A Novel Approach to Face Recognition and Expression Analysis using Local Directional Number Pattern

Gloria Nongthombam[1] Mrs. Kusuma S[2]
[1]M.Tech-II [2]Assistant Professor
[1,2]Department of Information Science
[1,2]RNSIT, Bangalore, India

Abstract— Image analysis and understanding has recently received significant attention, especially during the past several years. At least two reasons can be accounted for this trend: the first is the wide range of commercial and law enforcement applications, and the second is the availability of feasible technologies after nearly 30 years of research. In this paper we propose a novel local feature descriptor, called Local Directional Number Pattern (LDN), for face analysis, i.e., face and expression recognition. LDN characterizes both the texture and contrast information of facial components in a compact way, producing a more discriminative code than other available methods. An LDN code is obtained by computing the edge response values in 8 directions at each pixel with the aid of a compass mask. This directional information is then encoded them into a 6 bit binary number using the relative strength of these edge responses. To analyze a face, it is divided into several regions, and the distribution of the LDN features are extracted from them. Then, these features are concatenated into a feature vector, and it is used as a face descriptor. We perform several experiments in which our descriptor showed consistent results under age, illumination, expression, and noise variations.

Keywords: Expression recognition, face descriptor, face recognition, features extraction, image representation, local pattern.

I. INTRODUCTION

Face analysis is a task that we humans perform routinely and effortlessly in our daily lives. With the wide availability of powerful and low-cost desktop and embedded computing systems, an enormous interest has been created in automatic processing of digital images and videos. Face analysis has a varied range of applications, namely biometric authentication, surveillance, human-computer interaction, and multimedia management. Due to the endless possibility of its application and the interest generated, research and development in automatic face analysis which consist of face recognition and expression recognition follows naturally.

In face analysis, a key issue is the descriptor of the face appearance [1][2]. The efficiency of the descriptor depends on its representation and the ease of extracting it from the face. Ideally, a good descriptor should have a high variance among classes i.e. between different persons or expressions, but little or no variation within classes i.e. same person or expression in different conditions. These descriptors are used in several areas, such as, facial expression and face recognition.

Finding efficient facial features to represent the face appearance is the most critical aspect in face recognition. Facial features fall into two classes – global feature and local feature [3]. In global feature extraction process, the whole image is taken into account, but local feature considers only the local region within the image. There are many methods for local feature or the holistic class, such as, Eigenfaces [4] and Fisherfaces [5], which are built on Principal Component Analysis (PCA) [4]; the more recent 2D PCA [6] and Linear Discriminant Analysis [7] are also examples of holistic methods. Although, the global features are popular and studied widely, their performances deteriorate in changing environment. Hence local features are gaining more attention because of their robustness in uncontrolled environment like illumination and pose variations.

The local-feature methods compute the descriptor from parts of the face, and then gather the information into one descriptor. Local Features Analysis (LFA)[8], Gabor features [9], Elastic Bunch Graph Match[10], Local Binary Pattern (LBP) [11], [12] are popular among the local methods for locating the local face features. The LBP feature, which was originally designed for texture description [13], is applied to face recognition. LBP provides an illumination invariant description of face image so it gained popularity and is widely used. Newer methods tried to overcome the shortcomings of LBP, like Local Ternary Pattern (LTP) [14], and Local Directional Pattern (LDiP) [15]–[17]. These methods use other information, instead of intensity, to overcome noise and illumination variation problems. However these methods still suffer in non-monotonic illumination variation, random noise, and changes in pose, age, and expression conditions.

This paper describes a face descriptor, Local Directional Number Pattern (LDN) for robust face analysis that encodes the intensity variations thus distinguishing the face's textures. LDN considers the edge response values in all different directions with a compass mask instead of surrounding neighboring pixel intensities like LBP. This provides more consistency in the presence of noise; since edge response magnitude is more stable than pixel intensity. LDN is more compact with a six bit long code and conveys more information.

II. LOCAL DIRECTIONAL NUMBER PATTERN

The proposed Local Directional Number Pattern (LDN) is a six bit binary code assigned to each pixel of an input image that represents the structure of the texture and its intensity transitions. Edge magnitudes are used because they are insensitive to lighting changes. As a result, the LDN pattern is created by computing the edge response of the neighborhood using a compass mask, and by taking the top directional numbers, that is, the most positive and negative directions of those edge responses. The positive and negative responses provide valuable information about the structure of the neighborhood, because they reveal the gradient direction of bright and dark areas in the neighborhood. Thereby, this distinction, allows LDN to differentiate between blocks with the positive and the negative direction swapped by generating a different code for each instance, while other methods may mistake the swapped regions as one. Furthermore, these transitions occur often in the face, for example, the top and bottom edges of the eyebrows and mouth have different intensity transitions. Thus, it is important to differentiate among them; LDN can accomplish this task as it assigns a specific code to each of them.

