Abstract—The Local Binary Pattern (LBP) operator is a computationally efficient yet powerful feature for analyzing local texture structures. While the LBP operator has been successfully applied to tasks as diverse as texture classification, texture segmentation, face recognition and facial expression recognition, etc., it has been rarely used in the domain of Visual Object Classes (VOC) recognition mainly due to its deficiency of power for dealing with various changes in lighting and viewing conditions in real-world scenes. In this paper, we propose six novel multi-scale color LBP operators in order to increase photometric invariance property and discriminative power of the original LBP operator. The experimental results on the PASCAL VOC 2007 image benchmark show significant accuracy improvement by the proposed operators as compared with both the original LBP and other popular texture descriptors such as Gabor filter.

Keywords—local binary patterns; object recognition; feature extraction; PASCAL VOC challenge

I. INTRODUCTION

The Local Binary Pattern (LBP) operator was firstly introduced as a complementary measure for local image contrast [1]. It can be seen as a unified approach to statistical and structural texture analysis. Fig. 1 gives an example. For one pixel in a gray image, its eight neighboring pixels are considered — their values are operated by the value of the central pixel as threshold. The LBP code is computed by multiplying the thresholded results with weights given by powers of two, and summing up the results. Then, for each pixel in the image, the same process is followed to get its LBP code, and the final LBP operator is obtained by counting the histogram based on these codes.

```
<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>3</td>
<td>36</td>
</tr>
<tr>
<td>8</td>
<td>12</td>
<td>25</td>
</tr>
<tr>
<td>1</td>
<td>15</td>
<td>10</td>
</tr>
</tbody>
</table>
```

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

```
LBP = 4*32 + 36 = 44
```

Figure 1. Calculation of the original LBP operator

Because of its descriptive power for analyzing both micro and macro texture structures, and computational simplicity, the LBP operator has been successfully applied for many applications, such as texture classification [2,3,4], texture segmentation [5], face recognition [6], and facial expression recognition [7].

However, the LBP operator has been rarely used in the domain of Visual Object Classes (VOC) recognition. We hold that main reasons are as follows. On one hand, the LBP operator ignores all color information (its calculation is based on gray image), while color plays an important role for distinction between objects, especially in natural scenes. On the other hand, there can be various changes in lighting and viewing conditions in real-world scenes, leading to large variations of objects in surface illumination, scale, etc., which make recognition task more complicated and challenging. According to its definition, the LBP operator is only invariant to monotonic light changes in gray-level, and is deficient in power for dealing with these variations which mostly occur in natural scenes.

Therefore, in order to include color information, as well as to increase photometric invariance property and discriminative power of the original LBP operator, we propose, in this paper, six novel multi-scale color LBP operators which are proven more suitable for the VOC recognition tasks. The performances of the proposed operators are analyzed experimentally using the PASCAL VOC 2007 image benchmark.

II. MODEL ANALYSIS FOR ILLUMINATION CHANGES

Changes in the illumination can be modeled by the diagonal model (1) and the diagonal-offset model (2) [8], where u and c represent the values before and after illumination transformation respectively:

\[
\begin{align*}
R^c &= \begin{pmatrix} a & 0 & 0 \end{pmatrix} R^u \\
G^c &= \begin{pmatrix} 0 & b & 0 \end{pmatrix} G^u \\
B^c &= \begin{pmatrix} 0 & 0 & c \end{pmatrix} B^u 
\end{align*} \tag{1}
\]

