

 Open access • Proceedings Article • DOI:10.1109/ICIP.2010.5651633

## Object detection using Non-Redundant Local Binary Patterns — [Source link](#)

Duc Thanh Nguyen, Zhimin Zong, Philip Ogunbona, Wanqing Li

**Institutions:** University of Wollongong

**Published on:** 03 Dec 2010 - International Conference on Image Processing

**Topics:** Local binary patterns and Object detection

Related papers:

- [Multiresolution gray-scale and rotation invariant texture classification with local binary patterns](#)
- [Histograms of oriented gradients for human detection](#)
- [Enhanced Local Texture Feature Sets for Face Recognition Under Difficult Lighting Conditions](#)
- [Distinctive Image Features from Scale-Invariant Keypoints](#)
- [A Completed Modeling of Local Binary Pattern Operator for Texture Classification](#)

Share this paper:    

View more about this paper here: <https://typeset.io/papers/object-detection-using-non-redundant-local-binary-patterns-21bgupv2mu>

1-1-2010

## Object detection using non-redundant local Binary Patterns

Duc Thanh Nguyen

*University of Wollongong*, dtn156@uow.edu.au

zhimin Zong

*University of Wollongong*

Philip Ogunbona

*University of Wollongong*, philipo@uow.edu.au

Wanqing Li

*University of Wollongong*, wanqing@uow.edu.au

Follow this and additional works at: <https://ro.uow.edu.au/infopapers>



Part of the [Physical Sciences and Mathematics Commons](#)

---

### Recommended Citation

Nguyen, Duc Thanh; Zong, zhimin; Ogunbona, Philip; and Li, Wanqing: Object detection using non-redundant local Binary Patterns 2010, 4609-4612.

<https://ro.uow.edu.au/infopapers/2131>

---

# Object detection using non-redundant local Binary Patterns

## Abstract

Local Binary Pattern (LBP) as a descriptor, has been successfully used in various object recognition tasks because of its discriminative property and computational simplicity. In this paper a variant of the LBP referred to as Non-Redundant Local Binary Pattern (NRLBP) is introduced and its application for object detection is demonstrated. Compared with the original LBP descriptor, the NRLBP has advantage of providing a more compact description of object's appearance. Furthermore, the NRLBP is more discriminative since it reflects the relative contrast between the background and foreground. The proposed descriptor is employed to encode human's appearance in a human detection task. Experimental results show that the NRLBP is robust and adaptive with changes of the background and foreground and also outperforms the original LBP in detection task.

## Keywords

local, binary, patterns, object, detection, non, redundant

## Disciplines

Physical Sciences and Mathematics

## Publication Details

Nguyen, D., Zong, z., Ogunbona, P. & Li, W. (2010). Object detection using non-redundant local Binary Patterns. Proceedings of 2010 IEEE 17th International Conference on Image Processing (pp. 4609-4612). New York, NY, USA: IEEE.

# OBJECT DETECTION USING NON-REDUNDANT LOCAL BINARY PATTERNS

*Duc Thanh Nguyen, Zhimin Zong, Philip Ogunbona, and Wanqing Li*

Advanced Multimedia Research Lab, ICT Research Institute  
School of Computer Science and Software Engineering  
University of Wollongong, Australia

## ABSTRACT

Local Binary Pattern (LBP) as a descriptor, has been successfully used in various object recognition tasks because of its discriminative property and computational simplicity. In this paper a variant of the LBP referred to as Non-Redundant Local Binary Pattern (NRLBP) is introduced and its application for object detection is demonstrated. Compared with the original LBP descriptor, the NRLBP has advantage of providing a more compact description of object's appearance. Furthermore, the NRLBP is more discriminative since it reflects the relative contrast between the background and foreground. The proposed descriptor is employed to encode human's appearance in a human detection task. Experimental results show that the NRLBP is robust and adaptive with changes of the background and foreground and also outperforms the original LBP in detection task.

