ICIC Express Letters ICIC International ©2018 ISSN 1881-803X
Volume **12**, Number **7**, July **2018** pp. **637–644**

HUMAN DETECTION WITH LOG-POLAR TRANSFORM AND HOG-LBP FEATURES

Wanli Peng, Xiaoyan Yu and Xianwei Rong*

School of Physics and Electronic Engineering
Harbin Normal University
No. 1, Shida Road, Linmin Economic Development Zone, Harbin 150025, P. R. China
*Corresponding author: hunanpwl@163.com

Received December 2017; accepted March 2018

Abstract. *Human detection is a critical and active research topic in computer vision. At present, the detection performance is not satisfactory due to the fact that the detection performance is usually affected by rotation variance. This paper presents a rotation invariant detection algorithm based on LPT (Log Polar Transform) and HOG-LBP (Histograms of Oriented Gradients-Local Binary Patterns) features. The proposed algorithm consists of three main modules: the first aims to convert an image from Cartesian coordinate system into log-polar coordinates with LPT; the second takes responsibility for extracting HOG features and LBP features which will be fed to train two different SVM (Support Vector Machine) classifiers, respectively; the third realizes the human detection by using two different SVM classifiers. Experimental results on the INRIA dataset show the robustness and performance of our algorithm.*
Keywords: Human detection, Log-polar transform, Rotation invariance, HOG-LBP

1. **Introduction.** Human detection is an initial step for human segmentation, which can be applied in many fields such as intelligent transportation, UAV (Unmanned Air Vehicle) and mobile devices [1-3]. Over past two decades, many advanced works with the help of human detection approaches have been proposed and they can be divided into two categories, local-gradient-based approaches and local-texture-based methods. In these methods, local gradient or local texture features would be extracted frequently to realize the human detection.

Various local gradient based detection algorithms are presented, especially after Dalal and Triggs proposed Histograms of Oriented Gradients (HOG) to detect human [4]. For example, Lu and Little utilize the Principal Components Analysis (PCA) to reduce the dimension of the HOG features [5], and Lowe proposes Scale Invariant Feature Transform (SIFT) to address the scale transformation [6]. Also, with the help of SIFT, Bay et al. put forward Speeded Up Robust Features (SURF) to decrease the computational complexity of SIFT descriptor [7]. Meanwhile, many local texture based detection methods are introduced, such as Local Binary Pattern (LBP), circle LBP and uniform LBP [19,20]. However, the resultant detection performance of these algorithms is affected by the rotation variation, so how to tackle this problem is of great value in practice.

To address this issue, there is much literature focusing on keeping rotation of HOG invariant concerning the use of HOG descriptor. For example, Vashaee et al. take advantage of top-down searching technique and variable block configuration to tackle the rotation variance and scale variance, respectively [8]. Liu et al. combine HOG with Fourier analysis in polar and spherical coordinates [9]. Liu et al. propose a new feature Sector-Ring HOG (SRHOG) to realize the rotation-invariant human detection for disaster rescue [10]. At present, the rotation invariant LBP descriptors are proposed, such as $LBP^{ri}_{R,P}$ and

DOI: 10.24507/icicel.12.07.637

$LBP^{riu2}_{R,P}$, where superscript $riu2$ reflects the use of rotation invariant 'uniform' LBP descriptor, P controls the quantization of the angular space, and R is the radius of the circle LBP descriptor [20]. Even though these improved algorithms resolve the rotation variance efficiently, the detection accuracy is not enough. Thus, some literature combines LBP descriptor or its variants with an HOG descriptor to enhance the performance of HOG [11,12].

As for Log-Polar Transform (LPT), this transform coupled with a mapping template was also researched [13,14], in order to settle the rotation variation. However, suffering from the weakness arisen from data compression, LPT is often applied in image registration and authentication, while it is rarely used in human recognition and detection [15].

