

Robust Face Detection Using Circular Multi Block Local Binary Pattern and Integral Haar Features

Dr.P.K.Suri

Dean, Chairman, Professor, CS&A
Kurukshetra University
Kurukshetra
India

Er.Amit Verma

A.P, ECE Department
REBIET, Sahauran,
India

Abstract— In real world applications, it is very challenging to implement a good detector which gives best performance with great speed and accuracy. There is always a trade-off in terms of speed and accuracy, when we consider performance of a face detector. In the current work we have implemented a robust face detector which uses the new concept called integral Haar histograms with CMBLBP or CSMBLBP (circular multi block local binary operator). Our detector runs for real world applications and its performance is far better than any of the present detector. It works with good speed and enough accuracy with varying face sizes, varying illumination, varying angle, different face expressions, rotation, scaling like challenges which are mostly issues of concern in the domain of face detection. We use Matlab and Image processing tool box for the implementation of the above mentioned technique.

Keywords- CMBLBP; MBLBP; LBP; Gentle Boosting; Face Detection.

I. SUMMARY OF THE PAPER

This paper presents the novel face detection system using a new technique called circular multi block local binary operator. In the second section, we present basic introduction. Third and fourth sections include LBP and MLBP (Multi block local binary operator). Fifth section gives detail of Boosting. Section six describes CMBLBP (circular multi block local binary operator). Section seventh gives Data base, Experiments and Result. Section Eight and nine include conclusion and references

II. INTRODUCTION

Face detection has a wide range of applications in the areas such as automatic face recognition, object detection, face tracking, red-eye removal, face expression recognition, human-machine interaction, surveillance, skin detection[17] etc. In recent years, there has been a vital progress on detection schemes [10-11] based on appearance of faces. To build automatic and robust systems in term of speed and accuracy that can be executed on mobile products or in cameras, very efficient and robust face detection algorithms[1] are required. Most of the systems consider face detection as a class problem with the variable having two dimensions that may be either face or non-face. Some parameters like facial appearance, lighting, expressions, and other factors make this two class problem very complex for differentiating face and non-face. Therefore, there is always a need of such classifiers having

good performance characteristics. The most effective method for constructing face and non-face classifiers is learning based approach. For example, neural network-based methods [2], support vector machines [18], etc. In our previous work[20], there has been proposed a very good method using Gabor filter in which 40 Gabor filters are used for a face to be detected and correspondingly feature vector is constructed. The performance is very good in that case but speed is major factor under concern. Although it gave good result in variable illumination, different facial expressions, different skin tone and with different human features, the detector didn't give good result for variable size and rotated face (angle more than 40°). In addition to the good performance of detector, there always remains a need of fast detector which can be used for live image clips in the real world. The Haar-like features[6][22] (as from figure 1),

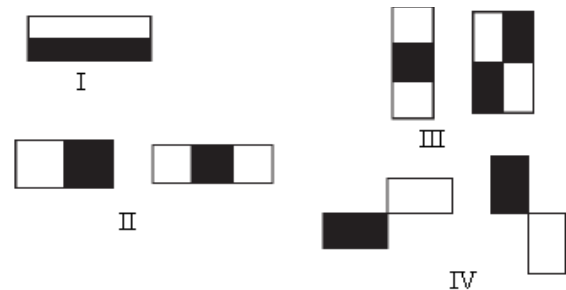


Figure 1. Haar features set.

Encode differences in average intensities between two rectangular regions and these can be calculated rapidly through integral image. Most of the proposed algorithms use pixel values as features. However, they are very sensitive to illumination conditions and noises. Papageorgiou et al. [19] also used Haar-like features. These features are able to extract texture without depending on absolute intensities. Many others [9] proposed variant of Haar features. Viola and Jones proposed [12] an efficient system for evaluating these features which is called an integral image [6] and, they also introduced an efficient scheme for constructing a strong classifier by cascading a small number of distinctive features using Adaboost [3] technique. It consequence comes in terms of increased robustness and higher computational efficiency.

Though Haar-like feature [19] and Gabor techniques as from our previous work[20], provides good performance in extracting textures and cascading architecture, it is still not

suitable for live clips or for real time analysis. Moreover, it is quite hard to detect faces of variable sizes with rotation in angle of face under test.

