2010 Seventh IEEE International Conference on Advanced Video and Signal Based Surveillance

Local Directional Pattern (LDP) – A Robust Image Descriptor for Object Recognition

Taskeed Jabid, Md. Hasanul Kabir, Oksam Chae
Department of Computer Engineering
Kyung Hee University, Republic of Korea
{taskeed, hasanul, oscahe}@khu.ac.kr

Abstract

This paper presents a novel local feature descriptor, the Local Directional Pattern (LDP), for describing local image feature. A LDP feature is obtained by computing the edge response values in all eight directions at each pixel position and generating a code from the relative strength magnitude. Each bit of code sequence is determined by considering a local neighborhood hence becomes robust in noisy situation. A rotation invariant LDP code is also introduced which uses the direction of the most prominent edge response. Finally an image descriptor is formed to describe the image (or image region) by accumulating the occurrence of LDP feature over the whole input image (or image region). Experimental results on the Brodatz texture database show that LDP impressively outperforms the other commonly used dense descriptors (e.g., Gabor-wavelet and LBP).

1. Introduction

Local image descriptors are employed in many real world applications like object detection and view matching using local invariant features [1], texture classification using micro textons [2], face detection and recognizing using local features [3], [4] etc. Every image descriptors attempt to describe the image robustly in adverse imaging condition like lighting variation, changed view point, alteration due to rotation, zooming etc. Descriptors found in literature can be classified into two groups: sparse descriptor and dense descriptor [5]. The sparse first detects the interest points from a given image for sampling local image patch around detected interest points then it generates a feature vector capable to describe the patch. On the other hand, the dense descriptors extract local image features pixel by pixel over the whole input image without identify the interest points.

Scale invariant feature transform (SIFT) [1] is the most notable descriptor in terms of distinctiveness [6] which generates the descriptors with a 3D histogram of gradient location and orientation. Several researchers make an effort to improve the SIFT descriptor and consequently proposed GLOH [7], PCA-SIFT [8], SURF [9] etc. PCA-SIFT use PCA to reduce large feature dimension whereas speeded up robust features (SURF) builds the descriptors by applying the integral image to compute image derivatives. Though sparse descriptor is much popular in view matching, it sometimes fails to resolve matching ambiguity that occurs from locally similar image patches. In addition to that error in interest point detection will lead to performance drop.

Gabor wavelet [10] and local binary pattern are two most popular dense descriptors. The first technique applies a number of Gabor filters on the image to capture small changes in frequency and orientation; finally statistics of these micro-features are used to describe the underlying texture. Recently LBP feature is gaining much attention as a local image descriptor due to its simplicity and excellent performance in texture analysis [11] and face image analysis [4]. Though LBP is robust to monotonic illumination change but it is sensitive to non-monotonic illumination variation and also shows poor performance in the presence of random noise [12]. Local Directional Pattern (LDP), a more robust facial feature, is proposed by Jabid et al. [12], which demonstrated better performance in different application of facial image analysis [13] [14].

This paper introduces LDP as a local image descriptor that can work as a dense descriptor. LDP feature considers relative edge response value in eight directions around a pixel to encode the local neighborhood property of image pixel with a binary bit sequence. In this work, we also proposed a rotation invariant LDP code which uses the largest edge response direction as a starting bit location to normalize the orientation change. It is nearly impossible that most significant edge response will corrupt in any kind of noise or photometric changes, hence generates a robust rotation invariant LDP code.

2. LDP image descriptor

In this section, after a brief review of local binary pattern (LBP) we introduce the local directional pattern (LDP) that encode the local micro pattern more efficiently. A rotational invariant edition of LDP is proposed and analyzed its effectiveness. Finally generation of image descriptor using LDP code is described which is used to model an image.

978-0-7695-4264-5/10 $26.00 © 2010 IEEE
DOI 10.1109/AVSS.2010.17

2.1. Local Binary Pattern (LBP)

Derived from a general definition of texture in a local neighborhood, LBP is defined as a grayscale invariant texture measure and is a useful tool to model texture images. LBP later has shown excellent performance in facial image analysis, in terms of speed and performance. The original LBP operator labels the pixels of an image by thresholding the 3x3 neighborhood of each pixel with the value of the central pixel and concatenating the results binomially to form a number. Fig. 1 shows an exemplary illustration for generation LBP code.

