IJREAT International Journal of Research in Engineering & Advanced Technology, Volume 2, Issue 2, Apr-May, 2014
**ISSN: 2320 – 8791 (Impact Factor: 1.479)**
**www.ijreat.org**

# Robust Facial Expression Recognition

**Mr. Mukund Kumar[1], Ms. D. Udaya[2]**

[1, 2]**Computer Science and Engineering, Dr. Pauls Engineering College,Villupuram**

**Abstract**

This paper proposes a novel local feature descriptor, Local Directional Number Pattern (LDN), for face analysis: face and expression recognition. LDN encodes the directional information of the face's textures (i.e., the texture's structure) in a compact way, producing a more discriminative code than current methods. We compute the structure of each micro-pattern with the aid of a compass mask, that extracts directional information, and we encode such information using the prominent direction indexes (directional numbers) and sign which allows us to distinguish among similar structural patterns that have different intensity transitions. We divide the face into several regions, and extract the distribution of the LDN features from them. Then, we concatenate these features into a feature vector, and we use it as a face descriptor. We perform several experiments in which our descriptor performs consistently under illumination, noise, expression, and time lapse variations. Moreover, we test our descriptor with different masks to analyze its performance in different face analysis tasks.

**Index Terms**—*Local pattern, directional number pattern, image descriptor, face descriptor, feature, face recognition, expression recognition*

## 1. Introduction

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 (between different persons or expressions), but little or no variation within classes (same person or expression in different conditions). These descriptors are used in several areas, such as, facial expression and face recognition. There are two common approaches to extract facial features: geometric-feature-based and appearance-based methods [3]. The former [4], [5] encodes the shape and locations of different facial components, which are combined into a feature vector that represents the face. An instance of these methods is the graph-based methods [6], which use several facial components to create a representation of the face and process it. Moreover, the Local-Global Graph algorithm [6] is an interesting approach that uses Voronoi tessellation and Delaunay graphs to segment local features and builds a graph for face and expression recognition. These features are mixed into a local graph, and then the algorithm creates a skeleton (global graph) by interrelating the local graphs to represent the topology of the face. Furthermore, facial features are widely used in expression recognition, as the pioneer work of Ekman and Friesen identifying six basic emotions produced a system to categorize the expressions, known as Facial Action Coding System, and later it was simplified to the Emotional Facial Action Coding System. However, the geometric-feature-based methods usually require accurate and reliable facial feature detection and tracking, which is difficult to accommodate in many situations. The appearance-based methods use image filters, either on the whole-face, to create holistic features, or some specific face-region, to create local features, to extract the appearance changes in the face image. The performance of the appearance-based methods is excellent in constrained environment but their performance degrades in environmental variation [1].

In the literature, there are many methods for the holistic class, such as, Eigenface and Fisher faces which are built on Principal Component Analysis (PCA) [2]; the more recent 2D PCA and Linear Discriminant Analysis are also examples of holistic methods. Although these methods have been studied widely, local descriptors have gained attention because of their robustness to illumination and pose variations. Heisele et al. showed the validity of the component-based methods, and how they outperform holistic methods. The local-feature methods compute the descriptor from parts of the face, and then gather the information into one descriptor. Among these methods are Local Features Analysis, Gabor features, Elastic Bunch Graph Matching, and Local Binary Pattern (LBP). The last one is an extension of the LBP feature that was originally designed for texture description applied to face recognition. LBP achieved better performance than previous methods, thus it gained popularity, and was studied extensively. Newer methods tried to overcome the shortcomings of LBP, like Local Ternary Pattern (LTP) [7], and Local Directional Pattern ($LD_iP$).The last method encodes the directional information in the neighborhood, instead of the intensity. Also, Zhang et al. explored the use of higher order local derivatives ($LD_eP$) to produce better results than LBP. Both 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. Although some methods, like Gradient faces have a high discrimination power under illumination variation, they still have low recognition capabilities for expression and age variation conditions. However, some methods explored different features, such as, infrared, near infrared and phase information to overcome the illumination problem while maintaining the performance under difficult conditions. In this paper, we propose a face descriptor, Local Directional Number Pattern (LDN), for robust face recognition that encodes the structural information and the intensity variations of the face's texture. LDN encodes the structure of a local neighborhood by analyzing its directional information. Consequently, we compute the edge responses in the neighborhood, in eight different directions with a compass mask.

