

## Research Article

# Compact Local Directional Texture Pattern for Local Image Description

Hui Zeng,<sup>1</sup> Rui Zhang,<sup>1</sup> Mingming Huang,<sup>1</sup> and Xiuqing Wang<sup>2</sup>

<sup>1</sup>School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing 100083, China

<sup>2</sup>Vocational & Technical Institute, Hebei Normal University, Shijiazhuang 050024, China

Correspondence should be addressed to Hui Zeng; hzeng@163.com

Received 7 May 2015; Revised 5 August 2015; Accepted 18 August 2015

Academic Editor: Chong Wah Ngo

Copyright © 2015 Hui Zeng et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper presents an effective local image feature region descriptor, called CLDTP descriptor (Compact Local Directional Texture Pattern), and its application in image matching and object recognition. The CLDTP descriptor encodes the directional and contrast information in a local region, so it contains the gradient orientation information and the gradient magnitude information. As the dimension of the CLDTP histogram is much lower than the dimension of the LDTP histogram, the CLDTP descriptor has higher computational efficiency and it is suitable for image matching. Extensive experiments have validated the effectiveness of the designed CLDTP descriptor.

## 1. Introduction

Local image descriptor construction is one of the fundamental problems in the fields of image analysis and pattern recognition. It has been widely used in image stitching, image matching, object recognition, visual tracking, robot localization, 3D reconstruction, and other applications. For an ideal local image feature, it should have high discriminative power and robustness to resist many kinds of image transformations, such as illumination, rotation, scale, and blur. Furthermore, it also should have low computational complexity [1]. In this paper, we will focus on robust and efficient local image feature descriptor construction and its application in image matching and object recognition.

Many local feature descriptor construction methods have been proposed in the literature, such as the Gaussian derivative based descriptor, the moment invariants based descriptor, the spatial frequency descriptor, the distribution of pixel gray values based descriptor, and the distribution of pixel gradient values based descriptor. Among these methods, the most widely used descriptor is SIFT (Scale Invariant Feature Transform) [2]. The SIFT descriptor is a 128-dimensional vector, which is built by a histogram of gradient locations and orientations where the contribution to bins is weighted

by the gradient magnitude and a Gaussian window overlaid over the region. It is invariant to image scale and rotation and robust to affine distortion, changes in 3D viewpoint, addition of noise, and changes in illumination. Inspired by the well performance of SIFT descriptor, many extensions of SIFT descriptor have been proposed. Ke and Sukthankar apply PCA on gradient maps to construct PCA-SIFT [3]. The dimension of the PCA-SIFT descriptor is 36 and it can be fast for matching. Bay proposes SURF (Speeded Up Robust Features) descriptor, which speeds up the gradient computations using integral images [4]. Some researchers improve the support region division method, such as GLOH (Gradient Location-Orientation Histogram) descriptor and RIFT (Rotation-Invariant Feature Transform) descriptor [5, 6]. The experimental results of several comparative studies on local image feature descriptor have shown that the SIFT-like descriptors perform best [5].

LBP (Local Binary Pattern) is one of the most popular texture features and has been widely used in face recognition, background extraction, image retrieval, and so on. It has been proved as a powerful means of texture analysis in 2D images, which uses local texture pattern as the texture primitive [7]. It has many advantages which are suitable for local feature region description, such as computational simplicity and

invariance to linear illumination. As the dimension of its histogram is high and it is not robust to flat image areas, it is not suitable to construct the local image feature descriptor directly. To address these problems, the CS-LBP (Center Symmetric Local Binary Pattern) descriptor has been proposed, which combines the strength of the SIFT descriptor and LBP operator [8]. The experimental results have shown that the CS-LBP descriptor performs better than the SIFT descriptor in the field of image matching. Though it shows robustness to monotonic illumination, it is sensitive to nonmonotonic illumination variation. LTP (Local Ternary Pattern), which is an improvement of LBP, describes the micropatterns by two thresholds and has better discriminative power and robustness than LBP [9]. But the dimension of the LTP histogram is extremely high and LTP is also not suitable to construct the local image feature descriptor directly. So CS-LTP (Center Symmetric Local Ternary Pattern) operator is proposed to alleviate the dimension problem [10]. However, the dimension of SIFT-like grid based CS-LTP descriptor is still very high. For example, the CS-LTP descriptor with 8 neighboring pixels and  $4 \times 4$  squared subregions is a 1296-dimensional ( $(16 \times 3^4)$ -dimensional) vector. Recently, Huang et al. proposed WOS-LTP (Weighted Orthogonal Symmetric Local Ternary Pattern) descriptor [11], which is an improvement of CS-LTP descriptor and achieves robustness against noise interference and discriminative ability for describing texture structure.

LDP (Local Directional Pattern) is another kind of local texture pattern, which is calculated by comparing the relative edge response value of a pixel in different directions [12]. It is insensitive to noise and nonmonotonic illumination variations, but it is sensitive to rotations and cannot describe the variety of intensity information. LDTP (Local Directional Texture Pattern) combines the advantages of CS-LTP and LDP, which includes both directional and intensity information [13]. The LDTP histogram is consistent against noise and illumination changes, and its dimension is 72. The SIFT-like grid based LDTP descriptor with  $4 \times 4$  squared subregions is a 1152-dimensional ( $(16 \times 72)$ -dimensional) vector. So the LDTP operator is not suitable to construct the local image feature descriptor for image matching.

In this paper, we propose a novel descriptor named as CLDTP (Compact Local Directional Texture Pattern), which not only reduces the dimension of LDTP descriptor but also retains the advantages of LDTP descriptor. Similar to the LDTP operator, CLDTP operator encodes the directional and contrast information in a local region by analyzing its principal directions and edge's responses. Compared with the LDTP histogram, the dimension of CLDTP histogram is reduced effectively. The dimension of SIFT-like grid based CLDTP descriptor with  $4 \times 4$  squared subregions is 320 ( $16 \times 20$ ). The performance of the CLDTP descriptor is evaluated for image matching and object recognition and the experimental results demonstrate its robustness and distinctiveness.