The key points of our proposed method are: 1) the coding scheme is based on directional numbers, instead of bit strings, which encodes the information of the neighbourhood in a more efficient way; 2) the implicit use of sign information, in comparison with previous directional and derivative methods we encode more information in less space, and, at the same time, discriminate more textures; and 3) the use of gradient information makes the method robust against illumination changes and noise.

III. ALGORITHM DESCRIPTION

A. Coding Scheme

In our coding scheme, we generate the code, LDN, by analyzing the edge response of each mask, $\{M^0,...,M^7\}$, that represents the edge significance in its respective direction, and by combining the dominant directional numbers. Given that the edge responses are not equally important; the presence of a high negative or positive value signals a prominent dark or bright area. Hence, to encode these prominent regions, we implicitly use the sign information, as we assign a fixed position for the top positive directional number, as the three most significant bits in the code, and the three least significant bits are the top negative directional number, as shown in Fig. 1. Therefore, the code is defined as:

$$LDN(x,y) = 8i_{x,y} + j_{x,y} \qquad (1)$$

where *(x, y)* is the central pixel of the neighborhood being coded, $i_{x,y}$ is the directional number of the maximum positive response, and $j_{x,y}$ is the directional number of the minimum negative response defined by:

$$i_{x,y} = \arg\max_i \{\Pi^i(x,y) \mid 0 \le i \le 7\} \qquad (2)$$

$$j_{x,y} = \arg\max_j \{\Pi^j(x,y) \mid 0 \le j \le 7\} \qquad (3)$$

where $\prod^i$ is the convolution of the original image, *I*, and the i^{th} mask, M^i , defined by:

$$\Pi^i = I * M^i \qquad (4)$$

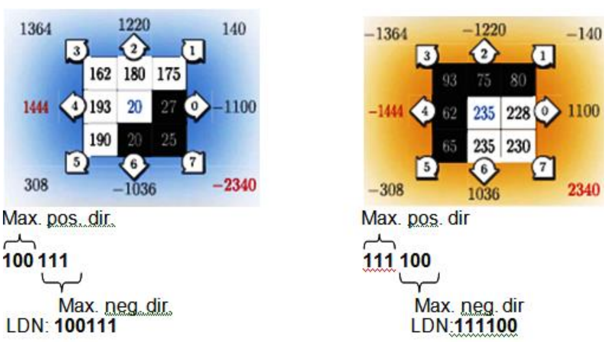

Fig.1: LDN Code Computation

B. Compass Mask

To produce the LDN code, a compass mask is needed to compute the edge responses. Two different asymmetric masks are used: Kirsch and derivative-Gaussian as shown in Fig. 2 and Fig. 3. Both masks operate in the gradient space, which reveals the structure of the face.

The Kirsch mask is rotated 45° apart to obtain the edge response in eight different directions, as shown in Fig. 2. The use of this mask to produce the LDN code is denoted by LDN^K. The derivative of a skewed Gaussian is used to create an asymmetric compass mask that is used to compute the edge response on the smoothed face. This mask is robust against noise and illumination changes, while producing strong edge responses. Hence, given a Gaussian mask defined by:

$$G_\sigma(x,y) = \frac{1}{2\pi\sigma^2}\exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) \qquad (5)$$

where *x*, *y* are location positions, and σ is the width of the Gaussian bell; we define the mask as:

$$M_\sigma(x,y) = G''_\sigma(x+k,y) * G_\sigma(x,y) \qquad (6)$$

where G'_σ is the derivative of G_σ with respect to *x*, σ is the width of the Gaussian bell, * is the convolution operation, and *k* is the offset of the Gaussian with respect to its centre. A compass mask is generated, $\{M^0\sigma ,...,M^7\sigma\}$, by rotating $M\sigma$ 45° apart, in eight different directions. Thus, a set of masks is obtained similar to those shown in Fig. 3.