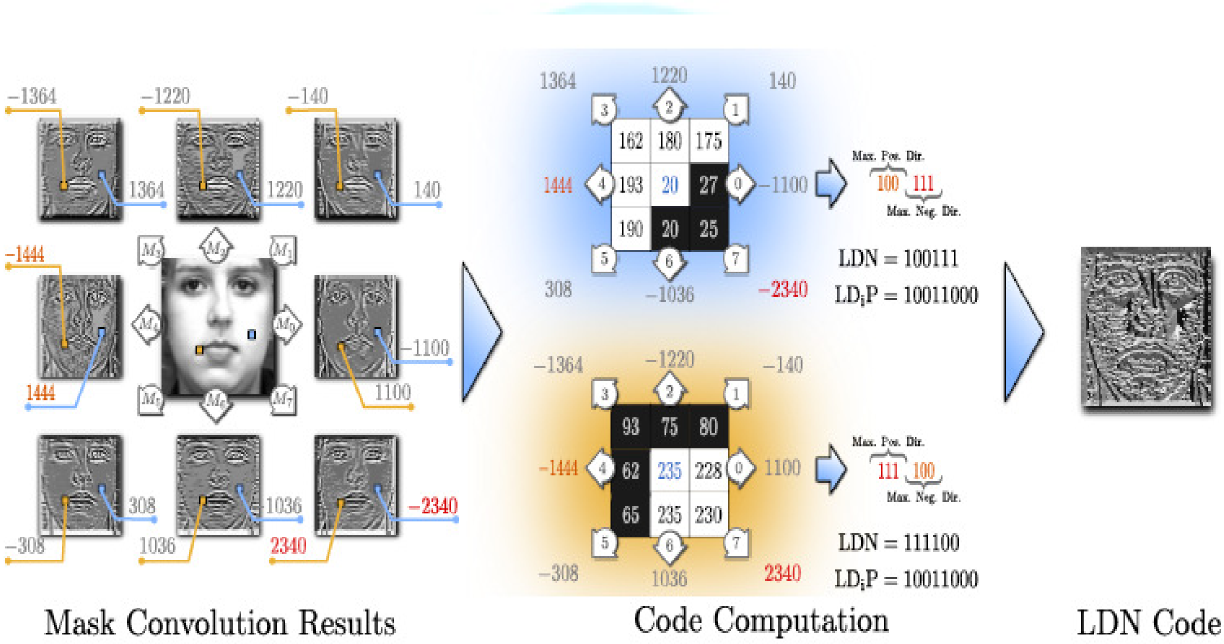


Fig.1:LDN code computation

Then, from all the directions, we choose the top positive and negative directions to produce a meaningful descriptor for different textures with similar structural patterns. This approach allows us to distinguish intensity changes (e.g., from bright to dark and vice versa) in the texture, that otherwise will be missed fig. 1. Furthermore, our descriptor uses the information of the entire neighborhood, instead of using sparse points for its computation like LBP. Hence, our approach conveys more information into the code, yet it is more compact as it is six bit long. Moreover, we experiment with different masks and resolutions of the mask to acquire characteristics that may be neglected by just one, and combine them to extend the encoded information. We found that the inclusion of multiple encoding levels produces an improvement in the detection process.

## 2. Existing System

There are two common approaches to extract facial features: geometric feature based and appearance-based methods. The former encodes the shape and locations of different facial components, which are combined into a feature vector that represents the face. An instance of these methods is the graph-based methods which use several facial components to create a representation of the face and process it. Moreover, the Local-Global Graph algorithm is an interesting approach that uses Voronoi tessellation and Delaunay graphs to segment local features and builds a graph for face and expression recognition. These features are mixed into a local graph and then the algorithm creates a skeleton (global graph) by interrelating the local graphs to represent the topology of the face.

Drawbacks in Existing System

1) The geometric-feature-based methods usually require accurate and reliable facial feature detection and tracking, which is difficult to accommodate in many situations.
2) The performance of the appearance-based methods is excellent in constrained environment but their performances degrade in environmental variation.

## 3. Scope Of The Project

A face descriptor Local Directional Number (LDN) for robust face recognition that encodes the structural information and the intensity variations of the face's texture is proposed in this method. LDN uses directional information that is more stable against noise than intensity to code the different patterns from the face's textures. Additionally, analyzed the use of two different compass masks (a derivative-Gaussian and Kirsch) to extract this directional information and their performance on different applications.

### 3.1 Problem Definition

The geometric feature based methods usually require accurate and reliable facial feature detection and tracking which is difficult to accommodate in many situations. The appearance based methods use image filters either on the whole face to create holistic features or some specific face region to create local features to extract the appearance changes in the face image. The performance of the appearance based methods is excellent in constrained environment but their performances degrade in environmental variation.

## 4. 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. As previous research showed, edge magnitudes are largely insensitive to lighting changes. Consequently, we create our pattern 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. We illustrate this coding scheme in fig. 1. The positive and negative responses provide valuable information of the structure of the neighborhood, as they reveal the gradient direction of bright and dark areas in the neighborhood. Thereby, this distinction, between dark and bright responses, allows LDN to differentiate between blocks with the positive and the negative direction swapped (which is equivalent to swap the bright and the dark areas of the neighborhood, as shown in the middle of fig. 1) 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.