The rest of the paper is organized as follows. In Section 2, the LDTP operator and the CLDTP operator are introduced. Section 3 gives the construction method of the CLDTP descriptor. The image matching and object recognition

experiments are conducted and their experimental results are presented in Section 4. Some concluding remarks are listed in Section 5.

## 2. LDTP Operator and CLDTP Operator

*2.1. Local Directional Texture Pattern (LDTP).* The LDTP operator is a powerful texture operator which extracts the texture information from the principal axis in each neighborhood. Compared with other operators that try to accommodate all available information, which sometimes may introduce errors into the code, the LDTP operator only includes the principal information. It has been used for facial expression recognition and scene recognition and exhibits good performance.

To obtain the LDTP code, the eight absolute edge response values  $F_i$  of each pixel are firstly calculated using Kirsch masks by

$$F_i = |I * M_i| \quad (i = 0, 1, \dots, 7), \quad (1)$$

where  $I$  is the image to be described,  $M_i$  is the  $i$ th Kirsch mask, and  $*$  denotes the convolution operation.

Then two principal directions are computed by sorting the response value  $F_i$  of each pixel. The first principal directional number of each pixel is determined by

$$P_{\text{dir}}^1 = \operatorname{argmax} \{F_i \mid 0 \leq i \leq 7\}. \quad (2)$$

The second principal directional number  $P_{\text{dir}}^2$  of each pixel can be determined in the same way, and it is the order number of the second maximum response.

As shown in Figure 1,  $m_c$  is the gray value of the center pixel and  $m_0, \dots, m_7$  are the gray value of its 8 neighborhood pixels. In each of the two principal directions, the difference of the gray value in the neighborhood is computed by

$$d_n = \begin{cases} m_{P_{\text{dir}}^n} - m_{P_{\text{dir}}^n+4} & P_{\text{dir}}^n \in \{0, 1, 2, 3\} \\ m_{P_{\text{dir}}^n} - m_{P_{\text{dir}}^n-4} & P_{\text{dir}}^n \in \{4, 5, 6, 7\}. \end{cases} \quad (3)$$

Then each difference can be encoded as

$$D_f(d_n) = \begin{cases} 0, & -\varepsilon \leq d_n \leq \varepsilon \\ 1, & d_n < -\varepsilon \\ 2, & d_n > \varepsilon, \end{cases} \quad (4)$$

where  $d_n$  is the difference of the gray value computed according (3),  $\varepsilon$  is the user-specified threshold, and  $D_f(d_n)$  is the encoded value of the difference  $d_n$ .

For each pixel, the code of LDTP can be calculated by concatenating the binary form of its first principal directional number  $P_{\text{dir}}^1$ , its first directional encoded difference  $D_f(d_1)$ , and its second directional encoded difference  $D_f(d_2)$ . As the first principal directional number has 8 possible values and the encoded difference has 3 possible values, the dimension of the LDTP histogram is 72 ( $8 \times 3 \times 3$ ). More detailed description about the LDTP operator can be found in [12].

$m_3$	$m_2$	$m_1$
$m_4$	$m_c$	$m_0$
$m_5$	$m_6$	$m_7$

FIGURE 1: The center pixel  $m_c$  and its 8 neighborhood pixels.

## 2.2. Compact Local Directional Texture Pattern (CLDTP).

Although the LDTP operator has high discriminative ability, the dimension of its histogram is high; it is not suitable to be directly used in image matching. To alleviate this problem, we propose CLDTP (Compact Local Directional Texture Pattern) operator, which can reduce the dimension of the histogram effectively while including both directional information and contrast information.

Unlike the LDTP operator, the CLDTP operator does not use the absolute edge response values of Kirsch masks. This is because of the fact that the sign of the responses values also includes some distinctive directional information. So we use the responses values to compute the CLDTP code directly. At first, we calculate the eight edge response values  $R_i$  of each pixel using Kirsch masks by

$$R_i = I * M_i \quad (i = 0, 1, \dots, 7). \quad (5)$$

As the LDTP operator uses the absolute edge response values to compute the principal directions, the local neighborhoods with different texture pattern often have the same principal directions. For example, as shown in Figure 2, the first principal directional numbers of the two sample neighborhoods are the same. That is to say, the principal directions of the LDTP operator cannot distinguish the sign of the edge response values. To solve this problem, we use the sign and the order number of the maximum absolute response value to determine the first principal directional number of each pixel. It can be computed as

$$P_{\text{dir}}^1 = \underset{i}{\operatorname{argmax}} \{ |R_i| \mid 0 \leq i \leq 7 \} + 8 \cdot S, \quad (6)$$

$$S = \begin{cases} 1, & R_i \geq 0 \\ 0, & R_i < 0. \end{cases}$$

From (6) we can see that the range of the first principal directional number is from 0 to 15.

