Haar Local Binary Pattern Feature for Fast Illumination Invariant Face Detection

Anindya Roy
Anindya.Roy@idiap.ch
École Polytechnique Fédérale de Lausanne
Lausanne, Switzerland

Sébastien Marcel
Sebastien.Marcel@idiap.ch
Idiap Research Institute
Martigny, Switzerland

Face detection systems based on boosted Haar features in a cascade architecture [3] perform very well both in terms of accuracy and speed. However, they are vulnerable to variations in illumination conditions. On the other hand, face detection systems based on the Local Binary Pattern (LBP) [2] or its variants [1] show robustness to illumination conditions but do not take advantage of the richness of the Haar feature set in efficiently modelling faces.

In this paper, we propose a face detection system based on a new type of feature called the Haar Local Binary Pattern (HLBP) feature which combines the advantages of both Haar and LBP and is boosted using AdaBoost as in [3]. This feature compares the LBP label counts in two adjacent image subregions similar to Haar masks, i.e. it indicates whether the number of times a particular LBP label occurs in one region is greater or lesser than the number of times it occurs in another region, offset by a certain threshold. They capture the region-specific variations of local texture patterns which is hypothesized to remain relatively stable across variations in illumination conditions. Experiments show that our features more robust to illumination variations, compared to Haar and LBP individually.

Let the input be an $N \times M$ gray-level image, represented as an $N \times M$ matrix I. Firstly, the LBP image I_{LBP} [2] is calculated from the original input image I using $LBP_{4,1}$, (ref. Fig. 1) found to be the optimal operator in our case. Next, the Integral Histogram set $\{I_k^H\}_{k=1}^{16}$ [4] of the LBP image I_{LBP} is calculated. The individual pixels $I_k^H(x,y)$ of the k-th Integral Histogram I_k^H is calculated as the number of pixels above and to the left of the pixel (x,y) in the LBP image I_{LBP} which have a label k, i.e. $I_k^H(x,y) = \sum_{u \le x, v \le y} \delta_k(u,v)$, where $\delta_k(u,v) = 1$ if the label of the pixel at location (u,v) in the LBP image I_{LBP} is k, and is zero otherwise. Using the following pair of references, for all $k \in \{1,16\}$:

$$i_k^H(x,y) = \quad i_k^H(x,y-1) + \delta_k(x,y) \tag{1}$$

$$I_k^H(x,y) = \quad I_k^H(x-1,y) + i_k^H(x,y) \tag{2}$$

where $i_k^H(x,0) = 0$ for any x and k, the Integral Histogram set can be calculated by only one pass over the LBP image. In the third and final stage, a particular HLBP feature $f_{T,k,x,y,w,h,\theta,p}(I)$, defined by the following parameters : mask type T (one out of five, ref. Fig. 2), LBP label k (one out of sixteen for $LBP_{4,1}$), position (x,y) of the mask inside the image plane, size (w,h) of the mask, a threshold θ and a direction p (either +1 or -1), is calculated by placing its corresponding mask of size (w,h) on the LBP image I_{LBP} at the location (x,y). Like in Viola and Jones' system, each mask type divides the mask region into two areas (ref. Fig. 2), a positive (A_+) and a negative (A_-) region. If we define,

$$S_{A+} = \quad \sum_{(u,v)\in A_+} \delta_k(u,v) \tag{3}$$

$$S_{A-} = \quad \sum_{(u,v)\in A_-} \delta_k(u,v) \tag{4}$$

with $\delta_k(u,v)$ as defined, then the HLBP feature value is given by,

$$f_{T,k,x,y,w,h,\theta,p}(I) = \begin{cases} 1 & \text{if } p.(S_{A_+} - S_{A_-}) > p.\theta, \\ -1 & \text{if } p.(S_{A_+} - S_{A_-}) \le p.\theta \end{cases} \tag{5}$$

Thus, the HLBP feature is a binary feature which indicates whether region A_+ (region A_-) has θ pixels more with the LBP label k compared to region A_- (region A_+), given $p = 1$ ($p = -1$), i.e. the spatial count differences of the LBP label k. However, to calculate S_{A+} and S_{A-} we need not use the above equations 3 and 4. They can each be calculated directly by only a few references to the corresponding Integral Histogram I_k^H calculated earlier, as follows. Let us denote by $(a_1,b_1),(a_2,b_2),(a_3,b_3),(a_4,b_4)$ the four corners of a generic rectangular region R, like A_+ or A_- (ref. Fig. 3). Then the sum S_R (as in Eqns.3 and 4) can be calculated directly as,

$$S_R = I_k^H(a_2,b_2) - I_k^H(a_3,b_3) - I_k^H(a_1,b_1) + I_k^H(a_4,b_4) \tag{6}$$

Thus the HLBP features involve counting the number of pixels in a region having a certain LBP label k, instead of summing over pixel intensities as with Haar features. Due to adverse illumination conditions, the pixel intensities in an image I may change. However, the LBP label of a pixel is much more robust to illumination changes. Thus, the number of pixels within a region having a particular LBP label will also remain more or less constant with varying illumination. Thus the final HLBP feature value, as defined in Eqn. 5, remains robust too. This is the advantage of the HLBP feature.

Face detection xperiments on several standard databases have shown that our system performs significantly better in adverse imaging conditions than normal Haar features and performs reasonably better than MCT features [1] with much less storage and computation time requirements.

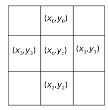

Figure 1: The $LBP_{4,1}$ label for a particular pixel (x_c, y_c) is calculated by comparing its intensity with each one of its four neighbors (vertical and horizontal only), $\{x_i, y_i\}_{i=0}^{3}$, and forming a 4-bit word.

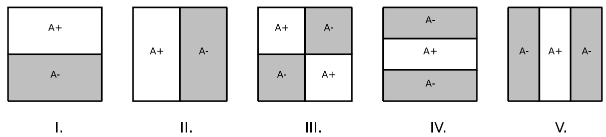

Figure 2: The five types of masks used for the calculation of both Haar and HLBP features, I. Bihorizontal, II. Bivertical, III. Diagonal, IV. Trihorizontal, V. Trivertical.

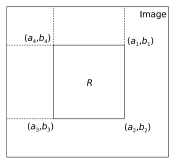

Figure 3: Calculation of the sum of LBP label counts within region R using Integral Histogram (ref. Eqn. 6).

[1] B. Froba and A. Ernst. Face detection with the modified census transform. In *Sixth IEEE International Conference on Face and Gesture Recognition*, pages 91–96, 2004.

[2] T. Ojala, M. Pietikainen, and D. Harwood. A comparative study of texture measures with classification based on feature distributions. *Pattern Recognition*, 29:51–59, 1996.

[3] P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features. In *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, pages 511–518, 2001.

[4] H. Wang, P. Li, and T. Zhang. Histogram Features-Based Fisher Linear Discriminant for Face Detection. In *Asian Conference on Computer Vision*, pages 521–530, 2006.