Single Image Face Recognition based on Gabor, Sobel and Local Ternary Pattern

Pronaya Prosun Das
Department of Computer
Science and Engineering,
Jahangirnagar University,
Dhaka, Bangladesh

Taskeed Jabid
Department of Computer
Science and Engineering,
East West University,
Dhaka, Bangladesh

S.M. Shariar Mahamud
Department of Computer
Science and Engineering,
East West University,
Dhaka, Bangladesh

ABSTRACT

Face recognition has been a fast growing, challenging and fascinating area in real time applications. A large number of face recognition algorithms have been developed in last decades. This paper presents a face recognition approach that utilizes the Sobel operator, Local Ternary Pattern (LTP) descriptor, Gabor features and Principal Component Analysis (PCA) algorithms to attain enhanced recognition accuracy where training set has only one image per person. In particular, the edge information of face image is enhanced using Sobel operator. Then we use LTP on the image to encode the micro-level information of spots, edges and other local characteristics. Finally Gabor-wavelets based features are then extracted and their histograms are concatenated together into a contiguous histogram to be used as a descriptor of the face. As Gabor features cause a very high dimensional histogram vector, therefore PCA is used to reduce the dimension. This approach is compared with the original LBP, LTP and Sobel LTP on gray-level images for face recognition. The experimental results exhibit that our method provides a remarkable performance under various conditions.

Keywords

Face recognition, Local Binary Pattern (LBP), Local Ternary Pattern (LTP), Sobel operator; Sobel LTP (S_LTP), Gabor features, Gabor Sobel LTP (G_S_LTP), Principal Component Analysis.

1. INTRODUCTION

Currently, face recognition has become one of the most popular research areas in the fields of computer vision, image processing, machine learning and pattern recognition due to its numerous applications [1, 2]. The practical applications include automated crowd surveillance, access control, design of human computer interface (HCI), content based image database management, criminal identification and so on. Actually a crucial concern in face analysis is to find an effective and discriminatory representation for face images. These are categorized as either holistic where faces are recognized using global features from faces, or feature based where faces are recognized using local features from faces. However, the features used in holistic and feature-based approaches are fundamentally different.

Features found from holistic approaches represent the optimal variances of pixel data in face images that are used to uniquely identify one individual to another. Various holistic approaches such as Principal Component Analysis (PCA) [4] and Linear Discriminant Analysis (LDA) [3] have been extensively inspected as a result of their capability to apprehend the variance in the training images. Fisherface and Eigenface are two long-familiar facial recognition techniques established on the above two methods respectively.

Alternatively, the features found from feature-based approaches represent face features like the eyes, noses and mouth, where these features are used to uniquely identify individuals. SIFT [5] or Histogram of Oriented Gradients (HOG) [6] are different sorts of successful descriptors using handcrafted encoding. The atomic component in these types of descriptors can be seen as the quantized code of the image gradients. In Local Feature Analysis (LFA) [7], a dense set of local-topological fields are produced to extract local features. Based on Gabor features, Elastic Bunch Graph Matching (EBGM) [8] accomplished a noticeable performance in the FERET evaluation test [9]. Ahonen et al. [10] proposed to use the histogram of Local Binary Pattern (LBP) [11] to express the micro-structures of the face. LBP encodes the relative intensity magnitude between every pixel and its surrounding pixels. It is invariant to monotonic photometric change and can be proficiently extracted. This representation outperform earlier approaches such as PCA, LDA and EBGM [12]. Two most essential properties of the LBP operator in real-world applications are its tolerance against illumination changes and computational simplicity.

In this paper, an extension of LBP called Local Ternary Patterns (LTP) is utilized. The original LBP method served as a local descriptor to capture only local information [12]. It thresholds all pixels in the neighborhood depending on the gray value of the center pixel. Therefore the original LBP turns out to be more sensitive to noise particularly in near uniform or flat areas [13]. One solution to this issue is to extract the features based on 3-valued texture operators as in [14]. Again we contend that the performance of LTP can be enormously enhanced via Sobel operator prior to feature extraction. Sobel operator offers the function of improving this gradient information accordingly. Compared with LBP and LTP, Sobel LTP (S_LTP) produces better performance. Moreover, S_LTP is extended to Gabor features. It is indicated that Gabor S_LTP (G_S_LTP) outperforms LBP, LTP and S_LTP, and accomplishes remarkable recognition accuracy in our comparative experiments. As this methodology is dissimilar to the learning-based methodologies, S_LTP and Gabor features are directly extracted from images without any training procedure. So it is suitable to the systems where few sample images per individual are available. It is known that Gabor features produce a high dimensional histogram vector. To assuage this problem PCA is used to find out the important dimensions and reduce the dimensionality.

