ICRMAT 2022
Proceedings of the International Conference on Research in Management & Technovation pp. 227–231
DOI: 10.15439/2022M6308
ISSN 2300-5963 ACSIS, Vol. 34

Face Recognition Technology Using the Fusion of Local Descriptors

Chi-Kien Tran*
AI Technology Center Group,
Samsung Display Vietnam Company
Bacninh, Vietnam
kientc.haui@gmail.com;
kien.tc@samsung.com

Do Van Tuan
AI Technology Center Group,
Samsung Display Vietnam Company
Bacninh, Vietnam

Phimvanh Khamphoui
Faculty of Information Technology,
Hanoi University of Industry
Hanoi, Vietnam

***Abstract*—Local phase quantization (LPQ) descriptor, first introduced by Ojansivu and Heikkila (2008), has successfully been applied in face recognition systems. In this paper, we combine local intensity area descriptor (LIAD), which was first introduced by Tran (2017), with LPQ descriptor to develop robust face recognition systems using LPQ descriptor. Face images were first encoded by LIAD as a noise and dimensionality reduction step. After that, the resulting images were presented through LPQ as a feature extraction step. A nearest neighbor method with chi-square measure is used in classification. Two famous datasets (the ORL Database of Faces and FERET) were used in experiments. The results confirmed that our proposed approach reached mean recognition accuracies that are 0.17% ÷ 7.7% better compared to five conventional descriptors (LBP, LDP, LDN, LTP, and LPQ).**

***Index Terms*—face recognition, fusion of descriptors, local intensity area descriptor, local phase quantization descriptor**

I. Introduction

Face recognition is among the few biometric recognition systems that produces a high level of accuracy and requires simple, straightforward implementation [1]. People can recognize thousands of faces based on distinctive facial features with a brief glance, even faces that they have not seen in a long time or those altered by the impacts of age, accidents, disguise, etc. For this reason, research into distinctive facial feature has garnered a high amount of attention by scholars across multiple disciplines over centuries [1, 2].

In the recent years, face recognition systems using deep learning methods have generated impressive results. However, deep learning methods usually require a large database and a high number of parameters, and as a result, a longer period for training and recognition [3]. Research papers [3, 4] introduced several face recognition methods based on local descriptors with results comparable to those obtained by deep learning methods, albeit with a much shorter implementation time.

Various recently developed descriptors can be placed in two classes: dense descriptors and sparse descriptors. For first type, local features were extracted from pixels of the input image. For second type, local patches of the points of interest are sampled and their invariant features are described[5]. There are several popular dense descriptors as local binary patterns (LBP) [6], local ternary patterns (LTP) [7], local directional pattern (LDP) [8, 9], and local phase quantization (LPQ) [10, 11]. For sparse descriptors, scale invariant feature transform (SIFT) [12], accelerated robust features (SURF) [13], and histogram of oriented gradients (HOG) [14] are among the most well-known. In face recognition systems, dense descriptors had been more widely applied compared to sparse descriptors.

Tran and his colleagues have recently proposed the use of a local descriptor for face recognition and called it a Local Intensity Area Descriptor (LIAD for short) [15]. LIAD is not only very effective with photographs which have controlled conditions and poses; it also works well with images that are prone to noise. So, we apply it to increase the accuracy of face recognition systems using LPQ method through several stages: face images were initially displayed by LIAD as a pre-processing step; the resulting images were subsequently displayed through LPQ as a feature extraction step. This approach was demonstrated as producing superior accuracy by comparing with five other methods, including: LBP, LDP, LDP, LTP, and LPQ. Two databases (ORL Database of Faces [16] and FERET [17]) were used for experiments.

II. Related Work

A. *Local Intensity Area Descriptor*

In recent study [15], Tran and his colleagues proposed a new texture analysis method based on intensity values of local pixels, in which each individual pixel of the image is replaced by a specific value calculated by applying trapezoidal numerical integration formula [18] to the values of eight neighboring pixels (Fig. 1). The proposed method is called Local intensity area descriptor (LIAD). The results of LIAD (basic LIAD) rounded to the nearest lower integers are called LIADdown. Inversely, LIADup are results of LIAD rounded to the nearest higher integers. Mathematically, the trapezoidal numerical integration formula is written as:

$$F = \int_{b_1}^{b_2} f(x)dx \approx \frac{h}{2}\left[y_0 + 2y_1 + 2y_2 + \ldots 2y_{n-1} + y_n\right], \quad (1)$$

where $[b_1, b_2]$ is partitioned into n subintervals of equal length. $h = (b_2 - b_1)/n$. $y_0 = f(b_1)$, $y_n = f(b_2)$, and $y_i = f(x_i)$.