In this paper, we propose a novel human detection algorithm by using of LPT and HOG-LBP features. The proposed algorithm shows several characteristics different from existing methods, which are summarized as follows. (1) To solve the rotation variance of HOG descriptor, we convert an image from Cartesian coordinate system to log-polar coordinates with LPT algorithm. (2) To cope with the data compression of LPT, the HOG-LBP features combination is utilized in the proposed algorithm, where the LBP feature is extracted from the rotated image rather than the transformed image. (3) Since the dimension of HOG feature vectors is much greater than that of LBP feature vectors, we train two different SVM classifiers using HOG and LBP feature vectors, respectively. The effectiveness of the proposed algorithm is verified by our experimental results on the general human detection dataset INRIA. Our proposal achieves better results in terms of Recall, Precision, F_1-score compared with other algorithms.

The rest of this paper is organized as follows. Section 2 provides the overview of the proposed algorithm, and then briefly describes the HOG descriptor and several LBP variant descriptors, followed by the elaboration about how to integrate two different SVM classifiers for human detection. A set of experiments and result analysis are illustrated in Section 3. Finally, Section 4 concludes this write-up.

2. **Proposed Algorithm.** The proposed algorithm makes use of rotation invariance of LPT to avoid the side effect derived from the rotation variance of HOG descriptor in human detection, which can be partitioned into three parts as shown in Figure 1. First is coordinate system transformation and HOG feature extraction. An original image is converted to log-polar coordinates from Cartesian coordinates by LPT, and then HOG features are extracted from the transformed image so as to train SVM classifier 1. Second is LBP feature extraction. The features such as $LBP^{riu2}_{R,P}$ and $LBP^{ri}_{R,P}$ illustrated in Figure 1, are extracted from the rotated image using the rotation invariant LBP descriptor to

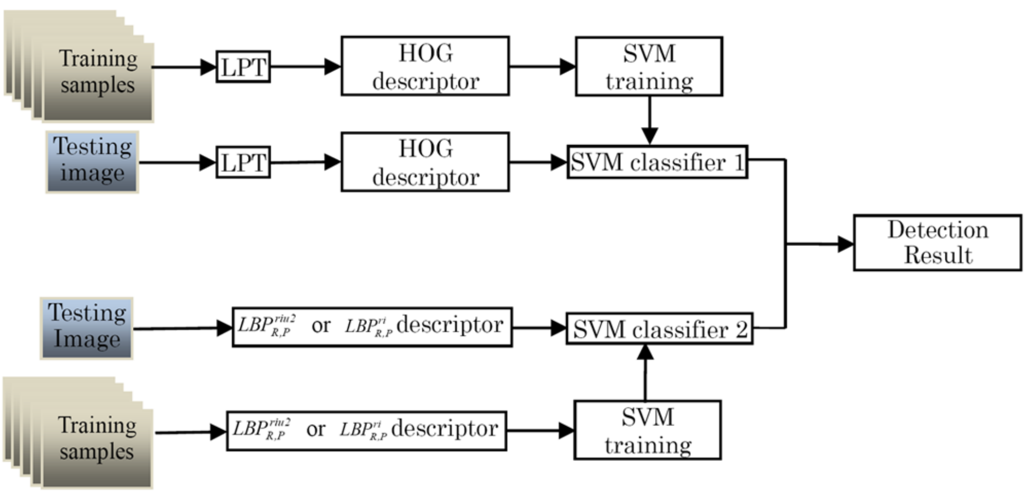

Figure 1. Block diagram of the proposed human detection algorithm

train SVM classifier 2. Third is human detection implementation. The human detection is implemented by combining the results from SVM classifier 1 and SVM classifier 2 to improve the detection accuracy.

2.1. **Log-polar transform.** LPT was used to realize transformation between Cartesian coordinates and log-polar coordinates due to its rotation and scale invariance. In mathematics, such coordinate conversion is given by

$$\begin{cases} x = e^{\rho}\cos\theta \\ y = e^{\rho}\sin\theta \end{cases} \qquad \begin{cases} \rho = \ln\sqrt{x^2+y^2} \\ \theta = \arctan y/x \end{cases}$$

where x and y are the Cartesian coordinates of a point, while ρ and θ are its log-polar coordinates. ρ is the logarithm of the Euclidean distance between a point (ρ, θ) and the origin, and θ is the angle between the x-axis and the straight line through the origin to a point (ρ, θ). It can be derived that the maximum of ρ is limited by the minimum of x and y.