$$\underset{M_0}{\begin{bmatrix} -3 & -3 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & 5 \end{bmatrix}} \underset{M_1}{\begin{bmatrix} -3 & 5 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & -3 \end{bmatrix}} \underset{M_2}{\begin{bmatrix} 5 & 5 & 5 \\ -3 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix}} \underset{M_3}{\begin{bmatrix} 5 & 5 & -3 \\ 5 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix}}$$

$$\underset{M_4}{\begin{bmatrix} 5 & -3 & -3 \\ 5 & 0 & -3 \\ 5 & -3 & -3 \end{bmatrix}} \underset{M_5}{\begin{bmatrix} -3 & -3 & -3 \\ 5 & 0 & -3 \\ 5 & 5 & -3 \end{bmatrix}} \underset{M_6}{\begin{bmatrix} -3 & -3 & -3 \\ -3 & 0 & -3 \\ 5 & 5 & 5 \end{bmatrix}} \underset{M_7}{\begin{bmatrix} -3 & -3 & -3 \\ -3 & 0 & 5 \\ -3 & 5 & 5 \end{bmatrix}}$$

Fig.2: Kirsch compass masks

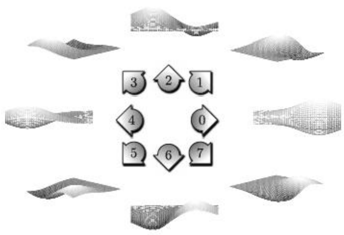

Fig.3: Derivative of Gaussian compass masks, computed by (6)

C. Histogram of LDP

Each face is represented by a LDN histogram (LH) as shown in Fig. 4. The LH contains fine to coarse information of an image, such as edges, spots, corners and other local texture features. Given that the histogram only encodes the occurrence of certain micro-patterns without location information, to aggregate the location information to the descriptor, we divide the face image into small regions $\{R^1,...,R^N\}$, and extract a histogram H^i from each region R^i . We create the histogram, H^i, using each code as a bin, and then accumulate all the codes in the region in their respective bin by:

$$H^i(c) = \sum_{\substack{(x,y)\in R^i \\ LDN(x,y)=c}} v, \forall c \quad (7)$$

where c is a LDN code, and (x, y) is a pixel position in the region R^i, LDN(x, y) is the LDN code for the position (x, y), and v is the accumulation value—commonly the accumulation value is one. Finally, the LH is computed by concatenating those histograms:

$$LH = \prod_{i=1}^{N} H^i \quad (8)$$

Where ∏is the concatenation operation, and N is the number of regions of the divided face. The spatially combined LH plays the role of a global face feature for the given face. We represent the face using a single-feature histogram, by using LH, or with a multi-feature histogram, by using MLH.

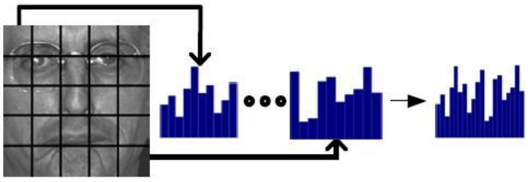

Fig. 4: Fig Facial image representation using spatially enhanced histogram.

IV. FACE ANALYSIS

A. Face Recognition

The LH and MLH are used during the face recognition process. The objective is to compare the encoded feature vector from one person with all other candidate's feature vector with the Chi-Square dissimilarity measure. This measure between two feature vectors, F1 and F2, of length N is defined as:

$$\chi^2(F_1, F_2) = \sum_{i=1}^{N} \frac{(F_1(i) - F_2(i))^2}{F_1(i) + F_2(i)} \quad (9)$$

The corresponding face of the feature vector with the lowest measured value indicates the match found.

B. Expression Recognition

We perform the facial expression recognition by using a Support Vector Machine (SVM) to evaluate the performance of the proposed method. SVM [41] is a supervised machine learning technique that implicitly maps the data into a higher dimensional feature space. Consequently, it finds a linear hyperplane, with a maximal margin, to separate the data in different classes in this higher dimensional space. Given a training set of M labeled examples $T = \{(x_i, y_i) \mid i = 1,...,M\}$, where $x_i \in \mathbb{R}^n$ and $y_i \in \{-1, 1\}$, the test data is classified by:

$$f(x) = sign(\sum_{i=1}^{M} \alpha_i y_i K(x_i, x) + b) \quad (10)$$

where αi are Lagrange multipliers of dual optimization problem, b is a bias, and $K(\cdot, \cdot)$ is a kernel function. Given that SVM makes binary decisions, multi-class classification can be achieved by adopting the one-against-one or one-against-all techniques.