Inspired by the WOS-LTP descriptor, we use the intensity differences of two orthogonal directions to construct the CLDTP descriptor. The two orthogonal directions are the first principal direction and its orthogonal direction. In the first

principal direction, we compute the intensity difference of the opposed pixels in the neighborhood. That is,

$$d_1 = |m_k - m_{k+4}|, \quad k = P_{\text{dir}}^1 - 4 \times \left\lfloor \frac{P_{\text{dir}}^1}{4} \right\rfloor, \quad (7)$$

where  $\lfloor x \rfloor$  is the floor function, which maps a real number to the nearest integer less than or equal to  $x$ . In the orthogonal direction of the first principal direction, the intensity difference of the opposed pixels in the neighborhood can be computed by

$$d_2 = \begin{cases} |m_{k+2} - m_{k+6}|, & \text{if } k = 0, 1 \\ |m_{k+2} - m_{k-2}|, & \text{if } k = 2, 3. \end{cases} \quad (8)$$

Like the LBP binary encoding, the above two intensity differences can be encoded as

$$D_n = \begin{cases} 0, & d_n > \varepsilon \\ 1, & d_n \leq \varepsilon, \end{cases} \quad (9)$$

where  $D_n$  is the encoded intensity difference,  $d_n$  is the actual intensity difference, and  $\varepsilon$  is a threshold value defined by experiments. It should be noted that although the LTP ternary encoding has better discriminative power than the LBP binary encoding, we still use binary encoding method. This is because of the fact that the principal directional number of the CLDTP operator has 16 possible values, and it already contains positive and negative information of intensity difference.

For example, consider the neighborhood shown in Figure 3. Assume the threshold  $\varepsilon = 15$ . Using (6) we can obtain that  $P_{\text{dir}}^1 = 5 + 8 = 13$ . Then the two-intensity difference in the first principal direction and the orthogonal direction of the first principal direction can be computed using (7) and (8). We can obtain that  $d_1 = |103 - 99| = 4$  and  $d_2 = |110 - 118| = 8$ . From (9), we can obtain the coded differences  $D_1 = 0$  and  $D_2 = 0$ .

To reduce the histogram of the CLDTP operator, we encode the directional number and the difference separately. The CLDTP operator can be defined as follows:

$$\begin{aligned} \text{CLDTP}^{(1)}(u, v) &= P_{\text{dir}}^1, \\ \text{CLDTP}^{(2)}(u, v) &= 2D_1 + D_2. \end{aligned} \quad (10)$$

As the directional number has 16 possible values and the encoded difference has 4 possible values, here the directional number represents the image gradient direction information and the encoded difference represents the image gradient magnitude information, so the CLDTP operator can describe the micropattern effectively.

## 3. Local Feature Descriptor Construction

**3.1. The CLDTP Histogram.** For the local image region, after the corresponding CLDTP code of each pixel has been computed, the corresponding histogram can be obtained

15	-16	49
-27	$x$	47
38	85	43

15	-16	49
-27	$x$	47
38	-85	43

FIGURE 2: The edge response values of two sample neighborhoods.

102	103	112
110	101	118
91	99	97

5	-78	47
-57	$x$	66
74	60	23

(a) The sample neighborhood      (b) The edge response values

FIGURE 3: A sample neighborhood and its edge response values.

by computing the number of the occurrences of each pattern. The CLDTP histogram includes the histogram of  $\text{CLDTP}^{(1)}(u, v)$  and the histogram of  $\text{CLDTP}^{(2)}(u, v)$ , and they can be computed as follows:

$$H_1(k_1) = \sum_{u=1}^W \sum_{v=1}^H f(\text{CLDTP}^{(1)}(u, v), k_1),$$

$$H_2(k_2) = \sum_{u=1}^W \sum_{v=1}^H f(\text{CLDTP}^{(2)}(u, v), k_2),$$

$$f(x, y) = \begin{cases} 1, & x = y \\ 0, & x \neq y, \end{cases}$$

$$k_1 \in K_1, k_2 \in K_2,$$

where the size of local image region is  $W \times H$ ,  $K_1$  is the maximal value of  $\text{CLDTP}^{(1)}(u, v)$ , and  $K_2$  is the maximal value of  $\text{CLDTP}^{(2)}(u, v)$ . The final CLDTP histogram can be obtained by connecting the histogram  $H_1(k_1)$  and the histogram  $H_2(k_2)$ , and the dimension of the CLDTP histogram is 20 (16 + 4).

**3.2. The Construction of CLDTP Descriptor.** In this section, the construction of CLDTP descriptor is presented. At first, the local feature regions are detected for calculating descriptors. In this paper, we use the Hessian-Affine detector to obtain the affine invariant region [14]. Then the detected regions are normalized. As shown in Figure 4, the elliptic region is rotated in order that the long axis of the ellipse is

the  $u$ -axis of the image coordinate, and it is mapped to the circular region. All the detected regions are normalized to the circular region with the same size to obtain scale and affine invariant [5]. In order to integrate the spatial structural information of the local image into the descriptor, we divide the normalized region into 16 ( $4 \times 4$ ) subregions using the grid division method of the SIFT descriptor. For each subregion, compute the CLDTP code of each pixel, respectively, and construct the corresponding histogram. Then connect the histograms together, and we can get a 320-dimensional ( $(16 \times 20)$ -dimensional) feature vector.

## 4. Experimental Results

**4.1. Image Matching.** In the image matching experiments, the Mikolajczyk dataset [15] is used to evaluate the performance of the SIFT, LDTP, WOS-LTP, and CLDTP descriptor. This dataset includes eight types of scene images with different illumination and geometric distortion transformations. This dataset has the ground-truth matches through estimated homography. As shown in Figure 5, we randomly select one image pair in each category from the dataset. Figures 5(a) and 5(b) show the image pairs with blur transformation, Figures 5(c) and 5(d) show the image pairs with viewpoint change, Figures 5(e) and 5(f) show the image pairs with scale and rotation transform, Figure 5(g) shows the image pair with lighting change, and Figure 5(h) shows the image pair with JPEG compression transformation.