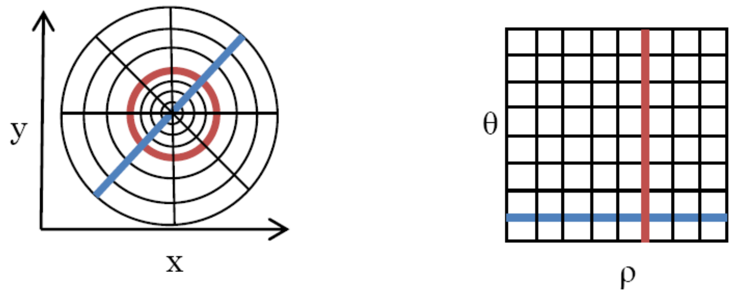

FIGURE 2. Approximate mapping from Cartesian coordinates to log-polar coordinates

2.2. **HOG and LBP features.** HOG descriptor is first introduced by Dalal and Triggs to detect pedestrian [4]. The process for extracting HOG features from transformed image can be divided into four steps: first, the color and gamma are normalized by an HOG descriptor in order to reduce the side effects arisen from local shadow and illumination; second, the detection windows are set and divided into several blocks with 2×2 cells consisting of 8×8 pixels; third, these gradients of nine orientation angles are calculated and the histogram of gradients for each block is accumulated by using a voted way; finally, the rate of block overlap is set as 0.5 to gather the HOG feature extracted from all blocks. Nevertheless, HOG features cannot adequately deal with the scale and rotation variation despite of their positive performance in human detection. Fortunately, scale variation has been tackled using a Pyramid-HOG (PHOG) descriptor [16,17]. To handle the rotation variation, we apply the LPT algorithm to realizing coordinate conversion.

In this work, an LBP descriptor is used to facilitate the performance of LPT-HOG. LBP initially proposed in [18,19] is a general texture descriptor. The LBP value can be obtained in a 3×3 pixel block by comparing the neighbor pixels with the center pixel. Suppose that one neighbor pixel value is greater than the center one, and then this pixel position is marked as 1, otherwise 0. Thus, the neighbors of the center pixel in the 3×3 neighborhood can produce an 8-bit binary number (usually converted into a decimal number regarded as LBP code, a total of 256). The value is regarded as the LBP value of the center pixel, which represents the local texture information of an image.

In order to improve the performance on object detection, several LBP variants are developed to improve the performance on object detection [12], such as circle LBP and rotation invariant LBP as well as uniform LBP [19,20].

Details on the procedure of feature extraction using LPT-HOG-LBP algorithm are presented in Algorithm 1. An example in which LPT is used in the extraction of rotation-invariant HOG features from images is shown in Figure 3. The HOG features extracted from a transformed image suffer from data compression, which has side effect on the performance of the human detection. To deal with data compression, this paper enhances the HOG descriptor by combining it with the rotation invariant uniform LBP descriptor ($LBP^{riu2}_{R,P}$) or rotation invariant LBP descriptor ($LBP^{ri}_{R,P}$). The histograms of LBP descriptor and its variants are illustrated in Figure 4.

Algorithm 1 Feature extraction of LPT-HOG-LBP algorithm

Input: $\mathrm{P_{train}}$: path of the training set
mapping: a structure containing a mapping table for LBP codes
cellsize: the size of cells at HOG descriptor
blocksize: the size of blocks at HOG descriptor
P: the quantization of the angular space at LBP descriptor
R: the radius of the LBP descriptor
for i = 0: the number of training samples
  I = imread(sample(i))
  $\mathrm{H_{lbp}}$ = LBP(I, R, P, mapping)
$\mathrm{F_{lbp}}$ = [$\mathrm{F_{lbp}}$; $\mathrm{H_{lbp}}$];
  $\mathrm{I_{log}}$ = Log-polar(I);
  $\mathrm{H_{hog}}$ = extractHOGFeatures($\mathrm{I_{log}}$, cellsize, blocksize);
$\mathrm{F_{hog}}$ = [$\mathrm{F_{hog}}$; $\mathrm{H_{hog}}$];
end
Output: $\mathrm{F_{lbp}}$: the feature of LBP and its variants descriptors
$\mathrm{F_{hog}}$: the feature of HOG descriptor