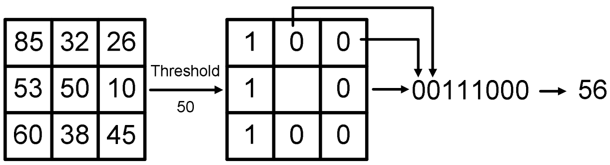


Figure 1: The basic LBP operator

2.2. Local Directional Pattern (LDP)

Recently researchers use change of gradient magnitude in a specific direction around pixels to encode local texture. [15], [16]. Instead of comparing neighboring intensity value these methods compare neighboring pixel's gradient magnitude along a specific direction and encode it like trivial LBP. Consequently, these are unable to encode the information which possibly achieved by analyzing different magnitude of edge responses in different directions of a particular pixel. Rather it considers only one directional edge magnitude. Motivated by this observation, we proposed the image feature Local Directional Pattern (LDP) that computes the edge response values in different directions and use these to encode the image texture.

$$\begin{bmatrix}-3 & -3 & 5\\ -3 & 0 & 5\\ -3 & -3 & 5\end{bmatrix} \quad \begin{bmatrix}-3 & 5 & 5\\ -3 & 0 & 5\\ -3 & -3 & -3\end{bmatrix} \quad \begin{bmatrix}5 & 5 & 5\\ -3 & 0 & -3\\ -3 & -3 & -3\end{bmatrix} \quad \begin{bmatrix}5 & 5 & -3\\ 5 & 0 & -3\\ -3 & -3 & -3\end{bmatrix}$$

East M_0 North East M_1 North M_2 North West M_3

$$\begin{bmatrix}5 & -3 & -3\\ 5 & 0 & -3\\ 5 & -3 & -3\end{bmatrix} \quad \begin{bmatrix}-3 & -3 & -3\\ 5 & 0 & -3\\ 5 & 5 & -3\end{bmatrix} \quad \begin{bmatrix}-3 & -3 & -3\\ -3 & 0 & -3\\ 5 & 5 & 5\end{bmatrix} \quad \begin{bmatrix}-3 & -3 & -3\\ -3 & 0 & 5\\ -3 & 5 & 5\end{bmatrix}$$

West M_4 South West M_5 South M_6 South East M_7

Figure 2: Kirsch edge response masks in eight directions

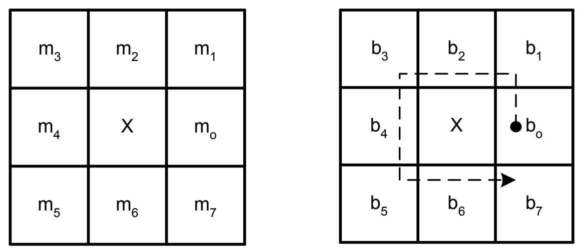
The proposed LDP feature is an eight bit binary code assigned to each pixel of an input image. This pattern is calculated by comparing the relative edge response value of a pixel in different directions. Kirsch edge detector, Prewitt edge detector, Sobel edge detector are the representative edge detectors which can be used in this regard [17]. Among them, the Kirsch edge detector has been known to detect different directional edge responses more accurately than others because the Kirsch edge detector considers all eight neighbors [18]. Given a central pixel in the image, the eight directional edge response values *{m_i}*, *i*=0,1,..,7 are computed by Kirsch masks M_i in eight different orientations centered on its position. These masks are shown in the fig. 2.

The response values are not equally important in all directions. The presence of corner or edge show high response values in some particular directions. Therefore, we are interested to know the k most prominent directions in order to generate the LDP. Here, the top k directional bit responses b_i are set to 1. The remaining (8-k) bits of 8-bit LDP pattern is set to 0. Finally, the LDP code is derived by (3). Fig. 3 shows the mask response and LDP bit positions, and fig. 4 shows an exemplary LDP code with k=3.

$$LDP_k = \sum_{i=0}^{7} b_i(m_i - m_k) \times 2^i \, . \tag{1}$$

$$b_i(a) = \begin{cases} 1 & a \geq 0 \\ 0 & a < 0 \end{cases} \tag{2}$$

where, m_k is the k-th most significant directional response.



(a) (b)

Figure 3: (a) Eight directional edge response positions. (b) LDP binary bit positions.