In the experiments, the affine invariant regions are firstly detected using the Hessian-Affine detector. Then the detected regions are normalized to the circular regions and the gray

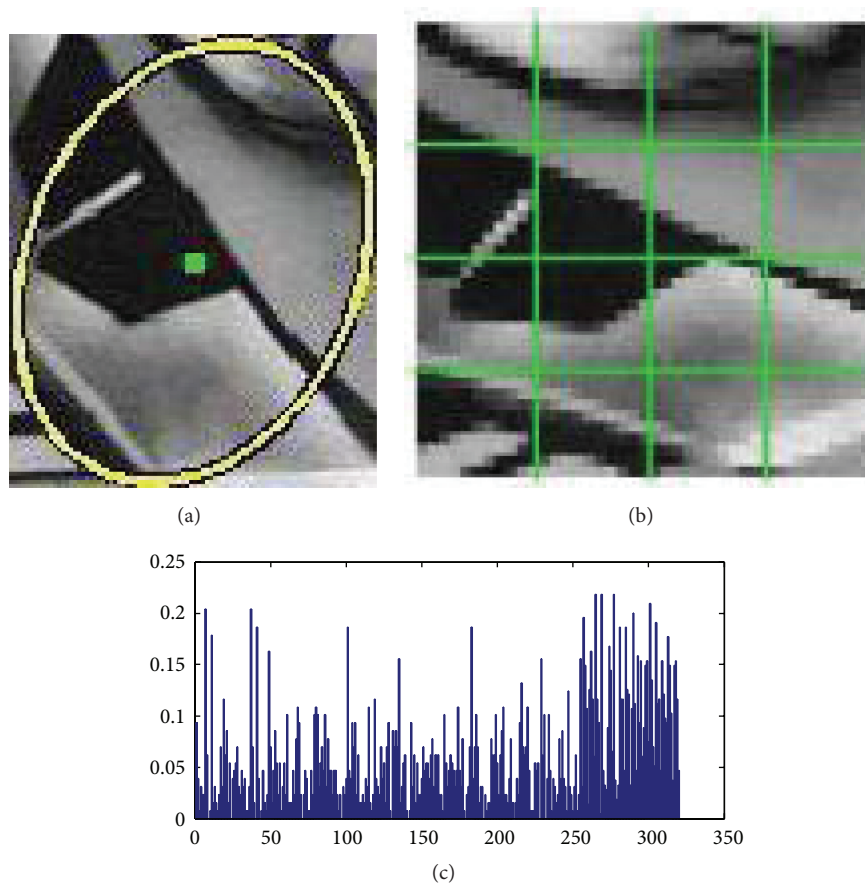


FIGURE 4: The construction of the CLDTP descriptor. (a) The detected affine invariant region. (b) The normalized region with grids. (c) The histogram of the CLDTP descriptor.

values of the regions are transformed to lie between 0 and 1. The descriptors are constructed based on the normalized regions. Finally, the nearest neighbor distance ratio (NNDR) matching algorithm is performed using the Euclidean distance as similarity measure [2, 5]. In our experiments, the size of the normalized image region is  $41 \times 41$  and set the parameter  $\varepsilon = 0.06$ . The parameter setting of the SIFT descriptor, the LDTP descriptor, and WOS-LTP descriptor is the same as the original proposed paper [2, 11, 13].

We use the Recall-Precision criterion to evaluate the matching results, which is similar to the criterion used in [5]. It is based on the number of the correct matches and the number of the false matches between a pair of images. The number of correct matches is determined by the overlap error [15]. Two regions are matched if the distance between their descriptors is below a threshold  $t$ . The Recall-Precision curve can be obtained by changing the distance threshold  $t$ . That is to say, different points on the curve correspond to the Recall-Precision results of an image pair under different values of  $t$ . A perfect descriptor would give a recall equal to 1 for any precision. More detailed description of the Recall-Precision criterion can be found in [3, 5].

The image matching results of the testing images are shown in Figure 6, and the corresponding images are displayed in Figure 5. Figures 6(a) and 6(b) show the results for

blur changes. Figure 6(a) is the results for the structured scene and Figure 6(b) for the textured scene. The results show that blur changes have largest influences on the performance of SIFT descriptor. For the structured scene, the performance of the LDTP, WOS-LTP, and CLDTP descriptors is similar. For the structured scene, the WOS-LTP and CLDTP descriptors perform better than other descriptors. Figures 6(c) and 6(d) show the performance of descriptors for viewpoint changes. Figure 6(c) is the results for the structured scene and Figure 6(d) for the textured scene. We can observe that the results for the structured scene are less influenced by viewpoint changes. The SIFT descriptor obtained the lowest score and the CLDTP descriptor performs better than other descriptors. Figures 6(e) and 6(f) show the results to evaluate the descriptors for combined image rotation and scale changes. We can see that the CLDTP descriptor gives the best results when  $1 - \text{precision}$  is larger. The ranking for LDTP descriptor and WOS-LTP descriptor is similar, and the SIFT descriptor obtains the worst matching score. Figure 6(g) shows the results for illumination changes. We can observe that the CLDTP descriptor obtains the best matching score, and the SIFT descriptor obtains worse results than the other three descriptors. Figure 6(h) shows the results to evaluate the influence of JPEG compression. From Figure 6(h) we can see that the four kinds of descriptors perform similarly,