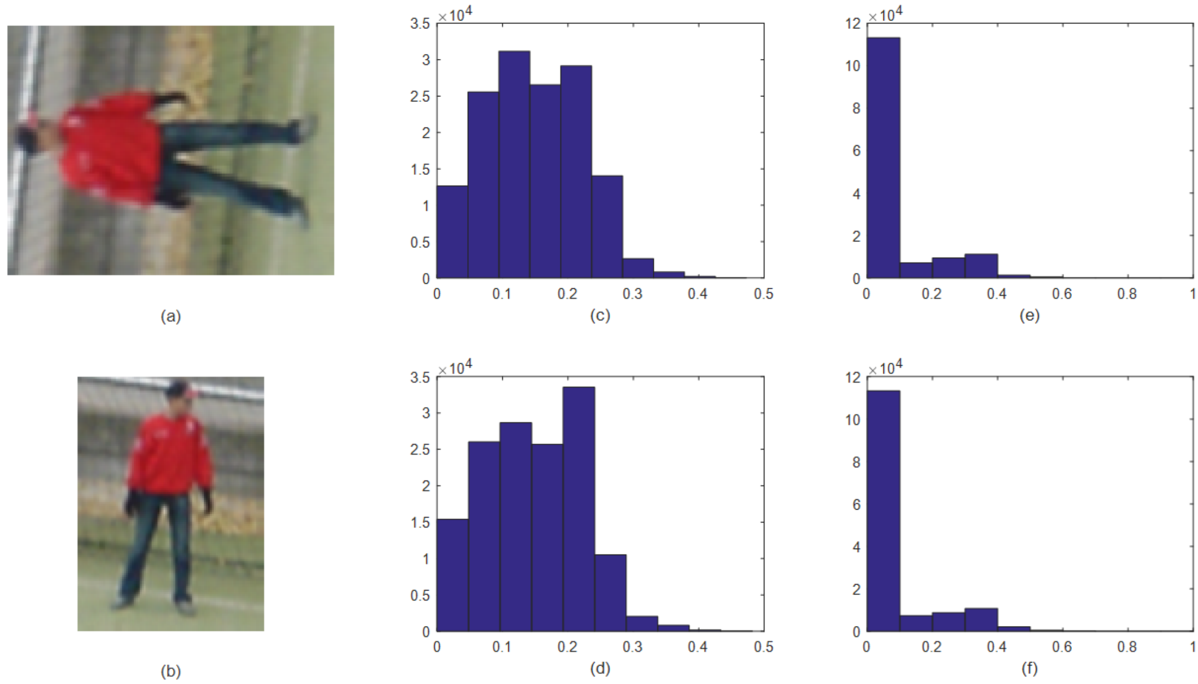

FIGURE 3. The example of the LPT utilized in human images to obtain rotation invariant HOG features which suffer from data compression: (a) original image, (b) rotated image, (c) HOG features histogram of original image, (d) HOG features histogram of rotated image, (e) LPT-HOG features histogram of the original image, and (f) LPT-HOG feature histogram of the rotated image