The rest of this paper is organized as follows. Section II presents the proposed methodology. In Section III,

comparative experiments on FERET database [9] are conducted to assess the performance of face recognition. Section IV concludes the paper.

2. METHODOLOGY

2.1 Local Ternary Patterns

As a local texture descriptor Local binary pattern (LBP) is an effective and also a simple technique. Ojala et al. [11] proposed it for grayscale images and rotation-invariant texture analysis. Afterward, it was used in many other research areas and exhibited remarkable performance in respect of both discriminative potential and speed [12, 15, 16]. The LBP operator can be presented as follows:

$$LBP_{P,R} = \sum_{p=0}^{P-1} s \times (i_{p,R} - i_c) \times 2^p, \tag{1}$$

$$s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \tag{2}$$

Here, the total number of neighbors are defined by P, the neighborhood radius is R, $i_{p,R}$ is the gray value of the surrounding neighbors and i_c is the gray value of the center pixel.

When calculating a LBP code for an image, the edge that do not have enough information are ignored. Besides its codes are sensitive to noise. So slight change in the intensity levels of the neighbors could be able to completely reshape the resulting binary code. To cope with this issue, local ternary pattern (LTP) was proposed by Tan and Triggs [14]. In this approach the binary LBP code is extended to a 3-valued ternary code in order to bring more steadiness in uniform and near-uniform regions. Here, gray values of the center pixel within a zone of width $\pm t$ are quantized to 0, and those below $-t$ and above $+t$ are quantized to -1 and $+1$, respectively. Therefore, the indicator s(x) in equation (2) is substituted by a 3-valued function:

$$s'(i_p, i_c, t) = \begin{cases} 1, & i_p - i_c \geq t \\ 0, & |i_p - i_c| < t \\ -1, & i_p - i_c \leq -t \end{cases} \tag{3}$$

Here, t is a user-defined threshold. The combination of these differentiating levels in a local neighborhood gives the final LTP value.

Because of using three-level encoding, the number of possible LTP patterns (3^8) is very greater than the number of possible LBP patterns(2^8), which produces a high dimensional feature vector. Here, we have chosen the method recommended in [14], where each ternary code is separated into its corresponding positive ($LTP_{P,R}^{pos}$) and negative ($LTP_{P,R}^{neg}$) parts and then concanated as followes:

$$LTP_{P,R} = \{LTP_{P,R}^{pos}, LTP_{P,R}^{neg}\} \tag{4}$$

Where,

$$LTP_{P,R}^{pos} = \sum_{p=0}^{P-1} s^{pos} \times (i_{p,R} - i_c) \times 2^p, \tag{5}$$

$$s^{pos}(i_p, i_c, t) = \begin{cases} 1, & if \ i_p - i_c \geq t \\ 0, & otherwise \end{cases} \tag{6}$$

and

$$LTP_{P,R}^{neg} = \sum_{p=0}^{P-1} s^{neg} \times (i_{p,R} - i_c) \times 2^p, \tag{7}$$

$$s^{neg}(i_p, i_c, t) = \begin{cases} 1, & if \ i_p - i_c \leq -t \\ 0, & otherwise \end{cases} \tag{8}$$

2.2 Sobel Operation

The Sobel operator is used to measure 2-D spatial gradient of an image. It gives priority to regions of high spatial frequency that correspond to edges. Generally it performs approximate absolute gradient magnitude estimation at each point in a grayscale image. The Sobel operator contains two 3×3 kernels (horizontal kernel S^x and vertical kernel S^y) which are convolved with the original image I to calculate gradient approximations:

$$I^x = S^x * I \ and \ I^y = S^y * I \tag{9}$$

Where,

$$S^x = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} and \ S^y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \tag{10}$$

Here, I^x and I^y represent the horizontally and vertically filtered results, respectively. They are combined to give the gradient magnitude I^G as equation (11).

$$I^G = |I^x| + |I^y| \tag{11}$$

We see that combining Sobel operator with LTP can enhance the local features, and thus more detailed information can be extracted from LTP operation. This operator is named as S_LTP.