*Index Terms*— Human detection, local binary patterns

## 1. INTRODUCTION

Object detection is an active research topic in computer vision and there are several techniques developed to accomplish the object detection task. In general object detection methods can be categorized as either global or local approach. Global approaches focus on detection of a full object in which the object is often encoded by a feature vector [1, 2, 3, 4]. Some image features such as HOG (Histogram of Oriented Gradients) [1] and Haar-like features [3] have been used widely to describe the appearance of the objects. Local approaches on the other hand detect objects by locating parts constituting the objects [5, 6, 7, 8, 9].

Compared with global detection, local detection has advantage of being able to detect objects with high articulation and cope with the problem of occlusion. However, determination of partial components is often based on prior knowledge of the structure of the objects. For example, for human detection, in [5, 6], human body was decomposed into a number of partial components corresponding to natural body parts such as head-shoulder, torso, and legs.

Recently, some methods [7, 8, 9] have employed interest points detectors, e.g. SIFT [10], to identify the locations

of the object's parts. Local appearances of these parts then were used as codewords to represent the object. Since parts of an object can be located automatically based on low-level features rather than prior knowledge, this approach can accommodate with variations in the pose of the object.

Motivated by the advantages of using a codebook of local appearance for object detection [7, 8, 9] and the discriminative power of local patterns in object recognition, this paper presents an object detection method with the following contributions. First, we identify the locations of the object's parts using interest points detectors and recognize the object by matching object's local appearances with predefined codewords. Second, to describe local appearance, we introduce a new image descriptor, namely, Non-Redundant Local Binary Pattern (NRLBP). This descriptor is based on the original LBP descriptor proposed for texture classification [11]. However, it is more discriminative and robust to illumination changes compared with the original LBP as well as efficient for computation and restoration. The proposed descriptor was evaluated by employing it to detect humans from static images. Experimental results show that the NRLBP is robust and adaptive to changes of the background and foreground and has a superior performance in comparison with the original LBP.

The rest of the paper is organized as follows. Section 2 reviews the original LBP and then describes the novel LBP descriptor. Section 3 presents an object detection method based on the proposed descriptor. Experimental results along with comparative analysis are presented in Section 4. Section 5 concludes the paper with remarks.

## 2. NON-REDUNDANT LBP (NRLBP)

### 2.1. LBP

The original LBP was developed for texture classification [11] and the success has been due to its robustness under illumination changes, computational simplicity and discriminative power. Fig. 1 represents an example of the original LBP in which the LBP code of the center pixel (in red color and value 20) is obtained by comparing its intensity with neighboring pixels' intensities. The neighbor pixels whose intensities are

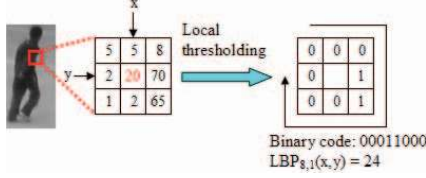


Fig. 1. An illustration of the original LBP descriptor.

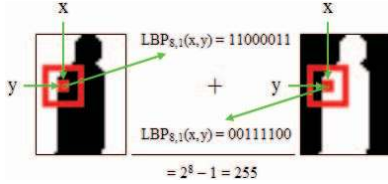


Fig. 2. The left and right image represents the same human with inverted changes of the background and foreground. The LBP code at  $(x, y)$  (in the left and right image) is complementary each other, i.e. the sum of these codes is  $2^P - 1$ .

equal or higher than the center pixel's are labeled as "1"; otherwise as "0".

We adopt the following notation. Given a pixel  $c = (x_c, y_c)$ , the value of the LBP code of  $c$  is defined by:

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

where  $n_p$  is a neighbor pixel of  $c$  and the distance from  $n_p$  to  $c$  does not exceed  $R$ .  $g_{n_p}$  and  $g_c$  are the gray values (intensities) of  $n_p$  and  $c$  respectively, and

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

In (1),  $R$  is the radius of a circle centered at  $c$  and  $P$  is the number of sampled points. For example, in Fig. 2,  $R$  and  $P$  are 1 and 8 respectively. The following are important properties of the LBP descriptor.

**Uniform and non-uniform LBP.** Uniform LBP is defined as the LBP that has at most two bitwise transitions from 0 to 1 and vice versa in its circular binary representation. LBPs which are not uniform are called non-uniform LBPs. As indicated in [11], an important property of uniform LBPs is the fact that they often represent primitive structures of the texture while non-uniform LBPs usually correspond to unexpected noises and hence are less discriminative.