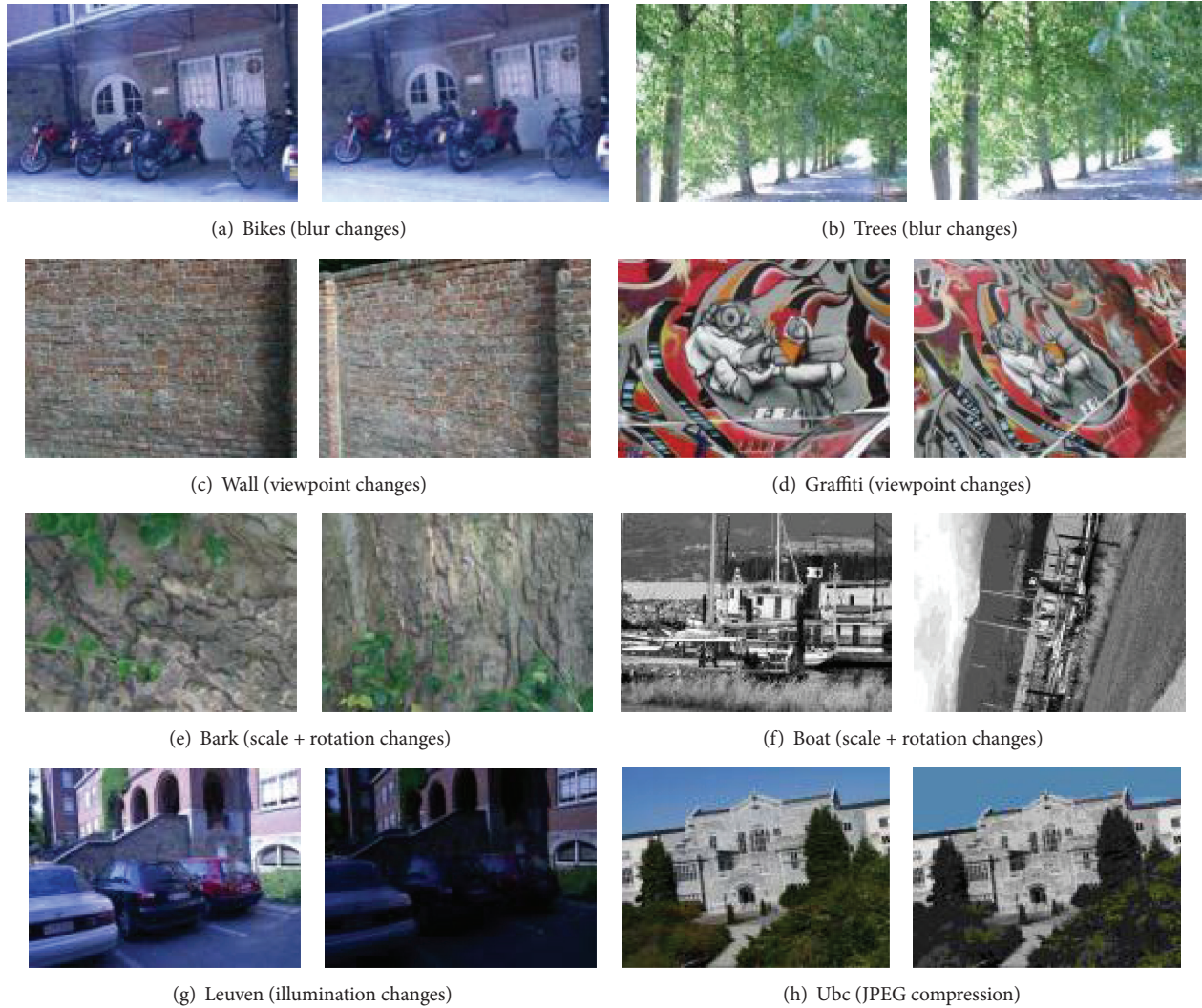


FIGURE 5: Testing image pairs.

and their performance is better than other cases. Based on the above analysis, we can conclude that the CLDTP descriptor performs much better than the well-known state-of-the-art SIFT descriptor especially under blur, viewpoint, and illumination changes, and it performs better than the LDTP descriptor and WOS-LTP descriptor especially for the textured scene. It is worth noting that the dimension of the CLDTP descriptor is much lower than the dimension of the LDTP descriptor.

**4.2. Object Recognition.** In this paper, the SIMPLcity dataset [16] and the Caltech-256 dataset [17] are used to conduct object recognition experiments for further evaluating the performance of our proposed descriptor. The SIMPLcity dataset is a subset of COREL image database, and it contains 10 different categories: African people, beach, building, bus, elephant, flower, food, horse, dinosaur, and mountain. Each category has 100 images. In the object recognition experiments, 50 images are randomly selected for training and the other 50 images for test. Some example images

are shown in Figure 7. The Caltech-256 dataset contains 29780 images falling into 256 categories with much higher intra-class variability and higher object location variability. Each category has a minimum of 80 images. We conduct the recognition algorithm on 30 and 45 training images per category, respectively, and the rest images of each category are used for test. Some example images are shown in Figure 8.

The steps of the object recognition are listed as follows. At first, the Hessian-Affine invariant regions are detected and normalized. Then, for each detected region, the descriptor is built. Each image can be represented by a set of descriptors. Finally the Sparse Coding Spatial Pyramid Matching (ScSPM) approach and linear Support Vector Machine (SVM) are used for object classification [18]. In this paper, we use the SIFT, LDTP, WOS-LTP, and CLDTP descriptors, respectively, to perform object recognition experiments. The parameter settings of the descriptors are the same as the image matching experiments. The codebook size of ScSPM is 256 and the recognition accuracy is used for evaluation. We repeat the experiments 10 times with different random