In [15], y_0, y_1, and y_i are values of pixels and the user decides the value of h. The LIAD is represented by three formulars as follows:

$$LIAD(x_c, y_c) = \frac{h}{2}(g_0 + 2g_1 + 2g_2 + \ldots + 2g_6 + g_7), \quad (2)$$

$$LIAD_{down}(x_c, y_c) = floor\left(\frac{h}{2}(g_0 + 2g_1 + 2g_2 + \ldots + 2g_6 + g_7)\right), (3)$$

$$LIAD_{up}(x_c, y_c) = ceil\left(\frac{h}{2}(g_0 + 2g_1 + 2g_2 + \ldots + 2g_6 + g_7)\right), (4)$$

where (x_c, y_c) denotes the central pixel and g_i $(i = 0, \ldots, 7)$ are the values of eight pixels. The results are rounded down (LIADdown) and up (LIADup) to an integer.

Fig. 1 illustrates the coding of pixels using LIAD, LIADup and LIADdown with h set at 0.1. The original and encoded images using LIADup and LIADdown with h set at 0.05 and 0.1. The original images and LIAD-images are shown in Fig. 2.

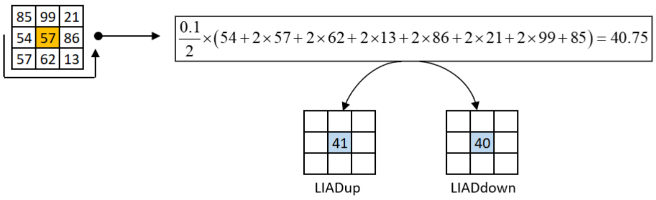

Fig. 1. Example of LIADup and LIADdown with $h = 0.1$.

Fig. 2. The original images and LIAD-images.

B. *Local phase quantization descriptor*

Ojansivu and Heikkila introduced the LPQ descriptor for texture description [10, 11]. Lately, it was widely applied in many facial recognition systems and proven to be a robust descriptor [19-22]. Fig. 3 presents LPQ-images.

a) *Fourier transform phase*

A discrete model of digital image for spastically shift-invariant blurring of an ideal image $u(z)$ is denoted as:

$$t(\mathrm{z}) = u(\mathrm{z}) * v(\mathrm{z}) + n(\mathrm{z}), \tag{5}$$

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

The PSF of a LPQ feature model is the centrally symmetric and the additive noise is ignored, so Formula 5 is expressed as:

$$t(z) = u(z) * v(z) \tag{6}$$

Formula 6 is computed in the Fourier domain as follows:

$$T(s) = U(s) * V(s), \tag{7}$$

where $T(s)$, $U(s)$, and $V(s)$ are the discrete Fourier transforms (DFT) of the observed image $t(\mathrm{z})$, the ideal image $u(\mathrm{z})$, and the point spread function $v(\mathrm{z})$. Formula 7 can be separated into the magnitude and phase parts, which can be expressed as follows:

$$\begin{aligned} |T(s)| &= |U(s)| \cdot |V(s)|, \\ \angle T(s) &= \angle U(s) + \angle V(s). \end{aligned} \tag{8}$$

In LPQ, the PSF $v(z)$ is considered centrally symmetric, namely, $v(z) = v(-z)$. The phase of it is a two-valued function as follows:

$$\angle V(s) = \begin{cases} 0 & if \quad V(s) \geq 0 \\ \pi & if \quad V(s) < 0 \end{cases} \tag{9}$$

As mentioned above, the LPQ method is designed by using the blur invariance property of the Fourier phase spectrum. A short-term Fourier transform (STFT) is applied to extract the local phase information, which is described as follows:

$$U(s,z) = \sum_{y \in N_z} u(z-y) e^{-j2\pi s^T y}, \tag{10}$$

where N_z is a local neighborhood at each pixel position z.

