https://doi.org/10.2352/ISSN.2470-1173.2017.10.IMAWM-170

Local Boosted Features for illumination Invariant Face Recognition

Almabrok Essa and Vijayan K. Asari; Department of Electrical and Computer Engineering, University of Dayton; Dayton, OH, USA

Abstract

This paper presents an illumination invariant face recognition technique that uses a combination of local edge gradient information from two different neighboring pixel configurations to represent face images. The proposed Local Boosted Features (LBF) is an oriented local descriptor that is able to encode various patterns of face images under different lighting conditions. It employs the local edge response values in different directions and multi-region histograms from each neighborhood size. Then concatenate these histograms to get one long LBF-feature vector for each image. Finally, we use a library for support vector machines (LIBSVM) classifier to define the similarity between a test feature vector and all other candidate feature vectors. The performance evaluation of the proposed LBF algorithm is conducted on several publicly available face databases and observed improvements in the recognition accuracy.

Introduction

Face recognition has received a great deal of attention in recent years and has spread in several applications such as biometric systems, access control and information security systems, surveillance systems, etc. The first automatic face recognition system was developed by Kanade [1], since then the performance of face recognition systems has improved significantly. The most fundamental way for face recognition is comparing the information from probe images with the one in the recognition system. Therefore, face recognition system can operate in either or both of two modes: face recognition (or identification) and face verification (or authentication).

- In an identification process, the face image of an unknown identity is compared with face images of known individuals from a large database.
- In a verification process, the face image is compared with one face image from a database with the claimed identity. The system measures the similarity between the two images and returns a value which needs to reach a predefined threshold to acknowledge or reject the claimed identity.

The most fundamental and important aspect in face recognition is representing and describing the face with efficient facial features. There are two common types of techniques to extract the features from a face image. The first one is subspace based holistic feature (geometric features) which can be represented by the shapes and locations of facial components such as eyes, nose, mouth, etc. The second one is local appearance based features that presents the appearance changes of the facial skin texture, such as wrinkles and furrows [2]. Two of the most successful local appearance based feature descriptors that are based on the concept of a spatial histogram model local pattern descriptors are local binary pattern (LBP) and local directional pattern (LDP).

The original LBP operator was introduced by Ojala et al. for texture analysis [3]. It is a nonparametric method which extracts local structures of images by comparing a pixel that is under consideration (central pixel of each 3×3 window) with its neighboring pixels. If a neighbor pixel has a higher gray value than the central pixel (or the same gray value) then a $'1'$ is assigned to that pixel, which is otherwise a $'0'$. Finally, the LBP binary code for the center pixel is produced by concatenating the eight $1s$ or $0s$ which can be converted to a decimal number to produce the new value of that central pixel.

The LDP descriptor encodes the directional information in each (3×3) neighboring pixels by convolving the image with Kirsch masks in eight different orientations [4]. It uses the relative strength magnitude to encode the image texture instead of the intensity as LBP does with higher computational cost. In order to generate the LDP code, we need to know the most prominent directions denoted as t. Then the top t directional bit responses are set to 1 and the other $(8-t)$ bits of the 8-bit LDP pattern are set to 0. Finally, the LDP binary code for the pixel under consideration (central pixel of each 3×3 window) is produced by concatenating the eight $1s$ or $0s$ which can be converted to a decimal number to produce the new value of that central pixel. These descriptors usually have been used in the field of face recognition and facial expression recognition, since they have quite important properties to be robust against uncontrolled environments such as illumination variation, random noise, and alignment error, as well as computational simplicity.

In this paper, we present a new technique that is based on the concept of local appearance based techniques for achieving the illumination invariant face recognition, named local boosted features (LBF). LBF extracts the facial features from two different neighborhood configurations, then a histogram is built from each neighboring configuration (layer). Finally, we concatenate these two histogram vectors to form the final LBF-feature vector for each image. In other words, the proposed technique is a combination of two eight bit binary patterns that come from two neighboring sizes, then map each 8-bit binary code to its own bin to get a histogram for each layer. Finally, concatenate these histograms to form a boosted feature vector, which is the descriptor for each image.