V. RESULTS

The performance of proposed LDN pattern is tested in the face recognition problem in accordance to the CSU Face Identification Evaluation System with images from the FERET database. Images are cropped and normalized to 100×100 pixels based on the ground truth positions of the two eyes and mouth. In our setup, every image is partitioned into 10x10 subblocks. We used *fa* image set as gallery image and other four sets of probe images, those are *fb* (expression variation), *fc* (illumination variation), *dupI* (age variation) and *dupII* (age variation). We compared performance of proposed LDN based method with LBP and PCA. The recognition rate is shown in Table I. Experimental results reflect that LDN texture description is more robust in lighting condition and aging effects. The proposed method cannot merely recognize face but can recognize with change in pose, age, expression.

Method	fb	fc	dupI	dupII
LDN	0.97	0.82	0.72	0.69
LBP	0.97	0.79	0.66	0.64
PCA	0.84	0.65	0.44	0.22

Table. 1: Result in FERET Database.

VI. CONCLUSIONS

This paper describes a novel encoding scheme called LDN which is computed from the edge response value using a compass mask. LDN takes advantage of the structure of the face's textures and encodes it efficiently into a compact code. It uses directional information that is more stable against noise than intensity, to code the different patterns from the face's textures. *LDN*, implicitly, uses the sign information of the directional numbers which allows it to distinguish similar texture's structures with different intensity transitions, *e.g.*, from dark to bright and vice versa. This new code performs with higher accuracy under difference expressions and aging conditions, making the system run reliably in uncontrolled environment.

References

[1] R. Chellappa, C. Wilson, and S. Sirohey, "Human and machine recognition of faces: A survey," *Proc. IEEE*, vol. 83, no. 5, pp. 705–741, May 1995.

[2] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, "Face recognition: A literature survey," *ACM Comput. Surv.*, vol. 35, no. 4, pp. 399–458, Dec. 2003

[3] [3] S. Z. Li, A. K. Jain, Y. L. Tian, T. Kanade, and J. F. Cohn, "Facial expression analysis," in *Handbook of Face Recognition*. New York: Springer-Verlag, 2005, pp. 247–275.

[4] M. Turk and A. Pentland, "Eigenfaces for recognition," *J. Cognit. Neurosci.*, vol. 3, no. 1, pp. 71–86, 1991.

[5] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 7, pp. 711–720, Jul. 1997.

[6] J. Yang, D. Zhang, A. F. Frangi, and J. Y. Yang, "Two-dimensional PCA: A new approach to appearance-based face representation and recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 1, pp. 131–137, Jan. 2004.
[7] K. Etemad and R. Chellappa, "Discriminant analysis for recognition of human face images," *J. Opt. Soc. Amer. A*, vol. 14, no. 8, pp. 1724–1733, 1997.
[8] P. Penev and J. Atick, "Local feature analysis: A general statistical theory for object representation," *Netw., Comput. Neural Syst.*, vol. 7, no. 3, pp. 477–500, 1996.
[9] D. Gabor, "Theory of communication," *J. Inst. Electr. Eng. III, Radio Commun. Eng.*, vol. 93, no. 26, pp. 429–457, Nov. 1946.
[10] L. Wiskott, J.-M. Fellous, N. Kuiger, and C. von der Malsburg, "Face recognition by elastic bunch graph matching," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 7, pp. 775–779, Jul. 1997.
[11] C. Shan, S. Gong, and P. W. McOwan, "Facial expression recognition based on local binary patterns: A comprehensive study," *Image Vis. Comput.*, vol. 27, no. 6, pp. 803–816, 2009.
[12] T. Ahonen, A. Hadid, and M. Pietikäinen, "Face description with local binary patterns: Application to face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 12, pp. 2037–2041, Dec. 2006.
[13] T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on feature distributions," *Pattern Recognit.*, vol. 29, no. 1, pp. 51–59, Jan. 1996.
[14] X. Tan and B. Triggs, "Enhanced local texture feature sets for face recognition under difficult lighting conditions," *IEEE Trans. Image Process.*, vol. 19, no. 6, pp. 1635–1650, Jun. 2010.
[15] T. Jabid, M. H. Kabir, and O. Chae, "Local directional pattern (LDP) for face recognition," in *Proc. IEEE Int. Conf. Consum. Electron.*, Mar. 2010, pp. 329–330.
[16] T. Jabid, M. H. Kabir, and O. Chae, "Robust facial expression recognition based on local directional pattern," *ETRI J.*, vol. 32, no. 5, pp. 784–794, 2010.
[17] M. Kabir, T. Jabid, and O. Chae, "A local directional pattern variance (LDPv) based face descriptor for human facial expression recognition," in *Proc. 7th IEEE Int. Conf. Adv. Video Signal Based Surveill.*, Sep. 2010, pp. 526–532.

