AN EFFICIENT LBP-BASED DESCRIPTOR FOR REAL-TIME OBJECT DETECTION

Tran Nguyen Ngoc
Department of Computer Science
FIT, Le Quy Don Technical University
236 Hoang Quoc Viet, Hanoi, Vietnam
Email: tnn1999@mail.ru

***Abstract*—This paper presents a novel method for interest region description which is based on the idea that the appearance of an interest region can be well characterized by the distribution of its local features. We introduce a new texture feature called Grid of Local Binary Patterns (GoLBP) that is a modified version of the well-known Local Binary Pattern (LBP) feature. Experiments show that, the proposed descriptor can be used for the real-time vehicle detection in Vietnam.**

I. INTRODUCTION

Real-time object detection has wide range of application such as human-machine interaction (gestures recognition, face detection), video surveillance systems (people tracking, video vehicle detection), robot localization, etc. One of the most used fast object detection techniques in images is the sliding window technique [3]. The key idea of the sliding window technique is using a window of predefined size, which is slided over the interest region of the input image with a certain step of displacement. At each slide step, the underlying sub-image is cropped and given as input to an image classifier. Thus, this type of processing is very time consuming. In order to overcome the problem of computing features at each step, integral representations have been proposed. Two types of integral representations usually employed are the Integral Image [2] and the Integral Histograms [16]. The basic characteristic of both is to precompute the features before sliding the windows in this way the features are computed only once.

In general, the majority of object detection systems have two main modules: feature extractor, and feature classifier. Feature extraction is special form of dimensionality reduction, when the input data (pixels of the image) will be transformed into a reduced representation set of features by some descriptor. On the other hand, good features should be discriminative, robust, easy to compute and efficient algorithms are needed for a variety of tasks such as detection, recognition and tracking.Various local descriptors such as SIFT[13], Haar-like [2], Histogram of Oriented Gradients (HOG) [12], Local Binary Pattern (LBP) and its variants [1] have been proposed over the recent years.

Feature classifiers may use Artificial Neural Networks, Support Vector Machines, Bayesian inference, or a combination of these, among other techniques. A very prominent face detection method is the combination of Adaboost and Integral Image representation, first proposed by Viola and Jones [2]

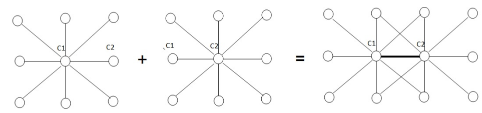

Fig. 1. Duplicate local neighborhood C1-C2 in the LBP operator

LBP is a computationally-efficient descriptor that is popular in texture classification [14], [15] and face detection [5], [6], [7]. LBP is illumination and contrast invariant as it only considers the sign of the difference between two pixels. Representing LBP in the form of histogram makes the descriptor resistant to translations within the neighbourhood of histogramming. This paper focuses only on designing a new descriptor, which modifies an existing LBP descriptor to capture more local information.

II. RELATED WORKS

The Local Binary Pattern operator has been widely used to represent, detect and recognize faces [7], [5], as well as other object classes [8]. The success of LBP in face description is due to the discriminative power and computational simplicity of the operator, and its robustness to monotonic gray scale changes caused by, for example, illumination variations. LBP was proposed by Ojala et al. [1] as a histogram with N levels. The LBP operator was extended in many ways, one of those extensions is the rotation invariant LBP [11]. In [17], the authors proposed using a concatenation of cell-structured LBP, similar to [12], and HOG to describe humans. Their feature showed a much better performance compared to [11]. Nguyen Duc Thanh et al. [8] proposed using NRLBP for human description by mapping an LBP code and its complement to the minimum of the two. So far many extensions of LBP have been proposed: boosting LBP [7], MultiBlock-LBP [5], CLBP [9] etc.

However, among these methods, LBP operators describe each pixel by the relative gray-levels of its neighboring pixels and the histogram of the binary pattern computed over a region of image. Hence, correlation between pixels can be obtained from duplicate local neighborhoods (Fig.1), obviously, the

978-1-4799-5431-5/14/$31.00

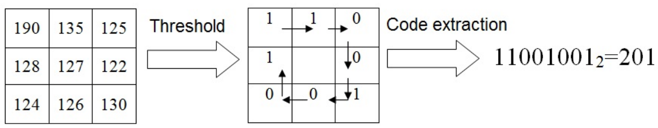

Fig. 2. LBP operator

number of the computational performance of the LBP operator can be reduced by decreasing number of the redundant neighborhoods. The next section presents how to localize redundant information between pixels and proposes an optimal descriptor.