The rest of the paper is organized as follows. In Section 2, the proposed LBF technique is illustrated in detail. Discussion on the datasets and experimental results are presented in Section 3. Finally, the conclusion is drawn in Section 4.

Local Boosted Features (LBF)

To capture sufficient detailed discriminative information, the proposed LBF technique encodes various distinctive spatial relationships from two neighborhood sizes of a pixel. Two main stages have to be done for the LBF. 1) Edge detection, where the input image magnitude is extracted by convolving the image with kernels. 2) Image encoding and decoding, where the relative strength is used to encode the image texture.

Edge detection

In this stage, we convolve the input image with Kirsch masks in eight different orientations. The Kirsch operator is a first derivate filter which is used to detect edges in all eight directions of a compass considering all eight neighbors [5]. Specifically, it takes a single mask, denoted as K_i for $i = 0, 1, ..., 7$, and rotates it in 45 degree increments through all 8 compass directions as shown in Fig. 1. The direction of the edge is determined by the mask that produces the maximum output value. Figure 2 shows an example of Kirsch kernels filtered images. All eight directional features of the face image are extracted with their corresponding masks.

$$\underset{K_0}{\begin{bmatrix}-3 & -3 & 5\\ -3 & 0 & 5\\ -3 & -3 & 5\end{bmatrix}} \underset{K_1}{\begin{bmatrix}-3 & 5 & 5\\ -3 & 0 & 5\\ -3 & -3 & -3\end{bmatrix}} \underset{K_2}{\begin{bmatrix}5 & 5 & 5\\ -3 & 0 & -3\\ -3 & -3 & -3\end{bmatrix}} \underset{K_3}{\begin{bmatrix}5 & 5 & -3\\ 5 & 0 & -3\\ -3 & -3 & -3\end{bmatrix}}$$

$$\underset{K_4}{\begin{bmatrix}5 & -3 & -3\\ 5 & 0 & -3\\ 5 & -3 & -3\end{bmatrix}} \underset{K_5}{\begin{bmatrix}-3 & -3 & -3\\ 5 & 0 & -3\\ 5 & 5 & -3\end{bmatrix}} \underset{K_6}{\begin{bmatrix}-3 & -3 & -3\\ -3 & 0 & -3\\ 5 & 5 & 5\end{bmatrix}} \underset{K_7}{\begin{bmatrix}-3 & -3 & -3\\ -3 & 0 & 5\\ -3 & 5 & 5\end{bmatrix}}$$

Figure 1. *Kirsch edge kernels in all eight directions.*

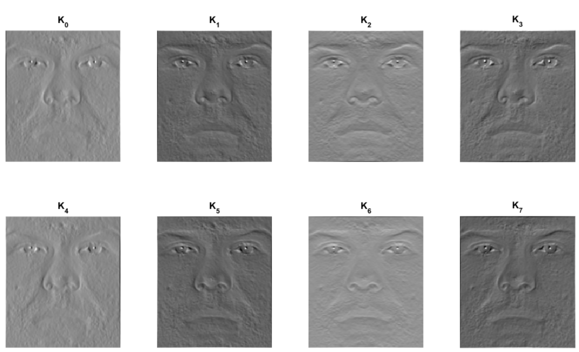

Figure 2. *Kirsch kernels filtered output images.*

Image Encoding and Decoding

After the edges are detected using the method mentioned above, the Image Encoding and Decoding (IED) strategy is applied [6]. To encode the local boosted features in the neighborhood, a binary coding strategy is applied to two neighborhood layers of each pixel in the image. Given a central pixel g_c in the image and its P circularly and evenly spaced R-radius neighbors g_p, $p = 0, 1, , P-1$ that can be seen in Fig. 3 [7]. We consider the first neighbor layer when $P = 8$ and $R = 1$, and the second layer when $P = 8$ and $R = 2$. Therefore, we have 8 neighboring pixels for each surrounding layer excluding the pixel under consideration (central pixel), the values of neighbors that are not in the center of grids can be estimated by interpolation. Then for each neighborhood layer, we compare the 8 neighboring pixels (excluding the central pixel) with their median. If a neighbor pixel has a higher edge value than the median value (or the same value) then a 1 is assigned to that pixel, which is otherwise a 0. Then a histogram is built for each layer. Finally, we form the proposed boosted features by concatenating these two histograms.