\[
\begin{align*}
R^c &= \begin{pmatrix} a & 0 & 0 \end{pmatrix} R^u + \begin{pmatrix} O_1 \\
G^c &= \begin{pmatrix} 0 & b & 0 \end{pmatrix} G^u + \begin{pmatrix} O_2 \\
B^c &= \begin{pmatrix} 0 & 0 & c \end{pmatrix} B^u + \begin{pmatrix} O_3 
\end{align*} \tag{2}
\]

```
Based on these two models, different kinds of illumination change can be expressed as follows [8]:

**light intensity change** Image values change by a constant factor in all channels \((a = b = c)\):

\[
\begin{bmatrix}
R^c \\
G^c \\
B^c
\end{bmatrix} = \begin{bmatrix}
a & 0 & 0 \\
0 & a & 0 \\
0 & 0 & a
\end{bmatrix}
\begin{bmatrix}
R^u \\
G^u \\
B^u
\end{bmatrix} + \begin{bmatrix}
O_1 \\
O_1 \\
O_1
\end{bmatrix}
\]

(3)

**light intensity shift** Image values change by an equal offset in all channels \((a = b = c = 1, O_1 = O_2 = O_3)\):

\[
\begin{bmatrix}
R^c \\
G^c \\
B^c
\end{bmatrix} = \begin{bmatrix}
R^u \\
G^u \\
B^u
\end{bmatrix} + \begin{bmatrix}
O_1 \\
O_1 \\
O_1
\end{bmatrix}
\]

(4)

**light intensity change and shift** Image values change by combining two kinds of change above:

\[
\begin{bmatrix}
R^c \\
G^c \\
B^c
\end{bmatrix} = \begin{bmatrix}
a & 0 & 0 \\
0 & a & 0 \\
0 & 0 & a
\end{bmatrix}
\begin{bmatrix}
R^u \\
G^u \\
B^u
\end{bmatrix} + \begin{bmatrix}
O_1 \\
O_1 \\
O_1
\end{bmatrix}
\]

(5)

**light color change** Image values change in all channels independently \((a \neq b \neq c)\), as (1).

**light color change and shift** Image values change in all channels independently with arbitrary offsets \((a \neq b \neq c\) and \(O_1 \neq O_2 \neq O_3\)), as (2).

### III. COLOR LBP OPERATORS AND PROPERTIES

In order to introduce color information into the original LBP operator, as well as to increase its photometric invariance properties of dealing with different kinds of illumination changes (as described in section II), six novel color LBP operators are proposed as follows:

**RGB-LBP** This operator is obtained by computing LBP over all three channels of the RGB color space independently, and then concatenating the results together. It is invariant to monotonic light intensity change due to the property of the original LBP, and has no additional invariance properties.

**nRGB-LBP** This normalized operator is obtained by computing LBP for both r and g channels of the normalized RGB color space (b channel is redundant because \(r + g + b = 1\)):

\[
\begin{bmatrix}
r \\
g \\
b
\end{bmatrix} = \begin{bmatrix}
R / (R + G + B) \\
G / (R + G + B) \\
B / (R + G + B)
\end{bmatrix}
\]

(6)

Due to the normalization, r and g channels are scale-invariant, which make this operator invariant to light intensity change as (3).

**Transformed color LBP** This operator is obtained by computing LBP over all three channels of the transformed color space \((\mu\) is the mean and \(\sigma\) is the standard deviation of each channel):

\[
\begin{bmatrix}
R' \\
G' \\
B'
\end{bmatrix} = \begin{bmatrix}
(R - \mu_R) / \sigma_R \\
(G - \mu_G) / \sigma_G \\
(B - \mu_B) / \sigma_B
\end{bmatrix}
\]

(7)

Due to the subtraction and the normalization, all three channels are scale-invariant and shift-invariant, which make this operator invariant to light intensity change and shift as (5). Furthermore, because each channel is operated independently, this operator is also invariant to light color change and shift as (2).

**Opponent-LBP** This operator is obtained by computing LBP over all three channels of the opponent color space:

\[
\begin{bmatrix}
O_1 \\
O_2 \\
O_3
\end{bmatrix} = \begin{bmatrix}
(R - G) / \sqrt{2} \\
(R + G - 2B) / \sqrt{6} \\
(R + G + B) / \sqrt{3}
\end{bmatrix}
\]

(8)

Due to the subtraction, \(O_1\) and \(O_2\) channels are invariant to light intensity shift as (4). \(O_3\) channel represents the intensity information, and has no invariance properties.

**nOpponent-LBP** This normalized operator is obtained by computing LBP over two channels of the normalized opponent color space:

\[
\begin{bmatrix}
O_1' \\
O_3 \\
O_2
\end{bmatrix} = \begin{bmatrix}
\sqrt{3}(R - G) \\
\sqrt{2}(R + G + B) \\
(R + G - 2B)
\end{bmatrix}
\]

(9)

Due to the normalization by intensity channel \(O_3\), \(O_1'\) and \(O_2'\) channels are not only shift-invariant, but also scale-invariant, which make this operator invariant to light intensity change and shift as (5).

**Hue-LBP** This operator is obtained by computing LBP for the Hue channel of the HSV color space:

\[
\text{Hue} = \arctan\left(\frac{O_1}{O_2}\right) = \arctan\left(\frac{\sqrt{3}(R - G)}{R + G - 2B}\right)
\]

(10)

Due to the subtraction and the division, Hue channel is scale-invariant and shift-invariant, therefore this operator is invariant to light intensity change and shift as (5).
IV. MULTI-SCALE COLOR LBP OPERATORS

Another big limitation of the original LBP operator is that it only covers a small neighborhood area (8 neighboring pixels), and can only get very limited local information. In order to obtain more local information by covering larger neighborhood area, and therefore to increase discriminative power of the original LBP, multi-scale LBP operator [2] is applied by combining different LBP operators which use a circular neighborhood with different radius and different number of neighboring pixels. Fig. 2 gives an example.

Formally, the LBP code of the pixel at \((x_c, y_c)\) is calculated according to the following equation:

\[
