Rotation Invariant Image Description with Local Binary Pattern Histogram Fourier Features

Timo Ahonen[1], Jiří Matas[2], Chu He[3,1], and Matti Pietikäinen[1]

[1] Machine Vision Group, University of Oulu, Finland
{tahonen,mkp}@ee.oulu.fi
[2] Center for Machine Percpetion, Dept. of Cybernetics,
Faculty of Elec. Eng., Czech Technical University in Prague
matas@cmp.felk.cvut.cz
[3] School of Electronic Information, Wuhan University, P.R. China
chuhe.whu@gmail.com

Abstract. In this paper, we propose Local Binary Pattern Histogram Fourier features (LBP-HF), a novel rotation invariant image descriptor computed from discrete Fourier transforms of local binary pattern (LBP) histograms. Unlike most other histogram based invariant texture descriptors which normalize rotation locally, the proposed invariants are constructed globally for the whole region to be described. In addition to being rotation invariant, the LBP-HF features retain the highly discriminative nature of LBP histograms. In the experiments, it is shown that these features outperform non-invariant and earlier version of rotation invariant LBP and the MR8 descriptor in texture classification, material categorization and face recognition tests.

1 Introduction

Rotation invariant texture analysis is a widely studied problem [1], [2], [3]. It aims at providing with texture features that are invariant to rotation angle of the input texture image. Moreover, these features should typically be robust also to image formation conditions such as illumination changes.

Describing the appearance locally, e.g., using co-occurrences of gray values or with filter bank responses and then forming a global description by computing statistics over the image region is a well established technique in texture analysis [4]. This approach has been extended by several authors to produce rotation invariant features by transforming each local descriptor to a canonical representation invariant to rotations of the input image [2], [3], [5]. The statistics describing the whole region are then computed from these transformed local descriptors.

Even though such approaches have produced good results in rotation invariant texture classification, they have some weaknesses. Most importantly, as each local descriptor (e.g., filter bank response) is transformed to canonical representation independently, the relative distribution of different orientations is lost. Furthermore, as the transformation needs to be performed for each texton, it must be computationally simple if the overall computational cost needs to be low.

In this paper, we propose novel Local Binary Pattern Histogram Fourier features (LBP-HF), a rotation invariant image descriptor based on uniform Local Binary Patterns (LBP) [2]. LBP is an operator for image description that is based on the signs of differences of neighboring pixels. It is fast to compute and invariant to monotonic gray-scale changes of the image. Despite being simple, it is very descriptive, which is attested by the wide variety of different tasks it has been successfully applied to. The LBP histogram has proven to be a widely applicable image feature for, e.g., texture classification, face analysis, video background subtraction, interest region description, etc[1].

Unlike the earlier local rotation invariant features, the LBP-HF descriptor is formed by first computing a non-invariant LBP histogram over the whole region and then constructing rotationally invariant features from the histogram. This means that rotation invariance is attained globally, and the features are thus invariant to rotations of the whole input signal but they still retain information about relative distribution of different orientations of uniform local binary patterns.

2   Rotation Invariant Local Binary Pattern Descriptors

The proposed rotation invariant local binary pattern histogram Fourier features are based on uniform local binary pattern histograms. First, the LBP methodology is briefly reviewed and the LBP-HF features are then introduced.

2.1   The Local Binary Pattern Operator

The local binary pattern operator [2] is a powerful means of texture description. The original version of the operator labels the image pixels by thresholding the 3x3-neighborhood of each pixel with the center value and summing the thresholded values weighted by powers of two.

