Fusion of motif co-occurrence matrix and local binary pattern based on intuitionistic fuzzy set for texture classification

C.-P. Yen¹

Department of Information Management, Central Police University

¹ ORCID: 0000-0002-1189-4922, peter@mail.cpu.edu.tw

Abstract

Texture classification plays an important role in computer vision and has a wide variety of applications. Based on intuitionistic fuzzy set (IFS) theory, this paper proposes a novel feature descriptor for texture classification by the fusion of motif co-occurrence matrix (MCM) and local binary pattern (LBP), namely IFS-MCMLBP. In this way, IFS is used to model vagueness or uncertainty, and the MCM method for extracting microtexture information, whilst the LBP method plays the role of a global feature. Intensive experiments conducted on many texture benchmarks such as CUReT, Outex, Brodatz and VisTex. The results show that the IFS-MCMLBP method can be remarkably superior to existing texture classification methods such as, LBP, GLCM, LTP, LDiP, LDeP and LTrP.

Keywords: Intuitionistic fuzzy set (IFS), texture classification, motif co-occurrence matrix (MCM), local binary pattern (LBP).

1. Introduction

Texture is a very important attribute in the field of computer vision and image processing. Numerous methods of texture analysis have been developed over the years. However, analyzing existing textures in the real world is a major challenge. A good real world texture classification method should have a highly discriminative and robust to variances such as rotation and scales. Among the many texture classification methods, local binary patterns (LBP) [1] and the gray-level co-occurrence matrix (GLCM) [2] are the most popular approaches. According to the research results of [3] [4], LBP is superior to GLCM in texture analysis performance. Therefore, there are many LBP-based methods have been proposed to enhance discrimination and improve robustness. Hafiane et al. [5] proposed median binary pattern (MBP) that uses the sign of the intensity difference compared to the local median within a neighborhood. Tan et al. [6] also proposed local ternary pattern (LTP), which extends original LBP to 3-valued codes. Liao et al. [7] employed the dominant local binary patterns (DLBP) to extract dominant patterns from textures. Jabid et al. [8] proposed local directional pattern (LDiP), which computes edge response values by using Kirsch masks in all eight directions at each pixel position and then generates a code, instead of the intensity. In [9], completed local binary pattern (CLBP) was proposed by Guo et al. to combine multiple LBP type features for texture classification by joint histogram. Zhang et al. [10] proposed local derivative pattern (LDeP), which extracts high-order local information by encoding various distinctive spatial relationships contained in a given local region. The local tetra pattern (LTrP) utilized first-order derivative calculations in the vertical and horizontal directions to encode the relationship between reference pixels and their neighbors by Muralla et al. [11]. Dubey et al. [12] developed a local bit plane decoding pattern (LBDP), which is generated by finding a binary pattern using the difference between the intensity value of the center pixel and the local bit-plane transformed value of each image pixel. Alaei et al. [13] showed the fast local binary pattern (F-LBP) method, which is derived from the vertical & horizontal, and diagonal & off-diagonal pixels in a 3x3 patch size, separately. Banerjee et al. [14] proposed a texture descriptor called
Local neighborhood intensity pattern (LNIP), which considers the relative intensity difference between a particular pixel and a central pixel by considering its neighbors and generates a sign and a magnitude pattern. In the above mentioned, six indicator texture descriptors — namely, LBP, GLCM, LTP, LDiP, LDeP and LTrP — will be compared with the proposed method IFS-MCMLBP.

The rest of this paper is organized as follows: in Section 2 briefly introduces LBP, MCM and IFS. Section 3 presents the proposed IFS-MCMLBP descriptor including IFS texture generation, motif, LBP and feature fusion. Section 4 describes similarity measure, classifier and cross-validation. Section 5 gives the experimental analysis on various datasets along with the other comparative methods. Finally, Section 6 concludes the whole paper.

2. Related work

2.1. Local Binary Patterns (LBP)

Ojala and colleagues [1] first proposed the local binary pattern (LBP) texture operator, which is just the first-order circular derivative of patterns and which works with the 3×3 neighborhood. In Fig. 1(a), each pixel is compared with its eight neighbors by taking the difference of the center pixel value; resulting strictly negative values are encoded with 0, and other values are encoded with 1. LBP is defined as

\[
LBP_{P, R} = \sum_{i=0}^{P-1} s(g_i - g_c) \cdot 2^i
\]

where \(g_i\) and \(g_c\) denote the gray values of the neighbor and central pixel, respectively; \(i\) is the index of the neighbor; \(P\) is the number of neighbors; and \(R\) is the radius of the circular neighborhood. Fig. 1(b) shows the neighbors for \(P = 8\) when \(R = 1\), \(P = 16\) when \(R = 2\), and \(P = 24\) when \(R = 3\). Fig. 1(c) shows an example of the LBP operator. For each given pixel, a binary number is obtained by concatenating all binary values in a clockwise direction starting with the binary value of the pixel’s top-left neighbor. The corresponding decimal value of the generated binary number is then used for labeling the given pixel.
2.2. Motif co-occurrence matrix (MCM)

