A Review of the Most Important Textual Descriptors Based on Local Binary Patterns for Facial Expression Recognition

MohammadReza Vatandoust\textsuperscript{a}, Tajedin Derikvand\textsuperscript{b}, Amin Keshavarzi\textsuperscript{c}

\textsuperscript{a} Main Author, Department of Computer, Fars Science and Research Branch, Islamic Azad University, Marvdasht, Iran.

\textsuperscript{b} Corresponding Author, Department of Computer, Marvdasht Branch, Islamic Azad University, Marvdasht, Iran.

\textsuperscript{c} Advisor, Department of Computer, Marvdasht Branch, Islamic Azad University, Marvdasht, Iran.

Abstract

Since human facial expression transmits many information, the proper detection is very important. In recent years, many studies have been conducted on the Human and Computer Interaction (HCI), but high-precision facial expression recognition is still a challenging issue and an important one in the field of image processing. Many of these studies have detected only the initial emotional states that include seven main categories of neural, anger, disgust, fear, happy, sadness and surprise. Extracting the attribute from incredible input images is important, and it is necessary to consider the features of the image which have a large effect on identifying the state.

Keywords: Feature Extraction, Local Binary Pattern, Facial Expression Recognition, Nearest Neighbor, Support Vector Machine
1- Introduction:

The local binary pattern has been one of the most popular and successful methods of extracting features in recent years. The advantages are, namely, easy to implement, the proper feature extraction, good separation, resistance to the uniform gray changes. It also summarizes the local structures of the images. Since the base LBP describes the image in a specific resolution, it is not appropriate to describe the overall texture of the image. The same weakness causes the features of the image to be similar to those of the other, and to weaken the rate of success in the grouping process.

Therefore, to improve accuracy, an optimized LBP is required. In this research, we first investigate some of the most important types of LBP operators and express each weaknesses and strengths.

In this study, we examine some of the most important types of LBP operators and express their weaknesses and strengths. Then, we examine a variety of local binary operators based on their operation and their impact on noise and rotation.

2- Local Binary Pattern (LBP)

The local binary pattern (LBP) is one of the easiest and most attractive approaches to texture described by Ojala et al. [1]. LBP has been one of the most popular and successful features extraction methods in recent years, due to its ease in the implementation and extraction of appropriate features with high classification accuracy, which has been One of the most popular and successful methods of extracting features in recent years, which is used in many machine vision problems.

The LBP compares each pixel with its neighbors, which lies on a circle around it (clockwise, starting from the top left), which is represented by a binary pattern [2]. This method uses statistical intensity of an image or local structure. So the process is like Figure 1.

![Figure 1: Basic LBP Operator](image-url)
Some of the most important features of the local binary pattern (LBP) are:

(1) is a powerful separator, (2) less computational complexity than other tissue methods such as GLCM, (3) is less sensitive to brightness variations than similar descriptors, (4) it can be easily walked (5) is resistant to uniform gray changes, (6) because it uses the statistical and structural characteristics of the texture, is a powerful tool for extracting the proper properties of the texture, and (7) is very flexible, so that it can be easily applied to various image processing issues and also with other descriptors [3]

The LBP operator was generalized to extract features from neighboring blocks with radius and number of pixels. Figure 2 shows an example of a developed LBP operator (ELBP), in which the symbol (P, R) represents the block of P of the sampling point in a circle to the radius R. [4]

![Figure 2: The ELBP Operator with Symmetric Circular Neighbors](image)

The corresponding LBP code for the pixel \((x_c, y_c)\) can be expressed as Equation 1 [4]

\[
LBP_{p,R}(x_c, y_c) = \sum_{p=0}^{P-1} S(i_p - i_c)2^p
\]  

(1)

Where \(i_c\) and \(i_p\) are values of the gray level of the pixel center and \(p\) pixels on a circle with radius \(R\), and the function \(S(x)\) can be calculated by equation 2.