The operator can also be extended to use neighborhoods of different sizes [2] (See Fig.1). To do this, a circular neighborhood denoted by (P, R) is defined. Here P represents the number of sampling points and R is the radius of the neighborhood. These sampling points around pixel (x, y) lie at coordinates $(x_p, y_p) = (x + R\cos(2\pi p/P), y - R\sin(2\pi p/P))$. When a sampling point does not fall at integer coordinates, the pixel value is bilinearly interpolated. Now the LBP label for the center pixel (x, y) of image $f(x, y)$ is obtained through
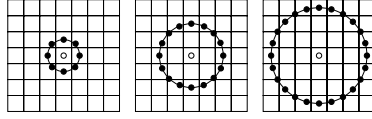
$$LBP_{P,R}(x, y) = \sum_{p=0}^{P-1} s(f(x, y) - f(x_p, y_p))2^p, \tag{1}$$

where $s(z)$ is the thresholding function

$$s(z) = \begin{cases} 1, z \geq 0 \\ 0, z < 0 \end{cases} \tag{2}$$

---

[1] See LBP bibliography at http://www.ee.oulu.fi/mvg/page/lbp_bibliography

Fig. 1. Three circular neighborhoods: (8,1), (16,2), (24,3). The pixel values are bilinearly interpolated whenever the sampling point is not in the center of a pixel.

Further extensions to the original operator are so called *uniform* patterns [2]. A local binary pattern is called uniform if the binary pattern contains at most two bitwise transitions from 0 to 1 or vice versa when the bit pattern is considered circular. In the computation of the LBP histogram, uniform patterns are used so that the histogram has a separate bin for every uniform pattern and all non-uniform patterns are assigned to a single bin. The 58 possible uniform patterns in neighborhood of 8 sampling points are shown in Fig. 2.

The original rotation invariant LBP operator, denoted here as LBP^{riu2}, is achieved by circularly rotating each bit pattern to the minimum value. For instance, the bit sequences 1000011, 1110000 and 0011100 arise from different rotations of the same local pattern and they all correspond to the normalized sequence 0000111. In Fig. 2 this means that all the patterns from one row are replaced with a single label.

2.2   Invariant Descriptors from LBP Histograms

Let us denote a specific uniform LBP pattern by $U_P(n, r)$. The pair (n, r) specifies an uniform pattern so that n is the number of 1-bits in the pattern (corresponds to row number in Fig. 2) and r is the rotation of the pattern (column number in Fig. 2).

Now if the neighborhood has P sampling points, n gets values from 0 to $P+1$, where $n = P + 1$ is the special label marking all the non-uniform patterns. Furthermore, when $1 \leq n \leq P - 1$, the rotation of the pattern is in the range $0 \leq r \leq P - 1$.

Let $I^{\alpha^\circ}(x, y)$ denote the rotation of image $I(x, y)$ by α degrees. Under this rotation, point (x, y) is rotated to location (x', y'). If we place a circular sampling neighborhood on points $I(x, y)$ and $I^{\alpha^\circ}(x', y')$, we observe that it also rotates by α°. See Fig. 3.

If the rotations are limited to integer multiples of the angle between two sampling points, *i.e.* $\alpha = a\frac{360^\circ}{P}, a = 0, 1, \ldots, P - 1$, this rotates the sampling neighborhood exactly by a discrete steps. Therefore the uniform pattern $U_P(n, r)$ at point (x, y) is replaced by uniform pattern $U_P(n, r + a \mod P)$ at point (x', y') of the rotated image.

Now consider the uniform LBP histograms $h_I(U_P(n, r))$. The histogram value h_I at bin $U_P(n, r)$ is the number of occurrences of uniform pattern $U_P(n, r)$ in image I.
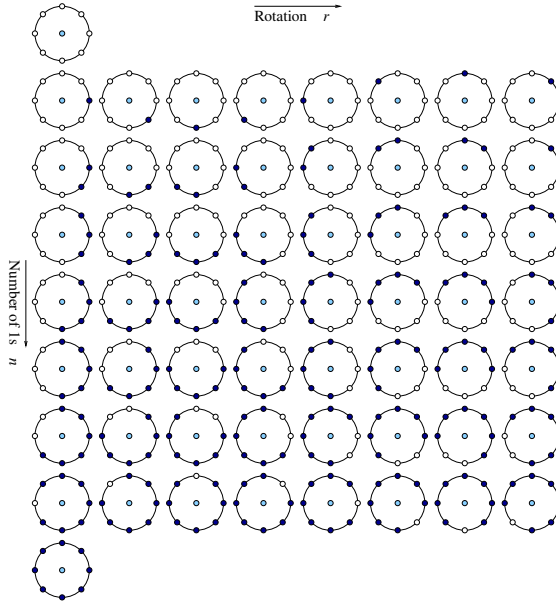
Fig. 2. The 58 different uniform patterns in (8,R) neighborhood

.