**LBP histogram.** Scanning a given image in pixel by pixel fashion, LBP codes are accumulated into a discrete histogram called LBP histogram. Intuitively, the number of  $LBP_{P,R}$  histogram bins is  $2^P$ .

## 2.2. Non-Redundant LBP

Although the robustness of the original LBP has been demonstrated in many applications, it has certain drawbacks when employed to encode object's appearance. Two notable disadvantages are:

- Storage requirement: as presented, the original LBP requires  $2^P$  bins of histogram.
- Discriminative ability: the original LBP is sensitive to the relative changes between the background and foreground (the region inside the object). It depends on the intensities of particular locations and thus varies based on object's appearance. For example, the LBP codes of the regions indicated by red rectangles (small box) in both the left and right images in Fig. 2 are quite different while they actually represent the same structure.

To overcome both of the above issues, we propose a novel LBP, namely, Non-Redundant LBP (NRLBP), as follows,

$$NRLBP_{P,R}(x_c, y_c) = \min \left\{ LBP_{P,R}(x_c, y_c), 2^P - 1 - LBP_{P,R}(x_c, y_c) \right\} \quad (2)$$

It can be seen that the NRLBP considers the LBP code and its complement as same. For example, the two LBP codes "11000011" and "00111100" in Fig. 2 are counted once. Obviously, with the NRLBP, the number of bins in the LBP histogram is reduced by half. Furthermore, compared with the original LBP, the NRLBP is more discriminative in this case since it reflects the relative contrast between the background and foreground rather than forcing a fixed arrangement of the background and foreground. Hence, the NRLBP is more robust and adaptive to changes of the background and foreground. The robustness of the NRLBP is also verified in experiments (see section 4).

## 3. PROPOSED OBJECT DETECTION METHOD

### 3.1. Training

Assume that we have a number of images containing objects of interest. We further assume that the object masks (silhouettes) are also given. The interest points can be detected using SIFT detector [10]. A set of interest points close to the object masks is then selected. This set is called positive interest points set. Other interest points not belonging to this set are called negative interest points. For each positive interest point, a  $(2L + 1) \times (2L + 1)$ -image patch centered at that point is collected. Some common appearance features such as HOG [1] or Haar-like features [5, 3] might be used to describe image patches. In this paper, NRLBP with  $P = 8$  and  $R = 1$  is employed. For each image patch, a NRLBP histogram accumulated from the NRLBP codes of all points in that image

patch is created and normalized using the  $L_1$ -square norm. NRLBP histograms are then clustered using a  $K$ -means algorithm in which the similarity between two histograms is measured using the Bhattacharyya distance (see Eq. (9)). If the spatial information of the image patches are considered, the clustering algorithm is extended to include the locations of the image patches. This step results in a codebook in which the codewords are means of the clusters.

Some codewords in the codebook might not be sufficiently discriminative to represent the foreground; thus a codeword filtering step is initiated. In particular, for each codeword  $c$ , we compute the relative frequency of matches with both the positive and negative interest points. The histograms of frequencies, (respectively, for the positive and negative interest point matches), can be considered as the conditional distributions of the codeword given positive ( $Obj$ ) and negative label ( $NonObj$ ) of a detected object. A codeword  $c$  is selected if the following condition is satisfied:

$$\log \left[ \frac{P(c|Obj)}{P(c|NonObj)} \right] \geq T \quad (3)$$

where  $T$  is a predefined threshold. The codeword filtering process yields a fine codebook  $\mathcal{C}$  including  $N$  codewords  $\mathcal{C} = \{c_1, c_2, \dots, c_N\}$ .

In addition to generating the codebook, similarly to [8], we compute the probability of object's location based on the votes casted by the codewords. In particular, since each codeword can vote for different object positions, the conditional probability  $P(l|c_i)$  can be computed ( $l$  denotes the location, e.g. center, of the object).