III. Object representation using LBP and GoLBP

The LBP operator is an image operator which transforms an image into an array or image of integer labels describing small-scale appearance of the image. These labels or their statistics, most commonly the histogram, are then used for further image analysis. The most widely used versions of the operator are designed for monochrome still images but it has been extended also for color images as well as videos and volumetric data. We adopt the following notation. Given a pixel $c = (x_c, y_c)$, the LBP code of c is defined by:

$$LBP_{(n-R)}(x_c, y_c) = \sum_{p=0}^{P-1} s(g_{n_p} - g_c)2^p \tag{1}$$

where P is a number of neighbor pixels, n_p is a neighbor pixel of c and the distance from n_p to c does not exceed R. g_{n_p}, g_c are the gray intensity of n_p and c respectively, and

$$s(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

However, Fig.5 represents the transformation of the original LBP operator on 9 pixels which the LBP code of the center pixel is obtained by comparing its intensity with 8 neighboring pixels' intensity. Obviously, the transformation is equivalents of 20 neighborhoods' information(C1-C2,C2-C3,C3-C6,C6-C9...). To capture more information of the object's pixels, we propose a novel LBP, namely, Grid of LBP, as follow,

$$GoLBP_{(n-R)}(x_c, y_c) = \sum_{i=0}^{[(P-1)/2]} s(g_{n_i} - g_c)2^i + \sum_{i=[(P-1)/2]+1}^{P-1} s(g_{n_j} - g_{n_k})2^i \tag{3}$$

where $[x]$ means integer part of x; n_i, n_j, n_k are neighboring pixels of c and their indexes j, k are selected by some structure (that is called Grid of LBP).
This paper focuses on GoLBP with $n = 1, P = 8$ (see Fig.3) as follow,

$$GoLBP = d_{0,c} + d_{2,c}2 + d_{3,c}2^2 + d_{5,c}2^3 + d_{1,4}2^4 + d_{1,6}2^5 + d_{7,2}2^6 + d_{7,4}2^7 \tag{4}$$

where $d_{x,y} = s(g_x, g_y)$ (see Fig.4).

It can be seen that the GoLBP captures more information than the LBP descriptor. For example, the GoLBP representation in Fig.6 reflects the relative contrast between 28 couples of pixels (against 20 in Fig. 5). Hence, the GoLBP is more robust and adaptive to changes of the gray intensity.

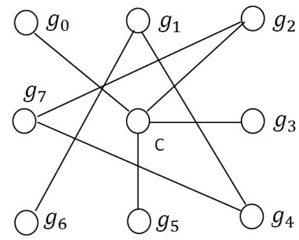

Fig. 3. GoLBP structure

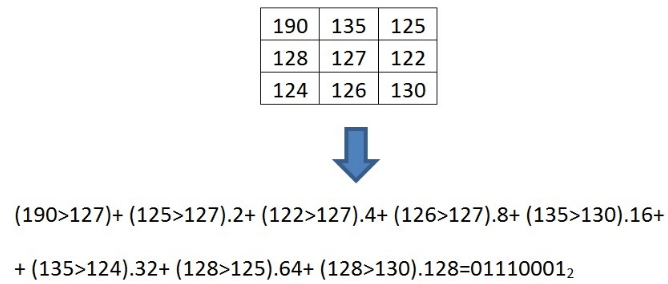

Fig. 4. GoLBP operator

IV. Real-time object detection using GoLBP

Visual object detection is under constant pressure to increase both its quality and speed. Such progress allows for new applications. A higher speed enables its inclusion into larger systems with extensive subsequent processing (e.g. as an initialization for segmentation or tracking), and its deployment in computationally constrained scenarios (e.g. embedded system, large scale data processing). In this paper we focus on improving the speed of vehicle detection (see Fig. 7), while providing state-of-the-art detection quality.