If the image I is rotated by $\alpha = a\frac{360°}{P}$, based on the reasoning above, this rotation of the input image causes a cyclic shift in the histogram along each of the rows,

$$h_{I^{a\circ}}(U_P(n, r + a)) = h_I(U_P(n, r)) \tag{3}$$

For example, in the case of 8 neighbor LBP, when the input image is rotated by 45°, the value from histogram bin $U_8(1,0) = 000000001b$ moves to bin $U_8(1,1) = 00000010b$, value from bin $U_8(1,1)$ to bin $U_8(1,2)$, etc.

Based on the property, which states that rotations induce shift in the polar representation (P, R) of the neighborhood, we propose a class of features that are invariant to rotation of the input image, namely such features, computed along the input histogram rows, that are invariant to cyclic shifts.

We use the Discrete Fourier Transform to construct these features. Let $H(n, \cdot)$ be the DFT of nth row of the histogram $h_I(U_P(n,r))$, *i.e.*

$$H(n, u) = \sum_{r=0}^{P-1} h_I(U_P(n,r))e^{-i2\pi ur/P}. \tag{4}$$

Now for DFT it holds that a cyclic shift of the input vector causes a phase shift in the DFT coefficients. If $h'(U_P(n,r)) = h(U_P(n, r - a))$, then

$$H'(n, u) = H(n, u)e^{-i2\pi ua/P}, \tag{5}$$
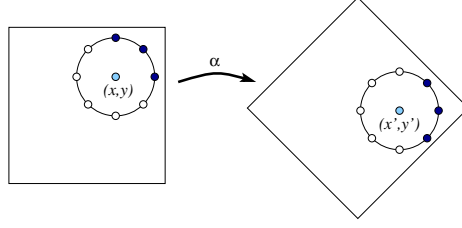
Fig. 3. Effect of image rotation on points in circular neighborhoods

and therefore, with any $1 \le n_1, n_2 \le P - 1$,

$$H'(n_1, u)\overline{H'(n_2, u)} = H(n_1, u)e^{-i2\pi ua/P}\overline{H(n_2, u)}e^{i2\pi ua/P} = H(n_1, u)\overline{H(n_2, u)}, \tag{6}$$

where $\overline{H(n_2, u)}$ denotes the complex conjugate of $H(n_2, u)$.

This shows that with any $1 \le n_1, n_2 \le P - 1$ and $0 \le u \le P - 1$, the features

$$\text{LBP}^{u2}\text{-HF}(n_1, n_2, u) = H(n_1, u)\overline{H(n_2, u)}, \tag{7}$$

are invariant to cyclic shifts of the rows of $h_I(U_P(n, r))$ and consequently, they are invariant also to rotations of the input image $I(x, y)$. The Fourier magnitude spectrum
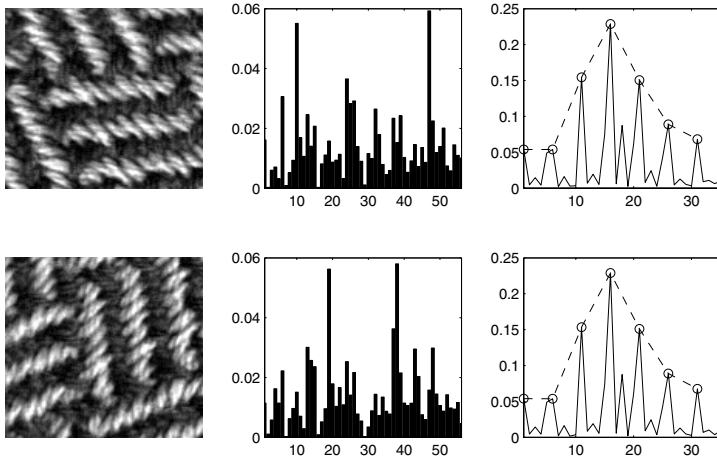
Fig. 4. 1st column: Texture image at orientations $0°$ and $90°$. 2nd column: bins 1–56 of the corresponding LBP^{u2} histograms. 3rd column: Rotation invariant features $|H(n, u)|, 1 \le n \le 7, 0 \le u \le 5$, (solid line) and LBP^{riu2} (circles, dashed line). Note that the LBP^{u2} histograms for the two images are markedly different, but the $|H(n, u)|$ features are nearly equal.