$$S\_LTP_{P,R} = \begin{cases} \sum_{p=0}^{P-1} s^{pos} \times (I_{p,R}^G - I_c^G) \times 2^p, \\ \sum_{p=0}^{P-1} s^{neg} \times (I_{p,R}^G - I_c^G) \times 2^p \end{cases} \tag{12}$$
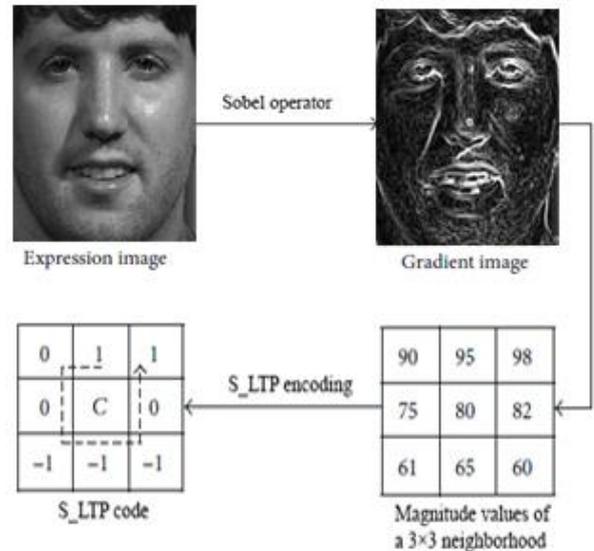

Figure 1: Illustration of the proposed S_LTP encoding scheme. Here, the S_LTP code is (1 0 0 −1 −1 −1 0 1) for threshold $t = 10$
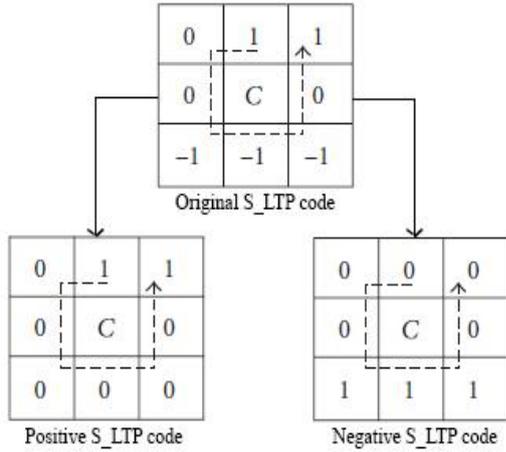
Figure 2: Generation of positive S_LTP and negative S_LTP codes from the original S_LTP code

2.3 S_LTP on Gabor Features

Gabor filter [18] capture the salient visual properties such as spatial characteristics in image space, scale and orientation. Here we use Gabor filters of multi-scale and multi-orientation which have been extensively and successfully used in face recognition to encode the local structure attributes embedded in face images. Gabor filters (kernels) with orientation μ, scale υ are defined as:

$$\psi_{\mu,\nu}(z) = \frac{k_{\mu,\nu}^2}{\sigma^2} exp\left(-\frac{k_{\mu,\nu}^2 z^2}{2\sigma^2}\right)\left[\exp(ik_{\mu,\nu}z) - \exp\left(-\frac{\sigma^2}{2}\right)\right] \quad (13)$$

Where, $z = (x, y)$, wave vector $k_{\mu,\nu}$ is defined as $k_{\mu,\nu} = k_\nu e^{i\varphi_\mu}$,
$k_\nu = \frac{k_{max}}{f^\nu}, k_{max} = \frac{\pi}{2}, f = \sqrt{2}, \varphi_\mu = \frac{\pi\mu}{8}.$ (14)

The indices μ and ν covers a discrete set of eight orientation $\mu = 0, \dots, 7$ and five different frequencies $\nu = 0, \dots, 4$ respectively, in total 40 kernels. The Gabor features $G_{\mu,\nu}$ of a given image *I(z)* are defined as its convolution with 40 Gabor kernels:

$$G_{\mu,\nu} = I(z) * \psi_{\mu,\nu}(z) \quad (15)$$
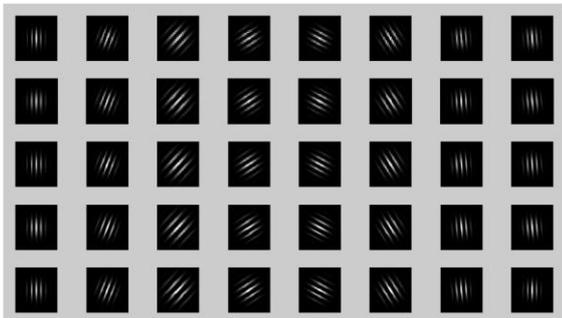
Gabor filters are illustrated in Fig. 3.