Jhanwar and colleagues [15] have proposed the motif co-occurrence matrix (MCM) for content-based image retrieval. The MCM is derived from the motif-transformed image, which is calculated by dividing the whole image into non-overlapping 2×2 pixel patterns. Each grid is replaced by a scan motif, as shown in Fig. 2, which minimizes the local gradient while traversing the 2×2 grid that forms the motif-transformed image. These motifs are defined over a 2×2 grid, each depicting a distinct sequence of pixels starting from the top left corner.

![Fig. 2. Illustration of scan motifs to traverse a 2×2 grid.](image)

We can derive a transformed image from an original image, as shown in Fig. 3(a). An 8×8 image, as in Fig. 3(b), shows the corresponding 4×4 transformed image.

![Fig. 3. (a) An 8×8 image (b) Motif transformed image from (a)](image)

3. Intuitionistic fuzzy set (IFS)

IFS theory, which is an extension of fuzzy set (FS), enhances images and helps to recover important structures that are not properly visible [16]. Atanassov [17] pioneered construction of IFS, which is defined by three feature functions as the degree of membership, degree of non-membership, and degree of hesitation or uncertainty.

Definition: An IFS $A$ in universe $X$ is an expression given by

$$A = \{x, \mu_A(x), \nu_A(x) | x \in X \}$$

(3)

where $\mu_A(x) : X \rightarrow [0, 1]$, $\nu_A(x) : X \rightarrow [0, 1]$ are the membership and non-membership degree of an element $x$ to the set $A$ with the condition:

$$0 \leq \mu_A(x) + \nu_A(x) \leq 1$$

for each $x \in X$  

(4)
For each IFS in $X$, we call $\pi_A(x) = 1 - \mu_A(x) - \nu_A(x)$ the degree of hesitancy of $x$ to $A$, $0 \leq \pi_A(x) \leq 1$ for each $x \in X$. The illustration of these degrees is shown in Fig. 4.

Fig. 4. Relationships between membership, non-membership, and hesitation degrees.

IFS is helpful in modeling vagueness or uncertainty, and important applications of IFS have been developed in many diverse areas, including medical diagnosis [18], pattern recognition [19], image processing [20], and decision making [21]. The proposed IFS-MCMLBP descriptor

Based on intuitionistic fuzzy set (IFS) theory, a novel descriptor IFS-MCMLBP is proposed by the fusion of Motif Co-occurrence Matrix (MCM) and Local Binary Pattern (LBP) as shown in Fig. 5. The IFS texture is used to model vagueness or uncertainty, and the MCM method for extracting microtexture information, whilst the LBP method plays the role of a global feature. And detailed information about it will be described in the following subsection.

3.1. IFS texture generation

In the beginning, a RGB (red, green, and blue) image is converted into an HSV (hue, saturation, and value) image. Then, intuitionistic fuzzy image processing of color texture is applied to each channel of HSV to generate the image’s membership, non-membership and hesitancy components.

Suppose channel image $A$ of size $M \times N$ pixels has $L$ gray levels ranging between 0 and $L - 1$. When applying IFS for image processing [22], an image can be considered as an array of fuzzy singletons. An intuitionistic fuzzy image is written as

$$A_{IFS} = \{x, \mu_A(x), \nu_A(x), \pi_A(x)\}, \quad 0 \leq x \leq L - 1$$

(5)
where \( x \) is the pixel value at \((i, j)\) point, \(i=0, 1, \ldots, N-1, \ j=0, 1, \ldots, M-1\).