\[
S(x) = \begin{cases} 
1, & \text{if } x \geq 0 \\
0, & \text{if } x < 0
\end{cases}
\]  

(2)

\[
H(k) = \sum_{i=1}^{I} \sum_{j=1}^{J} f(LBP_{p,R}(i,j), k), \text{ where } k \in [0,K], f(x,y) = \begin{cases} 
1, & x = y \\
0, & \text{otherwise}
\end{cases}
\]  

(3)

The \(LBP_{p,R}\) operator with neighboring pixel \(p\), produces different \(2P\) output values. Now, if the image rotates, the pixels surrounding each neighborhood will move equally along the circle's environment, and because...
the position of each pixel in the binary code is important, it will result in a different LBP code calculation. In order to remove the rotation effect, LBP was proposed to be resistant to rotation (Equation 4) [5]

$$LBP_{P,R}^i = \min\{\text{ROR}(LBP_{P,R} \cdot i) \mid i = 0,1,2, \ldots, P-1\}$$  \hspace{1cm} (4)

Where $\text{ROR}(x,i)$ performs a right-shift shift on the bit $x$, $i$ times. This work is done on all neighboring $P$ bits. The operator $LBP_{P,R}^i$ calculates the statistics of the unique rotation resistant patterns that correspond to the specific features in the image. Therefore, patterns can be used to extract features. However, it can be seen that such a rotation-resistant LBP operator does not necessarily provide discriminatory information [6]

3- Review On Prior Studies On LBP

Despite the impressive LBP success in machine visibility, the LBP operator has its disadvantages and limitations: (1) generates long histograms (2) histograms are sensitive to image rotation, (3) has a small spatial support , (4) base LBP operator with a small $3 \times 3$ neighborhood cannot extract large-scale features, (5) sensitive to noise, since the operator's threshold is exactly in the central pixel value [6] and searches for predefined textures so it may not contain textures in certain cases.

Recent LBP-based approaches have made significant changes to overcome these disadvantages. These changes focus on the various aspects of the LBP Operator, most notably: (1) improving the separation function, (2) improving its power, (3) selecting its neighborhood, (4) expanding to 3D data (5) Reducing the size of the histogram; and 6) Combining with other descriptors.

1. Improving The Resolution

H. Jin et al. [7] proposed an improved LBP (ILBP) that compares the intensity of all pixels (including the central pixel) with the average intensity of the block pixels (Fig. 3). Assuming that the neighborhood block is $3 \times 3$, ILBP generates a $511$ $(2^9 - 1)$ pattern. While $LBP$ $((8.1))$, it only leads to the production of $256$ $(2^8 - 1)$ patterns. They also provided ILBP with neighboring blocks of varying size. [8] G. Bai et al. [9] proposed an average LBP that is similar to ILBP but does not consider the central pixel.

![Figure 3: An Example of the ILBP Operator [14]](image-url)
H. Yang and his colleagues [10] proposed the Hammers LBP to re-classify non-uniform Hamming-based patterns instead of assembling them into a single rod, which was carried out at \([\text{LBP}}]^{U2}\). In LBP Hamming, these non-uniform patterns are incorporated into uniform patterns by minimizing the hamming distance between them. For example, the non-uniform pattern \(10001110_2\) becomes its uniform type \(10001111_2\), because their hamming distance is one.

When several uniform patterns are with the same hamming distance with one non-uniform pattern, one will be selected with a minimum Euclidean distance. They also developed the Extended LBP (ELBP) [11], which, in addition to binary comparison between pixel values, also expanded their gray-matter differences (GDs) by using additional binary digits of the code. In fact, the ELBP feature consists of several LBP codes in multiple and multiple layers, encoding the difference between gray values between the center pixels and adjacent pixels. As shown in Figure 4, the first layer of ELBP is basically the base LBP code that encodes the difference mark of pixel values.