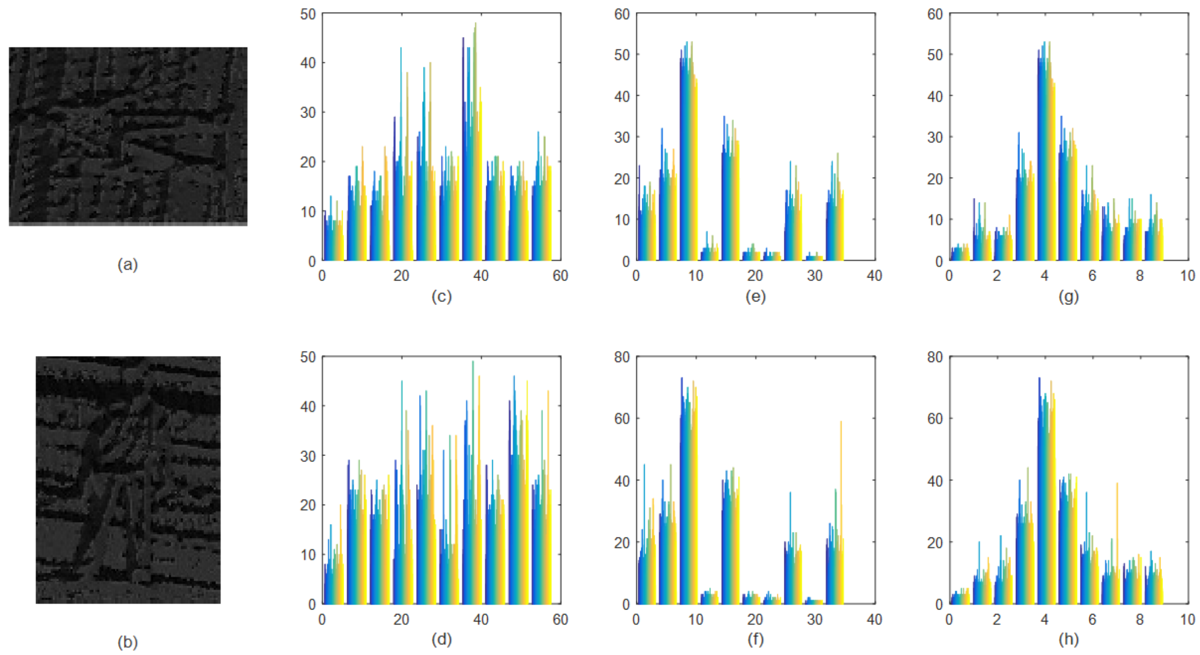

FIGURE 4. The histograms of LBP descriptor and its variants: (a) original image, (b) rotated image, (c) the features histogram of uniform LBP descriptor ($LBP^{u2}_{R,P}$) at original image, (d) the features histogram of uniform LBP descriptor ($LBP^{u2}_{R,P}$) at rotated image, (e) the features histogrom of rotation invariant uniform LBP descriptor ($LBP^{riu2}_{P,R}$) at original image, (f) the features histogrom of rotation invariant uniform LBP descriptor ($LBP^{riu2}_{P,R}$) at rotated image, (g) the features histogrom of rotation invariant LBP descriptor ($LBP^{ri}_{R,P}$) at original image, and (h) the features histogrom of rotation invariant LBP descriptor ($LBP^{ri}_{R,P}$) at rotated image

2.3. **SVM classifier.** In the experiment, the dimension of HOG feature vectors is 5670 (INRIA dataset). The dimension of LBP feature vectors is 36 ($LBP^{ri}_{1,8}$) or 10 ($LBP^{riu2}_{1,8}$), which indicates that the dimension of HOG feature vectors is much larger than LBP feature vectors. If these two feature vectors are connected directly to act as combined features vector, the LBP features may be ignored in the final results. Thus, to address this issue, we propose an SVM classifier by coupling with two different SVM classifiers, called SVM classifiers 1 and 2, trained by HOG feature vectors and LBP feature vectors, respectively. More specifically, the difference between two classifiers is feature vectors as input to two classifiers for classification, namely, HOG feature vectors are fed into classifier 1, while LBP feature vectors are fed into classifier 2. Moreover, both classifier 1 and classifier 2 are implemented by the same support vector machine model – LibSVM [21]. Finally, the final detection result is a combination of results from SVM classifier 1 and SVM classifier 2.