With the condition

\[
0 \leq \mu_A(x) + \nu_A(x) \leq 1
\]

then

\[
\pi_A(x) = 1 - \mu_A(x) - \nu_A(x)
\]

where

\[
\mu_A(x) \rightarrow [0, 1]: \text{membership degree}
\]

\[
\nu_A(x) \rightarrow [0, 1]: \text{non-membership degree}
\]

\[
\pi_A(x) \rightarrow [0, 1]: \text{hesitancy degree}
\]

Vlachos [23] represents the membership degree of image by

\[
\mu_A(x) = \frac{(x - x_{\text{max}})}{(x_{\text{max}} - x_{\text{min}})}
\]

where \( x_{\text{max}} \) and \( x_{\text{min}} \) are the maximum and the minimum gray levels of the image. Sugeno’s intuitionistic fuzzy generator (IFG) [24] constructs non-membership degree \( \nu_A(x) \) as

\[
\nu_A(x) = \frac{1 - \mu_A(x)}{1 + \lambda \cdot \mu_A(x)}
\]

where parameter \( \lambda > 0 \), and hesitancy degree

\[
\pi_A(x) = 1 - \mu_A(x) - \frac{1 - \mu_A(x)}{1 + \lambda \cdot \mu_A(x)}
\]

By varying the \( \lambda > 0 \) parameter, different intuitionistic fuzzy set can be obtained. As \( \lambda \) is not a fixed value for all images, the optimum value of \( \lambda \) is obtained by maximizing fuzzy entropy [25]. Given an examples, the texture image C57 of CURet texture database is used for representing IFS texture images when \( \lambda \) is 3.6. Fig. 6(a) depicts C57 image along with their corresponding membership \( \mu \), non-membership \( \nu \), and hesitancy \( \pi \) images.

### 3.2. Motif feature

According to Section 3.1, a color texture image is processed by IFS texture generation, which will generate three \( \mu \), three \( \nu \) and three \( \pi \) component images, respectively, for a total of nine. Then the nine IFS component images are individually converted into MCM. Next, we compute the histogram of each MCM. The histogram has 6 bins, representing the distribution of motif 1 to motif 6. Since motif 0 only represents a homogeneous texture, so motif 0 is not considered. Furthermore, the histogram of nine MCMs is concatenated in series to form a 54-dimensional feature vector of motif as shown Fig. 6(b). In brief, the motifs’ histograms can be easily implemented and computed from that image, and the histogram’s shape provides many clues to the image’s microtexture features.

### 3.3. LBP feature

As shown in Fig. 6(c), a 256-bin histogram of LBP is extracted out of the IFS texture images including three \( \mu \), three \( \nu \), and three \( \pi \) images one by one. Then, all the histograms are concatenated into a one 2304-dimensional feature vector for representing the global information of the texture image. This also makes this global feature more robustness for changes in rotation and scale.

### 3.4. Feature fusion

Since different features may be represented by different importance, it is possible to give an appropriate weight by experimental approaches when combining them. However, some studies have also pointed out that traditional feature fusion methods that simply concatenate several features may be better or more robust than using a single feature.
The aforementioned motif feature and LBP feature represent local texture information and global texture information, respectively. This paper uses traditional feature fusion methods to simply concatenate together as shown Fig. 6(d).

![Diagram](image)

**Fig. 6.** (a) Texture image C57 from the CURet texture database and showing IFS texture images including membership \( \mu \), non-membership \( \nu \), and hesitancy images. (b) Concatenate three \( \mu \), three \( \nu \), and three \( \pi \) motif histograms to form a 54-dimensional feature vector of motif. (c) Like (b), forming a 2304-dimensional feature vector of LBP. (d) Fusion local feature motif and global feature LBP.

### 4. Classification

#### 4.1. Similarity measure

Many measures have been proposed for discriminating the dissimilarity between two histograms, such as Manhattan, Euclidean, \( d \), Canberra, and \( \chi^2 \) (Chi-square) distance. In this paper, the \( \chi^2 \) distance function is chosen in the experiments due to its excellent performance in terms of good recognition rates which is calculated as

\[
D_{\chi^2}(Q, T) = \sum_{i=1}^{n} \frac{(q_i - t_i)^2}{q_i + t_i} 
\]

where \( Q=\{q_i\} \) and \( T=\{t_i\} \ (i=1...n) \) are two histograms with \( n \) bins, and the histogram is used as feature representation; \( q_i \) is \( i \)th bin (i.e. \( i \)th feature) of testing image \( Q \), \( t_i \) is \( i \)th bin of training image \( T \) in database.

#### 4.2. Classifier and cross-validation

The k-nearest neighbor (k-NN) classification [26] is one of the simplest but widely using pattern recognition algorithm. An object is classified by the distance from its neighbors, with the object being assigned to the class most common amongst its k nearest neighbors. This paper use 1-NN approach as a classifier in this work, i.e. the texture image is classified to the class of its nearest neighbor.
The sample data was split into training and test sets using the leave-one-out cross-validation (LOOCV) [27], for N samples, a total of N trials are conducted. In each trial a sample is taken out from the data set and kept for testing and the others are used for training. This procedure was repeated for all samples and the accuracy rate obtained as the percentage of classified samples out of the total number of samples. This methodology is superior to random partitioning of data to generate training and test sets as the resultant performance of the system may not reflect its true ability for texture classification.