Difference with previous work

Current methods have several shortcomings. For example, LBP encodes the local neighborhood intensity by using the center pixel as a threshold for a sparse sample of the neighboring pixels. The few number of pixels used in this method introduce several problems: First, it limits the accuracy of the method. Second, the method discards most of the information in the neighborhood. Finally, it makes the method very sensitive to noise. Moreover, these drawbacks are more evident for bigger neighborhoods. Consequently, to avoid these problems more information from the neighborhood can be used, as other methods do. Although the use of more information makes these methods more stable, they still encode the information in a similar way as LBP: by marking certain characteristics in a bit string. And despite the simplicity of the bit string coding strategy, it discards most information of the neighborhood.
In summary, the key points of our proposed method are:

(1) The coding scheme is based on directional numbers, in-stead of bit strings, which encodes the information of the neighborhood in a more efficient way.
(2) (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.

## 5. Proposed System

A face descriptor Local Directional Number (LDN) for robust face recognition is proposed, that encodes the structural information and the intensity variations of the face's texture. LDN encodes the structure of a local neighbourhood by analyzing its directional information. Consequently then compute the edge responses in the neighbourhood in eight different directions with a compass mask. Then from all the directions I choose the top positive and negative directions to produce a meaningful descriptor for different textures with similar structural patterns. This approach allows distinguishing intensity changes.

## Advantages of Proposed System

1) The proposed Local Directional Number (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.
2) It computes the edge responses in the neighborhood in eight different directions with a compass mask. Then, from all the directions we choose the top positive and negative directions to produce a meaningful descriptor for different textures with similar structural patterns.

## A. Coding scheme

In our coding scheme, we generate the code, LDN, by analyzing the edge response of each mask, $\{M^0 \ldots\ldots M^7\}$, 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, we define the code 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:

$$i_{x;y} = \arg\ {}_i \max fI^i(x; y)\ j\ 0\quad i\quad 7g; \qquad (2)$$

$$j_{x;y} = \arg\ {}_j \min fI^j(x; y)\ j\ 0\quad j\quad 7g; \qquad (3)$$

Where $I^i$ is the convolution of the original image, I, and the ith mask, $M^i$, defined by:

$$I^i = I\ \ M^i: \qquad (4)$$

.

## B. Compass masks

We use the gradient space, instead of the intensity feature space, to compute our code. The former has more information than the later, as it holds the relations among pixels implicitly (while the intensity space ignores these relations). Moreover, due to these relations the gradient space reveals the underlying structure of the image. Consequently, the gradient space has more discriminating power to discover key facial features. Additionally, we explore the use of a Gaussian to smooth the image, which makes the gradient computation more stable. These operations make our method more robust; similarly previous research used the gradient space to compute their code. Hence, our method is robust against illumination due to the gradient space, and to noise due to the smoothing. To produce the LDN code, we need a compass mask to compute the edge responses. In this paper, we analyze our proposed code using two different asymmetric masks: Kirsch and derivative-Gaussian (shown in figs. 2 and 3). Both masks operate in the gradient space, which reveals the structure of the face. Furthermore, we explore the use of Gaussian smoothing to stabilize the code in presence of noise by using the derivative-Gaussian mask. The Kirsch mask is rotated 45 apart to obtain the edge response in eight different directions, as shown in fig. 2. We denote the use of this mask to produce the LDN code by $LDN^K$.

## C. Derivative-Gaussian Mask

The use of the derivative-Gaussian mask allows to freely varying the size of the mask. The change in the size allows the coding scheme LDNG to capture different characteristics of the face. Hence, a fine to coarse representation is achieved by computing the LDNG and by concatenating the histogram of each region and can also merge the characteristics at different resolutions.

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

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

Fig. 2. Kirsch compass masks.

This mixture of resolutions is called as a multi-LDN histogram (MLH).

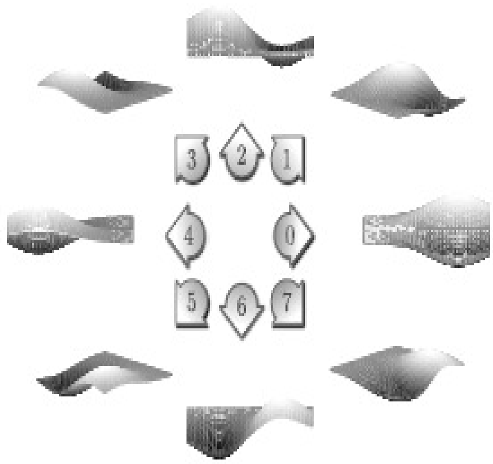

Fig. 3. Derivative of Gaussian compass masks, computed by Eq. (6).