\text{LBP}(x_c, y_c) = \sum_{p=0}^{P-1} S(g_p - g_c) \times 2^p, S(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \tag{11}
\]

where \(g_p\) is the value of the neighboring pixel, \(g_c\) is the value of the central pixel, and \(P\) is the total number of neighboring pixels.

Finally, the proposed multi-scale color LBP operators can be obtained by expanding each color LBP operator described in section III to the multi-scale one, and concatenating their histograms separately. By doing this, the proposed novel operators are not only invariant to different illumination changes, but also scale-invariant to a certain extent.

The same multi-scale configuration is applied for all the proposed operators (as Fig. 2): 8 neighboring pixels with radius 1, 12 neighboring pixels with radius 1.5, and 16 neighboring pixels with radius 2.

V. EXPERIMENTAL EVALUATION

The PASCAL Visual Object Classes (VOC) Challenge 2007 image benchmark [9] is used to evaluate the performances of the proposed operators. It contains nearly 10,000 images of 20 different object categories, such as bike, car, cat, table, person, sofa, train, etc. All the images are taken from real-world scenes, and under variant lighting conditions, which make it very suitable for evaluating the proposed operators. The database is divided into training set (5011 images) and test set (4952 images). The goal is to recognize the objects in images and to classify them into the correct categories. The mean average precision (MAP) is used as the evaluation criterion.

A. Experimental Setup

Three widely-used popular texture descriptors are chosen to make comparison, including Gabor filter [10], Texture Auto Correlation (TAC) [11], and Grey Level Co-occurrence Matrix (GLCM) [11].

For Gabor filter, 5 scales and 8 orientations are used. For TAC, 0 to 8 with step of 2 are applied as position difference in both \(x\) and \(y\) directions. For GLCM, 4 directions (horizontal, vertical and diagonal) with 1 offset between two pixels are considered.

The Support Vector Machine (SVM) algorithm is applied for classification. Once all the feature descriptors are extracted from the database, the \(\chi^2\) distance is computed to measure the similarity between each pair of the feature vectors \(F\) and \(F'\) (\(n\) is the size of the feature vector):

\[
dist_{\chi^2}(F, F') = \sum_{i=1}^{n} \frac{(F_i - F'_i)^2}{F_i + F'_i} \tag{12}
\]

Then, the kernel function based on this distance is used for SVM to train the classifier:

\[
K_{\chi^2}(F, F') = e^{-\frac{1}{D} \text{dist}_{\chi^2}(F, F')} \tag{13}
\]

where \(D\) is the parameter for normalizing the distances. Here \(D\) is set to the average value of the training set.

Finally, the precision-recall curve is plotted according to the output decision values of the SVM classifier, and the MAP is computed based on the proportion of the area under this curve.

B. Experimental Results

The proposed multi-scale color LBP operators are compared at first with the original LBP.

From the results shown in Fig. 3, it can be seen that the multi-scale LBP outperforms the original LBP by 14.1%, proving the importance of obtaining more local information and invariance to scaling. The proposed operators all further outperform the multi-scale LBP, with the improvements from 2.5% to 10.2% (17.0% ~ 25.8% if compared with the original LBP), which proves that the proposed operators truly have more discriminative power benefitting from the additional properties of illumination invariance.

Among these operators, Hue-LBP, Opponent-LBP and nOpponent-LBP have the best overall performance (improvement over 6% than the multi-scale LBP and over 20% than the original LBP), consistent with their strong properties of illumination invariance.

As one kind of texture feature, the best three multi-scale color LBP operators are also compared with some other widely-used popular texture descriptors.

![Figure 2. Multi-scale LBP operator](image-url)
VI. CONCLUSION

In this paper, we proposed six novel multi-scale color Local Binary Pattern (LBP) operators to deal with main shortcomings of the original LBP operator, namely deficiency of color information and sensitivity to non-monotonic lighting condition changes. The proposed operators not only have more discriminative power by obtaining more local information, but also possess invariance properties to different lighting condition changes. In addition, they keep the advantage of computational simplicity from the original LBP operator. The experimental results on the PASCAL Visual Object Classes (VOC) 2007 image benchmark showed that these proposed novel operators have gained significant accuracy improvement, and are more promising for real-world object recognition tasks.

<table>
<thead>
<tr>
<th>Combination</th>
<th>Mean Average Precision (MAP)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LBP (original)</td>
<td>0.283</td>
</tr>
<tr>
<td>LBP (multi-scale)</td>
<td>0.323</td>
</tr>
<tr>
<td>Hue-LBP (multi-scale)</td>
<td>0.356</td>
</tr>
<tr>
<td>Opponent-LBP (multi-scale)</td>
<td>0.351</td>
</tr>
<tr>
<td>nOpponent-LBP (multi-scale)</td>
<td>0.344</td>
</tr>
<tr>
<td>Combination</td>
<td>0.372</td>
</tr>
<tr>
<td>SIFT in [8]</td>
<td>≈ 0.38</td>
</tr>
</tbody>
</table>

TABLE I. COMPARISON OF THE PROPOSED MULTI-SCALE COLOR LBP OPERATORS AND THE SIFT DESCRIPTOR

ACKNOWLEDGMENT

This work was partly supported by the French ANR Omnia project under the grant ANR-07-MDCO-009-02.

REFERENCES