3. **Experimental Results.** Initially, this subsection measures the performance on standard INRIA dataset which is a popular dataset in the human detection. Afterwards, we take the 10-fold cross validation to distribute the training set and testing set. In the experiment, the training-set includes 2160 positive images and 2700 negative images, and then the testing-set contains 240 positive images and 300 negative images. Finally, we evaluate the performance of our proposed method with Recall, Precision and F_1-score, where:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$\text{F}_1\text{-score} = \frac{2 \times \text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}}$$

where TP (true positives) and FP (false positives) are the number of correctly and falsely detected humans respectively, while FN (false negatives) is the number of falsely detected non-humans. Our experiments comprise two parts: first, we compare the detection performance of our proposed algorithm with two LBP variants – $LBP^{ri}_{1,8}$, $LBP^{riu2}_{1,8}$ and two combined algorithms – HOG-$LBP^{ri}_{1,8}$, HOG-$LBP^{riu2}_{1,8}$ at INRIA dataset. Table 1 shows the performance of tested algorithms on human detection tasks. The Recall, Precision, and F_1-score and their corresponding deviations are illustrated in Figure 5, in which the deviations indicate their robustness indicators for the samples of the INRIA dataset. It can be derived from Table 1 that the best performance obtained by our proposed algorithm is LPT-HOG-$LBP^{ri}_{1,8}$. These results suggest that the LPT-HOG-$LBP^{ri}_{1,8}$ algorithm achieves better performance than other tested descriptors, since the proposed method addresses rotation-variant HOG features by using LPT. Moreover, the dimension of $LBP^{ri}_{1,8}$ is greater than $LBP^{riu2}_{1,8}$. The results on the INRIA dataset listed in Table 1 indicate that our proposal achieves better Recall, Precision and F_1-score, meaning that it is more robust than other tested methods. On the other hand, our proposed algorithm endures more computational load for feature extraction, due to the fact that LPT was used to facilitate the rotation invariance of the original HOG descriptor. In practice, as observed from Figure 6, the LPT-HOG-$LBP^{riu2}_{1,8}$ algorithm has the similar computational complexity with

Table 1. Recall, Precision and F_1-score for the tested algorithms with INRIA dataset (Best values show in bold)

	$LBP^{ri}_{1,8}$	$LBP^{riu2}_{1,8}$	HOG -$LBP^{ri}_{1,8}$	HOG -$LBP^{riu2}_{1,8}$	LPT-HOG -$LBP^{ri}_{1,8}$	LPT-HOG -$LBP^{riu2}_{1,8}$
Recall	0.92	0.89	0.95	0.93	**0.96**	**0.94**
Precision	0.45	0.44	0.49	0.49	**0.78**	**0.78**
F_1-score	0.60	0.59	0.65	0.64	**0.86**	**0.85**