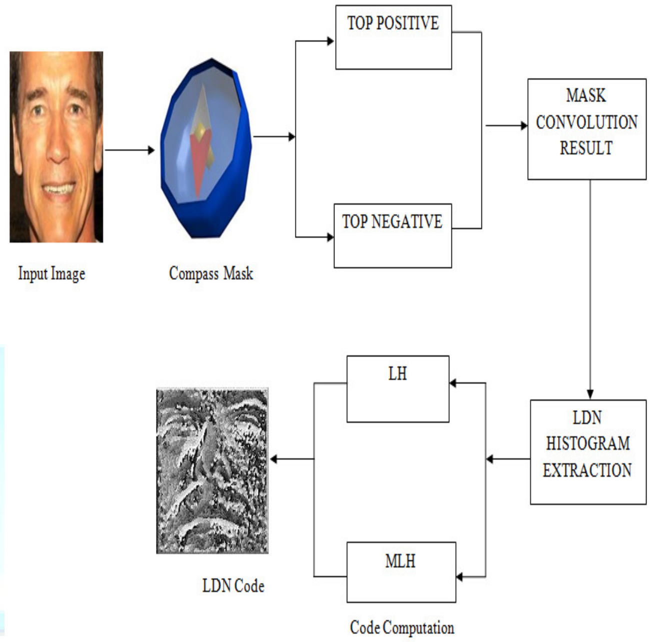


Fig.4. Proposed System model

## 6. Face Description

Each face is represented by a LDN histogram (LH) 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, this divides the face image into small regions $\{R^1 \ldots R^N\}$ and extract a histogram Hi from each region Ri. The histogram Hi is created using each code as a bin and then accumulating all the codes in the region in their respective bin. The spatially combined LH plays the role of a global face feature for the given face.

### 6.1. 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. The feature vector with the lowest measured value indicates the match found.

### 6.2. Expression Recognition

The facial expression recognition is performed by using a Support Vector Machine (SVM) to evaluate the performance of the proposed method. SVM is a supervised machine learning technique that implicitly maps the data into a higher dimensional feature space. Consequently, it finds a linear hyper plane with a maximal margin to separate the data in different classes in this higher dimensional feature space. Consequently, it finds a linear hyper plane, with a maximal margin, to separate the data in different classes in this higher dimensional space.

## 7. Conclusion

A novel encoding scheme has been introduced using Local Directional Number (LDN) that takes advantage of the structure of the face's textures and encodes it efficiently into a compact code. LDN uses directional information that is more stable against noise than intensity to code the different patterns from the face's textures. Additionally it analyzed the use of two different compass masks (a derivative-Gaussian and Kirsch) to extract this directional information and their performance on different applications. In general 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. To evaluate LDN under expression time lapse and illumination variations and found that it is reliable and robust throughout all these conditions unlike other methods. For example, Gradient faces had excellent results under illumination variation but failed with expression and time lapse variation. Also LBP and LDiP recognition rate deteriorates faster than LDN in presence of noise and illumination changes.

## References

[1] Hsu .C .W and Lin C. J (2002) "A Comparison of Methods for Multiclass Support Vector Machines" IEEE Trans. Neural Newts., vol. 13, no. 2, pp. 415–425.

[2] Hyeonjoon Moon and Jonathon Phillips .P (2011) "Computational and Performance Aspects of PCA- based face – recognition algorithms" Perception, volume 30, pages 303-321.

[3] Lee K .C, Ho .J and Krieg man D .J (2005) "Acquiring Linear Subspaces for Face Recognition under Variable Lighting" IEEE Transactions on Pattern Analysis and Machine Intelligence vol. 27, no. 5, pp. 684–698.

[4] Liao .S, Law .M and Chung .A (2009) "Dominant Local Binary Patterns for Texture Classification" IEEE Transactions on Image Processing, vol. 18, no. 5.

[5] Tan .X and Trigg's .B (2010) "Enhanced Local Texture Feature Sets for Face Recognition under Difficult Lighting Conditions" IEEE Trans. Image Process., vol. 19, no. 6, pp. 1635–1650.

[6] Taskeed Jabid, Md. Hasanul Kabir and Oksam Chae (2012) "Local Directional Pattern Variance (LDPv): A Robust Feature Descriptor for Facial Expression Recognition" International Journal of Innovative Computing, Information and Control Volume 8, Number 4.

[7] Zhang .B, Zhang .L, Zhang .D and Shen .L (2010) "Directional Binary Code with Application to PolyU Near-Infrared Face Database" Pattern Recognition Letters, vol. 31, issue 14, pp. 2337-2344.