In the LPQ method, only four complex coefficients are considered, including: $s_1 = [k, 0]^{\mathrm{T}}$, $s_2 = [0, k]^{\mathrm{T}}$, $s_3 = [k, k]^{\mathrm{T}}$, and $s_4 = [k, -k]^{\mathrm{T}}$, where k is a scalar frequency that satisfies $V(s) > 0$. For each pixel position, the resulting vector is as follows:

$$\begin{aligned} U_z &= [U(s_1,z), U(s_2,z), U(s_3,z), U(s_4,z)], \\ T_z &= [\mathrm{Re}\{U_z\}, \mathrm{Im}\{U_z\}]^T. \end{aligned} \tag{11}$$

where Re{.} is the real portion and Im{.} is the imaginary portion.

b) *Local phase quantization*

The local phase quantization is processed in 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 portions of each component in T_z (with or without the decorrelation process).

$$q_j(z) = \begin{cases} 1, & g_j(z) \geq 0 \\ 0, & otherwise \end{cases}, \tag{12}$$

(5)

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

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

$$f_{LPQ}(z) = \sum_{j=1}^{8} q_j(z) 2^{j-1} \tag{13}$$

(6)

(7)

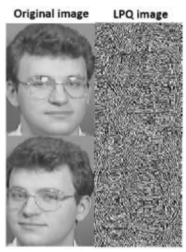

Fig. 3. Original images and LPQ-images.

III. PROPOSED APPROACH

First, we encode face images by the LIAD method to reduce noise and dimensionality. After that, the obtained images are extracted features using the LPQ method. The classification uses a nearest neighbor method with chi-square measure [15]. Fig. 4 presents this approach.

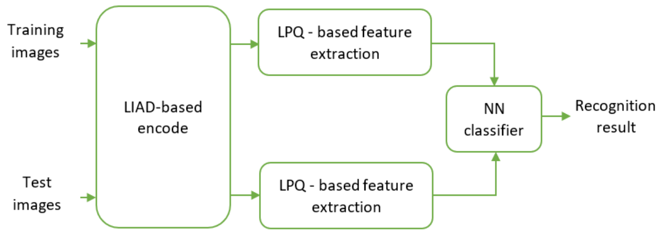

Fig. 4. Flowchart of the proposed approach.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

A. Datasets and Settings

a) FERET database

FERET *database* consists of 14051 grayscale images of 1196 individuals. It contains subsets: *fa* (training set), *fb*, *fc*, *dup* I and, *dup* II. *fa* and *fb* sets are images taken under the same illumination conditions. *fc* set contains images taken with different lighting conditions. *dup* I and *dup* II sets are taken on different day (aging variations).

Fig. 5. Example images of FERET data set. The first row is the images of *fa* set. The images of *fb* set are shown in the second row. Third row is the images of *fc* set. Last rows are the images of *dup* I and *dup* II sets.

In this study, the training data is the *fa* set and the testing data is the others. The experimental images (150 × 130 pixels) were cropped from the original images. To extract features of each image, a single histogram was concatenated from 10 × 10 block histograms. After this step, each image had been represented by a histogram. Sample images of it are introduced in Fig. 5.

b) ORL database

The ORL Database of Faces (ORL for short) has 400 8-bit images of 40 persons and resolution as 92 x 112 pixels (see Fig. 6). The database was split into training and test sets (see Table I).

TABLE I. SIX TRAINING SETS AND SIX CORRESPONDING TEST SETS

Training set	Test set
[1,2,3,4,5]	[6,7,8,9,10]
[2,3,4,5,6]	[1,7,8,9,10]
[3,4,5,6,7]	[1,2,8,9,10]
[4,5,6,7,8]	[1,2,3,9,10]
[5,6,7,8,9]	[1,2,3,4,10]
[6,7,8,9,10]	[1,2,3,4,5]

Fig. 6. Example images of three individuals in ORL dataset.

c) Experimental settings

The nearest-neighbor method with Chi-square measure is applied in classification [15]. For the LDP method, the value of parameter k is set at 3. The threshold t of LTP is set as 1. The parameter h is set at 0.12 for LIAD.