Figure 3: 40 Gabor kernels

Let, $G_{\mu,\nu}^{Re}$ and $G_{\mu,\nu}^{Im}$ denote the Gabor real and imaginary features of an image, respectively, where $\mu = 0, \dots, 7$ and $\nu = 0, \dots, 4$. So, we get magnitude by combining $G_{\mu,\nu}^{Re}$ and $G_{\mu,\nu}^{Im}$ as follows:

$$G_{\mu,\nu}^{Mag} = \sqrt{(G_{\mu,\nu}^{Re})^2 + (G_{\mu,\nu}^{Im})^2} \quad (16)$$

S_LTP operators on Gabor features can be defined as:

$$G\_S\_LTP_{\mu,\nu,P,R}^{Mag} =$$

$$\left\{\begin{array}{l}\sum_{p=0}^{P-1} s^{pos} \times (S^x \times G_{\mu,\nu,p,R}^{Mag} - S^x \times G_{\mu,\nu,c}^{Mag}) \times 2^p, \\ \sum_{p=0}^{P-1} s^{neg} \times (S^y \times G_{\mu,\nu,p,R}^{Mag} - S^y \times G_{\mu,\nu,c}^{Mag}) \times 2^p\end{array}\right\} \quad (17)$$

2.4 Histograms

The histograms contain the distribution information of the local features in an image. Spatial histograms computed from the whole encoded image do not reflect the spatial information of the micro-patterns, only their occurrence frequencies are represented [17]. In order to preserve spatial information, a face image is divided into several non-overlapping rectangular sub regions. A spatial histogram, which concatenates the histograms of all the sub-regions, is employed to represent the face. Figure 4 illustrates an example of face image divided into 4×4 sub-regions.
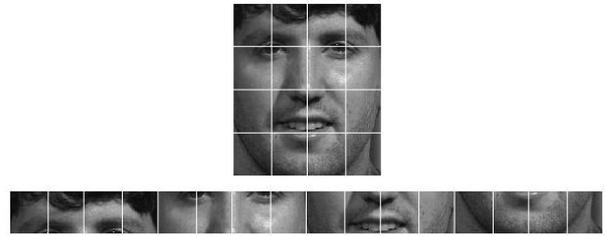

Figure 4: A face image divided into 4×4 sub-regions.

Here, histogram of the S_LTP patterns extracted from the sub-region R_i $(i = 1,2, \dots, N)$ denoted by $H_{S\_LTP}(R_i)$. The spatial histogram represented as,

$$SH_{S\_LTP}(I) = \{H_{S\_LTP}(R_i) \mid i = 1,2,3, \dots, N\} \quad (18)$$

Similar to that the spatial histograms of G_S_LTPs are represented as,

$$SH_{G\_S\_LTP_{\mu,\nu}^{Mag}}(I) = \{H_{G\_S\_LTP_{\mu,\nu}^{Mag}}(R_i) \mid$$

$$i = 1, \dots, N; \mu = 0, \dots, 7; \nu = 0, \dots, 4\} \quad (19)$$

It should be noted that S_LTP extracted from Gabor features produces 40 times longer histogram than that of image S_LBP. To alleviate this problem, we use PCA to reduce the dimensions of $SH_{G\_S\_LTP_{\mu,\nu}^{Mag}}$ histogram.

2.5 Principal Component Analysis

It is a way of identifying patterns in data, and expressing the data in such a way as to highlight their similarities and differences. Once patterns have been found, it is possible to reduce the number of dimensions, without much loss of information. Here, we use this technique on $SH_{G\_S\_LTP_{\mu,\nu}^{Mag}}$ histogram to reduce its dimensions. For each combination of μ and ν, we get a set of values which represents a G_S_LTP feature. We treat these values as a vector H of size N. So there are R number of vectors of size N. In our case R is $\mu \times \nu = 40$. First, we calculate the mean of each vector as follows:

$$\overline{H_j} = \frac{1}{N}\sum_{i=1}^{N} H_{i,j} \; ; \; j = 1,2,3, \dots, R \quad (20)$$