$$|H(n, u)| = \sqrt{H(n, u)\overline{H(n, u)}} \tag{8}$$

can be considered a special case of these features. Furthermore it should be noted that the Fourier magnitude spectrum contains LBP^{riu2} features as a subset, since

$$|H(n, 0)| = \sum_{r=0}^{P-1} h_I(U_P(n, r)) = h_{\text{LBP}^{riu2}}(n). \tag{9}$$

An illustration of these features is in Fig. 4

3   Experiments

We tested the performance of the proposed descriptor in three different scenarios: texture classification, material categorization and face description. The proposed rotation invariant LBP-HF features were compared against non-invariant LBP^{u2} and the older rotation invariant version LBP^{riu2}. In the texture classification and material categorization experiments, the MR8 descriptor [3] was used as an additional control method. The results for the MR8 descriptor were computed using the setup from [6].

In preliminary tests, the Fourier magnitude spectrum was found to give most consistent performance over the family of different possible features (Eq. (7)). Therefore, in the following we use feature vectors consisting of three LBP histogram values (all zeros, all ones, non-uniform) and Fourier magnitude spectrum values. The feature vectors are of the following form:

$$
\begin{aligned}
fv_{\text{LBP-HF}} = [&|H(1,0)|, \ldots, |H(1, P/2)|, \\
&\ldots, \\
&|H(P-1, 0)|, \ldots, |H(P-1, P/2)|, \\
&h(U_P(0,0)), h(U_P(P,0)), h(U_P(P+1,0))]_{1 \times ((P-1)(P/2+1)+3)}.
\end{aligned}
$$

In experiments we followed the setup of [2] for nonparametric texture classification. For histogram type features, we used the log-likelihood statistic, assigning a sample to the class of model minimizing the LL distance

$$\text{LL}(h^S, h^M) = -\sum_{b=1}^{B} h^S(b) \log h^M(b), \tag{10}$$

where $h^S(b)$ and $h^M(b)$ denote the bin b of sample and model histograms, respectively. The LL distance is suited for histogram type features, thus a different distance measure was needed for the LBP-HF descriptor. For these features, the L_1 distance

$$\text{L}_1(fv^S_{\text{LBP-HF}}, fv^M_{\text{LBP-HF}}) = \sum_{k=1}^{K} |fv^S_{\text{LBP-HF}}(k) - fv^M_{\text{LBP-HF}}(k)| \tag{11}$$

was selected. We derived from the setup of [2] by using nearest neighbor (NN) classifier instead of 3NN because no significant performance difference between the two was observed and in the setup for the last experiment we had only 1 training sample per class.

Table 1. Texture recognition rates on Outex_TC_0012 dataset

	LBP^{u2}	LBP^{riu2}	LBP-HF
$(8, 1)$	0.566	0.646	0.773
$(16, 2)$	0.578	0.791	0.873
$(24, 3)$	0.45	0.833	0.896
$(8, 1) + (16, 2)$	0.595	0.821	0.894
$(8, 1) + (24, 3)$	0.512	0.883	0.917
$(16, 2) + (24, 3)$	0.513	0.857	0.915
$(8, 1) + (16, 2) + (24, 3)$	0.539	0.87	**0.925**
MR8		0.761	

3.1   Experiment 1: Rotation Invariant Texture Classification

In the first experiment, we used the Outex_TC_0012 [7] test set intended for testing rotation invariant texture classification methods. This test set consists of 9120 images representing 24 different textures imaged under different rotations and lightings. The test set contains 20 training images for each texture class. The training images are under single orientation whereas different orientations are present in the total of 8640 testing images. We report here the total classification rates over all test images.

