicated that our approach had the highest performance.

TABLE II: RECOGNITION RATES (%) OF METHODS EXPERIMENTED ON THE FERET DATABASE

Method	Illumination preprocessing method						
	None	DoG	GF	HE	SQI	TT	WF
CS-LBP	66.42	67.87	(58.00)	(65.97)	(64.76)	**69.80**	(58.36)
LBP	68.51	72.37	(58.83)	(67.89)	(61.60)	**72.38**	(65.11)
LDP	64.03	68.30	(54.85)	(61.20)	(60.92)	**69.42**	(61.58)
LTP	68.26	**73.48**	(58.31)	(67.74)	(66.14)	72.83	(59.31)
RLBP	43.46	**67.87**	(37.68)	45.05	(34.91)	(39.54)	(34.53)

TABLE III: RECOGNITION RATES (%) OF LPQ AND DOG + LPQ METHODS EXPERIMENTED ON THE FERET DATABASE

Subset	Method	
	LPQ	DoG + LPQ
fb	90.37	**92.38**
fc	42.78	**87.62**
du1	64.68	**68.28**
du2	43.16	**57.26**
Average	60.25	**76.39**

Abbreviations: DoG + LPQ means LPQ combined with DoG.

D. Experimental Results on the ORL Database

Tables IV ÷ IX list the accuracy of the methods conducted on the ORL database. The results listed in Table IV indicated that for a combination of CS-LBP and illumination pretreatment methods, they obtained negative results for all most case. This problem also come to LBP (see Table V) and LTP (see Table VII). For the LDP method, it reached the highest average rate of 94.06% when was combined with HE but got a negative result for the case of 6 training images. Similarly to the LDP method, the RLBP method had the highest average rate of 92.24% when was combined with TT but got a negative result for the case of 6 training images.

Table IX displays the results of the LPQ and DoG + LPQ methods. Comparing the results of the LPQ method without using illumination pretreatment methods with that of other methods (see Tables IV ÷ VIII), it could be easily seen that the LPQ provided the highest rates. Comparing the results of the LPQ and DoG + LPQ methods (see Table IX), it could be indicated that the DoG + LPQ method produced recognition rates better than that of the LPQ method. This implied that proposed method (DoG + LPQ) reached the highest recognition rate compared to other methods.

TABLE IV: RECOGNITION RATES (%) OF THE CS-LBP METHOD CONDUCTED ON THE ORL DATABASE

No	Illumination preprocessing method						
	None	DoG	GF	HE	SQI	TT	WF
3	91.07	(89.64)	(86.42)	(90.71)	(89.28)	(88.92)	(87.14)
4	97.08	(91.25)	(91.66)	(96.25)	(93.75)	(91.66)	(91.66)
5	96.50	(93.00)	(92.00)	96.50	(92.00)	(93.00)	(93.00)
6	98.75	(92.50)	(94.37)	98.75	(95.62)	(94.37)	(92.50)
Average	**95.85**	(91.59)	(91.11)	(95.55)	(92.66)	(91.99)	(91.07)

TABLE V: RECOGNITION RATES (%) OF THE LBP METHOD CONDUCTED ON THE ORL DATABASE

No	Illumination preprocessing method						
	None	DoG	GF	HE	SQI	TT	WF
3	92.85	93.57	(86.78)	(91.78)	**93.92**	**93.92**	(91.78)
4	97.08	(96.66)	(95.41)	(95.83)	(96.66)	(95.41)	(96.25)
5	97.00	97.00	(95.50)	97.00	97.00	(96.50)	(96.50)
6	99.37	(97.50)	(98.75)	99.37	(97.50)	(97.50)	(98.75)
Average	**96.57**	(96.18)	(94.11)	(95.99)	(96.27)	(95.83)	(95.82)

TABLE VI: RECOGNITION RATES (%) OF THE LDP METHOD CONDUCTED ON THE ORL DATABASE

No	Illumination preprocessing method						
	None	DoG	GF	HE	SQI	TT	WF
3	88.21	**91.07**	(87.85)	90.00	90.35	90.35	89.64
4	92.50	**94.58**	**94.58**	93.75	94.16	94.16	93.33
5	94.50	94.50	(94.00)	**95.00**	(93.50)	94.50	94.50
6	98.12	(93.75)	98.12	(97.50)	(96.25)	(94.37)	(96.25)
Average	93.33	93.47	93.64	**94.06**	93.56	93.34	93.43

TABLE VII: RECOGNITION RATES (%) OF THE LTP METHOD CONDUCTED ON THE ORL DATABASE

No	Illumination preprocessing method						
	None	DoG	GF	HE	SQI	TT	WF
3	93.21	**93.92**	(86.42)	(92.50)	(92.85)	93.21	(92.14)
4	96.66	(95.83)	(94.58)	96.66	**97.08**	(94.58)	(95.83)
5	96.50	96.50	(95.00)	97.00	**97.50**	96.50	**97.50**
6	99.37	(96.25)	(96.87)	99.37	(98.12)	(96.87)	(98.12)
Average	**96.43**	(95.62)	(93.22)	(96.38)	(96.39)	(95.29)	(95.90)