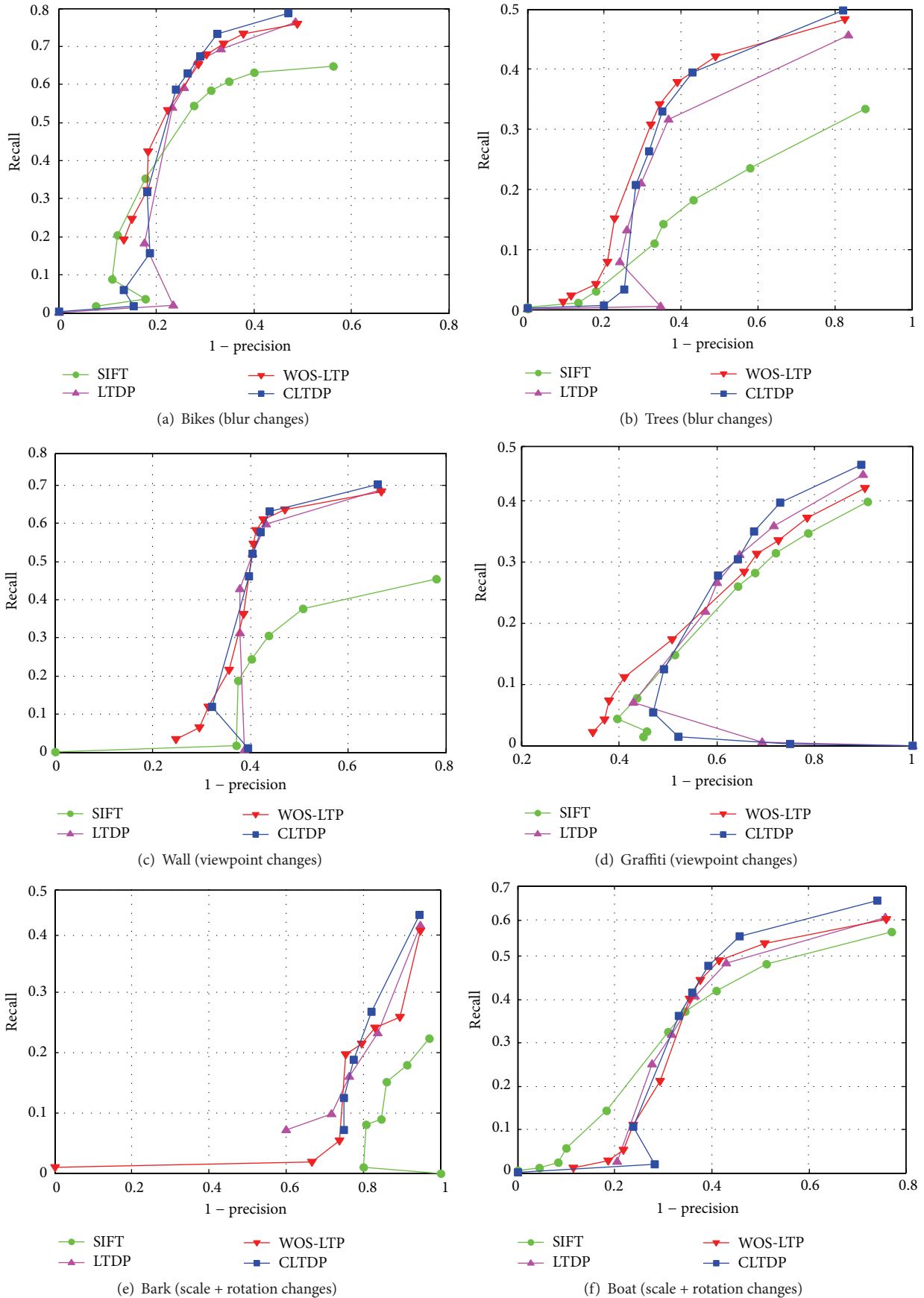


FIGURE 6: Continued.

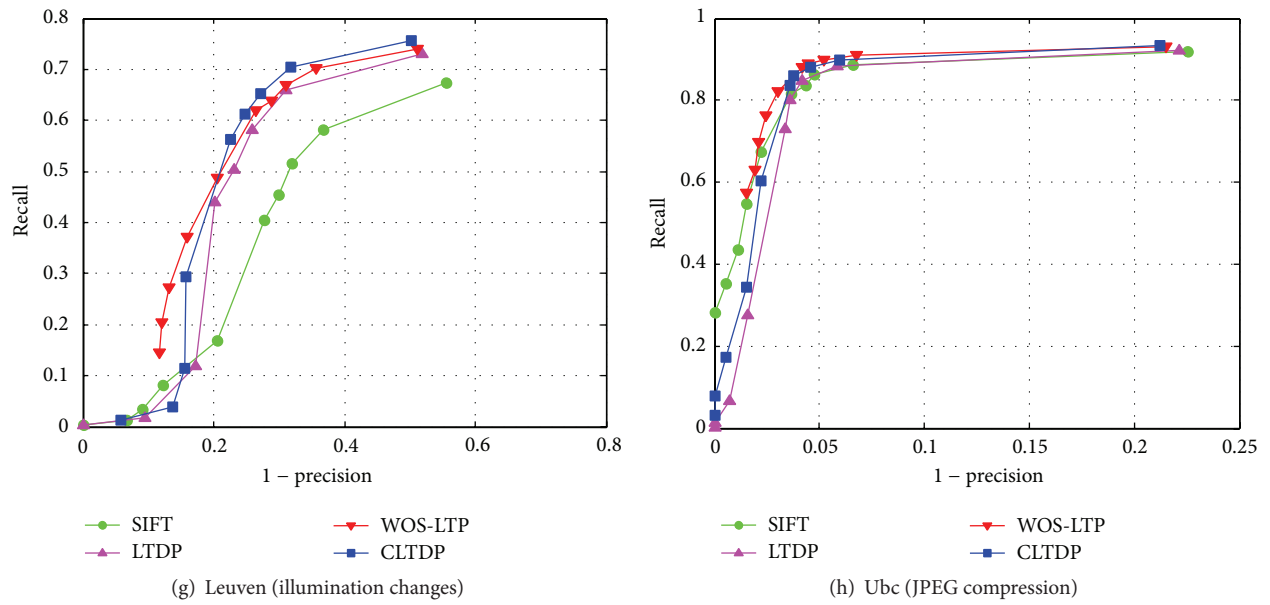


FIGURE 6: The matching results of the testing image pairs.