III. LBP

Local Binary Pattern (LBP)[7] features have performed very well in various applications, including texture classification and segmentation, image retrieval and surface inspection. Here we are using this pattern for face detection[25-26]. The original LBP operator labels the pixels of an image by keeping threshold of the 3x3 neighbourhood of each pixel with the centre pixel value and considering the result as shown in figure 2.

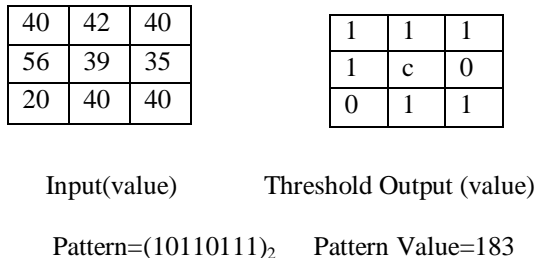


Figure 2. LBP Pattern

The 256-bin histogram of the labels computed over an image can be used as a texture descriptor. Each bin of histogram (LBP code) can be regarded as a micro-texton. Local primitives which are codified by these bins include different types of curved edges, spots, flat areas, etc. The LBP operator has been extended to consider different neighbor sizes. For example, the operator LBP_{4,1} uses 4 neighbors while LBP_{16,2} considers the 16 neighbors on a circle of radius 2 in case of circular local binary operator.

In general, the operator LBP_{P,R} refers to a neighborhood size of P equally spaced pixels on a circle of radius R that form a circularly symmetric neighbor set(as from figure 3). LBP_{P,R} produces 2^P different output values, corresponding to the 2^P different binary patterns that can be formed by the P pixels in the neighbor set. It has been shown that certain bins contain more information than those of others. Therefore, it is possible to use only a subset of the 2^P LBPs to describe the textured images. Ojala et al.[15] defined these fundamental patterns as those with a small number of bitwise transitions from 0 to 1 and vice versa. For example, 00000000 and 11111111 contain 0 transition while 00000110 and 01111110 contain 2 transitions and so on. Accumulating the patterns which have more than 2 transitions into a single bin yields an LBP descriptor.

It is observed that uniform patterns account for nearly 90% of all patterns in the (8,1) neighbourhood (as from figure 3) and for about 70% in the (16,2) neighbourhood in texture images. Hence, accumulating the patterns which have more than 2 transitions into a single bin yields an LBP operator, denoted as LBP_{P,R}^{U2} with less than 2^P bins. For example, the number of labels for a neighbourhood of 8 pixels is 256 for the standard LBP but 59 for LBP^{U2}. After labelling an image with the LBP

operator, a histogram of the labelled image f₁ (x,y) can be defined as follows:

$$H_i = \sum_{x,y} I(f_i(x,y) = i), i = 0 \dots n - 1 \quad (1)$$

where n is the number of different labels produced by the LBP operator and

$$I(X) = \begin{cases} 1 & A \text{ is true} \\ 0 & A \text{ is false} \end{cases} \quad (2)$$

This LBP histogram contains information about the distribution of the local micro-patterns, such as edges, spots and flat areas, over the whole image, so can be used to statistically describe image characteristics.

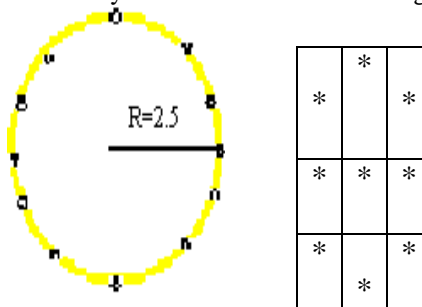


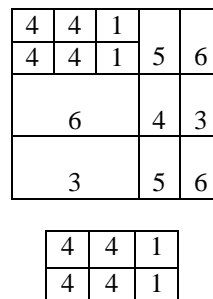
Figure. 3 local binary operator with P,R respectively 12,2.5 and 8,1

The most important out of all considerable properties of LBP features [23-24] is their tolerance against monotonic illumination changes and their computational simplicity at the same time.

IV. MBLBP

Traditional Haar-like rectangle features measure the difference between the average intensities of rectangular regions[8] (Figure 1).