The accuracy (ACC) of each method is computed as the ratio of the correctly labeled images (N) to the whole pool of testing images (T). Mathematically, this can be expressed as:

$$\mathrm{ACC}(\%) = \frac{\mathrm{N}}{\mathrm{T}} \times 100. \quad (14)$$

B. Experiments

Table II presents the experimental results of the conventional methods and the proposed approach on FERET database. The two last columns in Table II show the results of our approach using LIADup and LIADdown. Table III displays the mean rate of the conventional methods and the proposed approach on ORL database. The results of our approach are presented on two last columns. The best performances have been bolded in Tables II and III.

As shown in the first five columns (2 to 6) of Table II, the recognition rate of LTP on *fb* set were the highest, 91.38%. The recognition rate of LBP on *fc* set achieved the highest, 54.63%. The recognition rate of three methods: LTP, LDP, and LDN, on *dup* I set obtained the highest accuracy level: 55.95%. The recognition rates of LDP on *dup* II set were the highest: 45.29%. Comparing the results of classification obtained by all methods in Table II, the highest recognition accuracy belongs to our method (except LDP for subset *dup* II). The findings also showed that the LIADup + LPQ average recognition accuracy reached 62.76%, which was 5.72%, 7.09%, 7.70%, 7.69% and 5.28% higher than that of the LBP, LTP, LDP, LDN, and LPQ methods, respectively. The result is similar to LIADdown + LPQ. The average accuracy of the method, 62.44%, was 5.4%, 6.77%, 7.38%, 7.37%, and 4.96% higher than that of the LBP, LTP, LDP, LDN, and LPQ methods, respectively.

TABLE II. EXPERIMENTAL RESULTS OF METHODS ON THE FERET DATASET

Set	LBP	LTP	LDP	LDN	LPQ	PM1 + LPQ	PM2 + LPQ
fb	90.71	91.38	86.02	89.45	90.71	**92.88**	**92.97**
fc	54.63	43.29	32.98	38.14	47.93	**57.73**	**57.21**
dup I	52.90	55.95	55.95	55.95	56.23	**59.83**	**59.41**
dup II	29.91	32.05	**45.29**	36.75	35.04	40.59	40.17
Mean	57.04	55.67	55.06	55.07	57.48	**62.76**	**62.44**

PM1, PM2 are LIADup and LIADdown, respectively.

TABLE III. EXPERIMENTAL RESULTS OF METHODS ON THE ORL DATASET

LBP	LTP	LDP	LDN	LPQ	PM1 + LPQ	PM2 + LPQ
97.16	96.91	93.66	96.08	98.08	**98.25**	**98.25**

Table III indicates the mean recognition accuracies of the methods in ORL dataset. The highest accuracy belongs to LIADup + LPQ and LIADdown + LPQ, at 98.25%. The next highest accuracy is achieved by using LPQ, which yields 98.08% recognition rate. LDP produces the poorest recognition rate, at 93.66%.

By juxtaposing the results achieved through the proposed method with five other methods on FERET database, it is apparent that our method has significantly improved the recognition rates (see Table II). However, the level of improvement is more modest on ORL database. This can be attributed to a higher level of variance in facial expressions, angles in the ORL database. As a result, LIAD method does not work well with this particular type of image. The LIAD method's strength lies in its capacity to reduce noise and the variance in the distinctive feature of the same individual which can improve accuracy when used in conjunction with LPQ methods.

Through analyzing the results of the proposed method, it is apparent that there is a slight difference in the results of LIADup + LPQ and LIADdown + LPQ. However, this variance is negligible and users can opt for one of the two methods (LIADup + LPQ and LIADdown + LPQ) for face recognition systems.

Empirical evidence has indicated that parameter h should be set at a value in the $0.08 \div 0.14$ range for consistent recognition rates.

The above analyses have indicated a superior accuracy of our method compared to other methods on two given databases.

V. CONCLUSION

We implemented and validated a combination of LIAD and LPQ methods to boost the efficacy of LPQ-based face recognition systems. The face images are first presented by LIAD. After that, then LIAD-images are extracted LPQ-features. The experiments were conducted on FERET and ORL datasets. The classification is implemented using a nearest neighbor method with chi-square statistics as a measure. For FERET dataset, the best average rate of the proposed approach reached 62.76% whereas the corresponding average rate of other descriptors (LBP, LDP, LDN, LTP, and LPQ) are only 57.04%, 55.06%, 55.07%, 55.67%, and 57.48%, respectively. For ORL dataset, the average accuracy of our method achieved 98.25% whereas the five methods gave 97.16%, 93.66%, 96.08%, 96.91%, and 98.08%, respectively. These demonstrated the efficiency of the proposed approach and we recommend using it in face recognition systems.