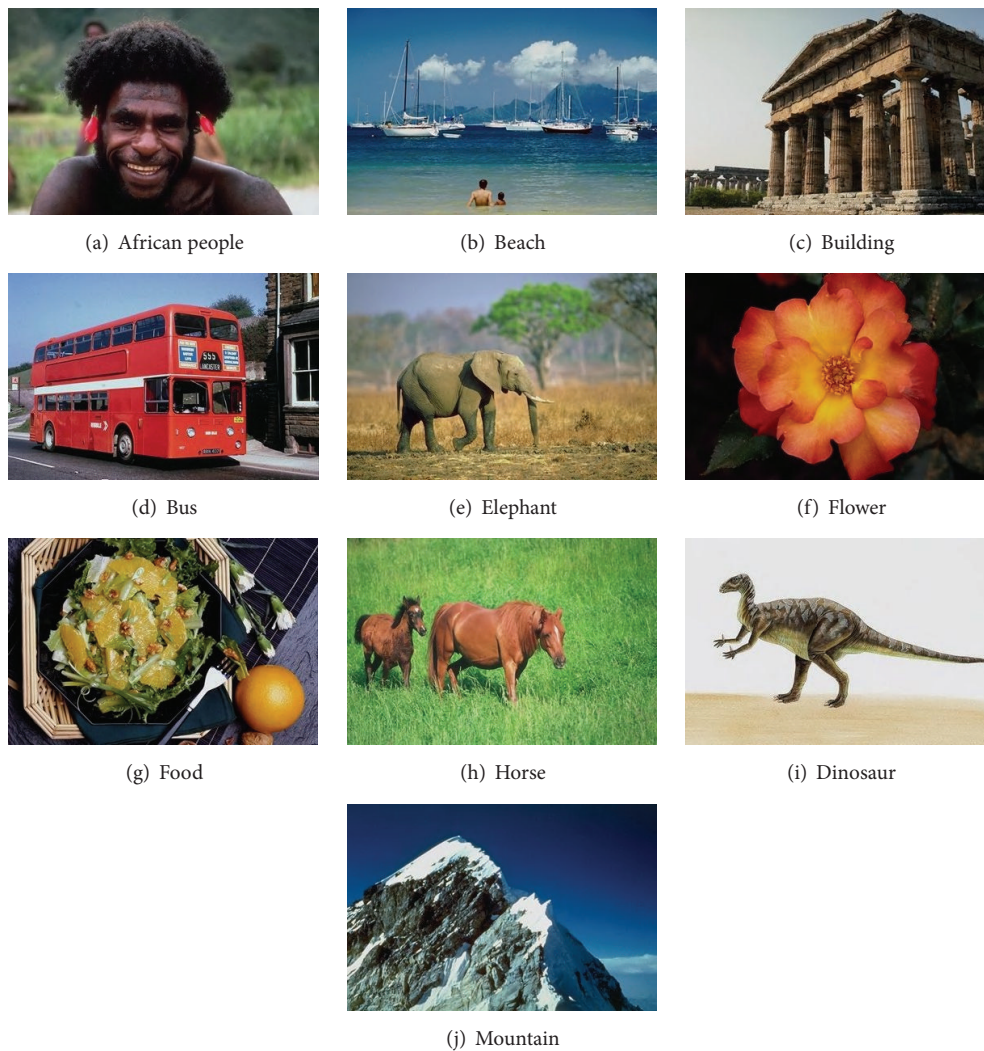


FIGURE 7: Example images of the SIMPLIcity dataset.



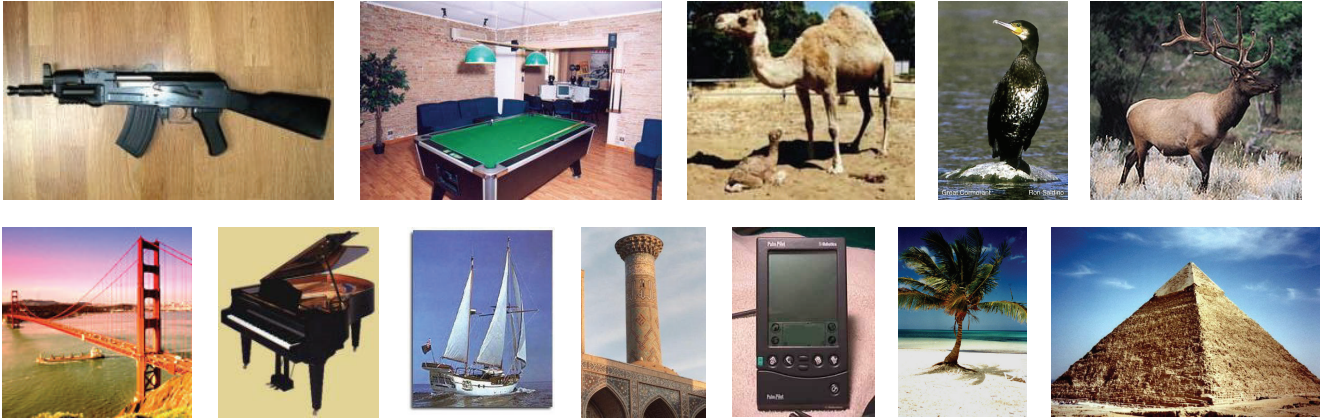


FIGURE 8: Example images of the Caltech-256 dataset.

TABLE 1: Object recognition results on the SIMPLIcity dataset.

Accuracy (%)	SIFT	LDTP	WOS-LTP	CLDTP
People	72.0	66.0	74.0	68.0
Beach	72.0	56.0	68.0	70.0
Building	68.0	80.0	66.0	84.0
Bus	96.0	100.0	98.0	100.0
Elephant	88.0	92.0	94.0	96.0
Flower	88.0	100.0	98.0	94.0
Food	68.0	78.0	66.0	86.0
Horse	92.0	98.0	94.0	100.0
Dinosaur	100.0	98.0	100.0	100.0
Mountain	80.0	74.0	80.0	74.0
Global	<b>82.4</b>	<b>84.2</b>	<b>83.8</b>	<b>87.2</b>

TABLE 2: Object recognition results on the Caltech-256 dataset.