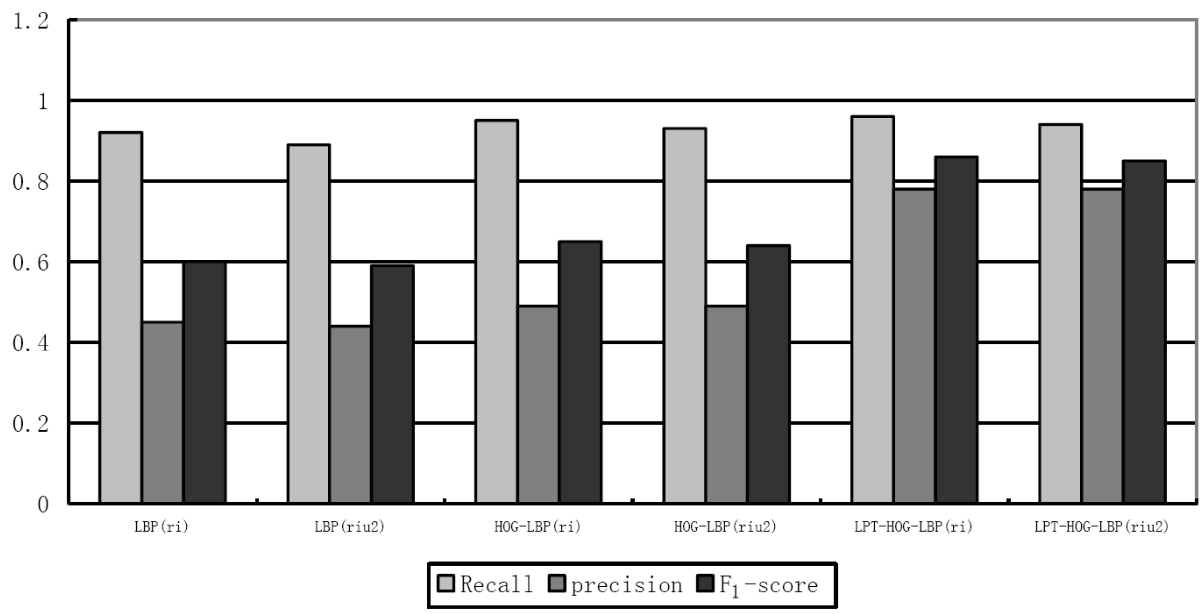

Figure 5. Values of Recall, Precision and F_1-score and their standard deviations for the tested algorithms with the INRIA dataset

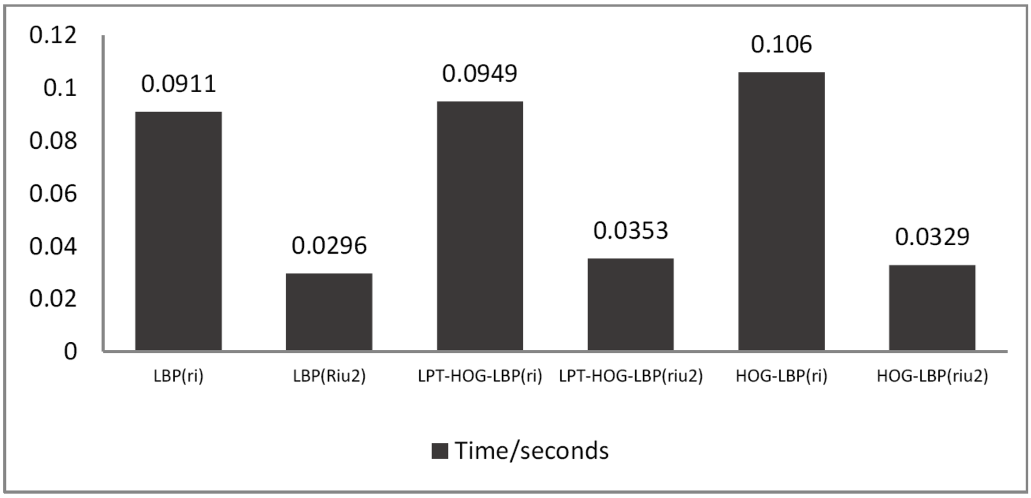

Figure 6. Feature extraction time (s) analysis for tested algorithms in rotated human detection

HOG-$LBP_{1,8}^{riu2}$, so do the LPT-HOG-$LBP_{1,8}^{ri}$ and HOG-$LBP_{1,8}^{ri}$ for human detection tasks in that the LBP variants are combined into the proposed algorithm. It is clear in Figure 6 that LPT-HOG-$LBP_{1,8}^{ri}$ endures computational complexity more than others even though it obtains the best performance among the tested algorithms. Compared to the LPT-HOG-$LBP_{1,8}^{ri}$, LPT-HOG-$LBP_{1,8}^{riu2}$ can achieve better balance between the computation cost and detection performance since its computation time is reduced significantly while obtaining the similar detection performance.

4. **Conclusion.** In this paper, a novel algorithm for the rotation invariant human detection (LPT-HOG-LBP) is developed by integrating LPT with HOG-LBP. Our method makes use of LPT and HOG-LBP features to feed the same SVM model – Lib-SVM, we first adopted LPT to address the rotation variance of the HOG features in a rotated testing set. Furthermore, the rotation invariant LBP descriptors are employed to deal with the data compression arisen from a log polar transform. Finally, experimental results demonstrate the proposed algorithm achieves better performance than single descriptors and popular combined descriptors in the INRIA dataset. Despite the lack of obvious advantages in speed, LPT-HOG-LBP is significantly superior in the detection performance. Other applications of the proposed algorithm maybe address the issue of rotation variance for object detection, image matching and object tracking. In addition, how to cope with other types of affine transformations in the human detection would be another important future work.

Acknowledgement. This work was supported by the National Natural Science Foundation of China under Grant 61401127.