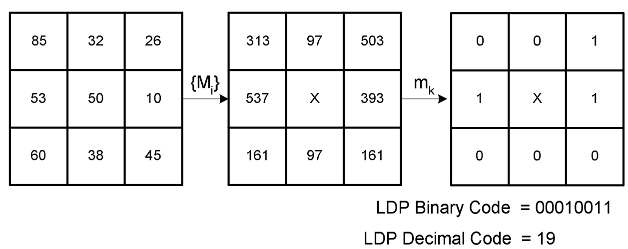


Figure 4: LDP code with k = 3.

2.3. Robustness of LDP

Since edge responses are more stable than intensity values, LDP pattern provides the same pattern value even presence of noise and non-monotonic illumination changes. For instance, fig. 5 shows an original image and the corresponding image after adding Gaussian white noise. After addition of noise, 5^{th} bit of LBP changed from 1 to 0, thus LBP pattern changed from uniform to a non-uniform code. Since edge response values are more stable than gray value, LDP pattern remains same even presence of that noise and non-monotonic illumination changes.

85	32	26
53	50	10
60	38	45

LBP = 00111000
LDP = 00010011

(a)

81	29	32
38	58	15
65	43	47

LBP = 00101000
LDP = 00010011

(b)

Figure 5: Stability of LDP vs. LBP (a) Original Image with LBP and LDP code, (b) Image with noise with new LBP and LDP code

2.4. Rotation invariant LDP

Rotational change of any image will lead to alter the spatial intensity distribution of that image. As a result edge response values of each direction will change and in consequence generate completely different LDP code. But observing the response values minutely we can conclude that relative position of these response values compared to strongest edge response will not affected by this rotation of the image. For example consider a particular image patch with eight edge response value indicated by fig. 6(a). If the image is rotated by 90^0 anticlockwise then intensity value of that image patch change which leads to change of edge response values. The rotated image patch along with edge responses are shown with fig 6(b). If we compare these two figures out that edge response values changes their corresponding direction by 90^0 anticlockwise. Observing this, we proposed a simple method for achieving rotation invariant LDP feature by applying a circular shift operation on the original binary code value showed with fig 7.

For achieving this shift operation, the direction of highest edge response is termed as the dominant direction of the LDP code. The bit value associated with the dominant direction is moved to the right most bit of the code. Then, the other bits are circularly shifted with the same number of bit position as the dominant direction bit shifted to get the right most position. For example, if the bit position original code is ''abcdefgh'' and dominant directional bit is at ''c'' location then bit value of 'c' position should be shifted 5 place to positioned into right most bit. So all other bit will also circularly shifted by 5 places and finally the normalized code will be ''defghabc''. This procedure generates rotation invariant LDP code which is denoted by LDP^{ri} and is shown with equation (3). This normalization method is based on the assumption that to compare similarity between two textures, they should be rotated so that their dominant directions are the same. It has been proved that image rotation in spatial domain is equivalent to circular shift of feature vector elements [19].

$$LDP^{ri} = ROR(LDP, d-1) \tag{3}$$

where d is the bit position of the strongest edge response.

The LDP operator produces C_k^8 different code values, corresponding to the eight bit binary patterns with exactly k bit value is 1. But after applying circular shift to get the rotation invariance number of LDP code reduces to C_{k-1}^7. In consequence histogram of rotation invariant LDP will have C_{k-1}^7 bins in comparison of C_k^8 bin in original LDP.

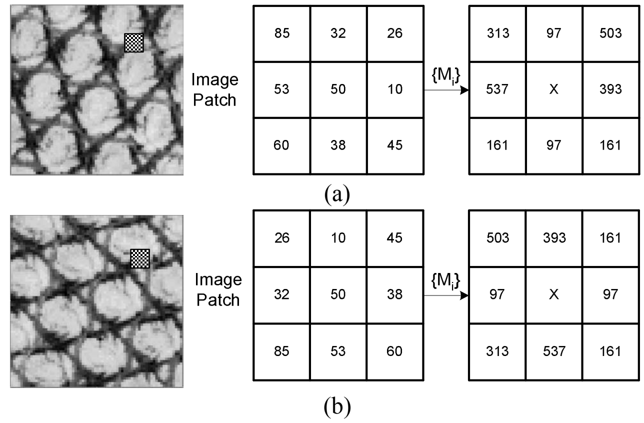
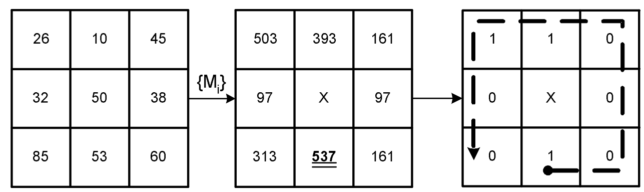


Figure 6: Changes in edge responses due to rotation of the image. (a) Original image and its edge response values (b) Rotated image along with changed edge responses values.