To evaluate the performance of the proposed method, we conducted some experiments. The proposed descriptor was evaluated by employing it to detect motorbike from video. The experiments were conducted on the data set including 604 images with 3378 labeled motorbike in frontal view. We use 3000 samples for training and 378 samples for testing (in 61 images). For the implementation of the training and detection experiments,in the MultiBlock LBP (that was implemented in

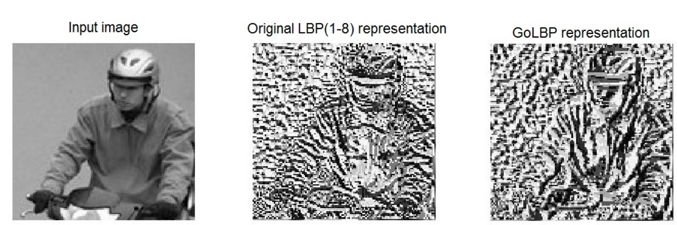

Fig. 7. LBP and GoLBP operators

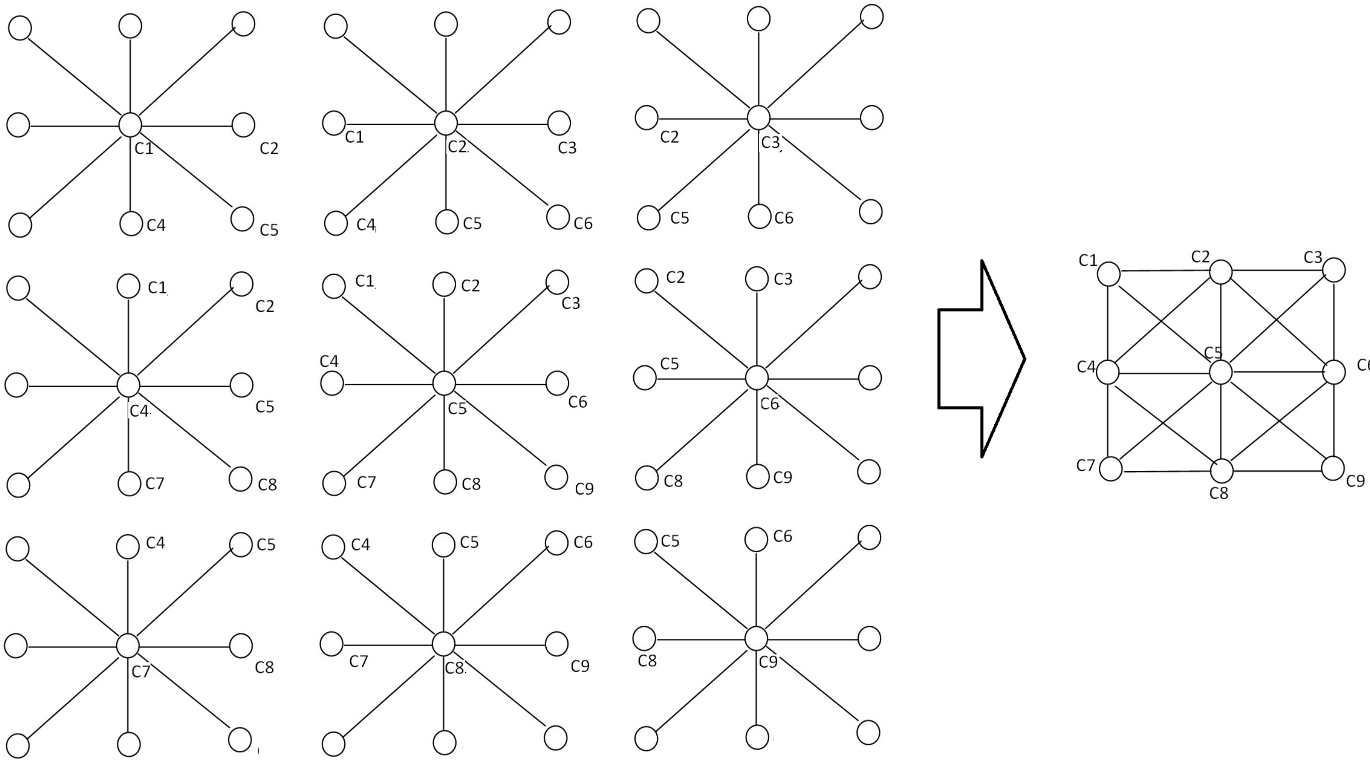

Fig. 5. LBP operator on 9 pixels and grid of their neighborhoods

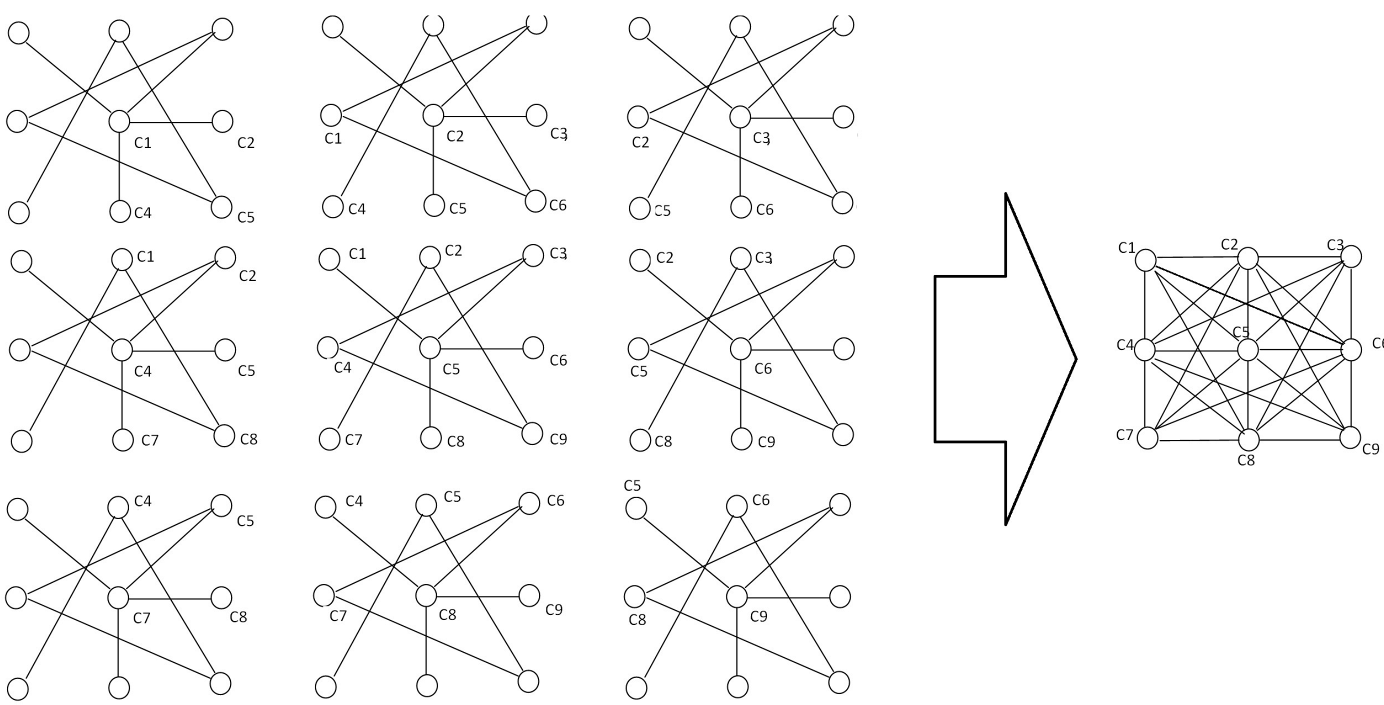

Fig. 6. GoLBP operator on 9 pixels and grid of their neighborhoods

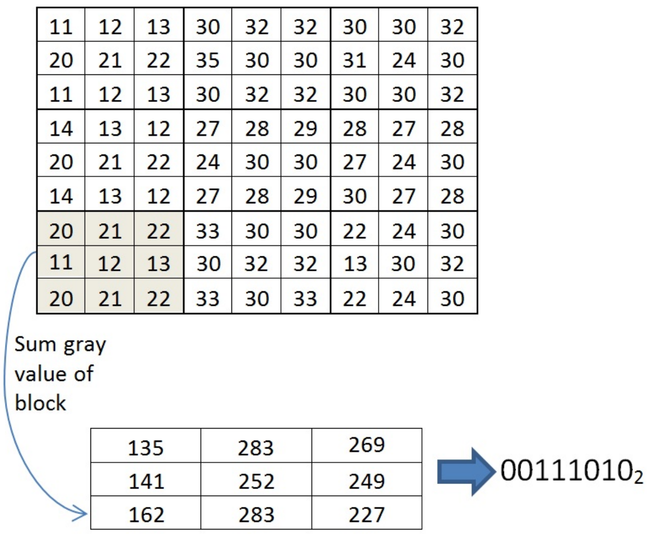

Fig. 8. MultiBlock GoLBP feature for image representation

OpenCV 2.4.6 library) original LBP descriptor was replaced by GoLBP descriptor and trained Gentle AdaBoost.