REFERENCES

[1] P. Y. Hsiao, S. Y. Lin and S. S. Huang, An FPGA based human detection system with embedded platform, *Microelectronic Engineering*, vol.138, pp.42-46, 2015.
[2] Y. Nizam and M. N. H. Mohd, Human fall detection from depth images using position and velocity of subject, *Procedia Computer Science*, vol.105, pp.131-137, 2017.
[3] K. Kim, C. Oh and K. Sohn, Personness estimation for real-time human detection on mobile devices, *Expert Systems with Applications*, vol.72, pp.130-138, 2017.
[4] N. Dalal and B. Triggs, Histograms of oriented gradients for human detection, *Computer Vision and Pattern Recognition*, vol.1, pp.886-893, 2005.
[5] W. L. Lu and J. J. Little, Simultaneous tracking and action recognition using the PCA-HOG descriptor, *The 3rd Canadian Conference on Computer and Robot Vision*, Quebec, Canada, pp.6-13, 2006.

[6] D. G. Lowe, Distinctive image features from scale-invariant key points, *International Journal of Computer Vision*, vol.60, pp.91-110, 2004.

[7] H. Bay, T. Tuytelaars and L. Van Gool, SURF: Speeded up robust features, *European Conference on Computer Vision*, pp.404-417, 2006.

[8] A. Vashaee, R. Jafari and D. Ziou, Rotation invariant HOG for object localization in web images, *Signal Processing*, pp.304-314, 2015.

[9] K. Liu, H. Skibbe and T. Schmidt, Rotation-invariant HOG descriptors using Fourier analysis in polar and spherical coordinate, *International Journal of Computer Vision*, vol.106, pp.342-364, 2014.

[10] B. Liu, H. Wu and W. Su, Sector-ring HOG for rotation-invariant human detection, *Signal Process: Image Communication*, pp.1-10, 2017.

[11] A. Wang, S. Dai, M. Yang and Y. Iwahori, A novel human detection algorithm combining HOG with LBP histogram flourier, *International Conference on Communications and Networking*, 2015.

[12] X. Wang, T. X. Han and S. Yan, An HOG-LBP human detector with partial occlusion handling, *Proc. IEEE Int. Conf. on Computer Vision*, pp.31-39, 2009.

[13] V. J. Traver and F. Pla, Log-polar mapping template design: From task-level requirement to geometry parameters, *Image and Vision Computing*, vol.26, no.10, pp.1354-1370, 2008.

[14] F. Solari, M. Chessa and S. P. Sabatini, Design strategies for direct multi-scale and multi-orientation feature extraction in the log polar domain, *Pattern Recognition Letters*, pp.41-51, 2011.

[15] J. N. Sarvaiya, S. Patnaik and K. Kothari, Image registration using log polar transform and phase correlation to recover higher scale, *Pattern Recognition Research*, pp.90-105, 2012.

[16] J. Ouyang, G. Coatrieux and H. Shu, Robust hashing for image authentication using quaternion discrete Fourier transform and log-polar transform, *Digital Signal Processing*, pp.98-109, 2015.

[17] A. Bosch, A. Zisserman and X. Munoz, Representing shape with a spatial pyramid kernel, *The ACM International Conference on Image and Video Retrieval*, 2007.

[18] T. Ojala, M. Pietikainen and D. Harwood, Performance evaluation of texture measures with classification based on Kullback discrimination of distributions, *Proc. of the Int. Conf. on Pattern Recognition*, vol.1, pp.582-585, 1994.

[19] T. Ojala, M. Pietikainen and D. Harwood, A comparative study of texture measures with classification based on featured distributions, *Pattern Recognition*, vol.29, no.1, pp.51-59, 1996.

[20] T. Ojala, M. Pietikainen and T. Maenpaa, Gray scale and rotation invariant texture classification with local binary patterns, *IEEE Trans. Pattern Analysis and Machine Intelligence*, 2002.

[21] C.-C. Chang and C.-J. Lin, LIBSVM: A library for support vector machines, *ACM Trans. Intelligent Systems and Technology*, pp.1-27, 2011.