REFERENCES

[1] M. O. Oloyede, G. P. Hancke, and H. C. Myburgh, "A review on face recognition systems: recent approaches and challenges," Multimedia Tools and Applications, vol. 79, no. 37, pp. 27891-27922, 2020.

[2] 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.

[3] J. Sun, Y. Lv, C. Tang, H. Sima, and X. Wu, "Face Recognition Based on Local Gradient Number Pattern and Fuzzy Convex-Concave Partition," IEEE Access, vol. 8, pp. 35777-35791, 2020.

[4] X. Wei, H. Wang, B. Scotney, and H. Wan, "Selective multi-descriptor fusion for face identification," International Journal of Machine Learning and Cybernetics, vol. 10, no. 12, pp. 3417-3429, 2019.

[5] F. Liu, Z. Tang, and J. Tang, "WLBP: Weber local binary pattern for local image description," Neurocomputing, vol. 120, pp. 325-335, 2013.

[6] T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on feature distributions," Pattern Recognition, vol. 29, no. 1, pp. 51-59, 1996.

[7] T. Xiaoyang and B. Triggs, "Enhanced Local Texture Feature Sets for Face Recognition Under Difficult Lighting Conditions," Image Processing, IEEE Transactions on, vol. 19, no. 6, pp. 1635-1650, 2010.

[8] T. Jabid, M. H. Kabir, and C. Oksam, "Facial expression recognition using Local Directional Pattern (LDP)," in Image Processing (ICIP), 2010 17th IEEE International Conference on, 2010, pp. 1605-1608.

[9] C. K. Tran, T. F. Lee, and P. J. Chao, "Improving face recognition performance using similarity feature-based selection and classification algorithm," Journal of Information Hiding and Multimedia Signal Processing, vol. 6, no. 1, 2015.

[10] T. Ahonen, E. Rahtu, V. Ojansivu, and J. Heikkilä, "Recognition of Blurred Faces Using Local Phase Quantization," in International Conference on Pattern Recognition, 2008, pp. 1-4.

[11] 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. (Lecture Notes in Computer Science: Springer Berlin Heidelberg, 2008, pp. 236-243.

[12] D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," Int. J. Comput. Vision, vol. 60, no. 2, pp. 91-110, 2004.

[13] H. Bay, A. Ess, T. Tuytelaars, and L. V. Gool, "Speeded-Up Robust Features (SURF)," Computer Vision and Image Understanding, vol. 110, no. 3, pp. 346-359, 2008.

[14] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, 2005, vol. 1, pp. 886-893 vol. 1.

[15] C. K. Tran et al., "Local intensity area descriptor for facial recognition in ideal and noise conditions," Journal of Electronic Imaging, vol. 26, no. 2, pp. 023011-1 - 023011-10, 2017.

[16] A. T. L. Cambridge. The Database of Faces [Online]. Available: http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html

[17] P. J. Phillips, H. Moon, S. A. Rizvi, and P. J. Rauss, "The FERET Evaluation Methodology for Face Recognition Algorithms," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 10, pp. 1090-1104, 2000.

[18] E. Kreyszig, Advanced Engineering Mathematics, 10 ed. John Wiley & Sons, 2010, p. 1264.

[19] F. Jafari and H. Rashidy Kanan, "Disguised Face Recognition by Using Local Phase Quantization and Singular Value Decomposition,"

(in en), Journal of Computer & Robotics, vol. 9, no. 1, pp. 51-60, 2016.

[20] J. Jiao and Z. Deng, "Deep combining of local phase quantization and histogram of oriented gradients for indoor positioning based on smartphone camera," International Journal of Distributed Sensor Networks, vol. 13, no. 1, 2017.

[21] L. Nanni, S. Brahnam, and A. Lumini, "Local phase quantization descriptor for improving shape retrieval/classification," Pattern Recognition Letters, vol. 33, no. 16, pp. 2254-2260, 2012.

[22] C. K. Tran, T. H. Ngo, C. N. Nguyen, and L. A. Nguyen, "SVM-Based Face Recognition through Difference of Gaussians and Local Phase Quantization," International Journal of Computer Theory and Engineering, vol. 13, no. 1, pp. 1-8, 2021.