A more detailed description of such GoLBP operator can be found in Fig.8. GoLBP (based on MultiBlock approach) can capture large scale structures that may be the dominant features of images. Totally, we can get 256 kinds of binary patterns. The experimental results show the GoLBP features are more distinctive than Haar-like features and original LBP features. Another advantage is that the number of exhaustive set of features (rectangles at various scales, locations and aspect ratios) is much smaller than Haar-like features. Given a sub-window size of 20x20, there are totally 2049 features, this amount is about 1/20 of Haar-like features (45891).

For real-time system vehicle detection as can be seen in Fig.9, the GoLBP outperforms the original Multiblock LBP and Haar-like with regard to accuracy.

V. Conclusion and future works

In this paper we presented the GoLBP descriptor and related algorithms for fast object detection. Superior performance of the GoLBP descriptor and algorithms were demonstrated on several examples with detailed comparisons to previous techniques and features. The method can be extended in several ways. For example, following automatical detection of an object in a video, it can be tracked in the following frames using this approach. As the object leaves the scene, the distance score will increase significantly which ends the tracking. Currently we are working on multiple classification algorithms which use the GoLBP features to recognize hand gestures.

Acknowledgment

This work has been done in cooperation and financially supported by the FPT Software JSC, through the FSOFT research program ”Real-time computer vision system for vehicle counting in Vietnam”.

References

[1] T. Ojala, M. Pietikainen, and T. Maenpaa, ”Multiresolution gray-scale and rotation invariant texture classification with local binary patterns”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, pp. 971-987, 2002.

[2] P. Viola and M. J. Jones, Robust real-time face detection, *International Journal of Computer Vision*, vol. 57, no. 2, pp.137-154, 2004.

[3] C. H. Lampert, M. B. Blaschko, and T. Hofmann, Beyond sliding windows: Object localization by efficient subwindow search, *in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (CVPR 2008), pp.1-8, 2008.

[4] A. Hadid, The local binary pattern approach and its applications to face analysis, *in First Workshop on Image Processing Theory, Tools and Applications*, 2008, pp. 1-9.

[5] Zhang, Lun, et al. ”Face detection based on multi-block lbp representation.” *Advances in Biometrics*. Springer Berlin Heidelberg, 2007. 11-18.

[6] H. Wang, P. Li, and T. Zhang, ”Proposal of novel histogram features for face detection”, in ICAPR, L*ecture Notes in Computer Science*, 2005, pp. 334-343.

[7] Do, Toan Thanh, et al. ”Boosted of Haar-like Features and Local Binary Pattern Based Face Detection.” *Computing and Communication Technologies*, 2009. RIVF’09. International Conference on. IEEE, 2009.

[8] Nguyen, Duc Thanh, et al. ”Object detection using non-redundant local binary patterns.” *Image Processing (ICIP)*, 2010 17th IEEE International Conference on. IEEE, 2010.

[9] Xue, Gengjian, et al. ”Hybrid center-symmetric local pattern for dynamic background subtraction.” *Multimedia and Expo (ICME)*, 2011 IEEE International Conference on. IEEE, 2011.

[10] Lee, Kwon, et al. ”Fast object detection based on color histograms and local binary patterns.” *TENCON 2012-2012 IEEE Region 10 Conference*. IEEE, 2012.

[11] Mu, Yadong, et al. ”Discriminative local binary patterns for human detection in personal album.” *Computer Vision and Pattern Recognition, 2008*. CVPR 2008. IEEE Conference on. IEEE, 2008.

[12] N. Dalal and B. Triggs, Histograms of oriented gradients for human detection, *in Proc IEEE Int. Conf. on Computer Vision and Pattern Recognition,* 2005, vol. 1, pp. 886-893.

[13] D. G. Lowe, Distinctive image features from scale-invariant keypoints, *Int. Journal of Computer Vision*, vol. 60, no. 2, pp.91-110, 2004.