For example, the value of a two-rectangle filter is the difference between the sums of the pixels within two rectangular regions. If we change the position, size, shape and arrangement of rectangular regions, the Haar-like features can capture the intensity gradient at different locations, spatial frequencies and directions. Viola and Jones [6] applied three kinds of such features for detecting frontal faces. By using the integral image, any rectangle filter types, at any scale or location, can be evaluated in constant time period. However, the Haar-like features seem too simple and show some limits [19].



Average value 18/6=3

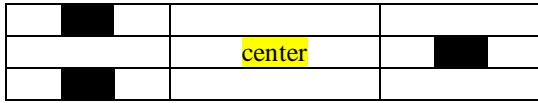


Figure 4. Multi-block LBP feature for image representation.

As shown in the figure 4, the MBLBP features encode rectangular region's intensities by local binary pattern. The resulting binary patterns describe diverse image structures. When compared with the original Local Binary Pattern calculated in a local 3x3 neighbourhood between pixels, we see that MBLBP can capture large scale structure.

V. BOOSTING

Boosting algorithm are used to reduce the number of distinctive rectangle features (to make the size of feature vector smaller) and construct a powerful classifier. It speeds up the computations thereby increases the speed of the detector to find the faces for the given image. But Haar-like rectangle features seem too simple, and the detector often contains thousands of rectangle features for considerable performance. The size of feature vector decides the computation costs both in training as well as test phase. More is the size of feature vector lesser is the speed of the detector. At the resulting stages, weak classifiers based on these features become too weak to improve the classifier's performance. For that, many other methods have been proposed like rotated Haar-like features [19], census transform [5][16], sparse features [4-5], etc. In this paper, we present a new distinctive feature, called Circular-Multi-block Local Binary Pattern (CMBLBP) feature, to represent facial image. The basic idea of CMB-LBP is to encode rectangular regions in the circle by local binary pattern operator [7]. Since it is circular and multi block, our method captures much more area than that by other methods proposed in [12][20]. Original Local Binary Pattern is calculated for 3x3 neighbourhoods, so the area captured by LBP in that case is small. On the other hand, the MB-LBP features can capture large scale structure that may include the dominant features of image structures. The problem with the value of LBP is that its output is just a symbol for representing the binary string. For this non-metric feature value, multi-branch operations are done to calculate weak classifiers. We implement Gentle Adaboost for feature selection and classifier construction. Then a cascade detector is built. Another advantage of CMBLBP is that the number of exhaustive set of CMBLBP features is much smaller. Boosting-based method use Adaboost algorithm to select a significant feature set from the large complete feature set and to enhance its speed, fast Adaboost is further used[13]. We here use gentle boost as a variant of Adaboost. The small feature set of CMBLBP can make this procedure even more simple.

VI. CMBLBP

In this paper, we propose a new distinctive rectangle feature from circular region, called CMBLBP. The basic idea behind that is simple difference rule as in Haar[14]. It is changed into encoding circular region from blocks of a rectangle by binary operator. In this method the value of central pixel is subtracted from the value of neighbour pixels. Then, the information is presented without any loss as a joint distribution(as from

equation 3) of the value of central pixel and the differences.

$$T = t(g_c, g_o - g_c, \dots, g_{p-1} - g_c) \quad (3)$$

Assuming the difference independent of central pixel value, the distribution can be factorized as:

$$T = t(g_c)t(g_o - g_c, \dots, g_{p-1} - g_c) \quad (4)$$

Since $T = t(g_c)$ described the overall luminance of an image, which is not related to local image.

$$\text{So, } T = t(g_o - g_c, \dots, g_{p-1} - g_c) \quad (5)$$

Although invariant against gray scale shifts, the differences are effected by scaling only. To achieve invariance with respect to any transformation of the gray scale, only the sign of difference is considered that is,

$$T \approx t(s(g_o - g_c), \dots, s(g_{p-1} - g_c)) \quad (6)$$

Here

$$s(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (7)$$

After binomial weight assignment (2^p) to each sign that is $s(g_p - g_c)$, transforming the difference in a neighborhood into a unique LBP code for given block of pattern. The code characterization of image $I(x,y)$

$$\text{LBP (for a given block of pattern)}_{P,R}(x_c, y_c) = \sum_{p=1}^{P-1} s(g_p - g_c) 2^p \quad (8)$$