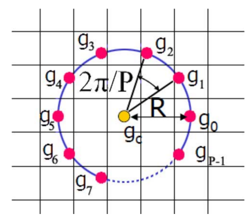

Figure 3. *Central pixel and its P circularly and evenly spaced neighbors with radius R.*

Given an input image $I(x,y)$, the eight different directional edge response values d_i can be computed by

$$d_i = I(x,y) * K_i, \qquad i = 0, 1, ..., 7. \tag{1}$$

where $*$ represents a convolution operation.

The initial step of IED computation after obtaining the eight different directional edge response values of the first and second neighborhood layers $d_{i,1}$ and $d_{i,2}$ for $i = 0, 1, ..., 7$ respectively, is to compute the median of $d_{i,1}$ and $d_{i,2}$. After that, we compare each pixel from each layer with its corresponding median to form the binary code. Then to retrieve the edge features map, we change that binary codes into the corresponding decimal codes D_1 and D_2, which can be computed by

$$\begin{aligned} D_1 &= \sum_{i=0}^{7} f(d_{i,1} - m_1) \times 2^i \\ D_2 &= \sum_{i=0}^{7} f(d_{i,2} - m_2) \times 2^i \end{aligned} \tag{2}$$

and

$$f(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases} \tag{3}$$

where m_1 and m_2 are the medians of each 8 neighboring pixels for the first and second neighborhood layers respectively. Figure 4 illustrates the whole procedure of the proposed LBF technique.

Experimental Results

To evaluate the effectiveness of the proposed LBF technique on face recognition applications, we tested it using tow publicly

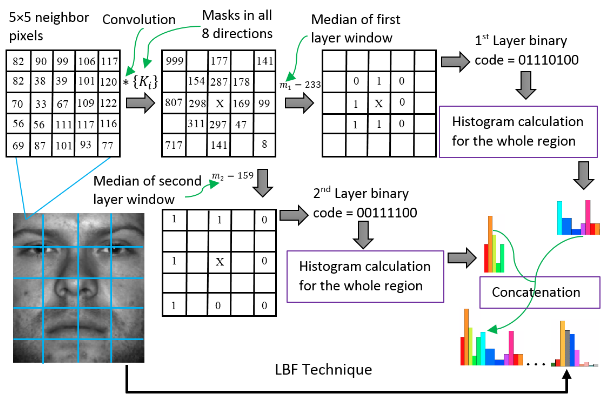

Figure 4. *An example of calculating the proposed LBF.*

available face datasets, namely extended Yale B database [8, 9] and AT&T dataset (ORL) [10]. These datasets include diversity of extreme illumination and lighting problems as in Yale B dataset and slight pose variations as in ORL dataset. The extended Yale B database has 2280 face images for 38 subjects representing 60 illumination conditions per subject under the frontal pose, Fig. 5 shows some samples of one subject of this dataset. For ORL database, a total of 400 face images corresponding to 10 different images of 40 distinct subjects are utilized. Some samples of a subject are shown in Fig. 6. The images are taken at different times with different specifications, including varying slightly in illumination and pose, different facial expressions such as open and closed eyes, smiling and not smiling, and facial details like wearing glasses and not wearing glasses.