[14] G. Zhenhua, Z. Lei, and D. Zhang, A completed modeling of local binary pattern operator for texture classification, *IEEE Trans. Image Process.*, vol. 19, no. 6, pp. 1657- 1663, June 2010.

[15] S. Liao, M.W.K. Law, and A.C.S. Chung, Dominant local binary patterns for texture classification, *IEEE Trans. Image Process.*, vol. 18, no. 5, pp. 1107-1118, May 2009.

[16] F. Porikli, Integral histogram: A fast way to extract histograms in cartesian spaces, in IEEE *International Conference on Computer Vision and Pattern Recognition (CVPR),* 2005, pp. 829-836.

[17] T.X. Wang, X. Han and S. Yan, An hog-lbp human detector with partial occlusion handling, *in Proc. IEEE Int. Conf.Comput. Vis.*, 2009, pp. 32-39.

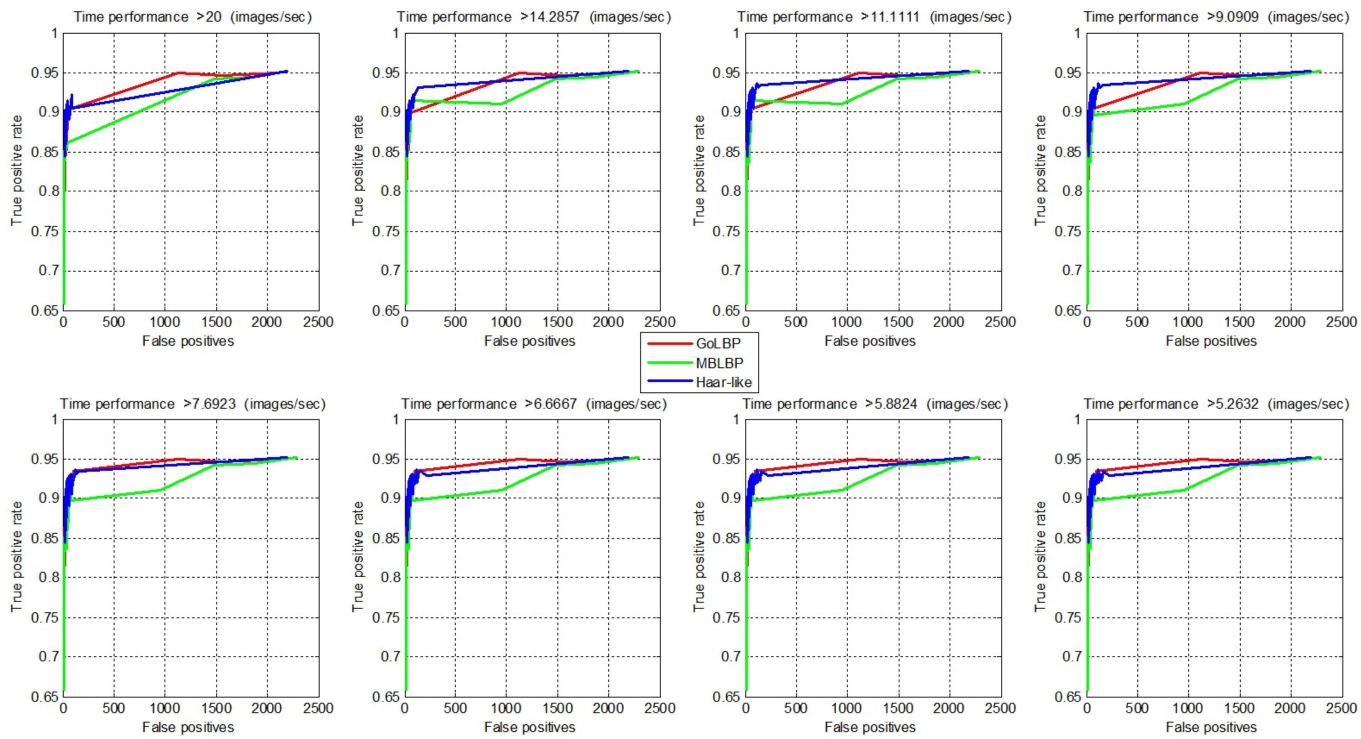

Fig. 9. Performance of GoLBP againts MB-LBP and Haar-like

Fig. 10. Some results of the motorbike detection process