TABLE VIII: RECOGNITION RATES (%) OF THE RLBP METHOD CONDUCTED ON THE ORL DATABASE

No	Illumination preprocessing method						
	None	DoG	GF	HE	SQI	TT	WF
3	82.50	86.42	(80.00)	86.07	82.50	87.14	**88.21**
4	90.00	90.83	(88.33)	91.66	90.00	**93.33**	**93.33**
5	91.00	92.00	(89.00)	91.50	(89.00)	**93.50**	93.00
6	97.50	(95.00)	(93.75)	(96.25)	(95.62)	(95.00)	(94.37)
Average	90.25	91.06	(87.77)	91.37	(89.28)	**92.24**	92.23

TABLE IX: RECOGNITION RATES (%) OF THE LPQ AND DOG + LPQ METHODS CONDUCTED ON THE ORL DATABASE

No	Method	
	LPQ	DoG + LPQ
3	94.28	**95.00**
4	97.50	97.50
5	97.00	**98.00**
6	98.75	**99.37**
Average	96.88	**97.46**

Comparing the results listed in Tables I ÷ IX, the following observations can be made from these results:

- The results of the CS-LBP, LBP, LDP, LTP, RLBP methods were often higher than that of them when were combined to the illumination pretreatment methods. This indicated that the mentioned illumination pretreatment methods were not suitable for improving the accuracy of classification of above methods under variations of pose, expression, occlusion, scale, and age.
- Comparing to the highest average accuracies of methods, the results of proposed approach conducted on three databases were higher than that of mentioned descriptors and a combination of them and illumination preprocessing methods from 0.89% to 17.50%.
- Comparing to the average accuracies of the LPQ method and the proposed method demonstrated that the proposed method increased the accuracy of the classification from 0.58% to 16.14% compared to that of the LPQ method.

In summary, it was clear that applying DoG to normalize the face images and SVM to classify enhanced the robustness and accuracy of face recognition based on LPQ. Thus, our method can be comparable in performance to all methods which mentioned in this study.

V. CONCLUSION

In this study, in order to improve performance of a face recognition system using LPQ based features, we proposed a new approach as follows: (i) normalize the face images using the difference of Gaussians (DoG) method; (ii) extract the features of the obtained images using the LPQ descriptor; and (iii) classify by support vector machines classifier. Experiments on the FEI, FERET, and ORL databases indicated that our approach got better results than that of mentioned descriptors (CS-LBP, LBP, LDP, LTP, and RLBP) and a combination of them and illumination preprocessing methods (DoG, GF, HE, SQI, TT, and WF). Thus, it could be postulated that the introduced approach was robust against variations in illumination, pose, expression, occlusion, scale, and age.

CONFLICT OF INTEREST

The authors have no conflict of interest to declare. Abstract of this study was presented on the 2018 International Symposium on Advanced Intelligent Informatics (SAIN), Yogyakarta, Indonesia.

AUTHOR CONTRIBUTIONS

Chi-Kien Tran proposed the idea, conducted the research, and wrote the paper; other authors analyzed the data and carried out the experiments; all authors had approved the final version.

REFERENCES

[1] C. K. Tran, D. T. Pham, C. D. Tseng, and T. F. Lee, "Face recognition under lighting variation conditions using tan-triggs method and local intensity area descriptor," in *Proc. the Eleventh International Conference on Genetic and Evolutionary Computing: Genetic and Evolutionary Computing*, November 6-8, 2017, Kaohsiung, Taiwan, J. C.-W. Lin, J.-S. Pan, S.-C. Chu, and C.-M. Chen, Eds. Singapore: Springer Singapore, pp. 84-92, 2018.

[2] C. K. Tran, C. D. Tseng, P. J. Chao, H. M. Ting, L. Chang, Y. J. Huang *et al.*, "Local intensity area descriptor for facial recognition in ideal and noise conditions," *Journal of Electronic Imaging*, vol. 26, pp. 1-10, 2017.

[3] T. Ahonen, E. Rahtu, V. Ojansivu, and J. Heikkilä, "Recognition of blurred faces using local phase quantization," in *Proc. International Conference on Pattern Recognition*, 2008, pp. 1-4.

[4] V. Ojansivu and J. Heikkilä, "Blur insensitive texture classification using local phase quantization," in *Image and Signal Processing*. vol. 5099, A. Elmoataz, O. Lezoray, F. Nouboud, and D. Mammass, Eds. Springer Berlin Heidelberg, 2008, pp. 236-243.

[5] T. Ahonen, A. Hadid, and M. Pietikäinen, "Face recognition with local binary patterns," in *Proc. Computer Vision - ECCV 2004*, vol. 3021, T. Pajdla and J. Matas, Eds. Springer Berlin Heidelberg, 2004, pp. 469-481.

[6] M. Dahmane and L. Gagnon, "Local phase-context for face recognition under varying conditions," *Procedia Computer Science*, vol. 39, pp. 12-19, 2014.

[7] B. Yuan, H. Cao, and J. Chu, "Combining local binary pattern and local phase quantization for face recognition," in *Proc. 2012 International Symposium on Biometrics and Security Technologies*, 2012, pp. 51-53.