Next, the difference ϕ of each vector from its mean is computed:

$$\phi_{i,j} = H_{i,j} - \overline{H_j} \; ; \; i = 1, \dots, N; \; j = 1, \dots, R \qquad (21)$$

Then the covariance matrix is estimated by:

$$C = AA^T \qquad (22)$$

Where,

$$A = \begin{bmatrix} \phi_{1,1}, \phi_{1,2}, \phi_{1,3}, \dots, \phi_{1,R} \\ \phi_{2,1}, \phi_{2,2}, \phi_{2,3}, \dots, \phi_{2,R} \\ \phi_{3,1}, \phi_{3,2}, \phi_{3,3}, \dots, \phi_{3,R} \\ . \\ . \\ . \\ \phi_{N,1}, \phi_{N,2}, \phi_{N,3}, \dots, \phi_{N,R} \end{bmatrix} \qquad (24)$$

The eigenspace can then be defined by computing the eigenvectors eig_i of C. Since C is very large $N \times N$, computing its eigenvectors will be very expensive. Instead, we can compute $veig_i$, the eigenvectors of $A^T A$, an $R \times R$ matrix. Then, eig_i can be computed from $veig_i$ as follows:

$$eig_i = \sum_{j=1}^{R} veig_{i,j} \, \phi_j \; ; \; j = 1,2,3,\dots, R \qquad (25)$$

From R eigenvectors eig_i, only R' should be chosen, which have highest Eigen values. The higher the Eigen value, the more characteristics of a histogram does the particular eigenvector describe. Eigen faces with low Eigen values can be omitted. The weights form a vector $\Omega = [w_1, w_2, w_3, \dots, w_{R'}]$ describes the contribution of each Eigen vector in representing the input G_S_LTP histogram. Weights can be calculated as,

$$W_k = eig_k^T \, \phi \; ; \; k = 1, \dots, R' \qquad (26)$$

2.6 Similarity Measurements

There are many similarity measures for histogram matching. In this paper, Chi square statistic χ^2 is utilized to compare two spatial histogram:

$$\chi^2(CH_1, CH_2) = \sum_{i=0}^{n} \frac{(CH_{1,i} - CH_{2,i})^2}{(CH_{1,i} + CH_{2,i})} \qquad (27)$$

Where, CH_1 and CH_2 are two slice histograms and n is the number of bin of the histogram.

3. EXPERIMENTAL RESULTS

The proposed method is assessed on the publicly available FERET database [5]. All the images are normalized and cropped to 88×88 pixels based on the positions of two eyes provided within the FERET database. The experiments are divided into four parts. Firstly, we apply LBP on gray-level images, Then experiment is performed with LTP and S_LTP. finaly, we apply S_LTP on Gabor features. We also compare all results produced by LBP, LTP, S_LTP and G_S_LTP. Here, same radius and sampling points (P = 8, R =1) are used in S-LTP, LTP and LBP operators.

To observe the performance of LBP, LTP, S_LTP and G_S_LTP under different conditions, the experiments are conducted on the original images from individual probe sets against the gallery set FA. Table 1 demonstrates the recognition accuracy of LBP, LTP, S_LTP and G_S_LTP on three probe sets with the sub-region size of 4×4 pixels for LBP, LTP and S_LTP, and sub-region size of 11×11 pixels for G_S_LTP. We use dimensionally reduced histogram for G_S_LTP. Table 1 is also illustrated using chart in figure 5.

Table 1. Recognition accuracy of LBP, LTP, S_LTP and G_S_LTP on three probe sets FB, DUPI and DUPII

Method	FB (%)	DUPI (%)	DUPII (%)
LBP	79	46	26
LTP	76	51	33
S_LTP	76	59	46
G_S_LTP	86	72	56

It can be seen from the Table 1 that on FB probe set LTP and S_LTP have same recognition of 76% whereas LBP has better accuracy of 79% than them. But for DUPI and DUPII probe sets, both LTP and S_LTP outperform LBP on recognition rates. This could be caused by the fact that LTP and S_LTP is sensitive to expression variations to a certain extent. It is interesting to note that when the test images are of significant detrimental quality (Dup II), S_LTP demonstrates a significantly better performance than LTP and LBP. It can also be induced from chart and table that G_S_LTP has achieved better performances on all the probe sets than images of LBP, LTP and S_LTP respectively.
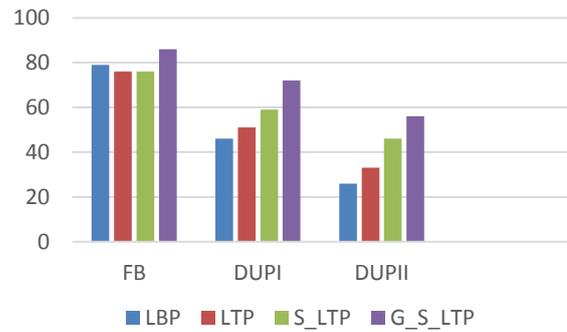