The local gray scale distribution can thus be approximately described with 2^p bin discrete distribution of LBP codes:

$$T \approx t(\text{LBP}(\text{given block of pattern})_{P,R}(x_c, y_c)) \quad (9)$$

In the calculation of the feature vector for given image ($M \times N$), only the central part is considered, as large neighborhood cannot be used on the borders. The distribution of the codes is used as feature vector, denoted by s .

$$T = t(\text{LBP}(\text{given block of pattern})_{P,R}(x, y)) \quad (10)$$

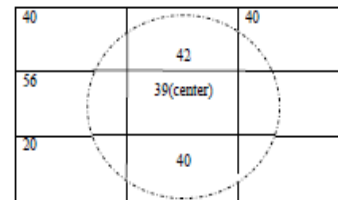
$$x \in \{[R] \dots \dots N - 1 - [R]\}, Y \in \{[R] \dots \dots M - 1 - [R]\}$$

$$\text{Thus CMBLBP (as from figure 5):} \quad \sum_1^n s(T_i - T_c) 2^n \quad (11)$$

Where T_c is the average intensity of the central pixel, $T_i=(1, \dots, n)$, are those in the neighborhood rectangles, n is number of blocks, and the output is computed as

$$s(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (12)$$

The detail description of such CMBLBP operator



(a)

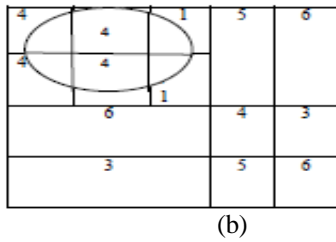


Figure 5. CMBLBP Feature selection

can be found in the above figure 5. In this method, the scaling problem of circular LBP is reduced by the property of multi block, so in our system the detector can detect diverse image structure (as edge, lines, spots, flats, corner areas) also at different scale and location with varying image size (better than [20]) and varying illumination.

Comparing to local binary pattern and MB pattern alone, our system captures larger area with greater accuracy. Collectively we get 256 bin patterns.

Actually in net we are getting LBP at first stage, in second stage we calculate integral histogram $(I_K^H)^n$ of the given test. The labels(n) depends on LBP operator used if $n=4$, so $2^4=16$, integral histograms will be created. In the final stage integral histogram enables us to calculate the feature vector as CHLBP or CHMBLBP. So the CHLBP features are the binary features as normal Haar features. Region separation is defined as R^+ if q pixels of region R^+ have more label compared to b pixels of region R^- .

A. Algorithm

- Take input image as $I(x,y)$
- For given block b_n
- Compute LBP operator O_p
- Compute LBP for given block of pattern $P_{P,R}(x_c, y_c) = \sum_{p=0}^{P-1} s(g_p - g_c) 2^p$
- Calculate $\sum_{i=1}^n s(T_i - T_c) 2^n$

Where T_i, T_c are values of i^{th} block and central block, n is number of blocks.

- Compute $(I_K^H)^n$

For given n (label), by $(I_K^H)^n$, where n depends on $\rightarrow LBP_{P,R}^u$

- Compute IH(the number of integral Haar histograms)
- Compute Feature vector from $s(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases}$

Perform classification

- Start with weight W_i
- Repeat for WL(weak learner)=1.....M

(a) Calculate least square error

(b) update additive model $F(x) \leftarrow F(x) + f_m(x)$

Here $f_m(x)$ is weak learner, $F(x)$ is strong learner.

(c) update $W_i \leftarrow W_i e^{-\gamma_i f_m(x)}$

(d) Normalization

- Output of the classifier $F(x) = \text{Sign}[\sum_{m=1}^M f_m(x)]$

In each step the weak classifier $f_m(x)$ are chosen so as to minimize the weighted squared error.