Rotation Invariant LDP Binary Code = 00110001
Rotation Invariant LDP Decimal Code = 49

Figure 7: Calculation of rotation invariant LDP code

2.5. LDP Descriptor

After generating LDP code in every pixel, we need to devise a method to generate an image descriptor using this LDP feature. In this regard, histogram has been widely used to represent, analyze, and characterize images [20]. Swain and Ballard [21] are pioneer in using histogram as image descriptor and following their work various recognition systems based on histograms were developed. Some of the reasons for their importance are that they can be computed easily and efficiently, they achieve significant data reduction, and they are robust to noise and local image transformations. In this work, we form histogram of LDP code to describe the image or image patch by accumulating the occurrence of LDP feature. So, after computing all the LDP code for every pixel (r,c), the input image I of size $M \times N$ is represented by a LDP histogram H using (3). The resultant histogram H is the LDP descriptor of that image.

$$H(\tau) = \sum_{r=1}^{M}\sum_{c=1}^{N} f\left(LDP_k(r,c),\tau\right) \tag{4}$$

$$f(a,\tau) = \begin{cases} 1 & a = \tau \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

where, τ is the LDP code value.

3. Texture classification using LDP descriptor

The LDP feature which is robust against different variations like non-monotonic changes in illumination and in random noise is used to represent the texture of the images for the application of texture classification. In this section, we present the method of using LDP histogram feature descriptor for texture classification and compare the performance with those of the state-of-the-art methods.

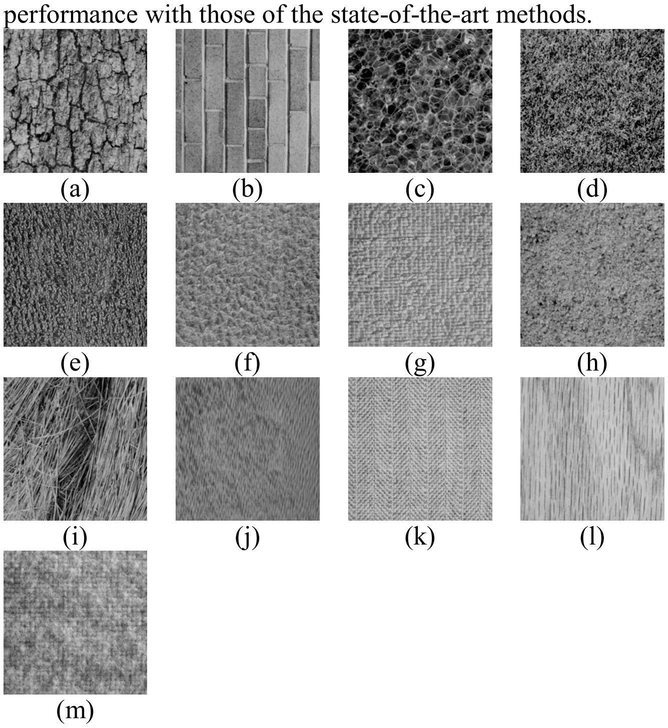


Figure 8: Primary pictures from Brodatz texture album: (a) Bark, (b) Brick, (c) Bubbles, (d) Grass, (e) Leather, (f) Pigskin, (g) Raffia, (h) Sand, (i) Straw, (j) Water, (k) Weave, and (l) Wood and (m) Wool.