Figure 5: Comparative recognition rates on images

4. CONCLUSION

In this paper, a face recognition approach is proposed based on Gabor wavelet. We extends LTP and presents S_LTP for facial image description and recognition. We used FERET database to conduct comparative experiments between LBP, LTP and S_LTP on both gray-level images and Gabor features, and training image contain only one image per person. As Gabor feature induces high dimensionality, PCA is used to reduce the dimension. The results demonstrate that LTP is slightly better than LBP, and S_LTP performs better than both LBP and LTP on all datasets with only one exception of FB probe set, where S_LTP performs slightly inferior to LBP and equal to LTP. Again G_S_LTP significantly outperforms S_LTP, LTP and LBP under various conditions including different illuminations and different time.

5. REFERENCES

[1] N. B. Kachare and V. S. Inamdar, Int. J. Comput. Appl., vol. 1, no. 1, (2010).

[2] T. Gong, "High-precision Immune Computation for Secure Face Recognition", International Journal of Security and Its Applications (IJSIA), vol. 6, no. 2, SERSC, (2012), pp. 293-298.

[3] P.N. Belhumeur, J.P. Hespanha, and D.J. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection," TPAMI, vol. 19, pp. 711-720, 1997.

[4] M.A. Turk and A.P. Pentland, "Face Recognition Using Eigenfaces," *Proceedings of CVPR*, pp. 586-591, 1991.

[5] D. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110, 2004.

[6] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In *Proc. CVPR*, 2005.

[7] P.S. Penev and J.J. Atick, "Local Feature Analysis: A General Statistical Theory for Object Representation," *Network: Computation in Neural Systems*, vol. 7, pp. 477-500, 1996.

[8] L. Wiskott, J.M. Fellous, N. Krüger, and C. von der Malsburg, "Face Recognition by Elastic Bunch Graph Matching," *TPAMI*, vol. 19, pp. 775-779, 1997.

[9] P.J. Phillips, H. Moon, S.A. Rizvi, and P.J. Rauss, "The FERET Evaluation Methodology for Face-Recognition Algorithms," *TPAMI*, vol. 22, pp. 1090-1104, 2000.

[10] T. Ahonen, A. Hadid, and M. Pietikainen. Face Recognition with Local Binary Patterns. *Lecture Notes in Computer Science*, pages 469–481, 2004.

[11] T. Ojala, M. Pietikainen, and T. 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.

[12] T. Ahonen, A. Hadid, and M. Pietikäinen, "Face Description with Local Binary Patterns: Application to Face Recognition," *TPAMI*, vol. 28, pp. 2037-2041, 2006.

[13] M. A. Akhloufi and A. Bendada, "Locally Adaptive Texture Features for Multispectral Face Recognition," IEEE International Conference on Systems Man and Cybernetics (SMC), Istanbul, Turkey, 2010, pp. 3308-3314.

[14] 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.

[15] G. Zhao and M. Pietikäinen, "Dynamic Texture Recognition Using Local Binary Patterns with an Application to Facial Expressions," TPAMI, vol. 29, pp. 915-928, 2007.

[16] M. Heikkilä and M. Pietikäinen, "A Texture-Based Method for Modeling the Background and Detecting Moving Objects," TPAMI, vol. 28, pp. 657-662, 2006.

[17] F. Ahmed, H. Bari, and E. Hossain, "Person-independent facial expression recognition based on Compound Local Binary Pattern (CLBP)," International Arab Journal of Information Technology, vol. 11, no. 2, 2013.

[18] Sanqiang Zhao; Yongsheng Gao; Baochang Zhang, "Sobel-LBP," Image Processing, 2008. ICIP 2008. 15th IEEE International Conference on, vol., no., pp.2144, 2147, 12-15 Oct. 2008.