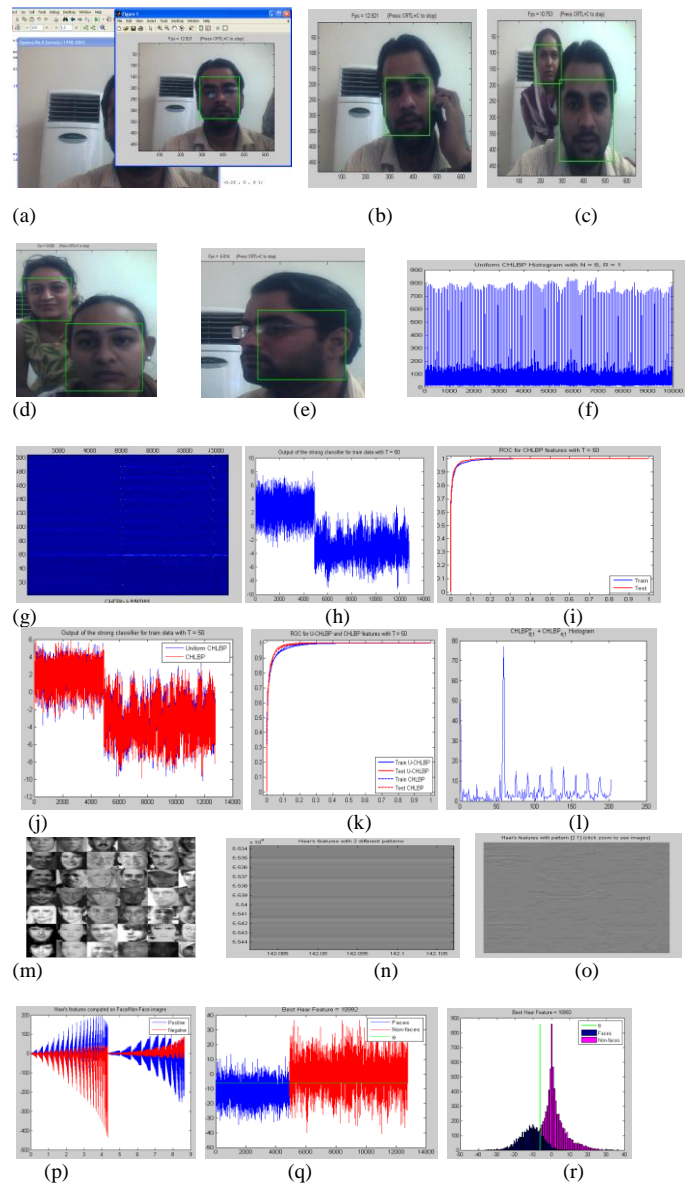
$$J_{wse} = \sum_{i=1}^N W_i (Y_i - f_m(x_i))^2$$

Our purpose of using gentle boosting is to minimize error as cost function using adaptive Newton step function as;

$$J = E[e^{-\gamma_i F(x)}]$$

VII. DATABASE, EXPERIMENTS AND RESULTS

In this section, we trained our system using two gray scale pixels frontal face database, Viola and Jones and Ole Jensen dataset. It consist of 4,916 images and Ole Jensen consist of 5,000 image. The negative sample was chosen from 10,000 images. LBP histogram features and are used for circular Haar like local binary operator as shown in the figure 6 below for full face description. of size 12x12 pixels and shifting of two pixels is used as scanning window.



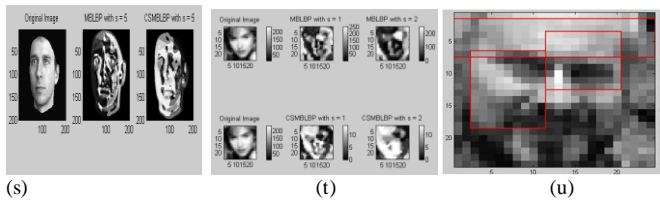


Figure 6. (a-b) test on real world,(c-d) variable face size with different illumination,(e) at angle and with spectacles,(f-g)CHLBP features and histogram (h)output of strong classifier,(i)ROC for uniform CHLBP,(j)strong classifier output(k)Uniform CHLBP and CHLBP (l) $LBP_{p,R}^u$ for $p=4$ and 8 with $R=1$,(m)database,(n-r)Haar features(s-t)MBLBP & CSMBLBP,(u) Three Haar features

TABLE I. COMPARATIVE RESULTS

False Alarm	6	10	21	400	Net Accuracy
our	90%	92%	93%	98%	98%
Using MB alone[21]	80.1%	-	85.6%	-	92%
Viola and jones[12]	-	78.3%	-	93.7%	88%

Figure 6 above gives output from our detector. Part(a-b) shows performance on real world life clips.

The processing time of our detector for a 320x240 image is far less than 0.1s. The compared results are as shown in tabular form in the