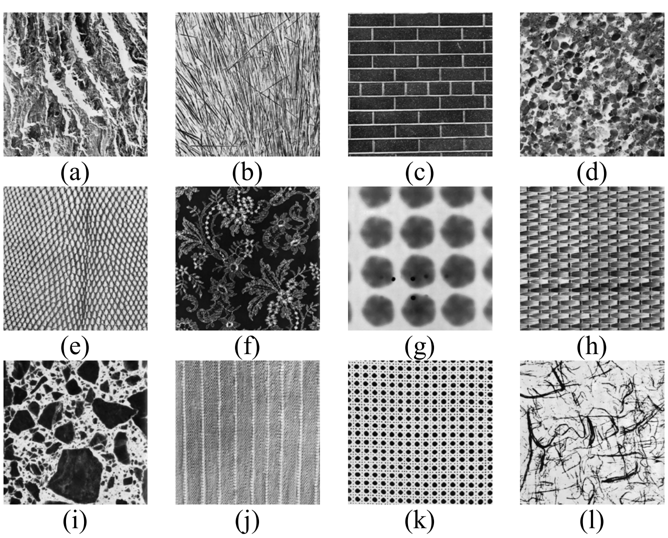


Figure 9: Some example picture from 112 Brodatz texture namely (D1-D112): (a) D13, (b) D15, (c) D26, (d) D27, (e) D35, (f) D40, (g) D48, (h) D56, (i) D62, (j) D85, (k) D101, (l) D108

3.1. Database

Experiments are carried out with two groups of textures; the first group consists of 13 primary textures, collected from Brodatz [22] texture album and the second group consists of all the 112 textures of Brodatz album. The first group of texture images are of size 512x512 used in our experiments and those texture of: (i) Bark, (ii) Brick, (iii) Bubbles, (iv) Grass, (v) Leather, (vi) Pigskin, (vii) Raffia, (viii) Sand, (ix) Straw, (x) Water, (xi) Weave, (xii) Wood and (xiii) Wool. Texture classification is done with a total of 91 (13x7=91) rotated textures of size 512x512 which are achieved by rotating the original image into seven different orientations of 0^0, 30^0, 60^0, 90^0, 120^0, 150^0 and 200^0. The rotated textures are of size 256x256, derived from the center portion of respective 512x512 size rotated textures.

The second group of textures comprises all the 112 textures of size 512x512 from the Brodatz album (D1–D112). They are rotated in steps of 10^0 up to 360^0 and used for classification, i.e., texture classification is done with a total of 4032 rotated texture images (112x36 = 4032). The rotated textures are of size 256x256, derived from the center portion of respective 512x512 size rotated textures.

3.2. Classification using LDP histogram

We choose rotation invariant LDP feature descriptor to represent the micro-textons of the given image. Extracted rotation invariant LDP features of each pixel of the image then combined to generate rotation invariant image descriptor using LDP histogram following equation 6. For the classification, we use the K-nearest neighbor, which has been successfully utilized in texture analysis arena. In our case, K=3. To compute the distance between two given images I_1 and I_2, we first obtain their LDP histogram features H_1 and H_2 then measure the similarity between those using histogram intersection [21]:

$$D(H^1, H^2) = \sum_{i=1}^{N} \min(H^1, H^2) \qquad (6)$$

where N is the number of bins in a histogram and it calculate the overlapping portion of two histograms.

3.3. Experimental result

Experimental results of two groups from Brodatz textures database are illustrated in Table 1. The accuracy of our method is calculated as a percentage of correct classifications which is computed as follows:

$$Accuracy = \frac{\#\, of\ correct\ classification}{\#\, of\ total\ images} \qquad (7)$$

As shown in Table 1, we compare our method with others popular dense descriptors like Gabor and LBP descriptor. We found that our approach performs better than those state-of-the-art methods.

Table 1: *The recognition result with Brodatz database*

Method	Texture of Group1	Texture of Group2
LDP	98.9	96.8
Gabor	96.7	95.9
LBP	93.4	91.2

4. Face recognition using LDP descriptor

The recent advancement of software and hardware technology has created more demand for personalized interaction in consumer products. This can be done by identifying the users through human face recognition and enabling appropriate services such as personalized TV program, intelligent digital photography, smart home, and may more. The LDP feature descriptor which is robust against different variations like non-monotonic changes in illumination and in random noise can be effectively used for face recognition.