The results of the first experiment are in Table 1. As it can be observed, the both rotation invariant features provide better classification rates than non-invariant features. The performance of LBP-HF features is clearly higher than that of MR8 and LBP^{riu2}. This can be observed at all tested scales, but the difference between LBP-HF and LBP^{riu2} is particularly large at the smallest scale $(8, 1)$.

3.2   Experiment 2: Material Categorization

In next experiments, we aimed to test how well the novel rotation invariant features retain the discriminativeness of the original LBP features. This was tested using two challenging problems, namely material categorization and illumination invariant face recognition

In Experiment 2, we tested the performance of the proposed features in material categorization using the KTH-TIPS2 database[2]. For this experiment, we used the same setup as in Experiment 1. This test setup resembles the most difficult setup used in [8].

The KTH-TIPS2 database contains 4 samples of 11 different materials, each sample imaged at 9 different scales and 12 lighting and pose setups, totaling 4572 images. Using each of the descriptors to be tested, a nearest neighbor classifier was trained with one sample (i.e. 9*12 images) per material category. The remaining 3*9*12 images were used for testing. This was repeated with 10000 random combinations as training and testing data and the mean categorization rate over the permutations is used to assess the performance.

---

[2] http://www.nada.kth.se/cvap/databases/kth-tips/

Table 2. Material categorization rates on KTH TIPS2 dataset

	LBP^{u2}	LBP^{riu2}	LBP-HF
$(8, 1)$	0.528	0.482	0.525
$(16, 2)$	0.511	0.494	0.533
$(24, 3)$	0.502	0.481	0.513
$(8, 1) + (16, 2)$	0.536	0.502	0.542
$(8, 1) + (24, 3)$	0.542	0.507	0.542
$(16, 2) + (24, 3)$	0.514	0.508	0.539
$(8, 1) + (16, 2) + (24, 3)$	0.536	0.514	**0.546**
MR8		0.455	

Results of material categorization experiments are in Table 2. LBP-HF reaches, or with most scales even exceeds the performance of LBP^{u2}. The performance of LBP^{riu2} is consistently lower than that of the other two, and the MR8 descriptor gives the lowest recognition rate. The reason for LBP-HF not performing significantly better then non-invariant LBP is most likely that different orientations are present in the training data so rotational invariance does not benefit much here. Unlike with LBP^{riu2}, no information is lost either, but a slight improvement over the non-invariant descriptor is achieved instead.

3.3   Experiment 3: Face Recognition

The third experiment was aimed to further assess whether useful information is lost due to the transformation making the features rotation invariant. For this test, we chose the face recognition problem where the input images have been manually registered, so rotation invariance is not actually needed.

The CMU PIE (Pose, Illumination, and Expression) database [9] was used for this experiment. Totally, the database contains 41368 images of 68 subjects taken at different angles, lighting conditions and with varying expression. For our experiments, we selected a set of 23 images of each of the 68 subjects. 2 of these are taken with the room lights on and the remaining 21 each with a flash at varying positions.

In obtaining a descriptor for the facial image, the procedure of [10] was followed. The faces were first normalized so that the eyes are at fixed positions. The uniform LBP operator at chosen scale was then applied and the resulting label image was cropped to size 128×128 pixels. The cropped image was further divided into blocks of size of 16×16 pixels and histograms were computed in each block individually. In case of LBP-HF descriptor, the rotation invariant transform was applied to the histogram, and finally the features obtained within each block were concatenated to form the spatially enhanced histogram describing the face.

Due to the sparseness of the resulting histograms, Chi square distance was used with histogram type features in this experiments. With LBP-HF descriptor, L1 distance was used as in the previous experiment.