### 3.2. Object Detection

The object detection method is based on the scanning-window approach. In other words we scan the input image by a detection window  $W$  at various scales and positions. Let  $I_W$  denote the image contained in  $W$ . The probability that a given image window  $I_W$  is considered as the object at location  $l$  is represented as  $P(Obj, l|I_W)$  and the problem of object detection is then formulated as verifying the following condition,

$$P(Obj, l|I_W) \geq \theta \quad (4)$$

where  $\theta$  is a predefined threshold.

Applying Bayesian rule, (4) can be rewritten as,

$$P(Obj, l|I_W) = \frac{P(I_W, l|Obj)P(Obj)}{P(I_W)} \quad (5)$$

We assume that  $P(Obj) = P(NonObj) = 0.5$  and  $P(I_W)$  has a uniform distribution. The problem is how to compute the likelihood  $P(I_W, l|Obj)$  that an image window  $I_W$  contains an object appearing at location  $l$ . This likelihood can be estimated as follows. Assuming that

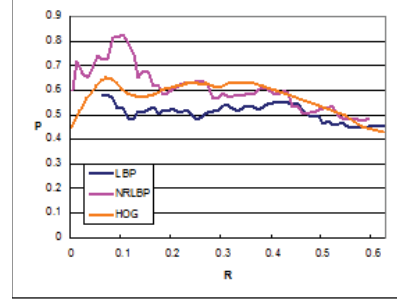


Fig. 3. PR Curves on the Penn-Fudan dataset.

$\forall i \in \{1, \dots, N\}, P(I_W, l|c_i, Obj) = P(I_W, l|c_i)$ , we have,

$$P(I_W, l|Obj) = \sum_{i=1}^N P(I_W, l|c_i)P(c_i|Obj) \quad (6)$$

Further assuming that  $I_W$  and  $l$  are independent conditioned on the codeword  $c_i$ , Eq. (6) can be rewritten as,

$$P(I_W, l|Obj) = \sum_{i=1}^N P(I_W|c_i)P(l|c_i)P(c_i|Obj) \quad (7)$$

where  $P(l|c_i)$  and  $P(c_i|Obj)$  are computed through training.  $P(I_W|c_i)$  represents the likelihood that  $I_W$  contains some image patch that matches  $c_i$  and can be computed as follows. Let  $E = \{e_1, e_2, \dots\}$  denote a set of the image patches located at interest points in  $I_W$ , we define,

$$P(I_W|c_i) = \max_{e_j \in \hat{E}_{c_i}} \rho(c_i, e_j) \quad (8)$$

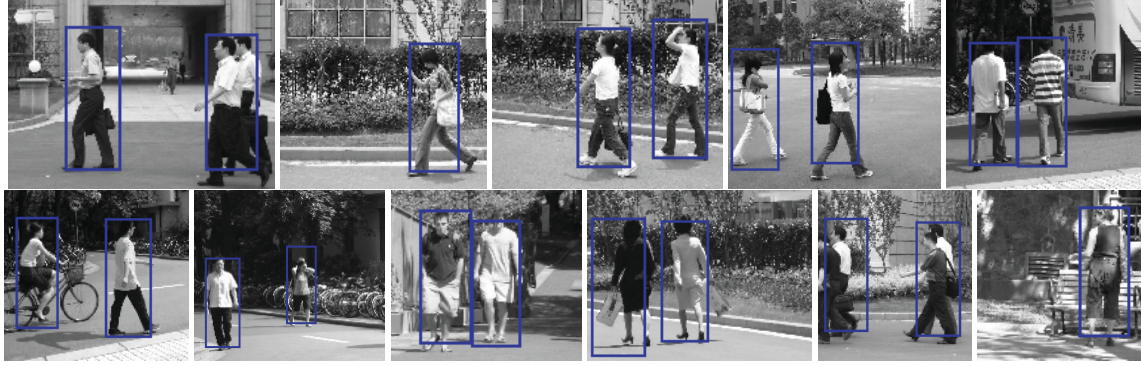
where  $\hat{E}_{c_i} = \{e_j \in E | c_i = \arg \max_{c_k \in \mathcal{C}} \rho(c_k, e_j)\}$  and  $\rho(c_k, e_j)$  represents the similarity between codeword  $c_k$  and image patch  $e_j$ . Intuitively,  $\hat{E}_{c_i}$  indicates the set of image patches which have the best matching codeword  $c_i$ . Since the NRLBP is employed to describe the codewords, i.e. each codeword is represented by a histogram of  $B$  bins,  $\rho(c_k, e_j)$  can be defined using the Bhattacharyya distance as,