4.1. Database

The performance of proposed LDP pattern is tested in the face recognition problem in accordance to the Colorado State University Face Identification Evaluation System [23] with images from the FERET [24] database. In this work, only frontal faces are considered with different lighting condition, different expression and with aging effects on the face image. These facial images can be divided into five sets which are known as: (i) ***fa*** set, used as a gallery set, contains frontal images of 1,196 people. (ii) ***fb*** set (1,195 images) with an alternative facial expression than in the *fa* photograph. (iii) ***fc*** set (194 images) taken under different lighting conditions. (iv) ***dupI*** set (722 images) taken later in time. (v) ***dupII*** set (234 images) subset of the *dup I* set containing images that were taken at least a year after the corresponding gallery image. Images from these five groups are shown in fig. 10. Images from the FERET database are cropped and normalized to 100×100 pixels based on the ground truth positions of the two eyes and mouth.

(a) (b) (c) (d) (e)

Figure 10: Example face image from FERET database. (a) Image from *fa* set, (b) Image from *fb* set, (c) Image from *fc* set, (d) Image from *dupI* set, (e) Image from *dupII* set.

4.2. Classification using LDP histogram

LDP histogram generated from the whole image will lose location information. Hence to incorporate some degree of location and spatial relationship, an extended LDP histogram is generated by dividing the image is divided into g number of regions $R_0, R_1, ..., R_{g-1}$ (shown in fig. 11) and building the LDP histogram from each region R_i. These histograms are concatenated to get the descriptor vector that represents the face image.

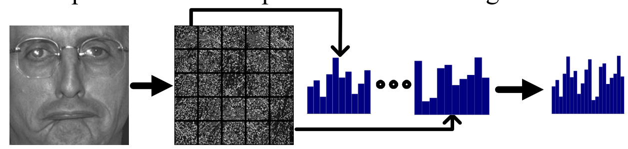


Figure 11: Facial image representation using spatially combined histogram

From the pattern classification perspective, a natural problem of face recognition is having a large number of classes but only a few, sometimes only one, number of training sample(s) are available for per class. In this situation, more sophisticated classifier is not applicable rather simple nearest-neighbor classifier is used in classify the face. Several dissimilarity measures have been proposed to compare closeness between two histograms named as – Histogram intersection, Log-likelihood statistics and Chi square statistics (χ^2). A weighted χ_w^2 statistics might be used to give more or less importance to particular regions such as eye, nose, and mouth regions. Literature shows that weights are set manually based on observations. In our case, we opted to use the χ^2 statistics for template matching.

$$\chi_w^2 = \sum_{i,\tau} w_i \frac{\left(S_i(\tau) - M_i(\tau)\right)^2}{S_i(\tau) + M_i(\tau)}. \qquad (9)$$

where, w_i is the weight of region R_i.

4.3. Experimental Result

In our setup, every image is partitioned into *7x6* sub-blocks. We used *fa* image set as gallery image and other four sets (*fb, fc, dupI* and *dupII*) as probe images. One image from probe set is compared using mentioned dissimilarity measure with all the images from gallery image set (*fa set*). The classification result is achieved through the nearest neighbor classification method. Table 2 shows recognition performance of proposed method along with other methods which ascertain the superiority of the proposed method.

Table 2: *The recognition result with FERET database*

Method	**fb**	**fc**	**dupI**	**dupII**
LDP, weighted	0.97	0.85	0.76	0.72
LDP, un-weighted	0.97	0.82	0.72	0.69
Gabor	0.97	0.80	0.67	0.64
LBP	0.97	0.79	0.66	0.64
PCA	0.85	0.65	0.44	0.22

5. Conclusions

This paper introduces a local feature descriptor LDP for object detection. LDP code which compute the edge response values in different directions and use this to encode the local image property. The discriminative power of the LDP descriptor mainly lies in the integration of local edge response into a single binary pattern that makes it robust and insensitive to noise and non-monotonous illumination changes. Experimental results show that LDP descriptor shows better classification accuracy on Brodatz textures database. It is also found that LDP descriptor provides better recognition accuracy in face recognition using FERET database.

6. References

[1] D. Lowe, Distinctive Image Features from Scale Invariant Key Points," International Journal of Computer Vision, vol. 60, no. 2, pp. 91-110, 2004.

[2] T. Ojala, M. Pietikäinen and T. Mäenpää, "Multiresolution Gray Scale and Rotation Invariant Texture Analysis with Local Binary Patterns, IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 24, no. 7, pp. 971 - 987, 2002.