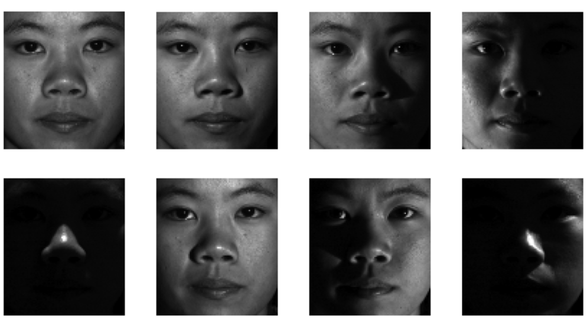

Figure 5. *Samples of one subject from the Extended Yale B database.*

Figure 6. *Samples of a subject from the ORL database.*

To consider the local information of face components, we divide each image into small blocks as can be seen in Fig. 4, then we extract the information of each block separately using our proposed LBF technique from each neighboring layer and represent it as a local LBF histogram. Finally, we concatenate these local histograms to form a global histogram for each layer of the input image. Then we concatenate these two global histograms to form a final feature vector that contains information about the distribution of the local micro-patterns of the image, and can be used to statistically describe the face image characteristics. When it comes to the face recognition process, the objective is to compare the encoded feature vector from one image with all other candidate feature vectors using a library for support vector machines (LIBSVM) [11]. To avoid any bias, we randomly select the data for training and testing and the experiments were repeated 10 times, then the average results are calculated for comparison.

For extended Yale B dataset, we summarize the highest recognition rates by randomly selecting half of the data for training (30 images per subject) and the other half for testing. Then the experiments were repeated 10 times and the average results are calculated for comparison. The performance results of well known face recognition algorithms like local ternary patterns (LTP) [12], Weber-face [13] and gradientface (GradFace) [14], as well as LBP and LDP [15], with the proposed method on extended Yale B dataset are presented in Table 1. Note that, the results we compared with are as we got from their original references which are mentioned in the table. Meanwhile, part of the extended Yale B dataset (standard Yale B dataset) was used in [13][14].

Table 1: Performance comparison of the proposed method with well known face recognition algorithms on extended Yale B dataset

Reference	Method	Recognition Accuracy
Proposed	LBF	**99.21 %**
[14]	GradFace	98.96 %
[12]	RLTP	98.71 %
[13]	Weber-face	98.30 %
[12]	LTP	98.25 %
Code Available	LBP	98.20 %
[14]	LTV	97.93 %
[15]	LDP+2D-PCA	96.43 %
[15]	LBP+2D-PCA	91.54 %
[15]	LDP+PCA	81.34 %

For AT&T (ORL) dataset, we follow the same database partition as in [16] and summarize the highest recognition rates by randomly selecting half of the data for training (5 images per subject) and the other half for testing, then the experiments were repeated 10 times and the average results are calculated for comparison. The performance results of well known face recognition algorithms like local binary patterns (LBP) [17], state preserving extreme learning machine (SPELM) [18], and a combined phase congruency and Gabor wavelet techniques (PC/GW) [16], with the proposed method on ORL dataset are presented in Table 2. Note that, the results we compared with are as we got from their original references which are mentioned in the table, since we do not have any original codes of of these algorithms, while the LBP Matlab code that has been used in this paper for both comparisons is publicly available at: http://www.cse.oulu.fi/CMV/Downloads/LBPMatlab.

Table 2: Performance comparison of the proposed method with well known face recognition algorithms on ORL dataset

Reference	Method	Recognition Accuracy
Proposed	LBF	**98.75 %**
[16]	GW+PC+PCA	98.00 %
[16]	GW+PC	98.00 %
[18]	Gabor+SPELM	97.97 %
Code Available	LBP	97.10 %
[18]	PHOG+SPELM	92.45 %
[16]	PCA	88.00 %

Conclusion

In this paper, we introduced a new feature descriptor technique named as local boosted features (LBF). Throughout the performance evaluation, we found that LBF is robust for face recognition regardless of extreme variations of illumination/lighting environments as in extended Yale B database, and slight changes of pose conditions as in AT&T dataset. In addition, compared to the other state-of-the-art methods, we can see that our method provides better accuracy in most test cases. In general, considering all comparison results, we can assess that LBF can be a promising candidate for face recognition applications.

References

[1] Takeo Kanade. Picture processing system by computer complex and recognition of human faces. *Doctoral dissertation, Kyoto University*, 3952:83–97, 1973.

[2] J Suneetha. A survey on video-based face recognition approaches. *International Journal of Application or Innovation in Engineering And Management (IJAIEM)*, 3(2), 2014.