5. Experiments and results

In this study, the proposed method compares performance with other well-known or state-of-the-art approaches, namely LBP$_{8,1}$, LBP$_{16,2}$, GLCM, LTP, LDIP, LDeP and LTrP. Experiments performed on four benchmark of texture images, the Colored Brodatz Texture (CBT) database, the Columbia-Utrecht Reflection and Texture (CUReT) database, the Outex database, the Vision Texture (VisTex) database in MIT university. All the experiments are run in Matlab environment.

5.1. Experiments on texture database

The CBT database consists of 112 periodic or non-periodic 640×640 texture images, parts of which are shown in Fig. 7(a). Each texture image is divided into 25 non-overlapping sub-images for experiments, resulting in 112 categories, each category contains 25 images, and 11,200 total samples for size of 128×128 pixels.

The CUReT texture image database contains 61 texture of real-world surfaces as shown in Fig. 7(b), 205 images per class, acquired at different viewpoints, illuminations, and orientations. There are 103 images shot from a viewing angle of 22.5 and 45 degrees, in total 6,283 (=61×103) images are selected. The size of each texture is 640×480 which is divided into nine 128×128 non-overlapping sub-images, and then a database of 56,547 samples is obtained.

This paper use a commonly test suite Outex_TC_00013 (TC13) of Outex database as shown in Fig. 7(c), containing 68 classes with 20 texture images per class, in total 1,360 (=68×20) images. Each texture class is collected under “inca” illuminations with resolution of images at 128×128.

The VisTex database includes 40 classes texture images, parts of which are shown in Fig. 7(d). The size of each texture is 512×512 which is divided into sixteen 128×128 non-overlapping sub-images, and then a database of 640 samples is obtained.

![Sample textures from the (a) CBT, (b) CUReT, (c) Outex, and (d) VisTex database.](image)
5.2. Comparative analysis of the accuracy

This paper employed the nearest neighbor (1-NN) classification and χ² (chi-square) distance, while the accuracy estimate uses leave-one-out cross-validation (LOOCV). Experiments were performed on aforementioned four texture datasets and compared to seven prior art LBP8,1, LBP16,2, GLCM, LTP, LDiP, LDeP and LTrP. Fig. 8 illustrates the accuracy rates of all these descriptors and following are noted.

1) Clearly, it is evident that the proposed IFS-MCMLBP descriptor provide better classification performance (average accuracy of 96.08%) compared to other state-of-the-art descriptors. However, for the VisTex database, the proposed descriptor shows a classification accuracy of 98.44%, which is the same as the LDeP descriptor but less than the LTrP descriptor.

2) The accuracy rate of all other descriptors are less than 85% on the Outex(TC13) texture. However, the IFS-MCMLBP has achieved a classification rate around 87%.

3) GLCM performance is very low on four texture databases when compared to other descriptors.

5.3. Comparative analysis of robust rotation

For the analysis of rotation robustness, all 112 texture images are rotated by 19 different angles (i.e., 0°, 10°, 20°, ..., 180°) from the CBT database. Here we use the nearest neighbor interpolation method for rotation of image. Each rotated image is divided into four non-overlapping 320×320 sub-images. Four 128×128 sub-images are then cropped from the center of the 320×320 sub-images. Hence, we obtain a rotation test set with a total of 8,512 (=112×19×4) sub-images. Fig. 9 shows an example of different rotation angles from 0 to 180 degrees using 10-degree steps.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>LBP 8,1</th>
<th>LBP 16,2</th>
<th>GLCM</th>
<th>LTP</th>
<th>LDiP</th>
<th>LDeP</th>
<th>LTrP</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outex(TC13)</td>
<td>83.89</td>
<td>84.34</td>
<td>48.82</td>
<td>80.29</td>
<td>74.78</td>
<td>82.13</td>
<td>82.87</td>
<td>87.79</td>
</tr>
<tr>
<td>VisTex</td>
<td>98.28</td>
<td>98.13</td>
<td>66.72</td>
<td>94.38</td>
<td>95.47</td>
<td>98.44</td>
<td>99.22</td>
<td>98.44</td>
</tr>
<tr>
<td>CURET</td>
<td>95.63</td>
<td>93.08</td>
<td>52.82</td>
<td>97.09</td>
<td>93.62</td>
<td>92.53</td>
<td>94.72</td>
<td>98.18</td>
</tr>
<tr>
<td>CBT</td>
<td>96.86</td>
<td>96.25</td>
<td>72.21</td>
<td>97.71</td>
<td>96.25</td>
<td>96.68</td>
<td>97.18</td>
<td>99.89</td>
</tr>
<tr>
<td>Average</td>
<td>93.67</td>
<td>92.95</td>
<td>60.14</td>
<td>92.37</td>
<td>90.03</td>
<td>92.45</td>
<td>93.50</td>
<td>96.08</td>
</tr>
</tbody>
</table>