The layers below ELBP encode the absolute value of the difference in pixel values. In Figure 4, the first layer of the LBP code is the original, which results in a binary number \(11010011_2\) or a decimal number of 211. The absolute difference between the values of pixels 1, 5, 3, 2, 1, 2, 3, and 0 is initially converted to binary form: \(001_2\), \(011_2\), \(001_2\), \(010_2\), and so on. By weighting binary bits similar to the LBP method, the ELBP code is generated corresponding to that layer, for example, L2 is composed of \(01000000_2\) and its decimal value is 64. L3 is composed of \(00110110_2\) and its decimal value is 54. L4 is composed of \(00110110_2\) and its decimal value is 234. Although this method increases resolution, it also greatly increases the dimensions of the features.

Z. Guo et al. [9] presented the complete LBP (CLBP) by combining base LBP with local intensity and central gray level criteria similar to that of ELBP, except that the CLBP contains both symbols and differences in the values of pixels. Be Unlike the binary coding strategy used by ELBP, the CLBP re-
compares the absolute magnitude of the difference between pixel values and the central pixel for generating LBP code.

### 3.2 Power Improvement (Strength)

X. Tan and colleagues introduced [12] local triangular patterns (LTPs), where the original binary code is replaced by a three-digit code (1.0.1) using a threshold. In addition to improving the LBP’s power against noise, this new operator is also robust, but the t-bar setting must be done by the user, which is a relatively difficult task. In LTP, the function $S(x)$ is replaced by relation 5 in relation 1.

\[
S(t_n, t_{ext}) = \begin{cases} 
1, & t_n \geq t_e + t \\
0, & |t_n - t_e| < t \\
-1, & t_n \leq t_e - t 
\end{cases} 
\]  

(5)

The triangular pattern is divided into two binary patterns with positive and negative components. The histogram of these components, which is calculated on an area, joined together. (Figure 5)

Figure 5: Local triangular pattern operator (LTP) [12]

T. Ahonen [13] proposed soft-soft LBP (SLBP), which uses two fuzzy membership functions (relations 6 and 7) instead of equation (2).

\[
S_{1d}(x) = \begin{cases} 
0, & x < -d \\
0.5 + 0.5 \frac{x}{d}, & -d \leq x \leq d \\
1, & x > d 
\end{cases} 
\]  

(6)
The parameter \( d \) controls the fuzzy value executed by the fuzzy function. When the neighborhood block is composed of \( P \) pixels, the histogram will have numbered rods \( 0, 1, \ldots, 2P-1 \). The positioning of the pixel form in the histogram is \( x_c \in [0,0.5] \) in the rod calculated from equation 8.

\[
S_{d}(x) = 1 - S_{1-d}(x)
\]

(7)

With SLBP, a pixel may be placed in two different bars. This method has a high computational complexity and is not resistant to uniform gray variations. Also, like LTP, the appropriate value of \( d \) must be set.

3- Neighborhood Selection

Choosing the right neighbor can have a significant impact on the final performance of LBP. Neighborhood selection includes the number, size, shape, and distribution of neighborhoods. The extraction of the base LBP properties occurs in a square or spherical neighborhood, which S. Liao et al. [14] describe as the main reason for defining the neighborhood in such a way, the invariance to rotation to describe the texture. In some applications, such as facial recognition, a rotating property is not required, but anisotropic data may be important. They proposed a drawn or elliptical LBP (ELBP) that used an elliptic neighborhood definition. ELBP, along with a local gradient (Contrast), provides improved results in face detection experiments. L. Nanni [15] examined various types of LBP using different neighborhoods. In their experiments, they have the best performance associated with elliptical neighborhoods. Figure 6 shows two examples of drawn LBP, in which \( A \) and \( B \) respectively represent the longitudinal axis and the transverse axis, and \( m \) is the number of adjacent pixels.