[8] H. T. Nguyen, "Contributions to facial feature extraction for face recognition," Dissertation, Université de Grenoble, 2014.

[9] S. R. Zhou, J. P. Yin, and J. M. Zhang, "Local binary pattern (LBP) and local phase quantization (LPQ) based on Gabor filter for face representation," *Neurocomputing*, vol. 116, pp. 260-264, 2013.

[10] X. Tan and B. Triggs, "Enhanced local texture feature sets for face recognition under difficult lighting conditions," *IEEE Transactions on Image Processing*, vol. 19, pp. 1635-1650, 2010.

[11] D. Marr and E. Hildreth, "Theory of edge detection," in *Proc. the Royal Society of London. Series B. Biological Sciences*, vol. 207, pp. 187-217, 1980.

[12] S. Wang, W. Li, Y. Wang, Y. Jiang, J. Shan, and R. Zhao, "An improved difference of Gaussian filter in face recognition," *Journal of Multimedia*, vol. 7, 2012.

[13] C. K. Tran, C. D. Tseng, L. Chang, and T. F. Lee, "Face recognition under varying lighting conditions: improving the recognition accuracy for local descriptors based on weber-face followed by difference of Gaussians," *Journal of the Chinese Institute of Engineers*, vol. 42, no. 7, pp. 593-601, 2019.

[14] C. J. C. Burges, "A tutorial on support vector machines for pattern recognition," *Data Mining and Knowledge Discovery*, vol. 2, pp. 121-167, Jun 1998.

[15] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, pp. 273-297, 1995.
[16] G. Guodong, S. Z. Li, and C. Kapluk, "Face recognition by support vector machines," in *Proc. the Fourth IEEE International Conference on Automatic Face and Gesture Recognition (Cat. No. PR00580)*, pp. 196-201, 2000.
[17] H. Li, S. Wang, and F. Qi, "Automatic face recognition by support vector machines," in *Proc. the Combinatorial Image Analysis: 10th International Workshop*, IWCIA 2004, Auckland, New Zealand, December 1-3, 2004, R. Klette and J. Žunić, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, pp. 716-725.
[18] M. Heikkilä, M. Pietikäinen, and C. Schmid, "Description of interest regions with center-symmetric local binary patterns," in *Proc. the 5th Indian Conference on Computer Vision, Graphics and Image Processing: ICVGIP 2006*, Madurai, India, December 13-16, 2006, P. K. Kalra and S. Peleg, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 58-69.
[19] T. Jabid, M. H. Kabir, and O. Chae, "Local directional pattern (LDP) for face recognition," *International Journal of Innovative Computing, Information and Control*, vol. 8, pp. 2423-2437, 2012.
[20] R. Mehta and K. Egiazarian, "Rotated local binary pattern (RLBP)-rotation invariant texture descriptor," in *Proc. the International Conference on Pattern Recognition Applications and Methods*, 2013, pp. 497-502.
[21] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 3rd ed. Prentice-Hall, Inc., 2006.
[22] T. Zhang, Y. Y. Tang, B. Fang, Z. Shang, and X. Liu, "Face recognition under varying illumination using gradientfaces," *IEEE Transactions on Image Processing*, vol. 18, pp. 2599-2606, 2009.
[23] H. Wang, S. Z. Li, Y. Wang, and J. Zhang, "Self quotient image for face recognition," in *Proc. the International Conference on Pattern Recognition*, 2004, pp. 1397-1400.
[24] B. Wang, W. Li, W. Yang, and Q. Liao, "Illumination normalization based on weber's law with application to face recognition," *IEEE Signal Processing Letters*, vol. 18, pp. 462-465, 2011.
[25] FEI face database. [Online]. Available: http://fei.edu.br/~cet/facedatabase.html
[26] C. E. Thomaz and G. A. Giraldi, "A new ranking method for principal components analysis and its application to face image analysis," *Image and Vision Computing*, vol. 28, pp. 902-913, 2010.
[27] P. J. Phillips, H. Moon, S. A. Rizvi, and P. J. Rauss, "The FERET evaluation methodology for face-recognition algorithms," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, pp. 1090-1104, 2000.
[28] ORL. (July 1, 2013). The ORL database of faces. [Online]. Available: http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html
[29] J. Platt, N. Cristianini, and J. Taylor, "Large margin {DAG}s for multiclass classification," *Advances in Neural Information Processing Systems*, pp. 547-553, 2000.
[30] V. N. Vapnik, *The Nature of Statistical Learning Theory*, Springer-Verlag New York, Inc., 1995.
[31] B. Schölkopf and A. J. Smola, *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*, MIT Press, 2001.
[32] V. Vapnik, *Statistical Learning Theory*, Wiley, 1998.

Chi-Kien Tran received the Ph.D. degree from the College of Electrical Engineering and Computer Science, National Kaohsiung University of Applied Sciences, Taiwan, in 2017. Currently, he is a lecturer of the Faculty of Information Technology, Hanoi University of Industry, Hanoi, Vietnam. His current research interests include computer vision, soft computing, and artificial intelligence.