Fig. 8. Classification accuracy (%) on CBT, CURET, Outex(TC13), and VisTex databases using different descriptors.
The results of classification accuracy with respect to rotation factors for seven tested descriptors (the GLCM classification accuracy is too low, not considered) are shown in Fig. 10 and Table 1 for clear comparison on CBT database. The proposed descriptor IFS-MCMLBP has the highly stable accuracy (average 95.68%) than the other descriptors under at each rotation angle, that confirms the robustness of the proposed descriptor. Thus, the other descriptors only obtain high accuracy at the rotation angles 0°, 90° and 180°, where the image distortion is minimal.

Table 1. Average classification rates (%) of different descriptor at each rotation angle on CBT database

<table>
<thead>
<tr>
<th>Descriptor</th>
<th>LBP 8,1</th>
<th>LBP 16,2</th>
<th>LTP</th>
<th>LDiP</th>
<th>LDeP</th>
<th>LTrP</th>
<th>IFS-MCMLBP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average</td>
<td>52.96</td>
<td>62.03</td>
<td>78.09</td>
<td>49.93</td>
<td>42.30</td>
<td>44.29</td>
<td>95.68</td>
</tr>
</tbody>
</table>
5.4. Comparative analysis of robust scale

For the analysis of scale robustness, all 112 texture images are scaled with scaling factors of 0.5 to 1.5 with 0.1 intervals (11 scales for each image) from the CBT database. The bicubic interpolation and antialiasing methods are used to scale image. Each scaled image is divided into four non-overlapping 320×320 sub-images. Four 128×128 sub-images are then cropped from the center of the 320×320 sub-images. Hence, a scaled test set was obtained with 4,928 (=112×11×4) sub-images. Fig. 11 shows an example in different scale factors: 0.5, 0.6, 0.8, 1.0, 1.2, 1.3, 1.4 and 1.5.

![Example scaled test set for image D66 from the CBT database; scale factors are 0.5, 0.6, 0.8, 1.0, 1.2, 1.3, 1.4 and 1.5.](image)

The classification accuracy results, with respect to scale factors for seven tested descriptors, are shown in Fig. 12 and Table 2. It is obvious that the proposed descriptor IFS-MCMLBP has the most stable accuracy (average 88.73%) of the seven tested descriptors under multi-scale factors. Only in the texture image with a scale factor of 1.5, the accuracy of the IFS-MCMLBP descriptor is 64.06%, which is slightly lower than 67.63% of the LTP descriptor. Moreover, the enlarged texture image could miss texture structures, and the shrinked texture image still remains original structures. Therefore, the accuracy of the former is worse than the latter.
Fig. 12. Texture classification results for seven tested descriptors on the scaled CBT database.

Table 2. Average classification rates (%) of different descriptor under multi-scale factors on CBT database

<table>
<thead>
<tr>
<th>Descriptor</th>
<th>LBP 8,1</th>
<th>LBP 16,2</th>
<th>LTP</th>
<th>LDiP</th>
<th>LDeP</th>
<th>LTrP</th>
<th>IFS-MCMLBP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average</td>
<td>55.19</td>
<td>66.88</td>
<td>85.41</td>
<td>66.48</td>
<td>58.24</td>
<td>55.93</td>
<td>88.73</td>
</tr>
</tbody>
</table>

6. Conclusions

This paper has presented a new method for texture classification by fusing the MCM and LBP features based on intuitionistic fuzzy set. The LBP describes local textures, while MCM emphasizes global microstructures and uses IFS to recover invisible important structures. As shown in the experimental results, the proposed IFS-MCMLBP method has highly stable accuracy and robustness to rotation and scale, and outperforms the existing methods. For the future possibility, the color information also can be incorporated by combining color histogram features on hue and saturation channels into the proposed global or local feature representation scheme for better performance. In addition, the proposed method can also be extended to other research fields such as content-based image retrieval (CBIR), object recognition, remote sensing and medical imaging.

References