Figure 6: Two examples of the LBP operator drawn [16]

L. Zhang and his colleagues proposed [16] a multi-part LBP (MB-LBP), which has been particularly popular in analyzing image-captured images. Instead of comparing pixels, the MB-LBP compares the average intensities of adjacent regions in order to extract features with a large structure in the image. In
fact, the base LBP can be considered as a specific case of MB-LBP. The area can be a rectangle or a square. For example, if the block size is $3 \times 3$ pixels, the effective area of the MB-LBP operator is $9 \times 9$ pixels. Figure 7 shows an example of MB-LBP, in which each area consists of six pixels.’

![Figure 7: Example of the MB-LBP Operator [17]](image)

L. Wolf et al. [18] considered the use of bit strings to encode the similarities between classes with pixels, and presented three types of LBPs of three blocks (TPLBP) and LBP of four blocks (FPLBP) that could able to obtain complete information for the description. For each pixel in TPLBP, a block is dimensioned $w \times w$ and pixel-centered, and $S$ is an additional block that is uniformly distributed in a loop with radius $r$ around it. Then, the values of each pair of blocks on a circle were compared at a given distance with those in the central block. The value of a unit bit is set according to the maximum similarity of the two blocks with the central block. The generated code per pixel will have $S$ bit. In FPLBP, two pixel-centric loops are used. This is while only one loop in TPLBP is considered (Fig. 8 and 9). However, A. Hadid et al. [19], by examining the 13 common species of LBP types, were the worst results obtained by TPLBP and FPLBP, and were used to process much more time than other types.

![Figure 8: a) TPLBP code with $\alpha = 2$ and $S = 8$. B) The TPLBP code calculated with $S = 8$, $w = 3$ and $\alpha = 2$. C) the result of application. [19]](image)
Y. Kaya et al. [20], LBP-based neighbors (nLBPd) and LBP-directed (dLBPα). nLBPd based on the relationship between neighbors around the central pixel by the distance parameter d and dLBPα, based on the comparison of the central pixel with its neighbors at the angle α (dLBPα) (Fig. 10 and 11).

Figure 10: Relationships Between Neighbors in nLBPd Based on Distance D [20]

Figure 11: How to scroll pixels based on different angles α in dLBPα [20]

The nLBPd code generated by $d = 1$ and $d = 2$ is expressed as $9$ and $10$

$$P_c = \{S(P_0 > P_1), S(P_1 > P_2), S(P_2 > P_3), S(P_3 > P_4), S(P_4 > P_5), S(P_5 > P_6), S(P_6 > P_7), S(P_7 > P_0)\}$$ \text{ for } d = 1

$$P_c = \{S(P_0 > P_2), S(P_1 > P_3), S(P_2 > P_4), S(P_3 > P_5), S(P_4 > P_6), S(P_5 > P_7), S(P_6 > P_0), S(P_7 > P_1)\}$$ \text{ for } d = 2

Where $S(P_i > P_j)$ is calculated from equation 11.
The parameter d greatly affects the amount of Pc. With different ds, different histograms are obtained. Only by testing can the optimal distance parameter be obtained.

dLBPα produces four different histograms with different α (Relationship 12). The function $S(P_i > P_j)$ is computed from equation 11. Only by testing one can get the desired angle for dLBPα.

$$S(P_i > P_j) = \begin{cases} 1, & \text{if } P_i > P_j \\ 0, & \text{if } P_i \leq P_j \end{cases}$$

The parameter d greatly affects the amount of Pc. With different ds, different histograms are obtained. Only by testing can the optimal distance parameter be obtained.

In many applications, the analysis and interpretation of 3D textures is required. For example, in medicine, a lot of imaging information, such as ultrasound and microscopic imagery, or in computer animation that uses 3D images. Recently, many researchers have tried to generalize LBP from a 2D page to a 3-D space [21]. This issue is simply not feasible and generally has two problems: first, sampling in spherical space is difficult, and secondly, the extraction of a feature A robust property of rotation in a three dimensional space is difficult task.