$$\rho(c_k, e_j) \propto \sum_{b=1}^B \sqrt{c_k(b)e_j(b)} \quad (9)$$

## 4. EXPERIMENTAL RESULTS

The proposed method was evaluated by employing it to detect humans from still images. The experiments were conducted on the Penn-Fudan dataset<sup>1</sup> including 170 images with 345 labeled pedestrians in multiple views and various cluttered

<sup>1</sup><http://www.cis.upenn.edu/~jshi/ped.html>



**Fig. 4.** Some detection results on the Penn-Fudan dataset.

backgrounds. We used 70 images for training and generating the codebook. The true and false positives are determined by comparing the detection results with true detections given in the ground truth using the criteria proposed in [8]. The detection performance was evaluated using the PR (Precision-Recall) curve shown in Fig. 3. Some detection results are shown in Fig. 4.

In addition to evaluation, we compared the proposed method with its variants and other state-of-the-art methods. In particular, we compared the performance of the NRLBP with original LBP (see Fig. 3). Since we use the NRLBP, we could reduce the number of bins in the histogram from 59 (as the original LBP) to 30 in which all non-uniform NRLBPs vote for one bin while each uniform NRLBP is casted into a unique bin corresponding to its NRLBP code. For fast computation of the NRLBP histograms, we employ the integral images proposed in [3].

As can be seen in Fig. 3, the NRLBP outperforms the original LBP with regard to accuracy. Furthermore, with the NRLBP we could reduce the dimensionality of the LBP histogram, thus saving memory and speeding up the detection task. We also compared the proposed method with the work of Dalal et al. [1] (using HOG to encode human's shape). The PR curves of these methods are presented in Fig. 3.

## 5. CONCLUSIONS

This paper presents an object detection method using Non-Redundant Local Binary Patterns (NRLBPs) to encode object's appearance. By considering the inverted changes of the background and foreground intensities similarly, the NRLBP descriptor provides a discriminative and compact description of object's appearance. We evaluated and compared the performance of the proposed descriptor with the original LBP and other descriptors for detecting humans from still images. By approaching the problem of object detection as local detection, we propose to solve the occlusion in future work.

## 6. REFERENCES

- [1] 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.
- [2] P. Viola and M. J. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proc IEEE Int. Conf. on Computer Vision and Pattern Recognition*, 2001, vol. 1, pp. 511–518.
- [3] P. Viola, M. J. Jones, and D. Snow, "Detecting pedestrians using patterns of motion and appearance," *Int. Journal of Computer Vision*, vol. 63, no. 2, pp. 153–161, 2005.
- [4] O. Tuzel, F. Porikli, and P. Meer, "Human detection via classification on riemannian manifolds," in *Proc IEEE Int. Conf. on Computer Vision and Pattern Recognition*, 2007.
- [5] A. Mohan, C. Papageorgiou, and T. Poggio, "Example-based object detection in images by components," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 23, no. 4, pp. 349–361, 2001.
- [6] B. Wu and R. Nevatia, "Detection of multiple, partially occluded humans in a single image by bayesian combination of edgelet part detectors," in *Proc Int. Conf. on Computer Vision*, 2005, pp. 90–97.
- [7] B. Leibe, A. Leonardis, and B. Schiele, "Combined object categorization and segmentation with an implicit shape model," in *Proc European Conference on Computer Vision*, 2004, pp. 17–32.
- [8] B. Leibe, E. Seemann, and B. Schiele, "Pedestrian detection in crowded scenes," in *Proc IEEE Int. Conf. on Computer Vision and Pattern Recognition*, 2005, vol. 1, pp. 878–885.
- [9] E. Seemann, B. Leibe, and B. Schiele, "Multi-aspect detection of articulated objects," in *Proc IEEE Int. Conf. on Computer Vision and Pattern Recognition*, 2006, vol. 2, pp. 1582–1588.
- [10] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.
- [11] T. Ojala, M. Pietikäinen, 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.