[3] P. Viola and M. Jones, " Rapid Object Detection Using a Boosted Cascade of Simple Features," Proc. IEEE Int'l Conf. on Computer Vision and Pattern Recognition, 2001.

[4] T. Ahonen, A. Hadid and M. Pietikainen, "Face description with Local Binary Patterns: Application to face recognition", *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 28, no. 12, pp. 2037 –2041, Dec. 2006.

[5] J. Chen, S. Shan, C. He, G. Zhao, M. Pietikäinen, X. Chen, and W. Gao, "WLD: A Robust Local Image Descriptor", *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 31, no. 8, pp. 1 –16, Dec. 2009.

[6] P. Moreno, A. Bernardino, and J. Santos-Victor, "Improving the SIFT descriptor with smooth derivative filters", *Pattern Recognition Letter,* vol. 30, no. 1, pp. 18 –26, Dec. 2009.

[7] K. Mikolajczyk and C. Schmid, "A Performance Evaluation of Local Descriptors," IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 27, no. 10, pp. 1615-1630, 2005.

[8] Y. Ke and R. Sukthankar, "PCA-SIFT: A More Distinctive Representation for Local Image Descriptors," Proc. IEEE Int'l Conf. on Computer Vision and Pattern Recognition, 2004.

[9] H. Bay, T. Tuytelaars and L. van Gool. "SURF: Speeded Up Robust Features," Proc. European Conf. on Computer Vision, 2006

[10] B. Manjunath and W. Ma, "Texture Features for Browsing and Retrieval of Image Data," IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 18, no. 8, pp. 837-842, 1996.

[11] H. Zhou, R. Wanga, and C. Wanga, "A novel extended local-binary-pattern operator for texture analysis", *Information Sciences,* vol. 178, no. 22, pp. 4314 – 4325, 2008.

[12] T. Jabid, M. H. Kabir, and O. S. Chae, "Local Directional Pattern (LDP) for Face Recognition," *IEEE International Conference on Consumer Electronics*, January 2010.

[13] T. Jabid, M. H. Kabir, and O. S. Chae, "Gender Classification using Local Directional Pattern (LDP)," *Accepted in International Conference on Pattern Recognition*, August 2010.

[14] T. Jabid, M. H. Kabir, and O. S. Chae, "Facial Expression Recognition using Local Directional Pattern (LDP)," *Accepted in IEEE International Conference on Image Processing*, September 2010.

[15] S. Zhao, Y. Gao, and B. Zhang, "Sobel-LBP," *International Conference on Image Processing*, 2008.

[16] R. Mattivi and L. Shao, "Human Action Recognition Using LBP-TOP as Sparse Spatio-Temporal Feature Descriptor," *International Conference on Compuetr Analysis of Image and Pattern*, Sep. 2009.

[17] W.K. Pratt, Digital Image Processing, Wiley, New York, 1978.

[18] S. W. Lee, Off-line recognition of totally unconstrained handwritten numerals using multilayer cluster neural network, *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 18, no. 6, pp. 648–652, Jun 1996.

[19] Zhang, D.S., Wong, A., Indrawan, M., Lu, G., 2000. Content based image retrieval using Gabor texture features. In: Proc. of 1st IEEE Pacific Rim Conference on Multimedia (PCM'00), pp. 392–395.

[20] E. Hadjidemetriou, M. D. Grossberg, and S. K. Nayar, "Multiresolution Histograms and Their Use for Recognition", *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 26, no. 7, pp. 831 – 847, Jul. 2004.

[21] M. Swain and D. Ballard, "Color indexing," International Journal of Computer Vision, vol. 7, no. 1, pp. 11-32, 1991.

[22] P. Brodatz. Textures: A Photographic Album for Artists and Designers, Dover Publications, New York, 1966.

[23] J.R. Beveridge, D. Bolme, B.A. Draper, and M. Teixeira, "The CSU Face Identification Evaluation System: Its Purpose, Features, and Structure," *Machine Vision and Applications*, vol. 16, no. 2, pp. 128–138, Feb. 2005.

[24] P.J. Phillips, H. Wechsler, J. Huang, and P. Rauss, "The FERET Database and Evaluation Procedure for Face Recognition Algorithms," Image and Vision Computing, vol. 16, no. 10, pp. 295–306, 1998.