Accuracy (%)	SIFT	LDTP	WOS-LTP	CLDTP
30 training images	34.1	35.8	35.2	37.2
45 training images	36.9	38.1	37.5	40.2

selected training and testing images. In this paper, the recognition accuracy is the ratio of the number of correctly classified test images and the number of test images. Tables 1 and 2 give the object recognition results on the SIMPLIcity dataset and the Caltech-256 dataset, respectively. From Table 1 we can see that, for the category of building, elephant, food, and horse, the CLDTP descriptor gives the best results. The global recognition accuracy of the proposed CLDTP descriptor is 87.2%, which is higher than that of other descriptors. From Table 2 we can observe that the CLDTP descriptor outperforms the SIFT descriptor by more than 3 percent and outperforms the LDTP and WOS-LTP descriptors by about 2 percent.

From both the image matching experimental results and the object recognition experimental results we can see that the CLDTP descriptor performs better than the LDTP descriptor. Although the dimension of the LDTP descriptor is higher than that of the CLDTP descriptor, the LDTP

descriptor cannot encode more information. This is because of the fact that both descriptors contain the gradient orientation information and the gradient magnitude information, and the difference between them is the encoding method. So the CLDTP descriptor is more effective than the LDTP descriptor.

## 5. Conclusions

This paper presents a novel CLDTP operator based image local feature descriptor construction method. The CLDTP descriptor combines the advantages of the SIFT descriptor and the LDTP descriptor. On this basis, the histogram of the first directional number and the histogram of the encoded difference are connected to compute the descriptor. The constructed CLDTP descriptor not only contains the gradient orientation information and the gradient magnitude information, but also contains the spatial structural information of the local image. Furthermore, the dimension of the CLDTP descriptor is much lower than LDTP descriptor. Our experimental results show that the CLDTP descriptor performs better than the other three descriptors. So the CLDTP descriptor is effective for local image description and it is more robust to image geometric distortions. In the future work, we will add the color invariant information into the descriptor to construct more robust and discriminative descriptors.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgments

This paper is supported by the National Natural Science Foundation of China (Grant no. 61375010, no. 61175059, and no. 61170116), Beijing Higher Education Young Elite Teacher Project (Grant no. YETP0375), and the Fundamental Research Funds for the Central Universities under Grant no. FRF-TP-14-120A2.

## References

- [1] X. Yang and K.-T. T. Cheng, "Local difference binary for ultrafast and distinctive feature description," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 1, pp. 188–194, 2014.
- [2] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.
- [3] Y. Ke and R. Sukthankar, "PCA-SIFT: a more distinctive representation for local image descriptors," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '04)*, pp. II506–II513, July 2004.
- [4] H. Bay, T. Tuytelaars, and L. Van Gool, "SURF: speeded up robust features," in *Proceedings of the 9th European Conference on Computer Vision*, vol. 1, pp. 404–417, May 2006.
- [5] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 10, pp. 1615–1630, 2005.
- [6] S. Lazebnik, C. Schmid, and J. Ponce, "A sparse texture representation using local affine regions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 8, pp. 1265–1278, 2005.
- [7] T. Ojala, M. Pietikäinen, and T. Mäenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971–987, 2002.
- [8] M. Heikkilä, M. Pietikäinen, and C. Schmid, "Description of interest regions with local binary patterns," *Pattern Recognition*, vol. 42, no. 3, pp. 425–436, 2009.
- [9] X. Tan and B. Triggs, "Enhanced local texture feature sets for face recognition under difficult lighting conditions," *IEEE Transactions on Image Processing*, vol. 19, no. 6, pp. 1635–1650, 2010.
- [10] R. Gupta, H. Patil, and A. Mittal, "Robust order-based methods for feature description," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '10)*, pp. 334–341, June 2010.
- [11] M. Huang, Z. Mu, H. Zeng, and S. Huang, "Local image region description using orthogonal symmetric local ternary pattern," *Pattern Recognition Letters*, vol. 54, pp. 56–62, 2015.
- [12] T. Jabid, M. H. Kabir, and O. Chae, "Facial expression recognition using local directional pattern (LDP)," in *Proceedings of the 17th IEEE International Conference on Image Processing (ICIP '10)*, pp. 1605–1608, September 2010.
- [13] A. R. Rivera, J. R. Castillo, and O. Chae, "Local directional texturepattern image descriptor," *Pattern Recognition Letters*, vol. 51, no. 1, pp. 94–100, 2015.
- [14] K. Mikolajczyk and C. Schmid, "Scale & affine invariant interest point detectors," *International Journal of Computer Vision*, vol. 60, no. 1, pp. 63–86, 2004.
- [15] K. Mikolajczyk, T. Tuytelaars, C. Schmid et al., "A comparison of affine region detectors," *International Journal of Computer Vision*, vol. 65, no. 1-2, pp. 43–72, 2005.
- [16] J. Z. Wang, J. Li, and G. Wiederhold, "SIMPLcity: semantics-sensitive integrated Matching for Picture Libraries," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 9, pp. 947–963, 2001.
- [17] G. Griffin, A. D. Holub, and P. Perona, "The Caltech 256," Caltech Technical Report, 2007.
- [18] J. Yang, K. Yu, Y. Gong, and T. Huang, "Linear spatial pyramid matching using sparse coding for image classification," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPR '09)*, pp. 1794–1801, June 2009.



**Hindawi**

Submit your manuscripts at  
<http://www.hindawi.com>