Table 3. Face recognition rates on CMU PIE dataset

	LBP^{u2}	LBP^{riu2}	LBP-HF
$(8,1)$	0.726	0.649	0.716
$(8,2)$	0.744	0.699	**0.748**
$(8,3)$	0.727	0.680	0.726

On each test round, one image per person was used for training and the remaining 22 images for testing. Again, 10000 random selections into training and testing data were used.

Results of the face recognition experiment are in Table 3. Surprisingly, the performance of rotation invariant LBP-HF is almost equal to non-invariant LBP^{u2} even though there are no global rotations present in the images.

4   Discussion and Conclusion

In this paper, we proposed rotation invariant LBP-HF features based on local binary pattern histograms. It was shown that rotations of the input image cause cyclic shifts of the values in the uniform LBP histogram. Relying on this observation we proposed discrete Fourier transform based features that are invariant to cyclic shifts of input vector and, when computed from uniform LBP histograms, hence invariant to rotations of input image.

Several other histogram based rotation invariant texture features have been discussed in the literature, e.g., [2], [3], [5]. The method proposed here differs from those since LBP-HF features are computed from the histogram representing the whole region, i.e. the invariants are constructed globally instead of computing invariant independently at each pixel location. The major advantage of this approach is that the relative distribution of local orientations is not lost.

Another benefit of constructing invariant features globally is that invariant computation needs not to be performed at every pixel location. This allows for using computationally more complex invariant functions still keeping the total computational cost reasonable. In case of LBP-HF descriptor, the computational overhead is negligible. After computing the non-invariant LBP histogram, only $P-1$ Fast Fourier Transforms of P points need to be computed to construct the rotation invariant LBP-HF descriptor.

In the experiments, it was shown that in addition to being rotation invariant, the proposed features retain the highly discriminative nature of LBP histograms. The LBP-HF descriptor was shown to outperform the MR8 descriptor and the non-invariant and earlier version of rotation invariant LBP in texture classification, material categorization and face recognition tests.

References

1. Zhang, J., Tan, T.: Brief review of invariant texture analysis methods. Pattern Recognition 35(3), 735–747 (2002)
2. Ojala, T., Pietikäinen, M., Mäenpää, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Transactions on Pattern Analysis and Machine Intelligence 24(7), 971–987 (2002)
3. Varma, M., Zisserman, A.: A statistical approach to texture classification from single images. International Journal of Computer Vision 62(1–2), 61–81 (2005)
4. Tuceryan, M., Jain, A.K.: Texture analysis. In: Chen, C.H., Pau, L.F., Wang, P.S.P. (eds.) The Handbook of Pattern Recognition and Computer Vision, 2nd edn., pp. 207–248. World Scientific Publishing Co., Singapore (1998)
5. Arof, H., Deravi, F.: Circular neighbourhood and 1-d dft features for texture classification and segmentation. IEE Proceedings - Vision, Image and Signal Processing 145(3), 167–172 (1998)
6. Ahonen, T., Pietikäinen, M.: Image description using joint distribution of filter bank responses. Pattern Recognition Letters 30(4), 368–376 (2009)
7. Ojala, T., Mäenpää, T., Pietikäinen, M., Viertola, J., Kyllönen, J., Huovinen, S.: Outex - new framework for empirical evaluation of texture analysis algorithms. In: Proc. 16th International Conference on Pattern Recognition (ICPR 2002), vol. 1, pp. 701–706 (2002)
8. Caputo, B., Hayman, E., Mallikarjuna, P.: Class-specific material categorisation. In: 10th IEEE International Conference on Computer Vision (ICCV 2005), pp. 1597–1604 (2005)
9. Sim, T., Baker, S., Bsat, M.: The cmu pose, illumination, and expression database. IEEE Transactions on Pattern Analysis and Machine Intelligence 25(12), 1615–1618 (2003)
10. Ahonen, T., Hadid, A., Pietikäinen, M.: Face description with local binary patterns: Application to face recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence 28(12), 2037–2041 (2006)