[3] Timo Ojala, Matti Pietikäinen, and David Harwood. A comparative study of texture measures with classification based on featured distributions. *Pattern recognition*, 29(1):51–59, 1996.

[4] Taskeed Jabid, Md Hasanul Kabir, and Oksam Chae. Local directional pattern (ldp) for face recognition. In *2010 Digest of Technical Papers International Conference on Consumer Electronics (ICCE)*, 2010.

[5] Russell A Kirsch. Computer determination of the constituent structure of biological images. *Computers and biomedical research*, 4(3):315–328, 1971.

[6] Almabrok Essa and Vijayan Asari. Face recognition based on modular histogram of oriented directional features. In *2016 National Aerospace and Electronics Conference (NAECON)*. IEEE, 2016.

[7] Zhenhua Guo, Lei Zhang, and David Zhang. A completed modeling of local binary pattern operator for texture classification. *IEEE Transactions on Image Processing*, 19(6):1657–1663, 2010.

[8] Athinodoros S Georghiades and Peter N Belhumeur. Illumination cone models for faces recognition under variable lighting. In *Proceedings of CVPR98*, 1998.

[9] Kuang-Chih Lee, Jeffrey Ho, and David J Kriegman. Acquiring linear subspaces for face recognition under variable lighting. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 27(5):684–698, 2005.

[10] Ferdinando S Samaria and Andy C Harter. Parameterisation of a stochastic model for human face identification. In *Applications of Computer Vision, 1994., Proceedings of the Second IEEE Workshop on*, pages 138–142. IEEE, 1994.

[11] Chih-Chung Chang and Chih-Jen Lin. Libsvm: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 2(3):27, 2011.

[12] Jianfeng Ren, Xudong Jiang, and Junsong Yuan. Relaxed local ternary pattern for face recognition. In *ICIP*, pages 3680–3684, 2013.

[13] Biao Wang, Weifeng Li, Wenming Yang, and Qingmin Liao. Illumination normalization based on weber's law with application to face recognition. *IEEE Signal Processing Letters*, 18(8):462–465, 2011.

[14] Taiping Zhang, Yuan Yan Tang, Bin Fang, Zhaowei Shang, and Xiaoyu Liu. Face recognition under varying illumination using gradientfaces. *IEEE Transactions on Image Processing*, 18(11):2599–2606, 2009.

[15] Dong-Ju Kim, Sang-Heon Lee, and Myoung-Kyu Sohn. Face recognition via local directional pattern. *International Journal of Security and Its Applications*, 7(2):191–200, 2013.

[16] E Bezalel and U Efron. Efficient face recognition method using a combined phase congruency/gabor wavelet technique. In *Optics & Photonics 2005*, pages 59081K–59081K. International Society for Optics and Photonics, 2005.

[17] Timo Ojala, Matti Pietikainen, and Topi Maenpaa. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on pattern analysis and machine intelligence*, 24(7):971–987, 2002.

[18] Md Zahangir Alom, Paheding Sidike, Vijayan K Asari, and Tarek M Taha. State preserving extreme learning machine for face recognition. In *2015 International Joint Conference on Neural Networks (IJCNN)*, pages 1–7. IEEE, 2015.

Author Biography

Almabrok Essa received his BS in electrical and electronic engineering from Sirt University, LY, and MS degree in electrical and computer engineering from University of Tripoli, LY. He is now working on a PhD program in electrical and computer engineering and serving as a graduate research assistant with the University of Dayton, Dayton, OH, USA. His research interests include computer vision, image processing, face recognition, hyper/multi-spectral imaging. He is a member of IEEE and SPIE.

Vijayan Asari is a Professor in electrical and computer engineering and Ohio Research Scholars endowed chair in wide area surveillance at the University of Dayton, Dayton, Ohio, USA. He received his BS in electronics and communication engineering from the University of Kerala, India, in 1978, M Tech and PhD degrees in electrical engineering from the Indian Institute of Technology, Madras in 1984 and 1994, respectively. He is a Senior Member of IEEE and SPIE.