G. Zhao et al. [22] provided a function called volumetric LBP (VLBP or 3D-LBP) that could be used to extract dynamic tissue information by combining motion and appearance information. The VLBP features are insensitive to movement and rotation (relative to the rotation around the Z axis) and are resistant to uniform variations in gray levels. In comparison with the VLBP considers the time domain, where L represents the length of the time interval. A small, three-dimensional local neighborhood around a central pixel yields a binary unit for which weights are given as a spiral line (Fig. 12). In order to easily implement VLBP, only features that are on the three orthogonal pages of X-Y, X-T, and Y-T are considered, and the textures are extracted from the pages by the LBP histogram. This version is easier to name than VLBP, LBP-TOP, and conventional circular sampling is replaced by an ellipse. Therefore, different radial parameters can be set in the field of space and time.
Later, D. Huang et al. [23] developed 3D-LBP into LBP (ELBP), which consisted of several LBP codes in multiple layers designed to differentiate gray values between the central pixel and its neighbors. G. Sandbach et al. [24] proposed a normal local binary pattern (LNBP), which used the angle between the normal at two points instead of the depth to obtain the local binary code. They further developed the idea of using surface normalization [25]. Similarly, H. Li and colleagues [26] extracted surface normal in three-dimensional and multi-scale scales to detect facial expressions. The value of the normal components along the three coordinate axes is interpreted as the depth of the interpretation, and the calculated LBP in these depth maps describes the value of the normal component.

J. Fehr et al. [26] used spherical harmonics to generate orthogonal fields, and then calculate LBP characteristics in the frequency domain. This method has a high computational complexity because of its robustness to rotation with complex techniques such as spherical correlation in the frequency domain.

L. Paulhac et al. [27] proposed another solution for LBP in 3D space. They used a number of circles to show the volume by adding the $S$ parameter. The descriptor $\text{VLBP}_{(P, R)}$ is displayed with the new parameter $S$ as $LBP_{(P, R)} \oplus LBP_{(S, P, R)}$ (Fig. 13).
They also defined the principle of uniform patterns, such as two-dimensional space. The problem with their method was that they presented the same description of different tissues. N. Werghi et al. [29] proposed a local mesh binary pattern (mesh-LBP) to compute local pseudo-binary patterns in triangular mesh structures and showed how to generalize two-dimensional LBP types to LBP meshes (Table 1). In fact, the mesh structure was introduced as a compact and flexible triangle for encoding 3D information.

Table 1: LBP Operator Types Grouped in 4 Categories and Patterns for Two-Dimensional LBP and Mesh LBP [29]

<table>
<thead>
<tr>
<th>Computational framework</th>
<th>2D LBP Edition</th>
<th>Two-dimensional LBP pattern</th>
<th>Mesh-LBP pattern</th>
</tr>
</thead>
<tbody>
<tr>
<td>Searching the route</td>
<td>Local linear binary pattern (LLBP)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Spiral</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Partition / Primary core</td>
<td>LBP drawn or elliptical (ELBP)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Based on comparison</td>
<td>Symmetric center</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Several operators have been introduced in the field of extracting the properties of images. One of the best local texture descriptors that has been widely used in various fields in recent years is LBP. Separation, resistant to uniform gray area changes, is an easy implementation of the benefits of LBP. Also, as a non-parametric method, the LBP summarizes the local structures of the images efficiently by comparing each pixel with its adjacent pixels. Since the base LBP method describes the image at a specific resolution, it is not appropriate to describe the overall texture of the image. In fact, the histogram distribution of neighborhoods is independent of the intensity of the central pixel brightness. The same weakness causes the features of the image to be similar to those of the other, and weakens the success rate at the classification stage.

Therefore, in order to increase the accuracy of the classification stage, an optimized LBP is required. Some recent changes, such as ELBP, VLBP, and LBP-based Gabor-wavelets, have even increased the length of the feature vector. It is observed that the LBP-based feature vector is optimized for the non-optimal version containing less waste and has a high resolution and is also more appropriate in real-time systems.

Acknowledgement

This article is derived by MA thesis at Islamic Azad University, Marvdasht Branch. The cooperation of this scientific branch is appreciated respectfully.

Reference
